MOD-ϕ CONVERGENCE: APPROXIMATION OF DISCRETE MEASURES AND HARMONIC ANALYSIS ON THE TORUS

REDA CHHAIBI, FREDDY DELBAEN, PIERRE-LOÏC MÉLIOT, AND ASHKAN NIKEGHBALI

Abstract. In this paper, we relate the framework of mod-ϕ convergence to the construction of approximation schemes for lattice-distributed random variables. The point of view taken here is that of Fourier analysis in the Wiener algebra, allowing the computation of asymptotic equivalents in the local, Kolmogorov and total variation distances. By using signed measures instead of probability measures, we are able to construct better approximations of discrete lattice distributions than the standard Poisson approximation. This theory applies to various examples arising from combinatorics and number theory: number of cycles in (possibly coloured) permutations, number of prime divisors (possibly within different residue classes) of a random integer, number of irreducible factors of a random polynomial, *etc.* One advantage of the approach developed in this paper is that it allows us to deal with approximations in higher dimensions as well. In this setting, we can explicitly see the influence of the correlations between the components of the random vectors in our asymptotic formulas.

Contents

Key words and phrases. Mod-ϕ convergence, Wiener algebra, Lattice distributions, Approximation of random variables.

1. Introduction

1.1. **Poisson approximation of lattice-valued random variables.** Consider a sequence $(B_i)_{i\geq 1}$ of independent Bernoulli random variables, with $\mathbb{P}[B_i = 1] = p_i$ and $\mathbb{P}[B_i = 0] = 1 - p_i$. We set

$$X_n = \sum_{i=1}^{n} B_i.$$

The *Poisson approximation* ensures that if the p_i's are small but their sum $\lambda_n = \sum_{i=1}^n p_i$ is large, then the distribution of X_n is close to the distribution of a Poisson random variable with parameter $\lambda_n = \sum_{i=1}^n p_i$. A first quantitative result in this direction is due to Prohorov and Kerstan (see [Pro53, Ker64]): if $p_i = \frac{\lambda}{n}$ for all $i \in [\![1, n]\!]$, then

$$d_{\mathrm{TV}}(X_n, \mathcal{P}(\lambda)) := \sum_{k\in\mathbb{N}} \left| \mathbb{P}[X_n = k] - \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \right| \leq \frac{2\lambda}{n}.$$

More generally, with parameters p_i that can be distinct and with $\lambda_n = \sum_{i=1}^n p_i$, Le Cam showed that

$$\sum_{k\in\mathbb{N}} \left| \mathbb{P}[X_n = k] - \mathrm{e}^{-\lambda_n} \frac{(\lambda_n)^k}{k!} \right| \leq 2 \sum_{i=1}^{n} (p_i)^2.$$

This is an immediate consequence of the inequality on total variation distances

$$d_{\mathrm{TV}}(\mu_1 * \mu_2, \nu_1 * \nu_2) \leq d_{\mathrm{TV}}(\mu_1, \nu_1) + d_{\mathrm{TV}}(\mu_2, \nu_2)$$

which holds for any probability measures $\mu_1, \mu_2, \nu_1, \nu_2$ on $\mathbb{Z}$ (*cf.* [Cam60]). By using arguments derived from Stein's method for the Gaussian approximation, Chen and Steele obtained improved versions of this inequality, *e.g.*,

$$\sum_{k\in\mathbb{N}} \left| \mathbb{P}[X_n = k] - \mathrm{e}^{-\lambda_n} \frac{(\lambda_n)^k}{k!} \right| \leq 2\,(1 - \mathrm{e}^{-\sum_{i=1}^n p_i}) \frac{\sum_{i=1}^n (p_i)^2}{\sum_{i=1}^n p_i},$$

see [Che74, Che75, Ste94]. We also refer to [AGG89] for an extension to possibly dependent Bernoulli random variables, and to [BHJ92] for a survey of the theory of Poisson approximations.

More generally, consider random variables $X_{n\geq 1}$ whose distributions are supported by the lattice $\mathbb{Z}^d \subset \mathbb{R}^d$, and which stem from a common probabilistic model. In many cases, the scaling properties of the model imply that when n is large, X_n can be approximated by a discrete infinitely divisible law ν_n, the Lévy exponents of these reference laws being all proportional:

$$\widehat{\nu_n}(\xi) := \sum_{k\in\mathbb{Z}^d} \nu_n(k)\, \mathrm{e}^{\mathrm{i}\langle \xi \,|\, k\rangle} = \mathrm{e}^{\lambda_n \phi(\xi)}, \quad \lambda_n \to +\infty.$$

To go beyond the classical Poisson approximation, one can try in this setting to write bounds on the total variation distance $d_{\mathrm{TV}}(X_n, \nu_n)$, or on another metric which measures the convergence in law (the local distance, the Kolmogorov distance, the Wasserstein

metric, *etc.*). A convenient framework for this program is the notion of *mod-ϕ convergence* developed by Barbour, Kowalski and Nikeghbali in [BKN09], see also the papers [KN10, DKN15, JKN11, FMN16, FMN17a]. The main idea is that, given a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ with values in a lattice $\mathbb{Z}^d$, and a reference infinitely divisible law ϕ on the same lattice, if one has sufficiently good estimates on the Fourier transform of the law μ_n of X_n and on the ratio

$$\frac{\widehat{\mu}_n(\xi)}{\widehat{\nu}_n(\xi)} = \mathbb{E}[\mathrm{e}^{\mathrm{i}\xi X_n}]\,\mathrm{e}^{-\lambda_n\phi(\xi)},$$

then one can deduce from these estimates the asymptotics of the distribution of X_n: central limit theorems [BKN09, JKN11], local limit theorems [KN10, DKN15], large deviations [FMN16], speed of convergence [FMN17a], *etc.* In this paper, we shall use the framework of mod-ϕ convergence to compute the precise asymptotics of the distance between the law of X_n and the reference infinitely divisible law, for various distances.

When approximating lattice-distributed random variables, another objective that one can pursue consists in finding better approximations than the one given by an infinitely divisible law. A general principle is that, if one allows signed measures instead of positive probability measures, then simple deformations of the infinitely divisible reference law can be used to get smaller distances, *i.e.* faster convergences. In the setting of the classical Poisson approximation of sums of independent integer-valued random variables, this idea was used in [BČ02, Theorem 5.1]. We also refer to [Pre83, Kru86, BP96, Ček97, Ček98, ČM99, BX99, JKK08] for other applications of the signed compound Poisson approximation (SCP). For mod-Poisson convergent random variables $(X_n)_{n\in\mathbb{N}}$, an unconditional upper bound on the distance between the law of X_n and a signed measure ν_n defined by means of Poisson–Charlier polynomials was proven in [BKN09, Theorem 3.1]. In this paper, we shall set up a general approximation scheme for sequences of lattice-valued random variables, which will allow us:

- to obtain signed measure approximations whose distances to the laws of the variables X_n are arbitrary negative powers of the parameter λ_n;
- to compute the asymptotics of these distances (instead of an unconditional upper bound).

Thus, this paper can be regarded as a complement to [BKN09], with an alternative approach. In [FMN16, Proposition 4.1.1] and [FMN17b], we also used signed measures in order to approximate mod-ϕ convergent random variables, but with respect to a continuous infinitely divisible distribution ϕ, and using a first-order deformation of the Gaussian distribution.

In the remainder of this introductory section, we recall the main definitions from the theory of mod-ϕ convergence, and we explain the general approximation scheme. In subsequent sections, this approximation scheme will yield the main hypotheses of our Theorems (Sections 2 and 3), and we shall apply it to various one-dimensional and multidimensional examples coming from probability, analytic number theory, combinatorics, *etc.* (Sections 4 and 5). We also present in this introduction the main tool that we shall use in this paper, namely, harmonic analysis in the Wiener algebra.

1.2. **Infinitely divisible distributions and distances between probability measures.**

Let us start by presenting the reference infinitely divisible laws and the distances

between probability distributions that we shall work with. Fix a dimension $d \geq 1$. If X is a random variable with values in $\mathbb{Z}^d$, we denote μ_X its probability law

$$\mu_X(k_1, k_2, \dots, k_d) = \mathbb{P}[X = (k_1, k_2, \dots, k_d)],$$

and $\widehat{\mu}_X$ its Fourier transform, which is defined on the torus $\mathbb{T}^d = (\mathbb{R}/2\pi\mathbb{Z})^d$:

$$\widehat{\mu}_X(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\langle \xi \,|\, X\rangle}] = \sum_{k_1,\dots,k_d \in \mathbb{Z}} \mu_X(k_1, \dots, k_d)\ \exp\left(\mathrm{i}\sum_{j=1}^{d} k_j \xi_j\right).$$

Assume that the law of X is infinitely divisible. Then the Fourier transform of X writes uniquely as

$$\widehat{\mu}_X(\xi) = \mathrm{e}^{\phi(\xi)},$$

where ϕ is a function that is $(2\pi\mathbb{Z})^d$-periodic:

$$\forall (n_1, \dots, n_d) \in \mathbb{Z}^d,\ \ \phi(\xi_1 + 2\pi n_1, \dots, \xi_d + 2\pi n_d) = \phi(\xi_1, \dots, \xi_d).$$

Moreover, the law μ_X is supported on $\mathbb{Z}^d$ and none of its sublattices if and only if the Lévy–Khintchine exponent ϕ is periodic with respect to $(2\pi\mathbb{Z})^d$ and none of its sublattices. We refer to [Sat99] and [SH04, Chapter 2] for details on lattice-distributed infinitely divisible laws (in the case $d = 1$); see also the discussion of [FMN16, §3.1]. Throughout this paper, we make the following assumptions on a reference infinitely divisible law μ_X:

- μ_X has moments of order 2;
- and μ_X is not supported on any sublattice of $\mathbb{Z}^d$.

Then, the corresponding Lévy–Khintchine exponent ϕ is twice continuously differentiable and one has the Taylor expansion around zero

$$\phi(\xi) = \mathrm{i}\,\langle m \,|\, \xi\rangle - \frac{\xi^t \Sigma \xi}{2} + o(|\xi|^2),$$

where $m = \mathbb{E}[X]$, ξ^t is the transpose of the column vector ξ, and Σ is the covariance matrix of X, which is non-degenerate. Moreover, $\mathrm{Re}(\phi(\xi))$ admits a unique non-degenerate global maximum on $[0, 2\pi]^d$ at $\xi = 0$.

Remark 1.1. Let $(\gamma \in \mathbb{R}^d, A \in \mathrm{M}(d \times d, \mathbb{R}), \Pi)$ be the triplet of the Lévy–Khintchine representation of the Fourier transform of an infinitely divisible random variable X (*cf.* [Sat99, Theorem 8.1]). Then, X is supported on $\mathbb{Z}^d$ and has a moment of order 2 if and only if:

(1) $A = 0$ and $\gamma \in \mathbb{Z}^d$;
(2) the Lévy measure Π is supported on $\mathbb{Z}^d$ and has a second moment.

Example 1.2. With $d = 1$, suppose that X follows a Poisson distribution $\mathcal{P}_{(\lambda)}$ with parameter λ. In this case,

$$\phi(\xi) = \lambda\left(\mathrm{e}^{\mathrm{i}\xi} - 1\right) = \mathrm{i}\lambda\xi - \lambda\frac{\xi^2}{2} + o(\xi^2).$$

Example 1.3. More generally, fix a random variable Z with values in $\mathbb{Z}$, and consider the compound Poisson distribution $X = \sum_{i=1}^{\mathcal{P}_{(\lambda)}} Z_i$, where the Z_i's are independent copies of Z, and $\mathcal{P}_{(\lambda)}$ is an independent Poisson variable. One obtains a new infinitely divisible law with values on the lattice $\mathbb{Z}$, and the corresponding Lévy–Khintchine exponent is

$$\phi(\xi) = \lambda\left(\mathbb{E}[\mathrm{e}^{\mathrm{i}\xi Z}] - 1\right) = \mathrm{i}\lambda\,\mathbb{E}[Z]\,\xi - \lambda\,\mathbb{E}[Z^2]\,\frac{\xi^2}{2} + o(\xi^2).$$

Example 1.4. The previous example is generic if one restricts oneself to infinitely divisible random variables with values in $\mathbb{N} = \{0, 1, 2, 3, \dots\}$; see [Kat67, Sam75]. Thus, a random variable X with values in $\mathbb{N}$ is infinitely divisible if and only it admits a representation

$$X = \sum_{i=1}^{\mathcal{P}_{(\lambda)}} Z_i$$

as a compound Poisson distribution. In this representation, the Z_i's are independent copies of a random variable Z with law

$$\mathbb{P}[Z = j] = \frac{\lambda_j}{\lambda} \quad \text{for all } j > 0;$$

and $\mathcal{P}_{(\lambda)}$ is an independent Poisson variable with parameter $\lambda = \sum_{j=1}^{\infty} \lambda_j$. Then, another representation of X is $X = \sum_{j=1}^{\infty} j\, U_j$, where the U_j's are independent Poisson variables with parameters λ_j. These parameters λ_j are the solutions of the system of equations

$$k\, \mathbb{P}[X = k] = \sum_{j=1}^{k} \lambda_j\, j\, \mathbb{P}[X = k - j],$$

and this system provides a numerical criterion of infinite divisibility: X is infinitely divisible if and only if the solutions λ_j are all non-negative.

Given a random variable X on $\mathbb{Z}^d$, the general question that we want to tackle is: how close is the law $\mu = \mu_X$ of X to an infinitely divisible law ν with exponent ϕ? For that purpose, one can choose different distances between probability measures, the typical ones being:

- the *local distance*:
$$d_{\mathrm{L}}(\mu, \nu) := \sup_{k \in \mathbb{Z}^d} |\mu(\{k\}) - \nu(\{k\})|\,.$$
- the *total variation distance*:
$$d_{\mathrm{TV}}(\mu, \nu) := 2 \sup_{A \subset \mathbb{Z}^d} |\mu(A) - \nu(A)| = \sum_{k \in \mathbb{Z}^d} |\mu(\{k\}) - \nu(\{k\})|\,.$$
- and in dimension 1, the *Kolmogorov distance*:
$$d_{\mathrm{K}}(\mu, \nu) := \sup_{k \in \mathbb{Z}} |\mu([\![-\infty, k]\!]) - \nu([\![-\infty, k]\!])| = \sup_{k \in \mathbb{Z}} |\mu([\![k, +\infty]\!]) - \nu([\![k, +\infty]\!])|\,,$$
where $[\![a, b]\!] = \{a, a+1, a+2, \dots, b\}$ denotes an integer interval.

Note that we multiplied the total variation distance by 2 in comparison to the standard definition. It is well known and immediate to check that

$$d_{\mathrm{L}}(\mu, \nu) \le 2\, d_{\mathrm{K}}(\mu, \nu) \le d_{\mathrm{TV}}(\mu, \nu)$$

for any pair of signed measures (μ, ν). The hardest distance to estimate is usually the total variation distance. Note that all the previous quantities metrize the convergence of signed measures on $\mathbb{Z}^d$ with respect to the weak or strong topology. As recalled before in the Poisson case, many approaches for the estimation of d_{L}, d_{K} and d_{TV} can be found in the literature, the most popular ones being Stein's method, coupling methods and semi-group methods, see [DP86, Ste94]. The use of characteristic functions has long been considered not so effective, until the remarkable series of papers of Hwang [Hwa96, Hwa98, Hwa99] (see also [FS09, Section IX.2]). In [Hwa99], Hwang explained how an effective Poisson

approximation can be achieved assuming the analyticity of characteristic functions. In a similar spirit, but with much weaker hypotheses, [BKN09] explored this question by using the notion of mod-ϕ convergence. Our paper revisits this notion, with the purpose of showing that the phenomenons at stake find their source in the harmonic analysis of the torus. While [BKN09] was devoted to the proof of *unconditional* upper bounds on the distances d_{L}, d_{K}, d_{TV}, here we shall be interested in the *asymptotics* of these distances.

1.3. **Mod-ϕ convergent sequences of random variables.** We consider an infinitely divisible law ν of exponent ϕ on $\mathbb{Z}^d$, and a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}^d$. Following [JKN11, DKN15, FMN16], we say that $(X_n)_{n\in\mathbb{N}}$ converges mod-ϕ with parameters $\lambda_n \to +\infty$ and limiting function ψ if

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle \xi \,|\, X_n\rangle}\right] = \mathrm{e}^{\lambda_n \phi(\xi)}\, \psi_n(\xi)$$

with

$$\lim_{n\to\infty} \psi_n(\xi) = \psi(\xi).$$

The precise hypotheses on the convergence $\psi_n \to \psi$ will be made explicit when needed; typically, we shall assume it to occur in some space $\mathscr{C}^r(\mathbb{T}^d)$ endowed with the norm

$$\|f\|_{\mathscr{C}^r} = \sup_{|\alpha|\leq r} \sup_{\xi\in\mathbb{T}^d} |(\partial^\alpha f)(\xi)|.$$

Let us provide a few examples which should enhance the understanding of this notion, as well as show its large scope of applications.

Example 1.5. Consider as in Section 1.1 a sum of independent random variables

$$X_n = \sum_{i=1}^{n} B_i,$$

with B_i following a law $\mathcal{B}(p_i)$: $\mathbb{P}[B_i = 1] = 1 - \mathbb{P}[B_i = 0] = p_i$. Let us assume that $\sum_{i=1}^{\infty} p_i = +\infty$ and $\sum_{i=1}^{\infty} (p_i)^2 < +\infty$. Then, setting $\lambda_n = \sum_{i=1}^{n} p_i$ and $\mu_n = \mu_{X_n}$,

$$\begin{aligned}\widehat{\mu}_n(\xi) &= \prod_{i=1}^{n} \left(1 + p_i(\mathrm{e}^{\mathrm{i}\xi} - 1)\right) = \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)} \prod_{i=1}^{n} \left(1 + p_i(\mathrm{e}^{\mathrm{i}\xi} - 1)\right) \mathrm{e}^{-p_i(\mathrm{e}^{\mathrm{i}\xi}-1)} \\ &= \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)} \prod_{i=1}^{n} \left(1 - \frac{(p_i(\mathrm{e}^{\mathrm{i}\xi} - 1))^2}{2}\,(1 + o(1))\right) = \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)}\, \psi_n(\xi)\end{aligned}$$

with $\psi_n(\xi) \to \psi(\xi) = \prod_{i=1}^{\infty} \left(1 + p_i(\mathrm{e}^{\mathrm{i}\xi} - 1)\right) \mathrm{e}^{-p_i(\mathrm{e}^{\mathrm{i}\xi}-1)}$. This infinite product converges uniformly on the circle because of the hypothesis $\sum_{i=1}^{\infty} (p_i)^2 < \infty$. So, one has mod-Poisson convergence with parameters λ_n.

The following two examples will later be generalized to the multidimensional setting.

Example 1.6. If σ is a permutation of the integers in $[\![1, n]\!]$, denote $\ell(\sigma)$ its number of disjoint cycles, including the fixed points. We then set $\ell_n = \ell(\sigma_n)$, where σ_n is taken at random uniformly among the $n!$ permutations of $\mathfrak{S}_n$. Using Feller's coupling (*cf.* [ABT03]), one can show that ℓ_n admits the following representation in law:

$$\ell_n = \sum_{i=1}^{n} \mathcal{B}\left(\frac{1}{i}\right),$$

where the Bernoulli random variables are independent. This representation will also be made clear by the discussion of §5.3 in the present paper. By the discussion of the

previous example, $(\ell_n)_{n\in\mathbb{N}}$ converges mod-Poisson with parameters $H_n = \sum_{i=1}^n \frac{1}{i}$ and limiting function

$$\prod_{i=1}^{\infty} \left(1 + \frac{\mathrm{e}^{\mathrm{i}\xi} - 1}{i}\right) \mathrm{e}^{-\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{i}} = \frac{1}{\Gamma(\mathrm{e}^{\mathrm{i}\xi})}\, \mathrm{e}^{\gamma\,(\mathrm{e}^{\mathrm{i}\xi}-1)},$$

where γ is the Euler–Mascheroni constant $\gamma = \lim_{n\to\infty}(H_n - \log n)$ (see [Art64] for this infinite product representation of the Γ-function, due to Weierstrass). Thus, one can also say that $(\ell_n)_{n\in\mathbb{N}}$ converges mod-Poisson with parameters $\lambda_n = \log n$ and limiting function

$$\psi(\xi) = \frac{1}{\Gamma(\mathrm{e}^{\mathrm{i}\xi})}.$$

Example 1.7. As pointed out in [KN10, JKN11], mod-ϕ convergence is a common phenomenon in probabilistic number theory. For instance, denote $\omega(k)$ the number of distinct prime divisors of an integer $k \geq 1$, and $\omega(N_n)$ the number of distinct prime divisors of a random integer N_n smaller than n, taken according to the uniform law. Then, it can be shown by using the Selberg–Delange method (*cf.* [Ten95, §II.5]) that

$$\mathbb{E}[\mathrm{e}^{z\omega(N_n)}] = \frac{1}{n}\sum_{k=1}^{n} \mathrm{e}^{z\omega(k)} = \mathrm{e}^{(\log\log n)(\mathrm{e}^z-1)} \left(\Psi(\mathrm{e}^z) + O\left(\frac{1}{\log n}\right)\right)$$

with

$$\Psi(z) = \frac{1}{\Gamma(z)} \prod_{p \text{ prime}} \left(1 + \frac{z-1}{p}\right) \mathrm{e}^{-\frac{z-1}{p}},$$

and where the remainder $O(\frac{1}{\log n})$ is uniform for z in a compact subset of $\mathbb{C}$. Therefore, one has mod-Poisson convergence at speed $\lambda_n = \log\log n$, and with limiting function $\psi(\xi) = \Psi(\mathrm{e}^{\mathrm{i}\xi})$. This limiting function involves two factors: the limiting function $\frac{1}{\Gamma(\mathrm{e}^{\mathrm{i}\xi})}$ of the number of cycles of a random permutation ("geometric" factor), and an additional "arithmetic" factor $\prod_{p\in\mathbb{P}}(1+\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{p})\exp(-\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{p})$. This apparition of two limiting factors is a common phenomenon in number theory [KN10, JKN11].

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables that converges mod-ϕ with parameters λ_n. Informally, $\psi_n(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\langle\xi\,|\,X_n\rangle}]\,\mathrm{e}^{-\lambda_n\phi(\xi)}$ measures a deconvolution residue, and mod-ϕ convergence means that this residue stabilizes, allowing the computation of equivalents of $d(\mu_n, \nu_n)$, where

$$\begin{aligned}\mu_n &= \text{law of } X_n;\\ \nu_n &= \text{law of the infinitely divisible law with exponent } \lambda_n\phi\end{aligned}$$

and d is one of the distances introduced in the previous paragraph. This setting will be called the *basic approximation scheme for a mod-ϕ convergent sequence.*

1.4. **The Wiener algebra and the general scheme of approximation.** As explained before, we shall be interested in more general approximation schemes, with signed measures ν_n that are closer to μ_n than the basic scheme. When $d = 1$, the apparition of signed measures is natural in the setting of the Wiener algebra of absolutely convergent Fourier series (see [Kah70, Kat04]):

Definition 1.8. *The Wiener algebra $\mathscr{A} = \mathscr{A}(\mathbb{T})$ is the algebra of continuous functions on the circle whose Fourier series converges absolutely. It is a Banach algebra for the pointwise product and the norm*

$$\|f\|_{\mathscr{A}} := \sum_{n\in\mathbb{Z}} |c_n(f)|,$$

where $c_n(f)$ denotes the Fourier coefficient $\int_0^{2\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}n\theta}\,\frac{d\theta}{2\pi}$.

Note that the characteristic function of any signed measure μ on $\mathbb{Z}$ belongs to the Wiener algebra, with

$$\|\widehat{\mu}\|_{\mathscr{A}} = \sum_{n\in\mathbb{Z}} |c_n(\widehat{\mu})| = \sum_{n\in\mathbb{Z}} |\mu(n)| = \|\mu\|_{\mathrm{TV}}$$

equal to the total variation norm of the measure. Thus, $\mathscr{A}$ is the right functional space in order to use harmonic analysis tools when dealing with (signed) measures. To the best of our knowledge, this interpretation of the total variation distance as the norm of $\mathscr{A}$ has not been used before. On the other hand, another important property of the Wiener algebra is:

Proposition 1.9 (Wiener's $\frac{1}{f}$ theorem)**.** *Let $f \in \mathscr{A}$. Then f never vanishes on the circle if and only if $\frac{1}{f} \in \mathscr{A}$.*

We refer to [New75] for a short proof of the Wiener theorem. As a consequence, if X_n is a $\mathbb{Z}$-valued random variable and if ν is an infinitely divisible distribution with exponent ϕ, then the deconvolution residue

$$\psi_n(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\xi X_n}]\,\mathrm{e}^{-\lambda_n\phi(\xi)}$$

always belongs to $\mathscr{A}$, so it is a convergent Fourier series $\psi_n(\xi) = \sum_{k=-\infty}^{\infty} a_{k,n}\,\mathrm{e}^{\mathrm{i}k\xi}$. The idea is then to replace this residue ψ_n by a simpler residue $\chi_n \in \mathscr{A}$, which after reconvolution by $\mathrm{e}^{\lambda_n\phi(\xi)}$ yields a signed measure approximating the law of X_n. We are thus lead to:

Definition 1.10. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables in $\mathbb{Z}^d$ that is mod-ϕ convergent with parameters $(\lambda_n)_{n\in\mathbb{N}}$. A* general approximation scheme *for $(X_n)_{n\in\mathbb{N}}$ is given by a sequence of discrete signed measures $(\nu_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}^d$, such that*

$$\widehat{\mu_n}(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\langle\xi\,|\,X_n\rangle}] = \mathrm{e}^{\lambda_n\phi(\xi)}\,\psi_n(\xi);$$
$$\widehat{\nu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\,\chi_n(\xi),$$

with $\lim_{n\to+\infty}\psi_n(\xi) = \psi(\xi)$ and $\lim_{n\to+\infty}\chi_n(\xi) = \chi(\xi)$. Here, the residues ψ_n, ψ, χ_n and χ are functions on the torus $\mathbb{T}^d = (\mathbb{R}/2\pi\mathbb{Z})^d$ that have absolutely convergent Fourier series, and with $\chi_n(0) = \chi(0) = 1$ (hence, $\nu_n(\mathbb{Z}^d) = 1$). The convergence of the residues $\psi_n \to \psi$ and $\chi_n \to \chi$ is assumed to be at least uniform on the torus, that is in the space of continuous functions $\mathscr{C}^0(\mathbb{T})$.

The basic approximation scheme is the case when $\chi_n(\xi) = \chi(\xi) = 1$. The residues χ_n and χ will typically be trigonometric polynomials, and they will enable us to enhance considerably the quality of our discrete approximations.

Example 1.11. An important case of approximation scheme in the sense of Definition 1.10 is the *approximation scheme of order $r \geq 1$*. Suppose that $d = 1$ and that the functions

$\psi_n(\xi)$ can be represented on the torus as absolutely convergent series

$$\psi_n(\xi) = 1 + \sum_{k=1}^{\infty} b_{k,n}\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^k + \sum_{k=1}^{n} c_{k,n}\,(\mathrm{e}^{-\mathrm{i}\xi} - 1)^k.$$

The coefficients $a_{k,n}$ of $\psi_n(\xi) = \sum_{k=-\infty}^{\infty} a_{k,n}\,\mathrm{e}^{\mathrm{i}k\xi}$ are related to the coefficients $b_{k,n}$ and $c_{k,n}$ by the equations

$$b_{k,n} = \sum_{l\geq k} \binom{l}{k} a_{l,n} \quad ; \quad c_{k,n} = \sum_{l\geq k} \binom{l}{k} a_{-l,n}$$

for all $k \geq 1$. Set then

$$\chi_n^{(r)}(\xi) = P_n^{(r)}(\mathrm{e}^{\mathrm{i}\xi}) = 1 + \sum_{k=1}^{r} b_{k,n}\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^k + \sum_{k=1}^{r} c_{k,n}\,(\mathrm{e}^{-\mathrm{i}\xi} - 1)^k.$$

The $\chi_n^{(r)}(\xi)$ are the Laurent polynomials of degree r that approximate the residues ψ_n around 0 at order $r \geq 1$. Then, if X_n is a random variable with law μ_n with characteristic function $\widehat{\mu_n}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)}\,\psi_n(\xi)$, the approximation scheme of order r of μ_n is given by the signed measures $\nu_n^{(r)}$, with

$$\widehat{\nu_n^{(r)}}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)}\,\chi_n^{(r)}(\xi).$$

We shall prove in Section 3 that $d_{\mathrm{L}}(\mu_n, \nu_n^{(r)})$, $d_{\mathrm{K}}(\mu_n, \nu_n^{(r)})$ and $d_{\mathrm{TV}}(\mu_n, \nu_n^{(r)})$ get smaller when r increases; in particular, $\nu_n^{(r\geq 1)}$ is asymptotically a better approximation of μ_n than the basic scheme $\nu_n^{(0)}$.

One can give a functional interpretation to the approximation schemes of order $r \geq 1$. Denote as before $P_n^{(r)}$ the Laurent approximation of order r of ψ_n, and introduce the shift operator S on functions $f : \mathbb{Z} \to \mathbb{C}$, defined by

$$(Sf)(k) = f(k+1).$$

If $\nu_n^{(r)}$ is the signed measure with Fourier transform $\widehat{\nu_n^{(r)}}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)}\,P_n^{(r)}(\mathrm{e}^{\mathrm{i}\xi})$, then for any square-integrable function $f : \mathbb{Z} \to \mathbb{C}$, if $\widehat{f}(\xi) = \sum_{k\in\mathbb{Z}} f(k)\,\mathrm{e}^{\mathrm{i}k\xi}$, then

$$\begin{aligned}
\nu_n^{(r)}(f) &= \sum_{k\in\mathbb{Z}} \nu_n^{(r)}(\{k\})\,f(k) = \frac{1}{2\pi}\int_0^{2\pi} \widehat{\nu_n^{(r)}}(\xi)\overline{\widehat{f}(\xi)}\,d\xi \\
&= \frac{1}{2\pi}\int_0^{2\pi} \widehat{\nu_n^{(0)}}(\xi)\,P_n^{(r)}(\mathrm{e}^{\mathrm{i}\xi})\overline{\widehat{f}(\xi)}\,d\xi = \frac{1}{2\pi}\int_0^{2\pi} \widehat{\nu_n^{(0)}}(\xi)\,\overline{\widehat{(P_n^{(r)}(S)f)}(\xi)}\,d\xi \\
&= \nu_n^{(0)}(P_n^{(r)}(S)f).
\end{aligned}$$

Moreover, the operator $P_n^{(r)}(S)$ is a linear combination of discrete difference operators:

$$P_n^{(r)}(S) = \mathrm{id} + \sum_{k=1}^{r} b_{k,n}\,(\Delta_+)^k + \sum_{k=1}^{r} c_{k,n}\,(\Delta_-)^k$$

with $(\Delta_+^k(f))(j) = \sum_{l=0}^{k}(-1)^{k-l}\binom{k}{l} f(j+l)$ and $(\Delta_-^k(f))(j) = \sum_{l=0}^{k}(-1)^{k-l}\binom{k}{l} f(j-l)$. Therefore:

Proposition 1.12. *In the previous setting, if $(\nu_n^{(r)})_{n\in\mathbb{N}}$ is the approximation scheme of order $r \geq 1$ of a sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ that is mod-ϕ convergent, then for any square-integrable function f,*

$$\nu_n^{(r)}(f) = \mathbb{E}[(P_n^{(r)}(S)(f))(Y_n)],$$

where Y_n follows an infinitely divisible law with exponent $\lambda_n\phi$. Thus, to estimate at order r the expectation $\mu_n(f) = \mathbb{E}[f(X_n)]$:

(1) one replaces X_n by an infinitely divisible random variable Y_n;

(2) and one also replaces the function f by $P_n^{(r)}(S)(f)$, which is better suited for discrete approximations.

Remark 1.13. Our notion of approximation scheme should be compared to the one of [Hwa99], which is the particular case where the reference infinitely divisible law is Poissonian, and the residues χ_n are constant equal to one. One of the main interest of our approach is that we are able to construct better schemes of approximation, by allowing quite general residues χ_n in the Fourier transform of the laws ν_n that approximate the random variables X_n. On the other hand, we consider only three distances among those studied in [Hwa99], but there should be no difficulty in adapting our results to the other distances, such as the Wasserstein metric and the Hellinger/Matusita metric.

Remark 1.14. The approximation scheme of order $r \geq 1$ can be considered as a discrete analogue of the Edgeworth expansion in the central limit theorem, which is with respect to the Gaussian approximation, and thus gives results for a different scale of fluctuations [Pet75, Chapter VI].

1.5. **Outline of the paper.** Placing ourselves in the setting of a general approximation scheme, one basic idea in order to evaluate the distances between the distributions μ_n and ν_n is to relate them to the distances between the two residues $\psi_n(\xi)$ and $\chi_n(\xi)$ in the Wiener algebra $\mathscr{A}(\mathbb{T}^d)$. In Section 2, we prove various concentration inequalities in this algebra, and we explain how to use them in order to compute distances between distributions. In Section 3, we apply these results to obtain asymptotic estimates for the distances in the general approximation scheme (see our main Theorems 3.3, 3.6 and 3.11). In these two sections, we shall restrict ourselves to the one-dimensional setting, postponing the more involved computations of the higher dimensions $d \geq 2$ to Section 5.

In Section 4, we apply our theorem to various one-dimensional examples of mod-Poisson convergent sequences:

- In Section 4.1, we explain how to use the general theory with the toy-model of sums of independent Bernoulli random variables (classical Poisson approximation). The combinatorics of the approximation schemes of order $r \geq 1$ can be encoded in the algebra of symmetric functions, and we explain this encoding in §4.2, by using the theory of formal alphabets. The formal alphabets provide a simple description of the approximation schemes for all the examples hereafter, although none of them (except the toy-model) come from independent Bernoulli variables.
- In Section 4.3, we study the number of disjoint cycles in a model of random permutations which generalises the uniform and the Ewens measure (Example

1.6). The mod-Poisson convergence of this model was proven in [NZ13], and we compute here the approximation schemes of this model.

- In Section 4.4, we show more generally how to use *generating series with algebraico-logarithmic singularities* to construct general approximation schemes of statistics of random combinatorial objects. We study with this method the number of irreducible factors of a random monic polynomial over the finite field $\mathbb{F}_q$ (counted with or without multiplicity), and the number of connected components of a random functional graph.
- In §4.5, we consider the number of distinct prime divisors of a random integer (Example 1.7). We explain how to use the form of the residue $\psi(\xi)$ of mod-Poisson convergence to construct explicit approximation schemes of the corresponding probability measures, using again the formalism of symmetric functions.

Our approach also allows us to measure the gain of the Poissonian approximation in comparison to the Gaussian approximation, when one has a sequence of random integers that behaves asymptotically like a Poisson random variable with a large parameter λ_n.

In Section 5, we extend our results to the multi-dimensional setting. One of the advantages of the Fourier approach to approximation of probability measures is that it allows one to deal with higher dimensions with the exact same techniques, although the computations are more involved. An important difference with the one-dimensional setting is that the dependence between the coordinates of a mod-ϕ convergent sequence $(X_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}^d$ can be read on the asymptotics of the distances from the approximation schemes to the laws of the random variables. Note that this happens even when the reference infinitely divisible distribution corresponds to independent coordinates. We provide two examples of this phenomenon, stemming respectively from the combinatorics of the wreath products $\mathfrak{S}_n \wr (\mathbb{Z}/d\mathbb{Z})$ and from the number theory of residue classes of prime numbers.

2. Concentration inequalities in the Wiener algebra

For convenience, until Section 5, we shall focus on the one-dimensional case ($d = 1$) and the corresponding torus $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, which we view as the set of complex numbers of modulus 1. Thus, a function on $\mathbb{T}$ will be a function of $e^{i\xi}$ with $\xi \in [0, 2\pi)$. For $p \in [1, +\infty)$, the space of complex-valued functions on $\mathbb{T}$ whose p-powers are Lebesgue integrable will be denoted $\mathscr{L}^p = \mathscr{L}^p(\mathbb{T})$; it is a Banach space for the norm

$$\|f\|_p := \left(\int_0^{2\pi} |f(e^{i\xi})|^p \, \frac{dt}{2\pi}\right)^{\frac{1}{p}}.$$

For $p = +\infty$, $\mathscr{L}^\infty(\mathbb{T})$ is the space of essentially bounded functions on the torus, with Banach space norm

$$\|f\|_\infty := \text{ess-sup}_{\xi\in[0,2\pi]} \left|f(e^{i\xi})\right|.$$

In the sequel, we abbreviate sometimes the Haar integral $\int_0^{2\pi} f(e^{i\xi}) \frac{d\xi}{2\pi}$ by $\int_{\mathbb{T}} f(e^{i\xi})$, or simply $\int_{\mathbb{T}} f$.

2.1. **Deconvolution residues in the Wiener algebra.** Recall that the Wiener algebra $\mathscr{A}(\mathbb{T})$ is the complex algebra of absolutely convergent Fourier series, endowed with the norm $\|f\|_{\mathscr{A}} = \sum_{n\in\mathbb{Z}} |c_n(f)|$. We shall use the following property of $\mathscr{A}(\mathbb{T})$, which is akin to Poincaré's inequality for Sobolev spaces (see [Kat04, §6.2]):

Proposition 2.1. *There is a constant $C_H = \frac{\pi}{\sqrt{3}} = 1.814\ldots$ such that*

$$\|f\|_{\mathscr{A}} \leq |c_0(f)| + C_H \|f'\|_{\mathscr{L}^2}$$

for all $f \in \mathscr{A}$.

Proof. Combining the Cauchy–Schwarz inequality and the Parseval identity, we obtain

$$\begin{aligned}
\|f\|_{\mathscr{A}} &= |c_0(f)| + \sum_{n\neq 0} \frac{1}{n} |n\, c_n(f)| \\
&\leq |c_0(f)| + \sqrt{\sum_{n\neq 0} \frac{1}{n^2}} \sqrt{\sum_{n\neq 0} |n\, c_n(f)|^2} \\
&\leq |c_0(f)| + \sqrt{\frac{\pi^2}{3}}\, \|f'\|_{\mathscr{L}^2}.
\end{aligned}$$

Note that the inequality is sharp, and it implies that the Sobolev space $\mathscr{W}^{1,2}(\mathbb{T})$ of $\mathscr{L}^2$ functions on $\mathbb{T}$ with weak derivative in $\mathscr{L}^2$ is topologically included in the Wiener algebra. □

Another important tool for the computation of total variation distances is the following Lemma 2.2. Call *deconvolution residue* of a signed measure μ on $\mathbb{Z}$ by another signed measure ν on $\mathbb{Z}$ the function

$$\delta(\xi) := \widehat{\mu}(\xi)\, (\widehat{\nu}(\xi))^{-1}.$$

By Wiener's theorem, if $\widehat{\nu}$ never vanishes, which is for instance the case when it is the Fourier transform of an infinitely divisible law ν [Sat99, Lemma 7.5], then $\widehat{\nu}^{-1} \in \mathscr{A}$ and the previous deconvolution happens in the Wiener algebra: $\delta \in \mathscr{A}$. Now, in order to measure the distance between the law μ of X and a reference law ν, one can adopt the following point of view, which is common in signal processing. For the deconvolution residue to be considered as a small noise, δ has to be close to the constant function 1 (the Fourier transform of the Dirac distribution at zero). In particular, suppose that we are given a general scheme of approximation of a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ by a sequence of laws $(\nu_n)_{n\in\mathbb{N}}$. Then, with the notation of Definition 1.10, one has a sequence of deconvolution residues

$$\delta_n(\xi) = \frac{\widehat{\mu_n}(\xi)}{\widehat{\nu_n}(\xi)},$$

and the quality of our scheme of approximation will be related to the speed of convergence of $(\delta_n)_{n\in\mathbb{N}}$ towards the constant function 1. More precisely:

Lemma 2.2 (Fundamental inequality for deconvolution residues)**.** *Let μ and ν be two signed measures on $\mathbb{Z}$ and $\delta = \frac{\widehat{\mu}}{\widehat{\nu}}$ be the deconvolution residue (we assume that $\widehat{\nu}$ does not vanish on $\mathbb{T}$). For any $c \in \mathscr{A}$,*

$$d_{\mathrm{TV}}(\mu,\nu) \leq \|c(\delta-1)\widehat{\nu}\|_{\mathscr{A}} + \|\delta - 1\|_{\mathscr{A}}\, \|(1-c)\widehat{\nu}\|_{\mathscr{A}}\,.$$

Proof.

$$\begin{aligned} d_{\mathrm{TV}}(\mu,\nu) &= \|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}} = \|(\delta-1)\widehat{\nu}\|_{\mathscr{A}} \\ &\leq \|c(\delta-1)\widehat{\nu}\|_{\mathscr{A}} + \|(1-c)(\delta-1)\widehat{\nu}\|_{\mathscr{A}} \\ &\leq \|c(\delta-1)\widehat{\nu}\|_{\mathscr{A}} + \|\delta-1\|_{\mathscr{A}}\,\|(1-c)\widehat{\nu}\|_{\mathscr{A}}\,. \end{aligned}$$

□

In practice, $c \in \mathscr{A}$ will be a carefully chosen cut function concentrated on regions where $\widehat{\nu}$ is large. So, informally, ν is a good approximation of μ if the deconvolution residue δ:

- is close to 1 where $\widehat{\nu}$ is large;
- has a reasonable norm in the Wiener algebra.

Note that the fundamental inequality *does* rely on the algebra norm of $\mathscr{A}(\mathbb{T})$. We now quantify this idea.

2.2. **Estimates of norms in the Wiener algebra.** In this paragraph, we fix two discrete measures μ and ν on $\mathbb{Z}$, both being convolutions of an infinitely divisible law with exponent $\lambda\phi$ by residues ψ and χ:

$$\begin{aligned} \widehat{\mu}(\xi) &= \mathrm{e}^{\lambda\phi(\xi)}\,\psi(\xi); \\ \widehat{\nu}(\xi) &= \mathrm{e}^{\lambda\phi(\xi)}\,\chi(\xi). \end{aligned}$$

In Section 3, we shall recover the setting of a general approximation scheme by adding indices n to this situation. We denote m and σ^2 the first two coefficients of the Taylor expansion of ϕ around 0:

$$\phi(\xi) = m\mathrm{i}\xi - \frac{\sigma^2\xi^2}{2} + o(\xi^2).$$

We then fix an integer $r \geq 0$ such that:

(1) the residues ψ and χ are assumed to be $(r+1)$ times continuously differentiable on $[-\varepsilon,\varepsilon]$ for a certain $\varepsilon > 0$;

(2) their Taylor expansion coincides at 0 up to the r-th order:

$$\forall s \in [\![0,r]\!]\,,\ \ (\psi-\chi)^{(s)}(0) = 0.$$

In this setting, we define the non-negative quantities

$$\begin{aligned} \beta_{r+1}(\varepsilon) &= \sup_{\xi\in[-\varepsilon,\varepsilon]} \left|(\psi-\chi)^{(r+1)}(\xi)\right|; \\ \gamma(\varepsilon) &= \sup_{\xi\in[-\varepsilon,\varepsilon]} \left|\phi''(\xi)+\sigma^2\right|; \\ M &= -\sup_{\xi\in[-\pi,\pi]} \left(\frac{\mathrm{Re}(\phi(\xi))}{\xi^2}\right). \end{aligned}$$

The strict positivity of M as soon as ϕ is not the constant distribution concentrated at 0 is proven in [FMN16, Section 3]. Note that for any $\xi \in [-\varepsilon,\varepsilon]$, one has

$$\begin{aligned} \mathrm{Re}(\phi(\xi)) &= \int_{\theta=0}^{\xi} (\xi-\theta)\,\mathrm{Re}(\phi''(\theta))\,d\theta = -\frac{\sigma^2\xi^2}{2} + \int_{\theta=0}^{\xi} (\xi-\theta)\,\mathrm{Re}(\phi''(\theta)+\sigma^2)\,d\theta \\ &\leq \frac{(\gamma(\varepsilon)-\sigma^2)\,\xi^2}{2}. \end{aligned}$$

On the other hand, outside the interval $[-\varepsilon, \varepsilon]$, one can use the inequality $\mathrm{Re}(\phi(\xi)) \leq -M\xi^2$.

Theorem 2.3 (Norm estimate)**.** *Take ε small enough so that $\gamma(\varepsilon) \leq \frac{\sigma^2}{2}$, and suppose that λ is large enough so that $\lambda \geq \frac{2}{\sigma^2}$. Then:*

$$\|\widehat{\mu} - \widehat{\nu}\|_{\mathscr{A}} \leq \|\psi - \chi\|_{\mathscr{A}} \left(1 + C_H \left(\sqrt{\frac{2}{\pi\varepsilon}} + \lambda\|\phi'\|_\infty\right)\right) \mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}} + \frac{C_{r+1}\,\beta_{r+1}(\varepsilon)\,(\varepsilon^{-1} + \sqrt{5(r+1)})}{(\frac{\sigma^2}{2}\,\lambda)^{\frac{r}{2}+\frac{1}{4}}},$$

where C_{r+1} is a constant depending only on r:

$$C_{r+1} = \frac{1}{(r+1)!}\sqrt{\frac{2\pi}{3}\,\Gamma\left(r + \frac{3}{2}\right)}.$$

In order to prove this theorem, we introduce the cut function $\xi \in \mathbb{T} \mapsto c(\xi)$ that is piecewise linear, vanishes outside of $[-\varepsilon, \varepsilon]$, and is equal to 1 on $[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]$, see Figure 1.

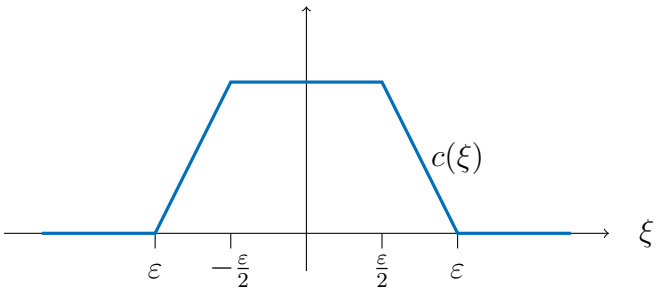

FIGURE 1. The cut-function $c(\xi)$.

Lemma 2.4. *Under the assumptions of Theorem 2.3,*

$$\left\|(1-c)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} \leq \left(1 + C_H\left(\sqrt{\frac{2}{\pi\varepsilon}} + \lambda\|\phi'\|_\infty\right)\right)\mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}}.$$

Proof. Note that $|(1-c(\xi))\,\mathrm{e}^{\lambda\phi(\xi)}|$ vanishes on $[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]$, and is smaller than $\mathrm{e}^{-\lambda M\xi^2} \leq \mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}}$ everywhere else. We then use Proposition 2.1:

$$\begin{aligned}
\left\|(1-c)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} &\leq \int_{\mathbb{T}} |1-c|\,\mathrm{e}^{\lambda\mathrm{Re}(\phi)} + C_H\left\|-c'\mathrm{e}^{\lambda\phi} + (1-c)\,\lambda\phi'\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{L}^2} \\
&\leq \mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}} + C_H\left\|c'\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{L}^2} + C_H\,\lambda\|\phi'\|_\infty\left\|(1-c)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{L}^2} \\
&\leq \mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}} + C_H\,\|c'\|_{\mathscr{L}^2}\,\mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}} + C_H\,\lambda\|\phi'\|_\infty\,\mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}} \\
&\leq \left(1 + C_H\left(\sqrt{\frac{2}{\pi\varepsilon}} + \lambda\|\phi'\|_\infty\right)\right)\mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}}.
\end{aligned}$$

□

Lemma 2.5. *Under the assumptions of Theorem 2.3, we have:*

$$\left\|c\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} \leq \frac{C_{r+1}\,\beta_{r+1}(\varepsilon)\,(\varepsilon^{-1} + \sqrt{5(r+1)})}{(\frac{\sigma^2}{2}\,\lambda)^{\frac{r}{2}+\frac{1}{4}}}.$$

Proof. Note that the norm $\|\cdot\|_{\mathscr{A}}$ is invariant after multiplication by $\mathrm{e}^{\mathrm{i}k\xi}$, $k \in \mathbb{Z}$. In particular,

$$\left\|c\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} = \left\|c(\xi)\,(\psi(\xi)-\chi(\xi))\,\mathrm{e}^{\lambda\phi(\xi)-\mathrm{i}\xi\lfloor\lambda m\rfloor}\right\|_{\mathscr{A}},$$

where $\lfloor\cdot\rfloor$ denotes the integer part of a real number. Using again Proposition 2.1, we obtain

$$\begin{aligned}\left\|c\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} &\le \int_{\mathbb{T}} |c\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}| + C_H\left\|\left(c(\xi)\,(\psi(\xi)-\chi(\xi))\,\mathrm{e}^{\lambda\phi(\xi)-\mathrm{i}\xi\lfloor\lambda m\rfloor}\right)'\right\|_{\mathscr{L}^2}\\ &\le \int_{\mathbb{T}} |c\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}| + C_H\left\|c'\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{L}^2}\\ &\quad + C_H\left\|c\,(\psi-\chi)'\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{L}^2} + C_H\left\|c\,(\psi-\chi)\,(\lambda\phi'-i\lfloor\lambda m\rfloor)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{L}^2}\\ &\le A+B+C+D.\end{aligned}$$

Let us bound separately each term. Recall that $\int_{-\infty}^{\infty}|t|^r\mathrm{e}^{-t^2}\,dt = \int_0^\infty u^{\frac{r-1}{2}}\,\mathrm{e}^{-u}\,du = \Gamma\left(\frac{r+1}{2}\right)$ for $r \ge 0$.

(A) For $\xi \in [-\varepsilon,\varepsilon]$,

$$|\psi(\xi)-\chi(\xi)| \le \frac{\beta_{r+1}(\varepsilon)\,|\xi|^{r+1}}{(r+1)!};$$
$$\mathrm{e}^{\lambda\mathrm{Re}(\phi(\xi))} \le \mathrm{e}^{-\frac{\lambda(\sigma^2-\gamma(\delta))\,\xi^2}{2}} \le \mathrm{e}^{-\frac{\lambda\sigma^2\xi^2}{4}}$$

the last inequality following from the assumptions on ε. Consequently,

$$\begin{aligned}A &\le \frac{\beta_{r+1}(\varepsilon)}{2\pi\,(r+1)!}\int_{-\varepsilon}^{\varepsilon}\mathrm{e}^{-\frac{\lambda\sigma^2\xi^2}{4}}\,|\xi|^{r+1}\,dt = \frac{\beta_{r+1}(\varepsilon)}{2\pi\,(r+1)!\,(\lambda\frac{\sigma^2}{4})^{\frac{r}{2}+1}}\,\Gamma\left(\frac{r}{2}+1\right)\\ &\le \frac{\beta_{r+1}(\varepsilon)}{2\pi\,(r+1)!\,(\lambda\frac{\sigma^2}{4})^{\frac{r}{2}+1}}\,\sqrt{\Gamma\left(\frac{r}{2}+\frac{3}{4}\right)\Gamma\left(\frac{r}{2}+\frac{5}{4}\right)}\\ &\le \frac{\beta_{r+1}(\varepsilon)}{(r+1)!\,(\lambda\frac{\sigma^2}{2})^{\frac{r}{2}+1}}\,\left(\frac{2}{\pi}\right)^{\frac14}\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{3}{2}\right)}.\end{aligned}$$

by using the log-convexity of the Gamma function, and the duplication formula $\Gamma(2z) = \frac{2^{2z-1}}{\sqrt{\pi}}\,\Gamma(z)\,\Gamma(z+\frac12)$.

(B) Since $|c'(\xi)| \le \frac{2}{\varepsilon}$ on $[-\varepsilon,\varepsilon]$ and vanishes outside $[-\varepsilon,\varepsilon]$, one obtains

$$\begin{aligned}B &\le \frac{2\,C_H\,\beta_{r+1}(\varepsilon)}{\varepsilon\,(r+1)!}\,\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{\lambda\sigma^2\xi^2}{2}}\,|\xi|^{2r+2}\,d\xi\right)^{\frac12}\\ &\le \frac{\beta_{r+1}(\varepsilon)}{(r+1)!\,(\lambda\frac{\sigma^2}{2})^{\frac{r}{2}+\frac34}}\,\frac{2\,C_H}{\varepsilon}\,\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{3}{2}\right)}.\end{aligned}$$

(C) For $\xi \in [-\varepsilon,\varepsilon]$,

$$|(\psi-\chi)'(\xi)| \le \frac{\beta_{r+1}(\varepsilon)\,|\xi|^r}{r!},$$

so as before,

$$C \leq \frac{C_H\,\beta_{r+1}(\varepsilon)}{r!}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{\lambda\sigma^2\xi^2}{2}}\,|\xi|^{2r}\,d\xi\right)^{\frac{1}{2}} = \frac{C_H\,\beta_{r+1}(\varepsilon)}{r!\,(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{1}{4}}}\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{1}{2}\right)}$$
$$\leq \frac{\beta_{r+1}(\varepsilon)}{(r+1)!\,(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{1}{4}}}\,\frac{C_H\,(r+1)}{\sqrt{r+\frac{1}{2}}}\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{3}{2}\right)}.$$

(D) Last, for $\xi\in[-\varepsilon,\varepsilon]$, $|\phi'(\xi)-\mathrm{i}\mu|\leq\gamma(\delta)\,|\xi|\leq\frac{\sigma^2|\xi|}{2}$, and then,

$$|(\lambda\phi'(\xi)-\mathrm{i}\lfloor\lambda\mu\rfloor)|\leq\frac{\lambda\sigma^2\,|\xi|}{2}+|\lambda\mu-\lfloor\lambda\mu\rfloor|\leq\frac{\lambda\sigma^2\,|\xi|}{2}+1.$$

Therefore,

$$D\leq\frac{C_H\,\beta_{r+1}(\varepsilon)}{(r+1)!}\left(\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{\lambda\sigma^2\xi^2}{2}}\,|\xi|^{2r+2}\,d\xi\right)^{\frac{1}{2}}+\frac{\lambda\,\sigma^2}{2}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{\lambda\sigma^2\xi^2}{2}}\,|\xi|^{2r+4}\,d\xi\right)^{\frac{1}{2}}\right)$$
$$\leq\frac{C_H\,\beta_{r+1}(\varepsilon)}{(r+1)!}\left(\frac{1}{(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{3}{4}}}\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{3}{2}\right)}+\frac{1}{(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{1}{4}}}\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{5}{2}\right)}\right).$$

To conclude, notice that since $\lambda\,\frac{\sigma^2}{2}\geq 1$, one can take in each denominator the smallest power of this quantity, namely, $(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{1}{4}}$. One thus obtains:

$$A+B+C+D\leq\frac{\beta_{r+1}(\varepsilon)}{(r+1)!\,(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{1}{4}}}\sqrt{\frac{1}{2\pi}\,\Gamma\left(r+\frac{3}{2}\right)}$$
$$\times\left(\left(\frac{2}{\pi}\right)^{\frac{1}{4}}+C_H\left(\frac{2}{\varepsilon}+\frac{r+1}{\sqrt{r+\frac{1}{2}}}+1+\sqrt{r+\frac{3}{2}}\right)\right)$$
$$\leq\frac{\beta_{r+1}(\varepsilon)}{(r+1)!\,(\lambda\,\frac{\sigma^2}{2})^{\frac{r}{2}+\frac{1}{4}}}\sqrt{\frac{2\pi}{3}\,\Gamma\left(r+\frac{3}{2}\right)}\left(\frac{1}{\varepsilon}+\sqrt{5(r+1)}\right)$$

by using on the last line $C_H=\frac{\pi}{\sqrt{3}}$ and

$$2\sqrt{5(r+1)}\geq\frac{r+1}{\sqrt{r+\frac{1}{2}}}+\sqrt{r+\frac{3}{2}}+1+\frac{\sqrt{3}}{\pi}\left(\frac{2}{\pi}\right)^{\frac{1}{4}}.$$ □

Proof of Theorem 2.3. We have

$$\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\leq\|\psi-\chi\|_{\mathscr{A}}\left\|(1-c)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}}+\left\|c\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}},$$

and Lemmas 2.4 and 2.5 allow us to control these two terms. □

Note that if ε is fixed and λ goes to infinity, then the dominant term in the norm inequality is the second one, because the first one decreases exponentially fast. In this case,

$$d_{\mathrm{TV}}(\mu,\nu)=\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}=O\left(\frac{1}{\lambda^{\frac{r}{2}+\frac{1}{4}}}\right).$$

The constant hidden in the $O(\cdot)$ depends only on ϕ, r, $\|\psi-\chi\|_{\mathscr{A}}$ and the parameter $\beta_{r+1}(\varepsilon)$ introduced before. In particular, if ϕ is fixed and we look at families of residues $(\psi_n)_{n\in\mathbb{N}}$ and $(\chi_n)_{n\in\mathbb{N}}$ such that the corresponding quantities $\|\psi_n-\chi_n\|_{\mathscr{A}}$ and $\beta_{r+1,n}(\varepsilon)$ stay bounded, then one can take a uniform constant in the $O(\cdot)$.

Remark 2.6. One can broaden slightly the scope of Theorem 2.3, if one notes that regarding the residues ψ and χ, one only used the bounds

$$|\psi(\xi)-\chi(\xi)| \leq \frac{\beta_{r+1}(\varepsilon)\,|\xi|^{r+1}}{(r+1)!}$$
$$\text{and}\quad |\psi'(\xi)-\chi'(\xi)| \leq \frac{\beta_{r+1}(\varepsilon)\,|\xi|^{r}}{r!}$$

for any $\xi\in[-\varepsilon,\varepsilon]$. In particular, if one assumes that ψ and χ belong to $\mathscr{C}^1(\mathbb{T})$ and that

$$\psi'(\xi)-\chi'(\xi) = \frac{\beta_{r+1}\,\xi^r\,(1+o_\xi(1))}{r!}$$

in a neighborhood $[-\varepsilon,\varepsilon]$ of 0, then the previous bounds are satisfied and the conclusions of Theorem 2.3 hold. We shall use this important remark in §3.3.

3. Asymptotics of distances for a general scheme of approximation

We now consider a general scheme of approximation of a sequence of $\mathbb{Z}$-valued random variables $(X_n)_{n\in\mathbb{N}}$ by a sequence of laws $(\nu_n)_{n\in\mathbb{N}}$, with

$$\widehat{\mu_n}(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\xi X_n}] = \mathrm{e}^{\lambda_n\phi(\xi)}\,\psi_n(\xi)\ ,$$
$$\widehat{\nu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\,\chi_n(\xi)\ ,$$

$\lim_{n\to\infty}\psi_n(\xi)=\psi(\xi)$ and $\lim_{n\to\infty}\chi_n(\xi)=\chi(\xi)$ as in Definition 1.10. We can immediately deduce from Theorem 2.3:

Proposition 3.1. *Consider a general scheme of approximation such that the convergences $\psi_n\to\psi$ and $\chi_n\to\chi$ occur in $\mathscr{C}^1(\mathbb{T})$. Then, $d_{\mathrm{TV}}(\mu_n,\nu_n)\to 0$ (and as well for $d_{\mathrm{L}}(\mu_n,\nu_n)$ and $d_{\mathrm{K}}(\mu_n,\nu_n)$).*

Proof. Taking $r=0$ in Theorem 2.3, we have

$$d_{\mathrm{TV}}(\mu_n,\nu_n) = \|\mu_n-\nu_n\|_{\mathscr{A}} = O\left(\frac{1}{(\lambda_n)^{\frac{1}{4}}}\right),$$

where the constant in the $O(\cdot)$ depends on $\|\psi_n-\chi_n\|_{\mathscr{A}}$ and $\|\psi_n-\chi_n\|_{\mathscr{C}^1}$. By Proposition 2.1, both quantities are bounded by a function of $\|\psi_n-\chi_n\|_{\mathscr{C}^1}$, which is itself bounded since $\psi_n-\chi_n\to\psi-\chi$ in $\mathscr{C}^1(\mathbb{T})$.

For the other distances, we use the inequality $d_{\mathrm{L}}\leq 2\,d_{\mathrm{K}}\leq d_{\mathrm{TV}}$. □

If we want to improve on the rate of convergence of the distances to 0, then we need an additional assumption similar to the hypothesis of Theorem 2.3, namely:

$$\forall n\in\mathbb{N},\ \psi_n(\xi)-\chi_n(\xi) = \beta_n\,(\mathrm{i}\xi)^{r+1}\,(1+o_\xi(1))$$
$$\text{and}\ \ \psi(\xi)-\chi(\xi) = \beta\,(\mathrm{i}\xi)^{r+1}\,(1+o_\xi(1)) \tag{H1}$$

with $\lim_{n\to\infty}\beta_n=\beta$, and the $o_\xi(1)$ that converges to 0 as ξ goes to 0, uniformly in n. We denote this condition by (H1). It is satisfied if, for instance, the convergences $\psi_n\to\psi$

and $\chi_n \to \chi$ occur in $\mathscr{C}^{r+1}(\mathbb{T})$, and if the scheme of approximation is the scheme of order r as defined in Example 1.11. However, this new condition (H1) is a bit more general, as we do not assume that the residues χ_n and χ are Laurent polynomials of degree r in $\mathrm{e}^{\mathrm{i}\xi}$.

In §3.1, §3.2 and §3.3, we shall consider the setting described above, and we shall establish some exact asymptotic formulas for $d_{\mathrm{L}}(\mu_n, \nu_n)$, $d_{\mathrm{K}}(\mu_n, \nu_n)$ and $d_{\mathrm{TV}}(\mu_n, \nu_n)$. These formulas generally follow from an application of the Laplace method to an integral representation of the distance, but we shall also use Theorem 2.3 in order to get rid of certain non dominant terms in the asymptotics. In §3.4, we introduce the notion of *derived approximation schemes* which involve simpler residues, and under appropriate assumptions, we extend the results of §3.1-3.3 to these approximation schemes.

3.1. **The local distance.** Until the end of this section, $(X_n)_{n\in\mathbb{N}}$ denotes a sequence that is mod-ϕ convergent with parameters $(\lambda_n)_{n\in\mathbb{N}}$, and $(\nu_n)_{n\in\mathbb{N}}$ is a scheme of approximation of it, which satisfies the hypothesis (H1). The main tool in the computation of the local distance $d_{\mathrm{L}}(\mu_n, \nu_n)$ is:

Proposition 3.2. *The error term being uniform in $k \in \mathbb{Z}$, one has*

$$\mu_n(\{k\}) - \nu_n(\{k\}) = \frac{(-1)^{r+1}\,\beta}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r}{2}+1}}\ \frac{\partial^{r+1}}{\partial\alpha^{r+1}}\left(\mathrm{e}^{-\frac{\alpha^2}{2}}\right)\Big|_{\alpha=\frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}} + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\right).$$

Proof. The computations hereafter are very similar to those of [FMN16, Section 3], but they are performed at a different scale for k. We combine the Fourier inversion formula

$$\mu_n(\{k\}) - \nu_n(\{k\}) = \int_{\mathbb{T}} (\psi_n(\xi) - \chi_n(\xi))\ \mathrm{e}^{\lambda_n\phi(\xi)}\,\mathrm{e}^{-\mathrm{i}k\xi}\,\frac{d\xi}{2\pi}$$

with the Laplace method. Note that (H1) implies that $\psi_n - \chi_n$ is bounded on the whole torus $\mathbb{T}$ by some constant C. As a consequence, for any $\varepsilon > 0$, since $\mathrm{Re}(\phi(\xi)) \leq -M\xi^2$ with $M > 0$, one has

$$\begin{aligned}\frac{1}{2\pi}\left|\int_{(-\pi,\pi)\setminus(-\varepsilon,\varepsilon)} (\psi_n(\xi) - \chi_n(\xi))\ \mathrm{e}^{\lambda_n\phi(\xi)}\,\mathrm{e}^{-\mathrm{i}k\xi}\,d\xi\right| &\leq C\,\exp\left(\lambda_n \sup_{\xi\in(-\pi,\pi)\setminus(-\delta,\delta)} \mathrm{Re}(\phi(\xi))\right)\\ &\leq C\,\exp(-\lambda_n M\varepsilon^2)\\ &= o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\right),\end{aligned}$$

this being uniform in k. Then, for $\varepsilon > 0$ small enough, one has the estimate:

$$\begin{aligned}&\int_{-\varepsilon}^{\varepsilon} (\psi_n(\xi) - \chi_n(\xi))\ \mathrm{e}^{\lambda_n\phi(\xi)}\,\mathrm{e}^{-\mathrm{i}k\xi}\,d\xi\\ &= \beta_n\,(1 + o_\varepsilon(1))\int_{-\varepsilon}^{\varepsilon} (\mathrm{i}\xi)^{r+1}\,\mathrm{e}^{\lambda_n\phi(\xi)-\mathrm{i}k\xi}\,d\xi\\ &= \beta\,(1 + o_\varepsilon(1))\int_{-\varepsilon}^{\varepsilon} (\mathrm{i}\xi)^{r+1}\,\mathrm{e}^{\lambda_n(\phi(\xi)-\mathrm{i}m\xi)}\,\mathrm{e}^{-\mathrm{i}(k-\lambda_n m)\xi}\,d\xi\\ &= \frac{\beta}{(\sigma^2\lambda_n)^{\frac{r}{2}+1}}\,(1 + o_\varepsilon(1))\int_{-\varepsilon\,\sigma\sqrt{\lambda_n}}^{\varepsilon\,\sigma\sqrt{\lambda_n}} (\mathrm{i}t)^{r+1}\,\mathrm{e}^{\lambda_n\left(\phi\left(\frac{t}{\sigma\sqrt{\lambda_n}}\right)-\frac{\phi'(0)\,t}{\sigma\sqrt{\lambda_n}}\right)}\,\mathrm{e}^{-\mathrm{i}t\frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}}\,dt\\ &= \frac{\beta}{(\sigma^2\lambda_n)^{\frac{r}{2}+1}}\,(1 + o_\varepsilon(1))\int_{-\infty}^{\infty} (\mathrm{i}t)^{r+1}\,\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{e}^{-it\frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}}\,dt.\end{aligned}$$

The last step uses the dominated convergence theorem. Indeed, we have the pointwise convergence:

$$\lambda_n \left(\phi\left(\frac{t}{\sigma\sqrt{\lambda_n}} \right) - \frac{\phi'(0)\, t}{\sigma\sqrt{\lambda_n}} \right) \longrightarrow_{n\to\infty} -\frac{t^2}{2}$$

and if ε is small enough so that $\mathrm{Re}(\phi''(\xi)) < -\frac{\sigma^2}{2}$ for all $\xi \in [-\varepsilon, \varepsilon]$, then for every $t \in [-\varepsilon\,\sigma\sqrt{\lambda_n}, \varepsilon\,\sigma\sqrt{\lambda_n}]$:

$$\left| \mathrm{e}^{\lambda_n \left(\phi\left(\frac{t}{\sigma\sqrt{\lambda_n}} \right) - \frac{\phi'(0)\, t}{\sigma\sqrt{\lambda_n}} \right)} \right| = \mathrm{e}^{\lambda_n \,\mathrm{Re}\left(\phi\left(\frac{t}{\sigma\sqrt{\lambda_n}} \right) - \frac{\phi'(0)\, t}{\sigma\sqrt{\lambda_n}} \right)} \leq \mathrm{e}^{\left(\sup_{\xi\in[-\varepsilon,\varepsilon]} \mathrm{Re}(\phi''(\xi)) \right) \frac{t^2}{2\sigma^2}} \leq \mathrm{e}^{-\frac{t^2}{4}}.$$

The integrand is therefore dominated by $\mathrm{e}^{-\frac{t^2}{4}}\, |t|^{r+1}$. Hence,

$$\mu_n(\{k\}) - \nu_n(\{k\}) = \frac{\beta}{2\pi(\sigma^2\lambda_n)^{\frac{r}{2}+1}} \left(\int_{-\infty}^{\infty} (\mathrm{i}t)^{r+1}\, \mathrm{e}^{-\frac{t^2}{2}}\, \mathrm{e}^{-\mathrm{i}t\frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}}\, dt \right) + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}} \right),$$

with a remainder uniform in $k \in \mathbb{Z}$. Last, we use the fact that

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{t^2}{2}}\, \mathrm{e}^{-\mathrm{i}t\alpha}\, \frac{dt}{\sqrt{2\pi}} = \mathrm{e}^{-\frac{\alpha^2}{2}},$$

is the Fourier transform of the standard Gaussian distribution evaluated at $\xi = -\alpha$, and by taking the $(r+1)$-th derivative with respect to α, we get:

$$\int_{-\infty}^{\infty} (\mathrm{i}t)^{r+1}\, \mathrm{e}^{-\frac{t^2}{2}}\, \mathrm{e}^{-\mathrm{i}t\alpha}\, \frac{dt}{\sqrt{2\pi}} = (-1)^{r+1} \frac{\partial^r}{\partial\alpha^{r+1}} \left(\mathrm{e}^{-\frac{\alpha^2}{2}} \right),$$

hence the claimed result. □

The computation of the local distance amounts then to find the maximum in $k \in \mathbb{Z}$ of the previous quantity. Recall that the r-th Hermite polynomial $H_r(\alpha)$ is defined by

$$G_r(\alpha) = \frac{\partial^r}{\partial\alpha^r} \left(\mathrm{e}^{-\frac{\alpha^2}{2}} \right) = (-1)^r\, H_r(\alpha)\, \mathrm{e}^{-\frac{\alpha^2}{2}}.$$

The local extremas of G_r correspond to the zeros of H_{r+1} since

$$G_r'(\alpha) = (-1)^r\, \left(H_r'(\alpha) - \alpha\, H_r(\alpha) \right) \mathrm{e}^{-\frac{\alpha^2}{2}} = (-1)^{r+1}\, H_{r+1}(\alpha)\, \mathrm{e}^{-\frac{\alpha^2}{2}} = G_{r+1}(\alpha).$$

Denote z_{r+1} the smallest absolute value of a zero of H_{r+1}; it is 0 when r is even, and it can be shown that in any case it corresponds to the global extrema of $|G_r|$; see [Sze39, Chapter 6], and Figure 2 for an illustration.

Theorem 3.3. *Under the assumptions stated at the beginning of this section (general scheme of approximation with condition (H1)), one has*

$$d_{\mathrm{L}}(\mu_n, \nu_n) = \frac{|\beta|\, |G_{r+1}(z_{r+2})|}{\sqrt{2\pi}\, (\sigma^2\lambda_n)^{\frac{r}{2}+1}} + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}} \right).$$

Proof. Denote $Z_n = \{\frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}},\ k \in \mathbb{Z}\}$. The previous discussion shows that

$$d_{\mathrm{L}}(\mu_n, \nu_n) = \frac{|\beta|}{\sqrt{2\pi}\, (\sigma^2\lambda_n)^{\frac{r}{2}+1}} \sup_{\alpha\in Z_n} |G_{r+1}(\alpha)| + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}} \right).$$

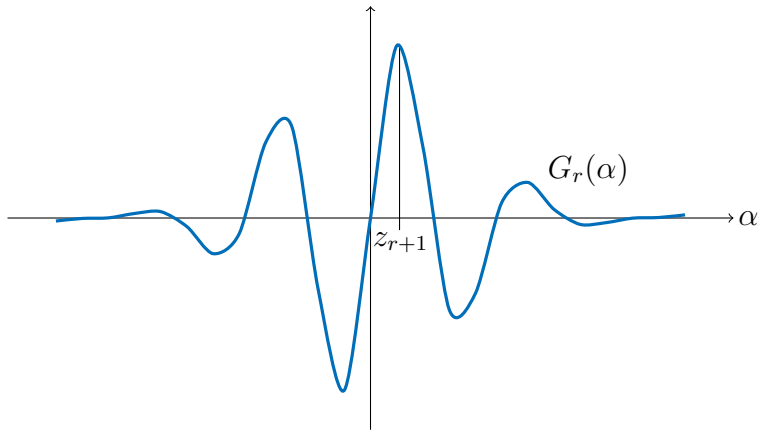

FIGURE 2. The Hermite function G_9; it attains its global extremas at the smallest zeroes of H_{10}.

As n goes to infinity, the set Z_n becomes dense in $\mathbb{R}$, and since G_{r+1} is a Lipschitz function, we have in fact

$$\begin{aligned} d_{\mathrm{L}}(\mu_n, \nu_n) &= \frac{|\beta|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r}{2}+1}} \sup_{\alpha\in\mathbb{R}} |G_{r+1}(\alpha)| + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\right) \\ &= \frac{|\beta|\,|G_{r+1}(z_{r+2})|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r}{2}+1}} + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\right). \end{aligned}$$ □

Remark 3.4. The first values of $M_r = |G_r(z_{r+1})|$ are

$$\begin{aligned} M_0 &= |G_0(0)| = 1; \\ M_1 &= |G_1(1)| = \mathrm{e}^{-\frac{1}{2}} = 0.60653\ldots; \\ M_2 &= |G_2(0)| = 1; \\ M_3 &= \left|G_3\left(\sqrt{3-\sqrt{6}}\right)\right| = \mathrm{e}^{-\frac{3-\sqrt{6}}{2}}\left(3\sqrt{6}-6\right) = 1.38012\ldots; \\ M_4 &= |G_4(0)| = 3. \end{aligned}$$

In particular, $M_{2r} = (2r-1)!! = \frac{(2r)!}{2^r\, r!}$.

3.2. **The Kolmogorov distance.** To evaluate the Kolmogorov distance between μ_n and ν_n, we use the classical integral formula

$$\mu_n([\![k,+\infty]\!]) - \nu_n([\![k,+\infty]\!]) = \int_{\mathbb{T}} (\psi_n(\xi) - \chi_n(\xi))\, \mathrm{e}^{\lambda_n\phi(\xi)}\, \frac{\mathrm{e}^{-\mathrm{i}k\xi}}{1-\mathrm{e}^{-\mathrm{i}\xi}}\, \frac{d\xi}{2\pi}.$$

Indeed, for any $l \geq k$, one has

$$\begin{aligned} \mu_n([\![k,l]\!]) - \nu_n([\![k,l]\!]) &= \sum_{j=k}^{l} \int_{\mathbb{T}} (\psi_n(\xi) - \chi_n(\xi))\, \mathrm{e}^{\lambda_n\phi(\xi)}\, \mathrm{e}^{-\mathrm{i}j\xi}\, \frac{d\xi}{2\pi} \\ &= \int_{\mathbb{T}} (\psi_n(\xi) - \chi_n(\xi))\, \mathrm{e}^{\lambda_n\phi(\xi)}\, \frac{\mathrm{e}^{-\mathrm{i}k\xi} - \mathrm{e}^{-\mathrm{i}(l+1)\xi}}{1-\mathrm{e}^{-\mathrm{i}\xi}}\, \frac{d\xi}{2\pi}. \end{aligned}$$

Since $\psi_n(\xi) - \chi_n(\xi) = \beta_n\,(\mathrm{i}\xi)^{r+1}\,(1+o(1))$, the quotient $\frac{\psi_n(\xi)-\chi_n(\xi)}{1-\mathrm{e}^{\mathrm{i}\xi}}$ is actually a continuous function on the torus, hence bounded. The same holds for $\mathrm{e}^{\lambda_n\phi(\xi)} = \mathbb{E}[\mathrm{e}^{\mathrm{i}\xi Y_n}]$, where Y_n is

an infinitely divisible random variable with exponent $\lambda_n\phi$. Thus,

$$f(\xi) = \frac{\psi_n(\xi) - \chi_n(\xi)}{1 - \mathrm{e}^{\mathrm{i}\xi}}\, \mathrm{e}^{\lambda_n \phi(\xi)}$$

is a bounded continuous function, hence in $\mathscr{L}^1(\mathbb{T})$, and by the Riemann–Lebesgue lemma, one concludes that

$$\lim_{l\to+\infty} \int_{\mathbb{T}} f(\xi)\, \mathrm{e}^{-\mathrm{i}(l+1)\xi}\, \frac{d\xi}{2\pi} = 0.$$

Since $\lim_{l\to+\infty} \mu_n([\![k,l]\!]) - \nu_n([\![k,l]\!]) = \mu_n([\![k,+\infty]\!]) - \nu_n([\![k,+\infty]\!])$, the integral formula is indeed shown. It leads to the analogue of Proposition 3.2 for probabilities of half-lines of integers $[\![k,+\infty]\!]$:

Proposition 3.5. *The error term being uniform in $k \in \mathbb{Z}$, one has*

$$\mu_n([\![k,+\infty]\!]) - \nu_n([\![k,+\infty]\!]) = \frac{(-1)^r\, \beta}{\sqrt{2\pi}\, (\sigma^2\lambda_n)^{\frac{r+1}{2}}}\, \frac{\partial^r}{\partial \alpha^r} \left(\mathrm{e}^{-\frac{\alpha^2}{2}} \right) \Big|_{\alpha = \frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}} + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}} \right).$$

Proof. One has the Taylor expansion

$$\frac{\psi_n(\xi) - \chi_n(\xi)}{1 - \mathrm{e}^{-\mathrm{i}\xi}} = \frac{\beta_n(\mathrm{i}\xi)^{r+1}}{\mathrm{i}\xi}\, (1 + o_\xi(1)) = \beta_n(\mathrm{i}\xi)^r\, (1 + o_\xi(1)).$$

Combined with the aforementioned integral formula and with the same Laplace method as in Proposition 3.2, it yields

$$\mu_n([\![k,+\infty]\!]) - \nu_n([\![k,+\infty]\!]) = \frac{\beta}{(\sigma^2\lambda_n)^{\frac{r+1}{2}}}\, (1 + o(1)) \int_{-\infty}^{\infty} (\mathrm{i}t)^r\, \mathrm{e}^{-\frac{t^2}{2}}\, \mathrm{e}^{-\mathrm{i}t\frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}}\, dt.$$

Therefore, one has the same asymptotics as in Proposition 3.2, but with r replacing $r+1$. □

The same argument as in §3.1 gives:

Theorem 3.6. *Under the assumptions stated at the beginning of this section (general scheme of approximation with condition (H1)), one has*

$$d_{\mathrm{K}}(\mu_n, \nu_n) = \frac{|\beta|\, |G_r(z_{r+1})|}{\sqrt{2\pi}\, (\sigma^2\lambda_n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}} \right).$$

Example 3.7. Consider the basic scheme of approximation of a mod-ϕ convergent sequence of random variables $(X_n)_{n\in\mathbb{N}}$:

$$\widehat{\nu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}.$$

Hence, ν_n is the law of an infinitely divisible random variable Y_n with Lévy–Khintchine exponent $\lambda_n\phi$. Then, if $\widehat{\mu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\, \psi_n(\xi)$ and $\psi_n(\xi) \to \psi(\xi)$ in $\mathscr{C}^1(\mathbb{T})$, the hypothesis (H1) is satisfied with $r = 0$ and $\beta_n \to \beta = \psi'(0)$. Therefore, in this setting,

$$d_{\mathrm{L}}(X_n, Y_n) = \frac{|\psi'(0)|\, \mathrm{e}^{-\frac{1}{2}}}{\sqrt{2\pi}\, \sigma^2\lambda_n} + o\left(\frac{1}{\lambda_n} \right);$$

$$d_{\mathrm{K}}(X_n, Y_n) = \frac{|\psi'(0)|}{\sqrt{2\pi\, \sigma^2\lambda_n}} + o\left(\frac{1}{\sqrt{\lambda_n}} \right).$$

The second result should be compared with a computation in [FMN16, Chapter 4], which ensures that

$$d_{\mathrm{K}}(X_n, Y_n) = \frac{|\psi'(0)|}{\sqrt{2\pi\,\lambda_n}} + o\left(\frac{1}{\sqrt{\lambda_n}}\right)$$

if $(X_n)_{n\in\mathbb{N}}$ is a sequence of random real numbers that converges mod-Gaussian with parameters λ_n, and if $(Y_n)_{n\in\mathbb{N}}$ is a sequence of random Gaussian variables with means 0 and variances λ_n.

3.3. **The total variation distance.** Before we estimate the total variation distance between the laws μ_n and ν_n, let us make some observations with Proposition 3.2. Assume again that one has a general scheme of approximation which satisfies the hypothesis (H1). For any fixed interval $I = [a, b]$, one can write

$$\begin{aligned} d_{\mathrm{TV}}(\mu_n, \nu_n) &= \sum_{k\in\mathbb{Z}} |\mu_n(\{k\}) - \nu_n(\{k\})| \geq \sum_{k=\lfloor \lambda_n m + a\,\sigma\sqrt{\lambda_n}\rfloor}^{\lfloor \lambda_n m + b\,\sigma\sqrt{\lambda_n}\rfloor} |\mu_n(\{k\}) - \nu_n(\{k\})| \\ &\geq \frac{|\beta_n|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r}{2}+1}} \left(\sum_{k=\lfloor \lambda_n m + a\,\sigma\sqrt{\lambda_n}\rfloor}^{\lfloor \lambda_n m + b\,\sigma\sqrt{\lambda_n}\rfloor} \left| \left(H_{r+1}(\alpha)\,\mathrm{e}^{-\frac{\alpha^2}{2}} \right)_{\alpha = \frac{k-\lambda_n m}{\sigma\sqrt{\lambda_n}}} \right| \right) + o\left(\frac{(b-a)}{(\sigma^2\lambda_n)^{\frac{r+1}{2}}} \right) \quad \text{(RS)} \\ &\geq \frac{|\beta_n|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r+1}{2}}} \int_a^b \left| H_{r+1}(\alpha)\,\mathrm{e}^{-\frac{\alpha^2}{2}} \right| d\alpha + o\left(\frac{(b-a)}{(\sigma^2\lambda_n)^{\frac{r+1}{2}}} \right), \end{aligned}$$

by identifying a Riemann sum in (RS). Since this is true for any a and b, and since $H_{r+1}(\alpha)\,\mathrm{e}^{-\frac{\alpha^2}{2}}$ is integrable, we conclude that in general, one has

$$\liminf_{n\to\infty} \left(d_{\mathrm{TV}}(\mu_n, \nu_n)\,(\sigma^2\lambda_n)^{\frac{r+1}{2}} \right) \geq |\beta| \int_{\mathbb{R}} |H_{r+1}(\alpha)|\,\mathrm{e}^{-\frac{\alpha^2}{2}}\,\frac{d\alpha}{\sqrt{2\pi}}.$$

The goal of this paragraph is to show that, under a slightly stronger hypothesis than (H1), this inequality is in fact an identity, and that the liminf above is actually a limit. To this purpose, it is convenient to introduce a third sequence $(\rho_n)_{n\in\mathbb{N}}$ of signed measures, defined by their Fourier transforms

$$\widehat{\rho_n}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)} \left(\psi_n(\xi) - \beta_n\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^{r+1} \right).$$

These signed measures have their values given by:

Lemma 3.8. *Denote $\nu_n^{(0)}$ the infinitely divisible discrete law on $\mathbb{Z}$ with exponent $\lambda_n\phi$. The signed measure ρ_n whose Fourier transform is $\widehat{\rho_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\,(\psi_n(\xi) - \beta_n\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^{r+1})$ is given by*

$$\rho_n(\{k\}) = \mu_n(\{k\}) - \beta_n \sum_{l=0}^{r+1} (-1)^{r+1-l} \binom{r+1}{l} \nu_n^{(0)}(\{k-l\}).$$

Proof. One expands the Fourier transform $\widehat{(\mu_n - \rho_n)}(\xi)$ in powers of $\mathrm{e}^{\mathrm{i}\xi}$:

$$\begin{aligned} \widehat{(\mu_n - \rho_n)}(\xi) &= \beta_n\,\left(\mathrm{e}^{\mathrm{i}\xi} - 1\right)^{r+1} \sum_{m=-\infty}^{\infty} \nu_n^{(0)}(\{m\})\,\mathrm{e}^{\mathrm{i}m\xi} \\ &= \beta_n \sum_{l=0}^{r+1} \sum_{m=-\infty}^{\infty} (-1)^{r+1-l} \binom{r+1}{l} \nu_n^{(0)}(\{m\})\,\mathrm{e}^{\mathrm{i}(l+m)\xi} \end{aligned}$$

$$= \sum_{k=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k\xi} \left(\beta_n \sum_{l=0}^{r+1} (-1)^{r+1-l} \binom{r+1}{l} \nu_n^{(0)}(\{k-l\}) \right)$$
$$= \sum_{k=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k\xi} \left(\mu_n(\{k\}) - \rho_n(\{k\}) \right).$$

The result follows by identification of the Fourier coefficients. □

Remark 3.9. Suppose in particular that $\phi(\xi) = \mathrm{e}^{\mathrm{i}\xi} - 1$ is the exponent of the Poisson law $\mathcal{P}_{(1)}$. We then have

$$\begin{aligned}\mu_n(\{k\}) - \rho_n(\{k\}) &= \beta_n \sum_{l=0}^{(r+1)\wedge k} (-1)^{r+1-l} \binom{r+1}{l} \frac{(\lambda_n)^{k-l}\,\mathrm{e}^{-\lambda_n}}{(k-l)!} \\ &= \beta_n\, \nu_n^{(0)}(\{k\}) \sum_{l=0}^{(r+1)\wedge k} (-1)^{r+1-l} \binom{r+1}{l} \frac{(\lambda_n)^{-l}\, k!}{(k-l)!} \\ &= \beta_n\, \nu_n^{(0)}(\{k\})\, c(r+1, k, \lambda_n),\end{aligned}$$

where $c(r+1, k, \lambda_n)$ is a Poisson–Charlier polynomial, see [Sze39, §2.8.1].

We now assume until the end of this paragraph that all the residues ψ_n, ψ, χ_n and χ are in $\mathscr{C}^1(\mathbb{T})$, and that

$$\begin{aligned}\forall n \in \mathbb{N},\ \psi_n'(\xi) - \chi_n'(\xi) &= \mathrm{i}(r+1)\,\beta_n\,(\mathrm{i}\xi)^r + \mathrm{i}(r+2)\,\gamma_n\,(\mathrm{i}\xi)^{r+1}\,(1+o_\xi(1)) \\ \text{and}\ \ \psi'(\xi) - \chi'(\xi) &= \mathrm{i}(r+1)\,\beta\,(\mathrm{i}\xi)^r + \mathrm{i}(r+2)\,\gamma\,(\mathrm{i}\xi)^{r+1}\,(1+o_\xi(1)) \qquad \text{(H2)}\end{aligned}$$

with $\lim_{n\to\infty} \beta_n = \beta$ and $\lim_{n\to\infty} \gamma_n = \gamma$. As in Equation (H1), the $o_\xi(1)$ tend to 0 as $\xi \to 0$, uniformly in n. We denote this new condition by (H2). Since $\psi_n(0) = \chi_n(0) = \psi(0) = \chi(0) = 1$, it implies by integration that

$$\begin{aligned}\forall n \in \mathbb{N},\ \psi_n(\xi) - \chi_n(\xi) &= \beta_n\,(\mathrm{i}\xi)^{r+1} + \gamma_n\,(\mathrm{i}\xi)^{r+2}\,(1+o_\xi(1)) \\ \text{and}\ \ \psi(\xi) - \chi(\xi) &= \beta\,(\mathrm{i}\xi)^{r+1} + \gamma\,(\mathrm{i}\xi)^{r+2}\,(1+o_\xi(1))\end{aligned}$$

so in particular (H2) is stronger than (H1). On the other hand, the hypothesis (H2) is satisfied for instance if the convergences $\psi_n \to \psi$ and $\chi_n \to \chi$ occur in $\mathscr{C}^{r+2}(\mathbb{T})$, and if the scheme of approximation $(\nu_n)_{n\in\mathbb{N}}$ is the scheme of order r as defined in Example 1.11.

Lemma 3.10. *Assume that $(\nu_n)_{n\in\mathbb{N}}$ is a general approximation scheme of a mod-ϕ convergent sequence $(X_n)_{n\in\mathbb{N}}$, with the new hypothesis (H2) satisfied. Then, with ρ_n defined as in Lemma 3.8,*

$$d_{\mathrm{TV}}(\rho_n, \nu_n) = o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}} \right).$$

Moreover, $(\rho_n)_{n\in\mathbb{N}}$ is a new general approximation scheme of $(X_n)_{n\in\mathbb{N}}$, with the hypothesis (H2) again satisfied, with the same parameter $r \geq 0$ and the same sequence $(\beta_n)_{n\in\mathbb{N}}$.

Proof. Denote $\widetilde{\psi_n}(\xi) = \psi_n(\xi) - \beta_n\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^{r+1}$, so that $\widehat{\rho_n}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)}\,\widetilde{\psi_n}(\xi)$. On the one hand, one has

$$\begin{aligned}(\widetilde{\psi_n} - \chi_n)'(\xi) &= (\psi_n'(\xi) - \chi_n'(\xi)) - \mathrm{i}(r+1)\,\beta_n\,\mathrm{e}^{\mathrm{i}\xi}(\mathrm{e}^{\mathrm{i}\xi} - 1)^r \\ &= \mathrm{i}(r+2)\left(\gamma_n - \frac{\beta_n(r+1)}{2} \right) (\mathrm{i}\xi)^{r+1}\,(1+o_\xi(1))\end{aligned}$$

by using the Taylor expansion $(e^{i\xi}-1)^r = (i\xi)^r + \frac{r(i\xi)^{r+1}}{2} + o(\xi^{r+1})$. By the remark following the proof of Theorem 2.3, this is a sufficient condition in order to use the norm inequality of Section 2, but with parameter $r+1$ instead of r. Moreover, the sequence

$$\left(\gamma_n - \frac{\beta_n(r+1)}{2}\right)_{n\in\mathbb{N}}$$

is convergent under the hypothesis (H2), hence bounded. As a consequence, the sequence $(\widetilde{\psi_n} - \chi_n)_{n\in\mathbb{N}}$ is bounded in $\mathscr{C}^1(\mathbb{T})$, and therefore in $\mathscr{A}$ by Proposition 2.1. For the same reason, one has a uniform bound on the parameters

$$\beta_{r+2,n}(\varepsilon) = \sup_{\xi\in[-\varepsilon,\varepsilon]} \frac{\left|(\widetilde{\psi_n} - \chi_n)'(\xi)\right|}{|\xi|^{r+1}}.$$

According to the discussion after Theorem 2.3, we can conclude that

$$d_{\mathrm{TV}}(\rho_n,\nu_n) = O\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}+\frac{1}{4}}}\right) = o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

This ends the proof of the first part of the lemma, and for the second part, we can write

$$\begin{aligned}\psi_n'(\xi) - \widetilde{\psi_n}'(\xi) &= i(r+1)\,\beta_n\, e^{i\xi}(e^{i\xi}-1)^r \\ &= i(r+1)\,\beta_n\,(i\xi)^r + i(r+2)\,\frac{\beta_n(r+1)}{2}\,(i\xi)^{r+1}\,(1+o_\xi(1)),\end{aligned}$$

hence the hypothesis (H2) is satisfied. □

The previous Lemma shows that, if under the hypothesis (H2) one wants to obtain an estimate of $d_{\mathrm{TV}}(\mu_n,\nu_n)$ which is a $O((\lambda_n)^{-\frac{r+1}{2}})$, then one can assume without loss of generality that

$$\psi_n(\xi) - \chi_n(\xi) = \beta_n\,(e^{i\xi}-1)^{r+1},$$

with $\lim_{n\to\infty}\beta_n = \beta$ (in other words, $\nu_n = \rho_n$ is given exactly by the formula of Lemma 3.8). This is now of course much easier, and one obtains:

Theorem 3.11. *Consider a general scheme of approximation $(\nu_n)_{n\in\mathbb{N}}$ of a mod-ϕ convergent sequence of random variables $(X_n)_{n\in\mathbb{N}}$. We also assume that the reference law ϕ has a third moment, so that in the neighborhood of 0, $\phi(\xi) = -\frac{\sigma^2\xi^2}{2} + O(|\xi|^3)$. If the hypothesis (H2) is satisfied, then*

$$d_{\mathrm{TV}}(\mu_n,\nu_n) = \frac{|\beta|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r+1}{2}}}\left(\int_{\mathbb{R}} |G_{r+1}(\alpha)|\,d\alpha\right) + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

Proof. According to the previous discussion, we have to compute the asymptotics of

$$\begin{aligned}d_{\mathrm{TV}}(\mu_n,\nu_n) &= \|\widehat{\mu_n} - \widehat{\nu_n}\|_{\mathscr{A}} = \beta_n\,\left\|e^{\lambda_n\phi(\xi)}\,(e^{i\xi}-1)^{r+1}\right\|_{\mathscr{A}} \\ &= \beta_n \sum_{k\in\mathbb{Z}} \left|\int_{\mathbb{T}} e^{\lambda_n\phi(\xi)}\,(e^{i\xi}-1)^{r+1}\,e^{-ik\xi}\,\frac{d\xi}{2\pi}\right|.\end{aligned}$$

We can here follow the arguments of [Hwa99, Proposition 1], but in the general case of a discrete reference measure with Lévy exponent ϕ (instead of a Poisson distribution).

Step 1: We remove all the terms k outside the range of the central limit theorem $\frac{Y_n-\lambda_n m}{\sigma\sqrt{\lambda_n}} \rightharpoonup \mathcal{N}_{\mathbb{R}}(0,1)$, where Y_n follows the infinitely divisible law with exponent $\lambda_n\phi$. We claim that the terms k outside the interval

$$I = I\Big((\lambda_n)^{\frac17}\Big) = \Big[\!\Big[\lfloor \lambda_n m - \sigma(\lambda_n)^{\frac17+\frac12} \rfloor, \lfloor \lambda_n m + \sigma(\lambda_n)^{\frac17+\frac12} \rfloor \Big]\!\Big]$$

give an exponentially small contribution to the sum. Indeed, note that

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} |\mu_n(\{k\}) - \nu_n(\{k\})| &= \beta_n \sum_{k\in\mathbb{Z}} \left| \sum_{l=0}^{r+1} (-1)^{r+1-l} \binom{r+1}{l} \nu_n^{(0)}(\{k-l\}) \right| \\
&= \sum_{k\in\mathbb{Z}} \left| \big((\Delta_{-1})^{\circ(r+1)} \nu_n^{(0)}\big)(\{k\}) \right| \\
&= \sup_{f:\mathbb{Z}\to[-1,1]} \left(\sum_{k\in\mathbb{Z}} \big((\Delta_{-1})^{\circ(r+1)} \nu_n^{(0)}\big)(\{k\})\ f(k) \right) \\
&= \sup_{f:\mathbb{Z}\to[-1,1]} \left(\sum_{k\in\mathbb{Z}} \nu_n^{(0)}(\{k\})\ \ \big\{(\Delta_1)^{\circ(r+1)} f\big\}(k) \right)
\end{aligned}$$

with the notations Δ_j of discrete derivatives introduced at the end of §1.3. Indeed, Δ_1 and Δ_{-1} are adjoint operators on $\ell^2(\mathbb{Z})$. By the central limit theorem, for every $f:\mathbb{Z}\to[-1,1]$, the sum over terms k outside the interval I is now bounded by

$$2^{r+1}\,\mathbb{P}\left[\left| \frac{Y_n-\lambda_n m}{\sigma\sqrt{\lambda_n}} \right| \geq (\lambda_n)^{\frac17} \right] = o\left(\frac{1}{(\lambda_n)^p} \right) \quad \text{for every exponent } p,$$

so it will not contribute in the asymptotics. Therefore,

$$d_{\mathrm{TV}}(\mu_n,\nu_n) = |\beta_n| \sum_{k\in I((\lambda_n)^{\frac17})} \left| \int_{\mathbb{T}} \mathrm{e}^{\lambda_n\phi(\xi)}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^{r+1}\,\mathrm{e}^{-\mathrm{i}k\xi}\,\frac{d\xi}{2\pi} \right| + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}} \right).$$

Step 2: We now estimate the terms with $k \in I((\lambda_n)^{1/7})$ with a more careful application of the Laplace method than in Proposition 3.2. Namely, outside the interval $(-(\lambda_n)^{\frac17-\frac12}, (\lambda_n)^{\frac17-\frac12})$, one can bound $(\mathrm{e}^{\mathrm{i}\xi}-1)^{r+1}$ by $|\xi|^{r+1}$ and $\mathrm{e}^{\lambda_n\phi(\xi)}$ by $\mathrm{e}^{-\lambda_n M\xi^2}$, whence

$$\begin{aligned}
&\left| \int_{(-\pi,\pi)\setminus\left(-(\lambda_n)^{\frac17-\frac12},(\lambda_n)^{\frac17-\frac12}\right)} \mathrm{e}^{\lambda_n\phi(\xi)}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^{r+1}\,\mathrm{e}^{-\mathrm{i}k\xi}\,d\xi \right| \\
&\leq \int_{(-\pi,\pi)\setminus\left(-(\lambda_n)^{\frac17-\frac12},(\lambda_n)^{\frac17-\frac12}\right)} \mathrm{e}^{-\lambda_n M\xi^2}\,|\xi|^{r+1}\,d\xi \\
&\leq \frac{2}{(\lambda_n)^{\frac{r+1}{2}}} \int_{(\lambda_n)^{\frac17}}^{\infty} \mathrm{e}^{-Mu^2}\,u^{r+1}\,du
\end{aligned}$$

which is exponentially small and will not contribute to the asymptotics, even after multiplication by the number of terms $O((\lambda_n)^{\frac12+\frac17})$ of the interval I. On the other

hand, if $k = \lambda_n m + x\,\sigma\sqrt{\lambda_n}$, then

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-(\lambda_n)^{\frac{1}{7}-\frac{1}{2}}}^{(\lambda_n)^{\frac{1}{7}-\frac{1}{2}}} \mathrm{e}^{\lambda_n\phi(\xi)}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^{r+1}\,\mathrm{e}^{-\mathrm{i}k\xi}\,d\xi \\
&= \frac{1}{2\pi}\int_{-(\lambda_n)^{\frac{1}{7}-\frac{1}{2}}}^{(\lambda_n)^{\frac{1}{7}-\frac{1}{2}}} \mathrm{e}^{-\mathrm{i}(k-\lambda_n m)\xi-\lambda_n\frac{\sigma^2}{2}\xi^2}\,(\mathrm{i}\xi)^{r+1}\,\left(1+O((\lambda_n)^{\frac{3}{7}-\frac{1}{2}})\right)\,d\xi \\
&= \frac{(1+o(1))}{2\pi\,(\sigma^2\,\lambda_n)^{\frac{r}{2}+1}}\int_{-\sigma(\lambda_n)^{\frac{1}{7}}}^{\sigma(\lambda_n)^{\frac{1}{7}}} \mathrm{e}^{-\mathrm{i}xu-\frac{u^2}{2}}\,(\mathrm{i}u)^{r+1}\,du = \frac{(1+o(1))}{\sqrt{2\pi}\,(\sigma^2\,\lambda_n)^{\frac{r}{2}+1}}\,(-1)^{r+1}\,H_{r+1}(x)\,\mathrm{e}^{-\frac{x^2}{2}},
\end{aligned}$$

the $o(1)$ being uniform in $k \in I((\lambda_n)^{1/7})$. Note that we used the hypothesis of third moment of the reference law ϕ in order to get the multiplicative error term $1 + O((\lambda_n)^{\frac{3}{7}-\frac{1}{2}})$.

The proof can now be completed by using the same argument of Riemann sums as in the beginning of this paragraph, and the convergence $\beta_n \to \beta$. □

Remark 3.12. The first values of $V_r = \int_{\mathbb{R}} |G_{r+1}(\alpha)|\,d\alpha$ are:

$$\begin{aligned}
V_0 &= 2; \\
V_1 &= 4\,\mathrm{e}^{-\frac{1}{2}}; \\
V_2 &= 2\,(1+4\,\mathrm{e}^{-\frac{3}{2}}).
\end{aligned}$$

Example 3.13. For the basic scheme of approximation ($\chi_n = \chi = 1$), assume that the residues ψ_n converge in $\mathscr{C}^2(\mathbb{T})$, which implies hypothesis (H2) with $r = 0$ and $|\beta| = |\psi'(0)|$. We get the following estimate for the total variation distance between X_n and an infinitely divisible random variable Y_n with exponent $\lambda_n\phi$:

$$d_{\mathrm{TV}}(X_n, Y_n) = \frac{2\,|\psi'(0)|}{\sqrt{2\pi\sigma^2\lambda_n}} + o\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

Note that this is twice the asymptotic formula for the *Kolmogorov* distance.

3.4. **Derived scheme of approximation with constant residue.** To conclude this section, let us simplify a bit the general setting of approximation of discrete measures presented in Definition 1.10. Until now, we have worked with measures ν_n such that

$$\widehat{\nu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\,\chi_n(\xi),$$

where the residues $\chi_n(\xi)$ converge to a residue $\chi(\xi)$, and on the other hand approximate the residues $\psi_n(\xi)$ associated to the random variables X_n. For instance, $\chi_n(\xi)$ can be the Laurent polynomial of degree r in $\mathrm{e}^{\mathrm{i}\xi}$ derived from $\psi_n(\xi)$ (scheme of approximation of order r). In this setting, the residues $\chi_n(\xi)$ are usually very good approximations of the residues $\psi_n(\xi)$, but on the other hand they vary with n, which is quite a complication for explicit computations. Hence, for applications, it is simpler to work with a *constant* residue $\chi(\xi)$ that does not depend on n. This leads to the following definition:

Definition 3.14. *Let $(\nu_n)_{n\in\mathbb{N}}$ be a general scheme of approximation of a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ that converges mod-ϕ with parameters λ_n. The* derived scheme of approximation with constant residue *is the new scheme $(\sigma_n)_{n\in\mathbb{N}}$, defined by the Fourier transforms*

$$\widehat{\sigma_n}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)}\, \chi(\xi),$$

where $\chi(\xi) = \lim_{n\to\infty} \chi_n(\xi)$ is the limit of the residues of the law ν_n (with the notations of Definition 1.10).

Example 3.15. Suppose that $\nu_n = \nu_n^{(r)}$ is the scheme of approximation of order r of $(X_n)_{n\in\mathbb{N}}$ (see Example 1.11), and that the convergence $\psi_n \to \psi$ occurs in $\mathscr{C}^r(\mathbb{T})$. Then, the derived scheme with constant residue $(\sigma_n)_{n\in\mathbb{N}}$ is defined by the Fourier transforms

$$\widehat{\sigma_n}(\xi) = \mathrm{e}^{\lambda_n \phi(\xi)}\, P^{(r)}(\mathrm{e}^{\mathrm{i}\xi}),$$

where $P^{(r)}(\mathrm{e}^{\mathrm{i}\xi})$ is the Laurent polynomial of degree r associated to the limit ψ of the mod-ϕ convergence. Thus, to compute $\sigma_n(f)$ with $f : \mathbb{Z} \to \mathbb{C}$, one can take a random variable Y_n with law of exponent $\lambda_n\phi$, and then calculate

$$\mathbb{E}[(P^{(r)}(S)f)(Y_n)]$$

where $T = P^{(r)}(S)$ is a *fixed* linear operator, which is a finite linear combination of discrete difference operators. This is clearly convenient for concrete applications.

Informally, if the derived scheme with constant residue $(\sigma_n)_{n\in\mathbb{N}}$ is sufficiently close to the initial scheme $(\nu_n)_{n\in\mathbb{N}}$, then the conclusions of Theorems 3.3, 3.6 and 3.11 hold also for $(\sigma_n)_{n\in\mathbb{N}}$ (under the appropriate condition (H1) or (H2) for $(\nu_n)_{n\in\mathbb{N}}$). Practically, this happens as soon as the size of $\chi_n - \chi$ is negligible in comparison to negative powers of λ_n. More precisely, we have:

Theorem 3.16. *Let $(\nu_n)_{n\in\mathbb{N}}$ be a general scheme of approximation of a sequence $(X_n)_{n\in\mathbb{N}}$ that is mod-ϕ convergent. We denote $(\sigma_n)_{n\in\mathbb{N}}$ the derived scheme with constant residue.*

(1) Suppose that $(\nu_n)_{n\in\mathbb{N}}$ satisfies the hypothesis (H1). If $\|\chi_n - \chi\|_\infty = o\Big(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\Big)$, then

$$d_{\mathrm{L}}(\mu_n, \sigma_n) = \frac{|\beta|\,|G_{r+1}(z_{r+2})|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r}{2}+1}} + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\right).$$

(2) Suppose that $(\nu_n)_{n\in\mathbb{N}}$ satisfies the hypothesis (H1). If $\|\chi_n' - \chi'\|_\infty = o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right)$, then

$$d_{\mathrm{K}}(\mu_n, \sigma_n) = \frac{|\beta|\,|G_r(z_{r+1})|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

(3) Suppose that $(\nu_n)_{n\in\mathbb{N}}$ satisfies the hypothesis (H2). If $\|\chi_n' - \chi'\|_\infty = o\Big(\frac{1}{(\lambda_n)^{\frac{r}{2}+\frac{5}{4}}}\Big)$, then

$$d_{\mathrm{TV}}(\mu_n, \sigma_n) = \frac{|\beta|}{\sqrt{2\pi}\,(\sigma^2\lambda_n)^{\frac{r+1}{2}}} \left(\int_{\mathbb{R}} |G_{r+1}(\alpha)|\, d\alpha\right) + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

Proof. In each case, we have to bound the distance between ν_n and the derived scheme σ_n. For the local distance, one has

$$d_{\mathrm{L}}(\nu_n, \sigma_n) = \sup_{k\in\mathbb{Z}} \left| \int_{\mathbb{T}} (\chi_n(\xi) - \chi(\xi))\, \mathrm{e}^{\lambda_n \phi(\xi)}\, \mathrm{e}^{-\mathrm{i}k\xi}\, \frac{d\xi}{2\pi} \right| \leq \|\chi_n - \chi\|_\infty = o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}} \right),$$

hence the estimate on $d_{\mathrm{L}}(\mu_n, \sigma_n)$. For the Kolmogorov distance, notice that

$$\left| \frac{\chi_n(\xi) - \chi(\xi)}{1 - \mathrm{e}^{-\mathrm{i}\xi}} \right| \leq \|\chi_n' - \chi'\|_\infty \left| \frac{\xi}{1 - \mathrm{e}^{-\mathrm{i}\xi}} \right| \leq \pi\, \|\chi_n' - \chi'\|_\infty$$

for any $\xi \in (-\pi, \pi)$. Therefore,

$$d_{\mathrm{K}}(\nu_n, \sigma_n) = \sup_{k\in\mathbb{Z}} \left| \int_{\mathbb{T}} \frac{\chi_n(\xi) - \chi(\xi)}{1 - \mathrm{e}^{-\mathrm{i}\xi}}\, \mathrm{e}^{\lambda_n \phi(\xi)}\, \mathrm{e}^{-\mathrm{i}k\xi}\, \frac{d\xi}{2\pi} \right| \leq \pi \|\chi_n' - \chi'\|_\infty = o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}} \right),$$

hence the estimate on $d_{\mathrm{K}}(\mu_n, \sigma_n)$. Last, for the total variation distance, we use Proposition 2.1:

$$\begin{aligned} d_{\mathrm{TV}}(\nu_n, \sigma_n) &= \left\| (\chi_n - \chi)(\xi)\, \mathrm{e}^{\lambda_n \phi(\xi)} \right\|_{\mathscr{A}} \\ &\leq \|\chi_n - \chi\|_\infty + C_H\, \|((\chi_n - \chi)'(\xi) + \lambda_n\, \phi'(\xi)\, (\chi_n - \chi)(\xi))\, \mathrm{e}^{\lambda_n \phi(\xi)}\|_{\mathscr{L}^2} \\ &\leq \pi \|\chi_n' - \chi'\|_\infty + C_H\, (\|\chi_n' - \chi'\|_\infty + \lambda_n\, \|\phi'\|_\infty\, \|\chi_n - \chi\|_\infty)\, \|\mathrm{e}^{\lambda_n \phi(\xi)}\|_{\mathscr{L}^2} \\ &\leq \|\chi_n' - \chi'\|_\infty\, \left(\pi + C_H (1 + \pi \lambda_n\, \|\phi'\|_\infty)\, \|\mathrm{e}^{\lambda_n \phi(\xi)}\|_{\mathscr{L}^2} \right). \end{aligned}$$

Since $|\mathrm{e}^{\phi(\xi)}| \leq \mathrm{e}^{-M\xi^2}$, we have the estimate $\|\mathrm{e}^{\lambda_n \phi(\xi)}\|_{\mathscr{L}^2} = O(\frac{1}{(\lambda_n)^{1/4}})$, so in the end

$$d_{\mathrm{TV}}(\nu_n, \sigma_n) = O\left(\|\chi_n' - \chi'\|_\infty\, (\lambda_n)^{\frac{3}{4}} \right) = o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}} \right),$$

whence the estimate on $d_{\mathrm{TV}}(\mu_n, \sigma_n)$. □

4. One-dimensional applications

In this section, we apply our main Theorems 3.3, 3.6 and 3.11 to various one-dimensional examples, most of them coming from number theory or combinatorics. In each case, we compute the exact asymptotics of the distances for the basic approximation scheme, and for some better approximation schemes, which are derived from the approximation schemes of order $r \geq 1$.

4.1. Poisson approximation of a sum of independent Bernoulli variables.

As in Example 1.5, we consider a sum $X_n = \sum_{j=1}^n \mathcal{B}(p_j)$ of independent Bernoulli random variables, with $\sum_{j=1}^\infty p_j = +\infty$ and $\sum_{j=1}^\infty (p_j)^2 < +\infty$. The corresponding Fourier transforms are:

$$\widehat{\mu_n}(\xi) = \mathrm{e}^{\lambda_n (\mathrm{e}^{\mathrm{i}\xi} - 1)}\, \psi_n(\xi) \quad \text{with } \psi_n(\xi) = \prod_{j=1}^n (1 + p_j(\mathrm{e}^{\mathrm{i}\xi} - 1))\, \mathrm{e}^{-p_j(\mathrm{e}^{\mathrm{i}\xi} - 1)}.$$

It is convenient to have an exact formula for the expansion in Laurent series of the deconvolution residues ψ_n. Denote

$$\mathfrak{p}_{1,n} = 0 \quad \text{and} \quad \mathfrak{p}_{k\geq 2,n} = \sum_{j=1}^n (p_j)^k.$$

One has

$$\begin{aligned}\psi_n(\xi) &= \left(\prod_{j=1}^{n}(1+p_j(\mathrm{e}^{\mathrm{i}\xi}-1))\right)\,\mathrm{e}^{-\sum_{j=1}^{n}p_j(\mathrm{e}^{\mathrm{i}\xi}-1)}\\ &= \exp\left(\sum_{j=1}^{n}\log(1+p_j(\mathrm{e}^{\mathrm{i}\xi}-1))-p_j(\mathrm{e}^{\mathrm{i}\xi}-1)\right)\\ &= \exp\left(\sum_{j=1}^{n}\sum_{k=2}^{\infty}\frac{(-1)^{k-1}}{k}\,(p_j(\mathrm{e}^{\mathrm{i}\xi}-1))^k\right)\\ &= \exp\left(\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\,\mathfrak{p}_{k,n}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^k\right).\end{aligned}$$

Denote $\mathfrak{P}$ the set of integer partitions, that is to say finite non-increasing sequences of positive integers

$$L=(L_1\geq L_2\geq\cdots\geq L_r>0).$$

If L is an integer partition, we denote $|L|=\sum_{i=1}^{r}L_i$, $\ell(L)=r$, and $m_k(L)$ the number of parts of L of size k. We also denote $z_L=\prod_{k\geq 1}k^{m_k(L)}\,m_k(L)!$, which is the size of the centralizer of a permutation of cycle-type L in the symmetric group $\mathfrak{S}_{|L|}$. The previous power series in $\mathrm{e}^{\mathrm{i}\xi}-1$ expands then as

$$\psi_n(\xi)=\sum_{L\in\mathfrak{P}}\frac{(-1)^{|L|-\ell(L)}}{z_L}\,\mathfrak{p}_{L,n}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^{|L|},$$

where $\mathfrak{p}_{L,n}=\prod_{i=1}^{\ell(L)}\mathfrak{p}_{L_i,n}$. Since $\mathfrak{p}_{n,1}=0$, the sum actually runs over integer partitions without part equal to 1. In particular, one has

$$\psi_n(\xi)=1-\frac{\mathfrak{p}_{2,n}}{2}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^2+\frac{\mathfrak{p}_{3,n}}{3}(\mathrm{e}^{\mathrm{i}\xi}-1)^3+o(|\xi|^3).$$

Consider then the basic approximation scheme $(\nu_n^{(0)})_{n\in\mathbb{N}}$ of the sequence of random variables $(X_n)_{n\in\mathbb{N}}$, which is mod-Poisson convergent with parameters $\lambda_n=\sum_{j=1}^{n}p_j$ and limit $\psi(\xi)=\prod_{j=1}^{\infty}(1+p_j(\mathrm{e}^{\mathrm{i}\xi}-1))\,\mathrm{e}^{-p_j(\mathrm{e}^{\mathrm{i}\xi}-1)}$. It satisfies the hypothesis (H1) with $r=1$ and

$$\beta_n=-\frac{1}{2}\,\mathfrak{p}_{2,n}\qquad;\qquad\beta=-\frac{1}{2}\,\mathfrak{p}_2=-\frac{1}{2}\sum_{j=1}^{\infty}(p_j)^2.$$

By Theorems 3.3 and 3.6, if $(Y_n)_{n\in\mathbb{N}}$ is a sequence of Poisson random variables with parameters λ_n, then

$$\begin{aligned}d_{\mathrm{L}}(X_n,Y_n)&=\frac{\mathfrak{p}_2}{2\sqrt{2\pi}(\lambda_n)^{\frac{3}{2}}}+o\left(\frac{1}{(\lambda_n)^{\frac{3}{2}}}\right);\\ d_{\mathrm{K}}(X_n,Y_n)&=\frac{\mathfrak{p}_2}{2\sqrt{2\pi\,\mathrm{e}}\,\lambda_n}+o\left(\frac{1}{\lambda_n}\right).\end{aligned}$$

The sequence of infinitely divisible laws $(\nu_n^{(0)})_{n\in\mathbb{N}}$ also satisfies the hypothesis (H2), since the expansion of $\psi_n(\xi)$ given before is exact and can be derived term by term. So, one can apply Theorem 3.11, and

$$d_{\mathrm{TV}}(X_n,Y_n)=\frac{2\,\mathfrak{p}_2}{\sqrt{2\pi\,\mathrm{e}}\,\lambda_n}+o\left(\frac{1}{\lambda_n}\right).$$

This last estimate should be compared with the unconditional bound

$$d_{\mathrm{TV}}(X_n, Y_n) \leq 2\,\frac{\mathfrak{p}_{2,n}}{\lambda_n} = 2\,\frac{\sum_{j=1}^n (p_j)^2}{\sum_{j=1}^n p_j}$$

stemming from the Chen–Stein method. On the other hand, we recover the asymptotic formula for $d_{\mathrm{TV}}(X_n, Y_n)$ established by Deheuvels and Pfeifer in [DP86], using only tools from harmonic analysis (instead of the semi-group method developed in *loc. cit.*).

We now look at a higher order approximation scheme, namely, the approximation scheme $(\nu_n^{(2)})_{n\in\mathbb{N}}$ of order $r = 2$. It is defined by the Fourier transforms

$$\widehat{\nu_n^{(2)}}(\xi) = \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)}\left(1 - \frac{\mathfrak{p}_{2,n}}{2}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^2\right).$$

Again, the hypotheses (H1) and (H2) are satisfied, this time with $r = 2$ and

$$\beta_n = \frac{1}{3}\,\mathfrak{p}_{3,n} \qquad ; \qquad \beta = \frac{1}{3}\,\mathfrak{p}_3 = \frac{1}{3}\sum_{j=1}^{\infty}(p_j)^3.$$

So, applying the main theorems of Section 3, we obtain

$$d_{\mathrm{L}}(X_n, \nu_n^{(2)}) = \frac{(\sqrt{6}-2)\,\mathfrak{p}_3}{\sqrt{2\pi}\,\mathrm{e}^{3-\sqrt{6}}\,(\lambda_n)^2} + o\left(\frac{1}{(\lambda_n)^2}\right);$$

$$d_{\mathrm{K}}(X_n, \nu_n^{(2)}) = \frac{\mathfrak{p}_3}{3\sqrt{2\pi}\,(\lambda_n)^{\frac{3}{2}}} + o\left(\frac{1}{(\lambda_n)^{\frac{3}{2}}}\right);$$

$$d_{\mathrm{TV}}(X_n, \nu_n^{(2)}) = \frac{2(1+4\mathrm{e}^{-\frac{3}{2}})\,\mathfrak{p}_3}{3\sqrt{2\pi}\,(\lambda_n)^{\frac{3}{2}}} + o\left(\frac{1}{(\lambda_n)^{\frac{3}{2}}}\right).$$

Thus, $(\nu_n^{(2)})_{n\in\mathbb{N}}$ is asymptotically a better approximation scheme than the simple Poisson approximation $(\nu_n^{(0)})_{n\in\mathbb{N}}$. Moreover, if

$$\sum_{j=n+1}^{+\infty}(p_j)^2 = o\left((\lambda_n)^{-\frac{9}{4}}\right) = o\left(\left(\sum_{j=1}^{n} p_j\right)^{-\frac{9}{4}}\right),$$

then all the results of Theorem 3.16 apply, and in the previous estimates one can replace $\nu_n^{(2)}$ by $\sigma_n^{(2)}$, defined by the Fourier transform

$$\widehat{\sigma_n^{(2)}}(\xi) = \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)}\left(1 - \frac{\sum_{j=1}^{\infty}(p_j)^2}{2}\,(\mathrm{e}^{\mathrm{i}\xi}-1)^2\right).$$

with fixed residue.

Remark 4.1. Performing the same computations as in Lemma 3.8, one can give an exact formula for $\nu_n^{(2)}(\{k\})$:

$$\nu_n^{(2)}(\{k\}) = \mathrm{e}^{-\lambda_n}\,\frac{(\lambda_n)^k}{k!}\left(1 + \beta\left(1 - \frac{2k}{\lambda_n} + \frac{k(k-1)}{(\lambda_n)^2}\right)\right),$$

here with $\beta = -\frac{\mathfrak{p}_{2,n}}{2}$. At first sight this might look like an important deformation of the Poisson measure with parameter λ_n, but note that in the range $k = \lambda_n(1 + o(1))$ where

the Poisson law $\nu_n^{(0)} = \mathcal{P}_{(\lambda_n)}$ has its mass concentrated, one has

$$1 - \frac{2k}{\lambda_n} + \frac{k(k-1)}{(\lambda_n)^2} = 1 - 2(1+o(1)) + (1+o(1)) = o(1).$$

More generally, assume that the remainder of the series $\sum_{j=n+1}^{+\infty}(p_j)^2$ is asymptotically smaller than any negative power of λ_n. If

$$\widehat{\sigma_n^{(r)}}(\xi) = \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)} \left(\sum_{k=0}^{r} \mathfrak{e}_k \, (\mathrm{e}^{\mathrm{i}\xi} - 1)^k \right),$$

with

$$\mathfrak{e}_k = \sum_{\substack{L \text{ partition of size } k \\ \text{without part of size 1}}} \frac{(-1)^{|L|-\ell(L)}}{z_L} \, \mathfrak{p}_L \qquad ; \qquad \mathfrak{p}_{k\geq 2} = \sum_{j=1}^{\infty} (p_j)^k,$$

then the sequence of signed measures $(\sigma_n^{(r)})_{n\in\mathbb{N}}$ forms a scheme of approximation of $(X_n)_{n\in\mathbb{N}}$, with

$$d_{\mathrm{L}}(X_n, \sigma_n^{(r)}) = \frac{|\mathfrak{e}_{r+1} \, G_{r+1}(z_{r+2})|}{\sqrt{2\pi}\,(\lambda_n)^{\frac{r}{2}+1}} + o\left(\frac{1}{(\lambda_n)^{\frac{r}{2}+1}}\right);$$

$$d_{\mathrm{K}}(X_n, \sigma_n^{(r)}) = \frac{|\mathfrak{e}_{r+1} \, G_{r}(z_{r+1})|}{\sqrt{2\pi}\,(\lambda_n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right);$$

$$d_{\mathrm{TV}}(X_n, \sigma_n^{(r)}) = \frac{\int_{\mathbb{R}} |\mathfrak{e}_{r+1} \, G_{r+1}(\alpha)| \, d\alpha}{\sqrt{2\pi}\,(\lambda_n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

Thus, one can in this setting approximate the law of X_n up to any order (negative power of λ_n).

4.2. **Poisson approximation schemes and formal alphabets.** The description of the approximation schemes and of the asymptotic estimates for the distances by means of the parameters $\mathfrak{p}_k$ and $\mathfrak{e}_k$ is actually possible for many other examples of discrete random variables. Thus, the combinatorics of approximation schemes can be encoded in the *algebra of symmetric functions* Sym. From an algebraic point of view, the language of λ-rings and *formal alphabets* make this description even more stunning. Denote Sym the algebra of symmetric functions, that is to say symmetric polynomials in an infinity of variables $\{x_1, x_2, x_3, \ldots\}$; we refer to [Mac95, Chapter 1] or [Mél17, Chapter 2] for the whole discussion of this paragraph. The algebra Sym admits for algebraic basis over $\mathbb{C}$ the power sums

$$\mathfrak{p}_{k\geq 1} = \sum_i (x_i)^k$$

and the elementary functions

$$\mathfrak{e}_{k\geq 1} = \sum_{i_1<i_2<\cdots<i_k} x_{i_1} x_{i_2} \cdots x_{i_k};$$

thus, $\mathrm{Sym} = \mathbb{C}[\mathfrak{p}_1, \mathfrak{p}_2, \ldots] = \mathbb{C}[\mathfrak{e}_1, \mathfrak{e}_2, \ldots]$. The two bases are related by the identity of formal series

$$1 = \left(\prod_i (1 - t x_i)\right) \exp\left(\sum_i \log \frac{1}{1 - t x_i}\right) = \left(1 + \sum_{k=1}^{\infty} \mathfrak{e}_k \, t^k\right) \exp\left(\sum_{k=1}^{\infty} \frac{\mathfrak{p}_k}{k} \, t^k\right),$$

which leads to the formula

$$\mathfrak{e}_k = \sum_{L \in \mathfrak{P}_k} \frac{(-1)^{|L|-\ell(L)}}{z_L} \mathfrak{p}_L$$

where $\mathfrak{P}_k$ is the set of integer partitions of size k, and $\mathfrak{p}_L = \prod_{i=1}^{\ell(L)} \mathfrak{p}_{L_i}$. This is a particular case of the Frobenius–Schur formula for Schur functions [Mél17, Theorem 2.32].

A formal alphabet, or specialisation is a morphism of algebras $\psi : \mathrm{Sym} \to \mathbb{C}$. In particular, if $A = \{a_1, a_2, \ldots\}$ is a set of complex numbers with $\sum_i |a_i| < +\infty$, then the map

$$\mathfrak{p}_k \mapsto \mathfrak{p}_k(A) = \sum_i (a_i)^k$$

can be extended to a morphism of algebras $\mathrm{Sym} \to \mathbb{C}$, hence a specialisation of Sym (in this case one can speak of a "true" alphabet for A). Beware that some specialisations are not of this form: indeed, a formal alphabet can send the power sums $\mathfrak{p}_k$ to an arbitrary set of values in $\mathbb{C}$. Still, we can agree to use the notation $f \mapsto f(A)$ for any specialisation $f \mapsto \psi(f)$ of the algebra Sym. In particular, suppose that $A = \{a_1, a_2, \ldots\}$ is a set of complex numbers, but this time with $\sum_i |a_i| = +\infty$ and $\sum_i |a_i|^2 < +\infty$. We then denote $f \mapsto f(A)$ the specialisation

$$\mathfrak{p}_1(A) = 0 \qquad ; \qquad \mathfrak{p}_{k\geq 2}(A) = \sum_i (a_i)^k.$$

This definition extends indeed to a unique morphism of algebras $\mathrm{Sym} \to \mathbb{C}$, although there is *a priori* no underlying "true" alphabet. Given two formal alphabets A and B, one can define new formal alphabets $\varepsilon(A)$ and $A + B$, as follows:

$$\mathfrak{p}_k(\varepsilon(A)) = (-1)^{k-1}\, \mathfrak{p}_k(A) \qquad ; \qquad \mathfrak{p}_k(A + B) = \mathfrak{p}_k(A) + \mathfrak{p}_k(B).$$

As before, these definitions extend uniquely to new morphisms of algebras $\mathrm{Sym} \to \mathbb{C}$. If A and B were true alphabets, then $A + B$ is the true alphabet which is the disjoint union of A and B. On the other hand, ε is related to a certain involution of the Hopf algebra of symmetric functions.

Let us now relate the theory of symmetric functions to the approximation schemes of discrete distributions. Consider a sequence of random integer-valued variables $(X_n)_{n\in\mathbb{N}}$, whose Fourier transforms write as:

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\xi X_n}] = \mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)} \exp\left(\sum_{k=2}^{\infty} \frac{c_{k,n}}{k!} (\mathrm{e}^{\mathrm{i}\xi} - 1)^k\right).$$

The coefficients $\lambda_n = c_{1,n}$ and $c_{k\geq 2,n}$ are called the *factorial cumulants* of X_n, and the Poisson random variables are characterized by the vanishing of their factorial cumulants of order $k \geq 2$ (see [DVJ02, Section 5.2]). We associate to these coefficients the following specialisation of Sym:

$$\mathfrak{p}_1(A_n) = 0 \qquad ; \qquad \mathfrak{p}_{k\geq 2}(A_n) = \frac{(-1)^{k-1}\, c_{k,n}}{(k-1)!}.$$

Then, the deconvolution residue $\psi_n(\xi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\xi X_n}]\, \mathrm{e}^{-\lambda_n(\mathrm{e}^{\mathrm{i}\xi}-1)}$ admits the expansion:

$$\psi_n(\xi) = \sum_{k=0}^{\infty} \mathfrak{e}_k(A_n)\, (\mathrm{e}^{\mathrm{i}\xi} - 1)^k,$$

with $\mathfrak{e}_0(A_n) = 1$ and $\mathfrak{e}_1(A_n) = 0$. The approximation scheme of order $r \geq 1$ is given by the residue

$$\chi_n^{(r)}(\xi) = \sum_{k=0}^{r} \mathfrak{e}_k(A_n)\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^k.$$

The reason why this formalism is convenient is that, for all the examples that we shall look at in this section, the limiting residue

$$\psi(\xi) = \sum_{k=0}^{\infty} \mathfrak{e}_k(A)\,(\mathrm{e}^{\mathrm{i}\xi} - 1)^k,$$

corresponds to a formal alphabet A which is explicit and which expresses in terms of "natural" parameters of the random model. These formal alphabets are given by Table 1.

X_n	formal alphabet A
sum of independent Bernoulli variables with parameters p_j (§4.1)	$\{p_1, p_2, p_3 \ldots\}$
number of cycles of a random uniform permutation (§4.3)	$\mathbb{N}^{*-1}$
number of cycles of a random Ewens permutation with parameter θ (§4.3)	Θ
number of connected components of a uniform random map (§4.4)	$(2\mathbb{N}+1)^{-1}$
number of distinct irreducible factors of a random monic polynomial (§4.4)	$\mathbb{N}^{*-1} + (q^{\deg \mathfrak{I}})^{-1}$
number of irreducible factors of a random monic polynomial (§4.4)	$\mathbb{N}^{*-1} + \varepsilon((q^{\deg \mathfrak{I}} - 1)^{-1})$
number of distinct prime divisors of a random integer (§4.5)	$\mathbb{N}^{*-1} + \mathbb{P}^{-1}$

TABLE 1. Formal alphabets associated to the mod-Poisson convergent sequences.

In this table we denote $A^{-1} = \{\frac{1}{a},\ a \in A\}$ if A is a true alphabet, and

$$\Theta = \left\{1, \frac{\theta}{\theta+1}, \frac{\theta}{\theta+2}, \ldots\right\};$$
$$q^{\deg \mathfrak{I}} = \left\{q^{\deg P},\ P \in \mathfrak{I}\right\};$$
$$(q^{\deg \mathfrak{I}} - 1) = \left\{q^{\deg P} - 1,\ P \in \mathfrak{I}\right\}$$

where $\mathfrak{I}$ is the set of irreducible polynomials over the finite field $\mathbb{F}_q$. Therefore, the asymptotic behavior of a sequence $(X_n)_{n\in\mathbb{N}}$ in Table 1 is entirely encoded in:

- a sequence of parameters $(\lambda_n)_{n\in\mathbb{N}}$,
- and a formal alphabet A which in some sense captures the geometric and arithmetic properties of the model.

In particular, the reader should notice the appearance of two contributions to the formal alphabet as soon as one studies the problem of factorization of a random element in a factorial ring.

4.3. **Number of disjoint cycles of a random permutation.** In this section, we study the number ℓ_n of disjoint cycles of a random permutation $\sigma_n \in \mathfrak{S}_n$. If the permutation σ_n is chosen according to the uniform probability measure (Example 1.6), then the discussion of Section 4.1 applies, with the parameters p_j taken equal to $\frac{1}{j}$, and

$$\lambda_n = \sum_{j=1}^{n} \frac{1}{j} = H_n \simeq \log n;$$

$$\sum_{j=n+1}^{+\infty} (p_j)^2 \simeq \frac{1}{n} = o\left(\frac{1}{(\lambda_n)^r}\right) \quad \text{for every } r \geq 0.$$

Therefore, if one looks for instance at the derived scheme with constant residue from the approximation scheme of order $r = 2$:

$$\widehat{\sigma_n^{(2)}}(\xi) = \mathrm{e}^{H_n(\mathrm{e}^{\mathrm{i}\xi}-1)} \left(1 - \frac{\pi^2}{12}(\mathrm{e}^{\mathrm{i}\xi}-1)^2\right)$$

then

$$d_{\mathrm{L}}(\ell_n, \sigma_n^{(2)}) = \frac{(\sqrt{6}-2)\,\zeta(3)}{\sqrt{2\pi\,\mathrm{e}^{3-\sqrt{6}}}\,(\log n)^2} + o\left(\frac{1}{(\log n)^2}\right),$$

and one has similar estimates for the Kolmogorov distance and for the total variation distance.

One can extend this result to more general models of random permutations, namely, random permutations $\sigma_n \in \mathfrak{S}_n$ under the so-called *generalized weighted measures*

$$\mathbb{P}_{\Theta,n}[\sigma_n] = \frac{1}{n!\,h_n(\Theta)} \prod_{k=1}^{n} (\theta_k)^{m_k(\sigma_n)},$$

where $\Theta = (\theta_k)_{k\geq 1}$ is a sequence of non-negative parameters, $m_k(\sigma)$ is the number of cycles of length k in $\sigma \in \mathfrak{S}_n$, and $h_n(\Theta)$ is the normalization constant so that each $\mathbb{P}_{\Theta,n}$ is a probability measure on $\mathfrak{S}_n$. If $\Theta = (\theta, \theta, \ldots)$ is a constant sequence, then one recovers the Ewens measures of parameter θ:

$$\mathbb{P}_n[\sigma_n] = \frac{\theta^{\ell(\sigma_n)}}{\theta(\theta+1)\cdots(\theta+n-1)}.$$

The generalized weighted measures have been studied previously in [BU09, BU11, BUV11, EU12, NZ13], and they are related to the quantum Bose gas of statistical mechanics. Denote as before $\ell_n = \ell(\sigma_n) = \sum_{k=1}^{n} m_k(\sigma_n)$ the number of disjoint cycles of a random permutation under the measure $\mathbb{P}_{\Theta,n}$. We also introduce the generating series of the parameter Θ:

$$g_\Theta(z) = \sum_{k=1}^{\infty} \frac{\theta_k}{k} z^k.$$

Note that $\exp(g_\Theta(z)) = \sum_{n=0}^{\infty} h_n(\Theta)\, z^n$ is the generating series of the partition functions of the models.

If Θ is the constant sequence equal to θ, then $g_\Theta(z) = \theta \log(\frac{1}{1-z})$. On the other hand, in this setting, one can represent ℓ_n as a sum of independent Bernoulli variables (Feller coupling):

$$\ell_n = \sum_{j=1}^{n} \mathcal{B}\left(\frac{\theta}{\theta+j-1}\right).$$

We can then use the discussion of §4.1, and the mod-Poisson convergence of $(\ell_n)_{n\in\mathbb{N}}$ with parameters $\lambda_n = \sum_{j=1}^{n} \frac{\theta}{\theta+j-1} \simeq \theta \log n$ and limiting residue

$$\psi(\xi) = \prod_{j=1}^{\infty} \left(1 + \frac{\theta}{\theta+j-1}\,(\mathrm{e}^{\mathrm{i}\xi}-1)\right) \mathrm{e}^{-\frac{\theta}{\theta+j-1}\,(\mathrm{e}^{\mathrm{i}\xi}-1)} = \mathrm{e}^{\gamma\theta(\mathrm{e}^{\mathrm{i}\xi}-1)} \frac{\Gamma(\theta)}{\Gamma(\theta \mathrm{e}^{\mathrm{i}\xi})}.$$

The last formula relies on the infinite product representation $\frac{1}{\Gamma(z+1)} = \mathrm{e}^{\gamma z} \prod_{k=1}^{\infty} (1+\frac{z}{k})\,\mathrm{e}^{-\frac{z}{k}}$.

More generally, the article [NZ13] shows that if one has a good understanding of the analytic properties of the generating series $g_\Theta(z)$, then one can establish the mod-Poisson convergence of $(\ell_n)_{n\in\mathbb{N}}$, with a limiting residue $\psi(\xi)$ similar to the previous expressions. Thus:

Theorem 4.2 ([NZ13], Lemma 4.1)**.** *Assume that $g_\Theta(z)$ is holomorphic on a domain*

$$\Delta_0(r, R, \phi) = \{z \in \mathbb{C},\ |z| < R,\ z \neq r,\ |\arg(z-r)| > \phi\}$$

with $0 < r < R$ and $\phi \in (0, \frac{\pi}{2})$, see Figure 3.

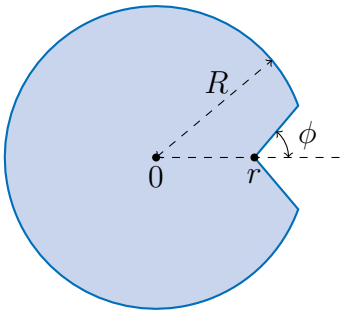

FIGURE 3. Domain of analyticity of the generating series $g_\Theta(z)$.

Suppose moreover that the singularity of g_Θ at $z = r$ is logarithmic:

$$g_\Theta(z) = \theta \log\left(\frac{1}{1-\frac{z}{r}}\right) + K + O(|z-r|).$$

Then, one has the asymptotics

$$\mathbb{E}_{\Theta,n}\left[\mathrm{e}^{z\ell_n}\right] = \mathrm{e}^{(\theta\log n + K)(\mathrm{e}^z-1)} \left(\frac{\Gamma(\theta)}{\Gamma(\theta \mathrm{e}^z)} + O\left(\frac{1}{n}\right)\right),$$

with a uniform remainder on compact subsets of $\mathbb{C}$. Hence, $(\ell_n)_{n\in\mathbb{N}}$ converges mod-Poisson with parameters $\lambda_n = \theta \log n + K$ and limit

$$\psi(\xi) = \frac{\Gamma(\theta)}{\Gamma(\theta \mathrm{e}^{\mathrm{i}\xi})}.$$

Moreover, the convergence of residues $\psi_n \to \psi$ happens in every space $\mathscr{C}^r$, with a norm $\|\psi_n - \psi\|_{\mathscr{C}^r}$ which is each time $O(\frac{1}{n})$.

Note that the last part of the theorem (convergence in every space $\mathscr{C}^r$) is an immediate consequence of the estimate of $\mathbb{E}_{\Theta,n}[\mathrm{e}^{z\ell_n}]$, and of the analyticity of all terms in

$$\mathbb{E}_{\Theta,n}[\mathrm{e}^{\mathrm{i}\xi\ell_n}]\,\mathrm{e}^{-(\theta\log n+K)(\mathrm{e}^{\mathrm{i}\xi}-1)} = \psi_n(\xi).$$

On the other hand, in the case of the Ewens measure, one has $K = 0$ and $r = 1$, and in the case of the uniform measure, one has moreover $\theta = 1$.

Because of the infinite product representation of $\frac{\Gamma(\theta)}{\Gamma(\theta e^{i\xi})}$, one can restate Theorem 4.2 as follows. If one looks at a generalized weighted measure with a generating series $g_\Theta(z)$ that satisfies the hypotheses of Theorem 4.2, then

$$\mathbb{E}[e^{i\xi\ell_n}] = e^{(\theta(\log n+\gamma)+K)(e^{i\xi}-1)} \left(\prod_{j=1}^{\infty} \left(1 + \frac{\theta(e^{i\xi}-1)}{\theta+j-1}\right) e^{-\frac{\theta(e^{i\xi}-1)}{\theta+j-1}} + O\left(\frac{1}{n}\right)\right).$$

Therefore, in this setting, one can deal with the asymptotics of $(\ell_n)_{n\in\mathbb{N}}$ in the same way as in §4.1, but with parameters $\lambda_n = \theta H_n + K$, and

$$\mathfrak{p}_1(\Theta) = 0 \qquad ; \qquad \mathfrak{p}_{k\geq 2}(\Theta) = \sum_{j=1}^{\infty} \left(\frac{\theta}{\theta+j-1}\right)^k. \tag{Θ}$$

Thus, one has the following:

Theorem 4.3. *We consider a generalized weighted measure $\mathbb{P}_{\Theta,n}$, with a generating series $g_\Theta(z)$ which satisfies the hypotheses of Theorem 4.2. For any $r \geq 1$, we introduce the scheme of approximation $(\sigma_n^{(r)})_{n\in\mathbb{N}}$, defined by the Fourier transforms*

$$\widehat{\sigma_n^{(r)}}(\xi) = e^{(\theta H_n+K)(e^{i\xi}-1)} \left(\sum_{k=0}^{r} \mathfrak{e}_k(\Theta)\,(e^{i\xi}-1)^k \right),$$

where Θ is the specialisation of Sym *specified by Equation (Θ). Note that the residue is equal to 1 for $r = 1$ (basic scheme of approximation). One has the asymptotics:*

$$d_{\mathrm{L}}(\ell_n, \sigma_n^{(r)}) = \frac{|\mathfrak{e}_{r+1}(\Theta)\, G_{r+1}(z_{r+2})|}{\sqrt{2\pi}\,(\theta\log n)^{\frac{r}{2}+1}} + o\left(\frac{1}{(\log n)^{\frac{r}{2}+1}}\right);$$

$$d_{\mathrm{K}}(\ell_n, \sigma_n^{(r)}) = \frac{|\mathfrak{e}_{r+1}(\Theta)\, G_r(z_{r+1})|}{\sqrt{2\pi}\,(\theta\log n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\log n)^{\frac{r+1}{2}}}\right);$$

$$d_{\mathrm{TV}}(\ell_n, \sigma_n^{(r)}) = \frac{\int_{\mathbb{R}} |\mathfrak{e}_{r+1}(\Theta)\, G_{r+1}(\alpha)|\, d\alpha}{\sqrt{2\pi}\,(\theta\log n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\log n)^{\frac{r+1}{2}}}\right).$$

4.4. **Mod-Poisson convergence and algebraico-logarithmic singularities.** The discussion of §4.3 can be extended to any statistic of a random combinatorial object, whose double generating series admits a certain form of singularity. Suppose given a combinatorial class, that is to say a sequence $\mathfrak{C} = (\mathfrak{C}_n)_{n\in\mathbb{N}}$ of finite sets. We write $\mathrm{card}\,\mathfrak{C}_n = |\mathfrak{C}_n|$. In this paragraph, we shall in particular study the following two examples, *cf.* [Hwa96, §5, Example 2] and [FS09, p. 449 and p. 671]:

(1) $\mathfrak{F}_n = \mathcal{F}(\llbracket 1,n\rrbracket, \llbracket 1,n\rrbracket)$ is the set of maps from the finite set $\llbracket 1,n\rrbracket$ to itself. It has cardinality n^n.

(2) $\mathfrak{P}_n = (\mathbb{F}_q[t])_n$ is the set of monic polynomials of degree n with coefficients in the finite field $\mathbb{F}_q$, where $q = p^e$ is some prime power. It has cardinality q^n.

This is mainly to fix the ideas, and the techniques hereafter can be applied to many other examples from the aforementioned sources. We consider a statistic $X : \bigsqcup_{n\in\mathbb{N}} \mathfrak{C}_n \to \mathbb{N}$, and we denote by X_n the random variable obtained by evaluating X on an element of $\mathfrak{C}_n$ chosen uniformly at random. We introduce the bivariate generating function

$$F(z,w) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{c\in\mathfrak{C}_n} w^{X(c)} = \sum_{n=0}^{\infty} \frac{z^n \,(\text{card }\mathfrak{C}_n)}{n!}\, \mathbb{E}[w^{X_n}]$$
$$\text{or } F(z,w) = \sum_{n=0}^{\infty} z^n \sum_{c\in\mathfrak{C}_n} w^{X(c)} = \sum_{n=0}^{\infty} z^n \,(\text{card }\mathfrak{C}_n)\, \mathbb{E}[w^{X_n}].$$

Suppose that for any w, $F(z,w)$ is holomorphic on a domain $\Delta_0(r(w), R(w), \phi)$, and admits an *algebraico-logarithmic singularity* at $z = r(w)$, that writes as

$$F(z,w) = K(w) \left(\frac{1}{1-\frac{z}{r(w)}}\right)^{\alpha(w)} \left(\log \frac{1}{1-\frac{z}{r(w)}}\right)^{\beta(w)} (1+o(1)).$$

Then, the coefficient $c_n(w)$ of z^n in $F(z,w)$ has for asymptotics:

$$c_n(w) = \frac{K(w)}{\Gamma(\alpha(w))}\, (r(w))^{-n}\, n^{\alpha(w)-1}\, (\log n)^{\beta(w)} \left(1 + O\left(\frac{1}{\log n}\right)\right),$$

see [FO90b]. Moreover, the $O(\frac{1}{\log n})$ is actually a $O(\frac{1}{n})$ if $\beta(w) = 0$, or if $\alpha(w) = 0$ and $\beta(w) = 1$; *cf.* [Hwa98]. If w stays in a sufficiently small compact subset of $\mathbb{C}$, then one can take a uniform constant in the $O(\cdot)$, which leads to an asymptotic formula for the expectations $\mathbb{E}[w^{X_n}]$:

$$\mathbb{E}[w^{X_n}] = \frac{c_n(w)}{c_n(1)} = \frac{K(w)}{K(1)}\, \frac{\Gamma(\alpha(1))}{\Gamma(\alpha(w))} \left(\frac{r(1)}{r(w)}\right)^n n^{\alpha(w)-\alpha(1)}\, (\log n)^{\beta(w)-\beta(1)} \left(1 + O\left(\frac{1}{\log n}\right)\right),$$

again with a smaller remainder $O(\frac{1}{n})$ if $\beta(w) = 0$ for all w. In this paragraph, we shall concentrate on this particular case; the case $\beta(w) \neq 0$ is more commonly observed in probabilistic number theory, see our next paragraph 4.5.

Example 4.4. If $f \in \mathfrak{F}_n = \mathcal{F}(\llbracket 1,n\rrbracket, \llbracket 1,n\rrbracket)$, the *functional graph* corresponding to f is the directed graph $G(f)$ with vertex set $\llbracket 1,n\rrbracket$, and with edges $(k, f(k))$, $k \in \llbracket 1,n\rrbracket$. Since f is a map, $G(f)$ is a disjoint union of cycles on which trees are grafted, see [FO90a] and Figure 4. We denote $X(f)$ the number of connected components of the functional graph $G(f)$. This generalizes the notion of cycle of a permutation. If $\mathfrak{T} = (\mathfrak{T}_n)_{n\in\mathbb{N}}$ is the combinatorial class of unordered rooted labeled trees, recall that its generating function $T(z) = \sum_{n=1}^{\infty} \frac{z^n}{n!} |\mathfrak{T}_n|$ is the solution of $T(z) = z\mathrm{e}^{T(z)}$, because a rooted tree is constructed recursively by taking a node and connecting to it a set of trees (also, Cayley's theorem gives card $\mathfrak{T}_n = n^{n-1}$).

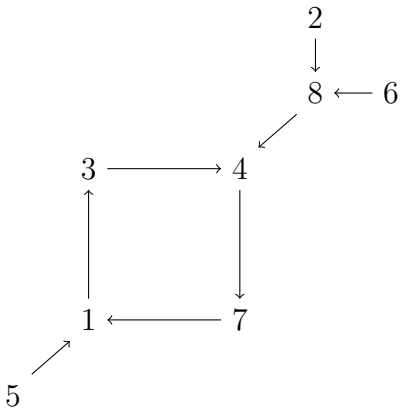

FIGURE 4. The functional graph associated to the function $f : [\![1,8]\!] \to [\![1,8]\!]$ with word 38471814, here with one connected component.

The class $\mathfrak{C} = (\mathfrak{C}_n)_{n\in\mathbb{N}}$ of cycles on such trees has generating series

$$\begin{aligned} C(z) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} |\mathfrak{C}_n| \\ &= \sum_{n=0}^{\infty} \sum_{k_1+\cdots+k_r=n} \frac{|\mathfrak{T}_{k_1}|\,|\mathfrak{T}_{k_2}|\cdots|\mathfrak{T}_{k_r}|}{r\,n!} \binom{n}{k_1,\ldots,k_r} z^n \\ &= \sum_{r=1}^{\infty} \frac{1}{r} \sum_{k_1,\ldots,k_r\geq 1} \frac{|\mathfrak{T}_{k_1}|\cdots|\mathfrak{T}_{k_r}|}{(k_1)!\cdots(k_r)!}\, z^{k_1+\cdots+k_r} \\ &= \sum_{r=1}^{\infty} \frac{(T(z))^r}{r} = \log\frac{1}{1-T(z)}. \end{aligned}$$

On the second line of this computation, the multinomial coefficient comes from the choice of the repartition of the integers of $[\![1,n]\!]$ into the r different trees, and the factor $\frac{1}{r}$ comes from the fact that r distinct cyclic permutations of the choices yield the same cycle of trees. Finally, a functional graph is a set of cycles of trees, so,

$$F(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} |\mathfrak{F}_n| = \exp(C(z)) = \frac{1}{1-T(z)}.$$

Moreover, counting connected components of functional graphs amounts to count a factor w for each cycle, hence,

$$F(z,w) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{f\in\mathfrak{F}_n} w^{X(f)} = \exp(wC(z)) = \left(\frac{1}{1-T(z)}\right)^w.$$

The generating series of rooted trees is classically known to have radius of convergence $\frac{1}{\mathrm{e}}$, and a square-root type singularity at $z = \frac{1}{\mathrm{e}}$, see [FO90a, Formula (11)]:

$$T(z) = 1 - \sqrt{2(1-\mathrm{e}z)} + O(1-\mathrm{e}z).$$

Therefore,

$$F(z,w) = 2^{-\frac{w}{2}} \left(\frac{1}{1-\mathrm{e}z}\right)^{\frac{w}{2}} (1+o(1)),$$

and the previous discussion applies with $K(w) = 2^{-\frac{w}{2}}$, $\alpha(w) = \frac{w}{2}$ and $\beta(w) = 0$. Consequently,

$$\mathbb{E}[w^{X_n}] = \mathrm{e}^{\frac{1}{2}\log\frac{n}{2}\,(w-1)}\,\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{w}{2})}\left(1+O\left(\frac{1}{n}\right)\right),$$

with a remainder that is uniform if w remains on the unit circle. We shall next interpret this result as a mod-Poisson convergence, and deduce from it the construction of approximation schemes.

Example 4.5. If $P \in \mathfrak{P}_n = (\mathbb{F}_q[t])_n$, then it writes uniquely as a product of monic irreducible polynomials, each of these monic irreducible polynomials appearing with a certain multiplicity. In terms of generating series, introducing the combinatorial class $\mathfrak{I} = (\mathfrak{I}_n)_{n\in\mathbb{N}}$ of monic irreducibles, this can be rewritten as:

$$\begin{aligned}P(z) &= \sum_{P\in\mathfrak{P}} z^{\deg P} = \prod_{P\in\mathfrak{I}}(1+z^{\deg P}+z^{2\deg P}+\cdots) = \prod_{P\in\mathfrak{I}}\frac{1}{1-z^{\deg P}}\\ &= \exp\left(\sum_{P\in\mathfrak{I}}\log\left(\frac{1}{1-z^{\deg P}}\right)\right) = \exp\left(\sum_{k=1}^{\infty}\sum_{P\in\mathfrak{I}}\frac{z^{k\deg P}}{k}\right) = \exp\left(\sum_{k=1}^{\infty}\frac{I(z^k)}{k}\right),\end{aligned}$$

where $P(z) = \sum_{n\geq 0}|\mathfrak{P}_n|\,z^n$ and $I(z) = \sum_{n\geq 1}|\mathfrak{I}_n|\,z^n$. Similarly, if $Y(P)$ and $Z(P)$ are respectively the number of distinct irreducible factors and the number of irreducible factors counted with multiplicity for P, then

$$\begin{aligned}P_Y(z,w) &= \sum_{P\in\mathfrak{P}} w^{Y(P)}\,z^{\deg P} = \prod_{P\in\mathfrak{I}}\frac{1+(w-1)z^{\deg P}}{1-z^{\deg P}}\\ &= \exp\left(\sum_{k=1}^{\infty}\sum_{P\in\mathfrak{I}}\frac{z^{k\deg P}}{k} - \frac{((1-w)z^{\deg P})^k}{k}\right) = \exp\left(\sum_{k=1}^{\infty}\frac{I(z^k)\,(1-(1-w)^k)}{k}\right)\end{aligned}$$

and

$$\begin{aligned}P_Z(z,w) &= \sum_{P\in\mathfrak{P}} w^{Z(P)}\,z^{\deg P} = \prod_{P\in\mathfrak{I}}\frac{1}{1-wz^{\deg P}}\\ &= \exp\left(\sum_{k=1}^{\infty}\sum_{P\in\mathfrak{I}}\frac{w^k z^{k\deg P}}{k}\right) = \exp\left(\sum_{k=1}^{\infty}\frac{w^k\,I(z^k)}{k}\right).\end{aligned}$$

Note that $P(z)$ is simply equal to $\sum_{n=0}^{\infty} q^n z^n = \frac{1}{1-qz}$. On the other hand, the number $|\mathfrak{I}_n|$ of irreducible monic polynomials of degree n is given by Gauss' formula

$$|\mathfrak{I}_n| = \frac{1}{n}\sum_{d|n}\mu\left(\frac{n}{d}\right)q^d,$$

as can be seen by gathering the elements of $\mathbb{F}_{q^n}$ according to their irreducible polynomials over $\mathbb{F}_q$ (here, μ is the arithmetic Möbius function). Consequently,

$$I(z) = \sum_{n\geq 1}|\mathfrak{I}_n|\,z^n = \sum_{k\geq 1}\sum_{d\geq 1}\mu(k)\,\frac{q^d\,z^{dk}}{dk} = \sum_{k\geq 1}\frac{\mu(k)}{k}\,\log\left(\frac{1}{1-qz^k}\right).$$

We split $I(z)$ into two parts: the first term $k=1$ and the remainder

$$R(z) = \sum_{k\geq 2}\frac{\mu(k)}{k}\,\log\left(\frac{1}{1-qz^k}\right).$$

The remainder $R(z)$ is an analytic function in z on the open disk of radius $q^{-1/2}$, whereas the first term $\log \frac{1}{1-qz}$ is analytic on the smaller open disk of radius q^{-1}. It follows that

$$P_Y(z,w) = \left(\frac{1}{1-qz}\right)^w \exp\left(wR(q^{-1}) + \sum_{k\geq 2} \frac{I(q^{-k})\,(1-(1-w)^k)}{k}\right)(1+o(1))$$
$$P_Z(z,w) = \left(\frac{1}{1-qz}\right)^w \exp\left(wR(q^{-1}) + \sum_{k\geq 2} \frac{I(q^{-k})\,w^k}{k}\right)(1+o(1))$$

in the neighborhood of the singularity $z = q^{-1}$. These algebraic singularities with exponents $\alpha(w) = w$ and $\beta(w) = 0$ lead to the asymptotic formulas

$$\mathbb{E}[w^{Y_n}] = \mathrm{e}^{(\log n + R(q^{-1}))(w-1)}\,\frac{1}{\Gamma(w)} \exp\left(\sum_{k\geq 2} \frac{(-1)^{k-1}\,I(q^{-k})}{k}\,(w-1)^k\right)\left(1+O\left(\frac{1}{n}\right)\right);$$
$$\mathbb{E}[w^{Z_n}] = \mathrm{e}^{(\log n + R(q^{-1}))(w-1)}\,\frac{1}{\Gamma(w)} \exp\left(\sum_{k\geq 2} \frac{I(q^{-k})}{k}\,(w^k-1)\right)\left(1+O\left(\frac{1}{n}\right)\right).$$

Thus, if a statistic X of a random combinatorial object has a double generating series that admits an algebraico-logarithmic singularity, then in many cases one can deduce from it the mod-Poisson convergence of $(X_n)_{n\in\mathbb{N}}$. We restate hereafter the previous computations in this framework. We note $\varphi(D)$ the Euler function (number of integers in $[\![1,D]\!]$ that are coprime with D), which satisfies the identity $\varphi(D) = \sum_{m|D} \mu(m)\,\frac{D}{m}$; and

$$S(z) = \sum_{k\geq 2} \frac{\varphi(k)}{k}\,\log\left(\frac{1}{1-qz^k}\right),$$

which is analytic on the open disk of radius $q^{-1/2}$.

Theorem 4.6. *Let X_n be the number of components of a random map in $\mathcal{F}([\![1,n]\!],[\![1,n]\!])$, and Y_n and Z_n be the numbers of irreducible factors of a random monic polynomial in $(\mathbb{F}_q[t])_{n\in\mathbb{N}}$, counted respectively without and with multiplicity. The sequences of random variables $(X_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ converge mod-Poisson with parameters*

$$\lambda_n^X = \frac{1}{2}\,(\log 2n + \gamma);$$
$$\lambda_n^Y = \log n + R(q^{-1}) + \gamma;$$
$$\lambda_n^Z = \log n + S(q^{-1}) + \gamma.$$

and limiting functions $\psi^{X/Y/Z}(\xi) = \sum_{k=0}^{\infty} \mathfrak{e}_k(X/Y/Z)\,(\mathrm{e}^{\mathrm{i}\xi}-1)^k$, where the parameters $\mathfrak{e}_k(X)$, $\mathfrak{e}_k(Y)$ and $\mathfrak{e}_k(Z)$ correspond to the specialisations of Sym*:*

$$\mathfrak{p}_{k\geq 2}(X) = \sum_{n\geq 1} \frac{1}{(2n-1)^k} \quad ; \quad \mathfrak{p}_1(X) = 0 \quad ;$$
$$\mathfrak{p}_{k\geq 2}(Y) = \zeta(k) + I(q^{-k}) \quad ; \quad \mathfrak{p}_1(Y) = 0 \quad ;$$
$$\mathfrak{p}_{k\geq 2}(Z) = \zeta(k) + (-1)^{k-1} \sum_{n\geq 1} \frac{|\mathfrak{I}_n|}{(q^n-1)^k} \quad ; \quad \mathfrak{p}_1(Z) = 0.$$

Moreover, in each case, one has $\psi_n(\xi) - \psi(\xi) = O(\frac{1}{n})$ uniformly on the cycle.

Proof. Each time we set $w = \mathrm{e}^{\mathrm{i}\xi}$ in the previous calculations, and we also rewrite the residue so that it does not involve a power of $w-1$ equal to 1.

(1) Number of connected components of a random map. One can write the infinite products

$$\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{w}{2})} = \mathrm{e}^{\frac{\gamma}{2}(w-1)} \prod_{n=1}^{\infty} \frac{1+\frac{\frac{w}{2}-1}{n}}{1-\frac{1}{2n}} \mathrm{e}^{-\frac{w-1}{2n}} = \mathrm{e}^{\frac{\gamma}{2}(w-1)} \prod_{n=1}^{\infty} \left(1+\frac{w-1}{2n-1}\right) \mathrm{e}^{-\frac{w-1}{2n}}$$
$$= \mathrm{e}^{(\frac{\gamma}{2}+\log 2)(w-1)} \prod_{n=1}^{\infty} \left(1+\frac{w-1}{2n-1}\right) \mathrm{e}^{-\frac{w-1}{2n-1}}$$

by using the identity $\sum_{n=1}^{\infty} \frac{1}{2n-1} - \frac{1}{2n} = \log 2$. In the last infinite product, one recognizes the same form as in §4.1, with parameters $p_j = \frac{1}{2j-1}$.

(2) Number of distinct irreducible factors of a random monic polynomial. One rewrites the asymptotic formula for $\mathbb{E}[w^{Y_n}]$:

$$\mathrm{e}^{(\log n + R(q^{-1})+\gamma)(w-1)} \left(\prod_{n=1}^{\infty} \left(1+\frac{w-1}{n}\right) \mathrm{e}^{-\frac{w-1}{n}}\right) \left(\prod_{k\geq 2} \frac{(-1)^{k-1} I(q^{-k})}{k} (w-1)^k\right)$$
$$= \mathrm{e}^{(\log n + R(q^{-1})+\gamma)(w-1)} \left(\prod_{k\geq 2} \frac{(-1)^{k-1}(w-1)^k}{k} \left(\sum_{n=1}^{\infty} \frac{1}{n^k} + I(q^{-k})\right)\right).$$

Again we recognize the formula $\exp\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1}(w-1)^k}{k} \mathfrak{p}_k\right) = \sum_{k=0}^{\infty} \mathfrak{e}_k (w-1)^k$.

(3) Number of irreducible factors of a random monic polynomial, counted with multiplicity. The factor $\Gamma(w)^{-1}$ is dealt with exactly as in the previous case, so we only have to deal with $\exp(\sum_{k\geq 2} \frac{I(q^{-k})}{k}(w^k-1))$. However,

$$\sum_{k\geq 2} \frac{I(q^{-k})}{k}(w^k-1) = \sum_{k\geq 2} \sum_{l=1}^{k} \frac{I(q^{-k})}{k} \binom{k}{l} (w-1)^l$$
$$= \left(\sum_{k\geq 2} I(q^{-k})\right)(w-1) + \sum_{l=2}^{\infty} \left(\sum_{k\geq l} \binom{k}{l} \frac{I(q^{-k})}{k}\right)(w-1)^l.$$

The coefficient of $(w-1)$ can be rewritten as follows:

$$\sum_{k\geq 2} I(q^{-k}) = \sum_{\substack{k\geq 2\\ m\geq 1}} \frac{\mu(m)}{m} \log\left(\frac{1}{1-q^{1-km}}\right) = \sum_{D=km\geq 2} \log\left(\frac{1}{1-q^{1-D}}\right) \sum_{\substack{m|D\\ m\neq D}} \frac{\mu(m)}{m}$$
$$= \sum_{D\geq 2} \frac{\varphi(D)-\mu(D)}{D} \log\left(\frac{1}{1-q^{1-D}}\right) = S(q^{-1}) - R(q^{-1}).$$

On the other hand, if $l \geq 2$, then the coefficient of $(w-1)^l$ rewrites as

$$\begin{aligned}\sum_{k\geq l}\binom{k}{l}\frac{I(q^{-k})}{k} &= \frac{1}{l}\sum_{k\geq l}\sum_{m\geq 1}\frac{\mu(m)}{m}\binom{k-1}{l-1}\log\left(\frac{1}{1-q^{1-km}}\right)\\ &= \frac{1}{l}\sum_{m,n\geq 1}\frac{\mu(m)}{nm}q^n\sum_{k\geq l}\binom{k-1}{l-1}q^{-k(mn)}\\ &= \frac{1}{l}\sum_{D\geq 1}\frac{|\mathfrak{I}_D|}{(q^D-1)^l}\end{aligned}$$

by setting $D = nm$ on the last line. We conclude that the Fourier transform $\mathbb{E}[w^{Z_n}]$ has asymptotics

$$\mathrm{e}^{(\log n + S(q^{-1})+\gamma)(w-1)}\exp\left(\sum_{l=2}^{\infty}\left((-1)^{l-1}\zeta(l)+\sum_{D\geq 1}\frac{|\mathfrak{I}_D|}{(q^D-1)^l}\right)\frac{(w-1)^l}{l}\right),$$

and the term of degree 1 in $(w-1)$ is separated from the other terms. □

From Theorem 4.6, one can easily construct approximation schemes of the sequences $(X_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$, which yield distances that are arbitrary large negative powers of $\log n$. For instance, suppose that one wants to approximate the law of Y_n by a signed measure ν_n, such that $d_{\mathrm{TV}}(\mu_n,\nu_n) = O(\frac{1}{(\log n)^2})$. By Theorem 3.11, we can take the derived scheme of the approximation scheme of order $r=3$, that is to say the sequence of measures $(\sigma_n)_{n\in\mathbb{N}}$ with Fourier transforms

$$\widehat{\sigma_n}(\xi) = \mathrm{e}^{(\log n + R(q^{-1})+\gamma)(\mathrm{e}^{\mathrm{i}\xi}-1)}\left(1-\frac{\zeta(2)+I(q^{-2})}{2}(\mathrm{e}^{\mathrm{i}\xi}-1)^2+\frac{\zeta(3)+I(q^{-3})}{3}(\mathrm{e}^{\mathrm{i}\xi}-1)^3\right).$$

One has in this situation

$$d_{\mathrm{TV}}(X_n,\sigma_n) \simeq \frac{1}{(\log n)^2}\left(\frac{(\zeta(2)+I(q^{-2}))^2}{8}+\frac{\zeta(4)+I(q^{-4})}{4}\right)\left(\int_{\mathbb{R}}|G_4(\alpha)|\,\frac{d\alpha}{\sqrt{2\pi}}\right).$$

Remark 4.7. A central limit theorem for the random variables Y_n has been proved by Rhoades in [Rho09]. Our results immediately yield the speed of convergence of this analogue on function fields of the Erdös–Kac central limit theorem.

4.5. **Number of distinct prime divisors of a random integer.** As pointed out in the introduction, number theory provides another area where the phenomenon of mod-Poisson convergence is prominent; see in particular the discussion of [FMN16, §7.2]. Consider as in Example 1.7 the number $\omega(N_n)$ of distinct prime divisors of a random integer chosen uniformly in $[\![1,n]\!]$. This quantity satisfies the celebrated Erdös–Kac central limit theorem:

$$\frac{\omega(N_n)-\log\log n}{\sqrt{\log\log n}} \rightharpoonup \mathcal{N}_{\mathbb{R}}(0,1),$$

where $\mathcal{N}_{\mathbb{R}}(0,1)$ denotes a standard one-dimensional Gaussian distribution of mean 0 and variance 1 (*cf.* [EK40]). This Gaussian approximation actually comes from a more precise Poissonian approximation, which can be stated in the framework of mod-Poisson convergence:

Theorem 4.8 (Rényi–Turán)**.** *Locally uniformly in* $z \in \mathbb{C}$,

$$\mathbb{E}[\mathrm{e}^{z\omega(N_n)}] = \mathrm{e}^{(\log\log n+\gamma)(\mathrm{e}^z-1)} \left(\prod_{n\in\mathbb{N}} \left(1+\frac{\mathrm{e}^z-1}{n}\right) \mathrm{e}^{-\frac{\mathrm{e}^z-1}{n}} \prod_{p\in\mathbb{P}} \left(1+\frac{\mathrm{e}^z-1}{p}\right) \mathrm{e}^{-\frac{\mathrm{e}^z-1}{p}} + O\left(\frac{1}{\log n}\right)\right).$$

This formula first appeared in [RT58], and it can be obtained by an application of the Selberg–Delange method, see [Ten95, Chapter II.5]. It is actually stronger than the prime number theorem, as it requires a larger zero-free region of the Riemann zeta function than the line $\mathrm{Re}(s) = 1$. The problem of the rate of convergence in the Erdös–Kac central limit theorem was already studied by Harper in [Har09] for the Kolmogorov distance to the Gaussian approximation; and by Barbour, Kowalski and Nikeghbali in [BKN09] for the three distances to the Poissonian (or signed measure) approximation.

Note that the residue $\psi(\xi) = \prod_{n\in\mathbb{N}^*} \left(1+\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{n}\right) \mathrm{e}^{-\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{n}} \prod_{p\in\mathbb{P}} \left(1+\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{p}\right) \mathrm{e}^{-\frac{\mathrm{e}^{\mathrm{i}\xi}-1}{p}}$ can be rewritten as

$$\exp\left(\sum_{k\geq 2} (-1)^{k-1} \frac{\mathfrak{p}_k(\mathbb{N}^{*-1}+\mathbb{P}^{-1})}{k} (\mathrm{e}^{\mathrm{i}\xi}-1)^k\right),$$

where $\mathfrak{p}_k(\mathbb{N}^{*-1}+\mathbb{P}^{-1}) = \sum_{n=1}^{\infty} \frac{1}{n^k} + \sum_{p\in\mathbb{P}} \frac{1}{p^k}$. Therefore:

Theorem 4.9. *For any* $r \geq 1$*, we introduce the scheme of approximation* $(\sigma_n^{(r)})_{n\in\mathbb{N}}$*, defined by the Fourier transforms*

$$\widehat{\sigma_n^{(r)}}(\xi) = \mathrm{e}^{(\log\log n+\gamma)(\mathrm{e}^{\mathrm{i}\xi}-1)} \left(\sum_{k=0}^{r} \mathfrak{e}_k(\mathbb{N}^{*-1}+\mathbb{P}^{-1}) (\mathrm{e}^{\mathrm{i}\xi}-1)^k\right),$$

where the parameters $\mathfrak{e}_k(\mathbb{N}^{*-1}+\mathbb{P}^{-1})$ *correspond to the specialisation of* Sym*:*

$$\mathfrak{p}_1(\mathbb{N}^{*-1}+\mathbb{P}^{-1}) = 0 \quad ; \quad \mathfrak{p}_{k\geq 2}(\mathbb{N}^{*-1}+\mathbb{P}^{-1}) = \zeta(k) + \sum_{p\in\mathbb{P}} \frac{1}{p^k}.$$

One has the asymptotic formulas:

$$d_{\mathrm{L}}(\omega(N_n), \sigma_n^{(r)}) = \frac{|\mathfrak{e}_{r+1}(\mathbb{N}^{*-1}+\mathbb{P}^{-1})\, G_{r+1}(z_{r+2})|}{\sqrt{2\pi}\,(\log\log n)^{\frac{r}{2}+1}} + o\left(\frac{1}{(\log\log n)^{\frac{r}{2}+1}}\right);$$

$$d_{\mathrm{K}}(\omega(N_n), \sigma_n^{(r)}) = \frac{|\mathfrak{e}_{r+1}(\mathbb{N}^{*-1}+\mathbb{P}^{-1})\, G_r(z_{r+1})|}{\sqrt{2\pi}\,(\log\log n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\log\log n)^{\frac{r+1}{2}}}\right);$$

$$d_{\mathrm{TV}}(\omega(N_n), \sigma_n^{(r)}) = \frac{\int_{\mathbb{R}} |\mathfrak{e}_{r+1}(\mathbb{N}^{*-1}+\mathbb{P}^{-1})\, G_{r+1}(\alpha)|\, d\alpha}{\sqrt{2\pi}\,(\log\log n)^{\frac{r+1}{2}}} + o\left(\frac{1}{(\log\log n)^{\frac{r+1}{2}}}\right).$$

This theorem allows one to quantify how much better the Poisson approximation is in comparison to the Gaussian approximation (Erdös–Kac theorem). Indeed, since the case $r = 1$ of Theorem 4.9 is the basic scheme of approximation of $\omega(N_n)$ by a Poisson random

variable Y_n of parameter $\lambda_n = \log\log n + \gamma$, one has:

$$d_{\mathrm{K}}\left(\frac{\omega(N_n) - \log\log n - \gamma}{\sqrt{\log\log n + \gamma}}, \frac{\mathcal{P}_{(\log\log n+\gamma)} - \log\log n - \gamma}{\sqrt{\log\log n + \gamma}}\right)$$
$$= d_{\mathrm{K}}(\omega(N_n), \mathcal{P}_{(\log\log n+\gamma)}) = O\left(\frac{1}{\log\log n}\right).$$

On the other hand, the classical Berry–Esseen estimate for sums of identically distributed random variables ensures that

$$d_{\mathrm{K}}\left(\frac{\mathcal{P}_{(\log\log n+\gamma)} - \log\log n - \gamma}{\sqrt{\log\log n + \gamma}}, \mathcal{N}_{\mathbb{R}}(0,1)\right) = O\left(\frac{1}{\sqrt{\log\log n}}\right).$$

Thus, in terms of Kolmogorov distance, the Poisson approximation of the sequence $(\omega(N_n))_{n\in\mathbb{N}}$ is at distance $O(\frac{1}{\log\log n})$, whereas the Gaussian approximation is at distance $O(\frac{1}{\sqrt{\log\log n}})$. Similar remarks can be made for the statistics of random combinatorial objects previously studied.

5. The multi-dimensional case

In this last section, we extend the theoretical results of Sections 2 and 3 to the multi-dimensional case, that is to say that we are going to approximate the laws of random variables with values in $\mathbb{Z}^{d\geq 2}$. Thus, we consider a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}^d$ that converges mod-ϕ with parameters λ_n, where ϕ is an infinitely divisible distribution with minimal lattice $\mathbb{Z}^d$:

$$\mu_n(\xi) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\sum_{i=1}^d \xi_i X_{n,i}}\right] = \mathrm{e}^{\lambda_n\,\phi(\xi)}\,\psi_n(\xi_1,\dots,\xi_d),$$

with $\lim_{n\to\infty}\psi_n = \psi$ uniformly on the torus. In all the examples that we shall look at, the law with exponent ϕ enjoys the factorization property:

$$\phi(\xi) = \phi(\xi_1,\dots,\xi_d) = \sum_{i=1}^d \phi_i(\xi_i),$$

where the ϕ_j's are infinitely divisible laws on $\mathbb{Z}$. In other words, if Y is a reference random variable with law on $\mathbb{Z}^d$ with Lévy–Khintchine exponent ϕ, then its coordinates are independent random variables on the lattice $\mathbb{Z}$. This does not mean at all that the X_n's will have independent coordinates. Hence, though we shall see independence of coordinates by looking at the first order asymptotics of X_n (captured by the infinitely divisible law ϕ), the higher order asymptotics will shed light on the non-independence of the coordinates (this being captured by the limiting residue $\psi(\xi)$, which does not factorize in general).

As in Section 3, we consider a general approximation scheme $(\nu_n)_{n\in\mathbb{N}}$ of $(X_n)_{n\in\mathbb{N}}$ with Fourier transforms

$$\widehat{\nu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\,\chi_n(\xi_1,\dots,\xi_d),$$

and $\lim_{n\to\infty}\chi_n = \chi$. The goal will be to compute the asymptotics of

$$d_{\mathrm{L}}(\mu_n,\nu_n) = \sup_{k\in\mathbb{Z}^d} |\mu_n(\{k\}) - \nu_n(\{k\})|;$$
$$d_{\mathrm{TV}}(\mu_n,\nu_n) = \sum_{k\in\mathbb{Z}^d} |\mu_n(\{k\}) - \nu_n(\{k\})|.$$

In dimension $d \geq 2$, there is no interesting analogue of the Kolmogorov distance (at least for discrete measures; see [FMN17b] for the case of continuous measures). In §5.1, we study the multi-dimensional Wiener algebra $\mathscr{A}(\mathbb{T}^d)$, and we establish an estimate of norms that is similar to our Theorem 2.3 (see Theorem 5.3). In §5.2, we deduce from this theorem the multi-dimensional analogue of the results of Section 3. Finally, in §5.3 and §5.4, we study two examples which generalize the ones of Section 4, and stem respectively from the combinatorics of coloured permutations, and from arithmetic progressions in number theory.

Let us fix a few standard notation relative to the multi-dimensional setting. The coordinates of a parameter $\xi \in \mathbb{T}^d = (\mathbb{R}/2\pi\mathbb{Z})^d$ (respectively, $k \in \mathbb{Z}^d$) will be denoted $(\xi_1, \dots, \xi_d)$ (resp., $(k_1, \dots, k_d)$). If $\alpha \in \mathbb{N}^d$ is a multi-index, we denote $|\alpha| = \sum_{i=1}^d \alpha_i$ and

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f(\xi_1, \dots, \xi_d)}{\partial \xi_1^{\alpha_1} \cdots \partial \xi_d^{\alpha_d}}.$$

In the same setting, if $n = |\alpha|$, then the multinomial coefficient $\binom{n}{\alpha}$ is $\frac{n!}{(\alpha_1)! \cdots (\alpha_d)!}$. On the other hand, if α and β are two d-tuples, we write

$$\langle \alpha \mid \beta \rangle = \sum_{i=1}^d \alpha_i \beta_i.$$

We also write $\alpha \leq \beta$ if $\alpha_i \leq \beta_i$ for all $i \in [\![1, d]\!]$.

5.1. **The multi-dimensional Wiener algebra.**

Definition 5.1. *The multi-dimensional Wiener algebra $\mathscr{A} = \mathscr{A}(\mathbb{T}^d)$ is the algebra of continuous functions on the d-torus whose Fourier series converges absolutely. It is a Banach algebra for the norm*

$$\|f\|_{\mathscr{A}} := \sum_{n \in \mathbb{Z}^d} |c_n(f)|,$$

where $c_n(f)$ denotes the Fourier coefficient

$$c_n(f) = \int_{\theta_1=0}^{2\pi} \cdots \int_{\theta_d=0}^{2\pi} f(\mathrm{e}^{\mathrm{i}\theta_1}, \dots, \mathrm{e}^{\mathrm{i}\theta_d})\, \mathrm{e}^{-\mathrm{i}\sum_{i=1}^d n_i\theta_i}\, \frac{d\theta}{(2\pi)^d},$$

where $n = (n_1, \dots, n_d)$.

As in the case $d = 1$, if μ is a (signed) measure on $\mathbb{Z}^d$, then its total variation norm is equal to $\|\mu\|_{\mathrm{TV}} = \|\widehat{\mu}\|_{\mathscr{A}}$. On the other hand, one has the analogue of Proposition 2.1:

Proposition 5.2. *There is a constant $C_{H,d}$ such that, for any $f \in \mathscr{A}(\mathbb{T}^d)$,*

$$\|f\|_{\mathscr{A}} \leq C_{H,d} \sup_{|\alpha| \leq K_d} \|\partial^\alpha f\|_{\mathscr{L}^2},$$

where $K = K_d = \lfloor \frac{d}{2} \rfloor + 1$, and $\mathscr{L}^2 = \mathscr{L}^2(\mathbb{T}^d)$ is the space of square-integrable functions on the torus, endowed with the norm

$$\|f\|_{\mathscr{L}^2} = \sqrt{\int_{\theta_1=0}^{2\pi} \cdots \int_{\theta_d=0}^{2\pi} |f(\mathrm{e}^{\mathrm{i}\theta_1}, \dots, \mathrm{e}^{\mathrm{i}\theta_d})|^2\, \frac{d\theta}{(2\pi)^d}}.$$

Proof. Note that $K_d \geq \frac{d+1}{2}$ for any $d \geq 1$. We use again the Cauchy–Schwarz inequality and the Parseval identity:

$$\begin{aligned}
\|f\|_{\mathscr{A}} &= \sum_{n\in\mathbb{Z}^d} |c_n(f)| = \sum_{n\in\mathbb{Z}^d} \left(\frac{1+\sum_{i=1}^d (n_i)^2}{1+\sum_{i=1}^d (n_i)^2}\right)^{\frac{K_d}{2}} |c_n(f)| \\
&\leq \sqrt{\sum_{n\in\mathbb{Z}^d} \left(\frac{1}{1+\sum_{i=1}^d (n_i)^2}\right)^{K_d}} \sqrt{\sum_{n\in\mathbb{Z}^d} \left(1+\sum_{i=1}^d (n_i)^2\right)^{K_d} |c_n(f)|^2} \\
&\leq \sqrt{\sum_{n\in\mathbb{Z}^d} \left(\frac{1}{1+\sum_{i=1}^d (n_i)^2}\right)^{\frac{d+1}{2}}} \sqrt{\sum_{\alpha_0+\alpha_1+\cdots+\alpha_d=K_d} \binom{K_d}{\alpha_0,\ldots,\alpha_d} \sum_{n\in\mathbb{Z}^d} |c_n(\partial^{(\alpha_1,\ldots,\alpha_d)} f)|^2} \\
&\leq \sqrt{\sum_{n\in\mathbb{Z}^d} \left(\frac{1}{1+\sum_{i=1}^d (n_i)^2}\right)^{\frac{d+1}{2}} \sum_{\alpha_0+\alpha_1+\cdots+\alpha_d=K_d} \binom{K_d}{\alpha_0,\ldots,\alpha_d}} \sup_{|\alpha|\leq K_d} \|\partial^\alpha f\|_{\mathscr{L}^2},
\end{aligned}$$

hence the result since the series $(C_{H,d})^2$ under the square root sign is finite. □

This inequality leads to the analogue of Theorem 2.3 in a multi-dimensional setting. We consider two measures μ and ν on $\mathbb{Z}^d$ with Fourier transforms

$$\begin{aligned}
\widehat{\mu}(\xi) &= \mathrm{e}^{\lambda\phi(\xi)}\,\psi(\xi); \\
\widehat{\nu}(\xi) &= \mathrm{e}^{\lambda\phi(\xi)}\,\chi(\xi)
\end{aligned}$$

where $\phi(\xi) = \sum_{i=1}^d \phi_i(\xi_i)$. We assume ϕ to have moments at least up to order K_d, and denote

$$\phi_i(\xi_i) = \mathrm{i}\, m_i \xi_i - \frac{\sigma_i^2\,\xi_i^2}{2} + o(\xi_i^2)$$

the Taylor expansion around 0 of each law on $\mathbb{Z}$ with Lévy exponent ϕ_i. We fix an integer $r \geq \lfloor \frac{d}{2} \rfloor$ such that

(1) the residues ψ and χ are $(r+1)$ times continuously differentiable on $\mathbb{T}^d$;

(2) their Taylor expansions at 0 coincide up to order r:

$$\forall |\alpha| \leq r,\ \ (\partial^\alpha(\psi-\chi))\,(0) = 0.$$

A parameter $\varepsilon \in (0,1)$ being fixed, we also introduce the non-negative quantities

$$\begin{aligned}
\beta_{r+1}(\varepsilon) &= \sup_{|\alpha|=r+1} \sup_{\xi\in[-\varepsilon,\varepsilon]^d} |\partial^\alpha(\psi-\chi)(\varepsilon)|; \\
\gamma(\varepsilon) &= \sup_{i\in[\![1,d]\!]} \sup_{\theta\in[-\varepsilon,\varepsilon]} |\phi_i''(\theta) + \sigma_i^2|; \\
M &= -\sup_{i\in[\![1,d]\!]} \sup_{\theta\in[-\pi,\pi]} \left(\frac{\mathrm{Re}(\phi_i(\theta))}{\theta^2}\right).
\end{aligned}$$

The quantity $\beta_{r+1}(\varepsilon)$ allows one to bound any derivative up to order $(r+1)$ of $\psi-\chi$ on $[-\varepsilon,\varepsilon]^d$. Indeed, this is obvious for exponents α with $|\alpha| = r+1$, and otherwise,

$$|\partial^\alpha(\psi-\chi)(\xi)| = \left|\int_{t=0}^1 \frac{d}{dt}(\partial^\alpha(\psi-\chi)(t\xi))\,dt\right| \leq \sum_{i=1}^d |\xi_i| \left|\int_{t=0}^1 \frac{\partial}{\partial\alpha_i}(\partial^\alpha(\psi-\chi)(t\xi))\,dt\right|,$$

which leads by descending induction on $|\alpha|$ to the bound

$$|\partial^\alpha(\psi-\chi)(\xi)| \le \left(\sum_{i=1}^d |\xi_i|\right)^{r+1-|\alpha|} \beta_{r+1}(\varepsilon) \le (d\varepsilon)^{r+1-|\alpha|}\,\beta_{r+1}(\varepsilon)$$

for any $\xi \in [-\varepsilon,\varepsilon]^d$.

Theorem 5.3. *Fix $\varepsilon \in (0,1)$ such that $\gamma(\varepsilon) \le \min_{i\in[\![1,d]\!]} \frac{\sigma_i^2}{2}$. There exist some positive constants $C_1(d,c,\phi)$ and $C_2(d,c,\phi)$, that depend only on d, c and ϕ, such that*

$$\begin{aligned}&\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\\ &\le C_1(d,c,\phi)\left(\|\psi-\chi\|_{\mathscr{A}}\,\frac{\lambda^K\,\mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}}}{\varepsilon} + \frac{\varepsilon^{r+1}\,\beta_{r+1}(\varepsilon)}{\lambda^{\frac{d}{4}}}\left(\sqrt{\lambda}+\frac{1}{\varepsilon}\right)^K\right)\left(1+\frac{C_2(d,c,\phi)}{\sqrt{\lambda}}\right)\end{aligned}$$

for any $\varepsilon \in (0,1)$ and any $\lambda > 0$.

Since Proposition 5.2 involves higher order derivatives than Proposition 2.1, in order to prove Theorem 5.3, we shall use a smoother cut-function than in the proof of Theorem 2.3. Thus, in the following, $c_{d,\varepsilon}(\xi) = \prod_{i=1}^d c(\frac{\xi_i}{\varepsilon})$, where $c(x)$ is a function of one variable with values in $[0,1]$, that is of class $\mathscr{C}^{K_d}$ on $\mathbb{R}$, with $c(x)=0$ outside $[-1,1]$ and $c(x)=1$ on $[-\frac12,\frac12]$. In particular, $c_{d,\varepsilon}$ (respectively, $1-c_{d,\varepsilon}$) vanishes on $[-\pi,\pi]^d \setminus [-\varepsilon,\varepsilon]^d$ (resp., on $[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]^d$).

Lemma 5.4. *Under the assumptions of Theorem 5.3, there exists a constant $C(d,c,\phi)$ such that*

$$\left\|(1-c_{d,\varepsilon})\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} \le C(d,c,\phi)\,\frac{\lambda^K}{\varepsilon}\,\mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}}\,\left(1+O\left(\frac{1}{\lambda}\right)\right),$$

where the constant hidden in the $O(\cdot)$ also depends only on d, c and ϕ.

Proof. We combine Proposition 5.2 with the rules of differentiation of functions of several variables:

$$\begin{aligned}\left\|(1-c_{d,\varepsilon})\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} &\le C_{H,d}\sup_{|\alpha|\le K_d}\|\partial^\alpha((1-c_{d,\varepsilon})\,\mathrm{e}^{\lambda\phi})\|_{\mathscr{L}^2([-\pi,\pi]^d\setminus[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]^d}\\ &\le C_{H,d}\sup_{|\alpha|\le K_d}\sum_{\beta\le\alpha}\prod_{i=1}^d\binom{\alpha_i}{\beta_i}\|\partial^{\alpha-\beta}(1-c_{d,\varepsilon})\|_{\mathscr{L}^2}\,\|\partial^\beta(\mathrm{e}^{\lambda\phi})\|_{\mathscr{L}^\infty([-\pi,\pi]^d\setminus[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]^d)}\\ &\le 2^{K_d}\,C_{H,d}\left(\sup_{|\gamma|\le K_d}\|\partial^\gamma(1-c_{d,\varepsilon})\|_{\mathscr{L}^2}\right)\left(\sup_{|\beta|\le K_d}\|\partial^\beta(\mathrm{e}^{\lambda\phi})\|_{\mathscr{L}^\infty([-\pi,\pi]^d\setminus[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]^d)}\right).\end{aligned}$$

Since the Lévy–Khintchine exponent ϕ is assumed to have the factorization property, for any multi-index β,

$$\partial^\beta(\mathrm{e}^{\lambda\phi}) = \prod_{i=1}^d \frac{\partial^{\beta_i}\mathrm{e}^{\lambda\phi_i(\xi_i)}}{\partial\xi_i^{\beta_i}},$$

and

$$\frac{\partial^n \mathrm{e}^{\lambda\phi_i(\theta)}}{\partial\theta^n} = \mathrm{e}^{\lambda\phi_i(\theta)}\sum_{k=1}^n B_{n,k}(\lambda\phi_i'(\theta),\lambda\phi_i''(\theta),\dots,\lambda\phi_i^{(n-k+1)}(\theta)),$$

where $B_{n,k}(x_1, \dots, x_{n+k-1})$ is the incomplete Bell polynomial of parameters n and k, see [Bel27]. More precisely,

$$B_{n,k}(x_1, \dots, x_{n+k-1}) = \sum_{\substack{m_1+m_2+\cdots+m_{n-k+1}=k \\ m_1+2m_2+\cdots+(n-k+1)m_{n-k+1}=n}} \frac{n!}{\prod_{i=1}^{n-k+1}(i!^{m_i}\, m_i!)} (x_1)^{m_1} \cdots (x_{n-k+1})^{m_{n-k+1}}$$

$$|B_{n,k}(x_1, \dots, x_{n+k-1})| \leq \left(\sup_{i \in [\![1, n+k-1]\!]} |x_i| \right)^k S(n,k)$$

where $S(n,k) = B_{n,k}(1, 1, \dots, 1)$ is the Stirling number of the second kind. As a consequence,

$$\left| \frac{\partial^n \mathrm{e}^{\lambda\phi_i(\theta)}}{\partial\theta^n} \right| \leq |\mathrm{e}^{\lambda\phi_i(\theta)}|\, T_n(\lambda \|\phi_i\|_{\mathscr{C}^n(\mathbb{T})}) \leq |\mathrm{e}^{\lambda\phi_i(\theta)}|\, T_n(\lambda \|\phi\|_{\mathscr{C}^n(\mathbb{T}^d)}),$$

where $T_n(x) = \sum_{k=1}^n S(n,k)\, x^k$ is the n-th Touchard polynomial, with leading term $S(n,n)\, x^n = x^n$. Thus,

$$\begin{aligned} |\partial^\beta \mathrm{e}^{\lambda\phi}| &\leq |\mathrm{e}^{\lambda\phi}| \prod_{i=1}^d \left| T_{\beta_i}(\lambda \|\phi\|_{\mathscr{C}^{\beta_i}(\mathbb{T}^d)}) \right| \\ &\leq |\mathrm{e}^{\lambda\phi}|\, (\lambda \|\phi\|_{\mathscr{C}^K(\mathbb{T}^d)})^{|\beta|} \left(1 + O\left(\frac{1}{\lambda \|\phi\|_{\mathscr{C}^K(\mathbb{T}^d)}} \right) \right) \end{aligned}$$

if $|\beta| \leq K_d$. Now, if $\xi \notin [-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]^d$, then one of the coordinates ξ_i has modulus larger than $\frac{\varepsilon}{2}$, so $|\mathrm{e}^{\phi(\xi)}| \leq \mathrm{e}^{-\frac{M\varepsilon^2}{4}}$. We conclude that

$$\begin{aligned} &\left\| (1 - c_d)\, \mathrm{e}^{\lambda\phi} \right\|_{\mathscr{A}} \\ &\leq 2^{K_d}\, C_{H,d} \left(\sup_{|\gamma| \leq K_d} \|\partial^\gamma (1 - c_{d,\varepsilon})\|_{\mathscr{L}^2} \right) (\lambda\, \|\phi\|_{\mathscr{C}^K(\mathbb{T}^d)})^K\, \mathrm{e}^{-\frac{\lambda M\varepsilon^2}{4}} \left(1 + O\left(\frac{1}{\lambda} \right) \right), \end{aligned}$$

where the constant in the $O(\cdot)$ only depends on d and the exponent ϕ. Finally, as soon as $\gamma \neq 0$, $\partial^\gamma (1 - c_{d,\varepsilon}) = -\frac{1}{\varepsilon^{|\gamma|}} (\partial^\gamma c_{d,1})(\frac{\xi}{\varepsilon})$, so,

$$\begin{aligned} \|\partial^\gamma (1 - c_{d,\varepsilon})\|_{\mathscr{L}^2} &= \frac{1}{\varepsilon^{|\gamma|}} \sqrt{\int_{[-\varepsilon,\varepsilon]^d} \left| (\partial^\gamma c_{d,1}) \left(\frac{\xi}{\varepsilon} \right) \right|^2 \frac{d\xi}{(2\pi)^d}} \\ &= \frac{1}{\varepsilon^{|\gamma| - \frac{d}{2}}} \sqrt{\int_{[-1,1]^d} |(\partial^\gamma c_{d,1})(\theta)|^2 \frac{d\theta}{(2\pi)^d}} \\ &= O\left(\frac{1}{\varepsilon^{|\gamma| - \frac{d}{2}}} \right). \end{aligned}$$

We conclude by taking the maximal possible value $|\gamma| = K_d$, which gives $K_d - \frac{d}{2} = 1$ if d is even, and $\frac{1}{2}$ if d is odd. □

Lemma 5.5. *Under the assumptions of Theorem 5.3, there exists another constant $C(d, c, \phi)$ such that*

$$\left\| c_{d,\varepsilon}\, (\psi - \chi)\, \mathrm{e}^{\lambda\phi} \right\|_{\mathscr{A}} \leq C(d, c, \phi)\, \frac{\varepsilon^{r+1}\, \beta_{r+1}(\varepsilon)}{\lambda^{\frac{d}{4}}} \left(\sqrt{\lambda} + \frac{1}{\varepsilon} \right)^K \left(1 + O\left(\frac{1}{\sqrt{\lambda}} \right) \right).$$

Again, the constant hidden in the $O(\cdot)$ depends only on d, c and ϕ.

Proof. The hypothesis on $\gamma(\varepsilon)$ ensures that, for any index $i \in [\![1,d]\!]$ and any $\theta \in [-\varepsilon,\varepsilon]$, $\mathrm{Re}(\phi_i''(\theta)) \le -\sigma_i^2/2$ and $|\phi_i''(\theta)| \le 2\sigma_i^2$. On the other hand, as in the case $d=1$ dealt with by Lemma 2.5, we can shift phases by $S(\xi) = \mathrm{i}\sum_{i=1}^d \xi_i \lfloor \lambda m_i \rfloor$:

$$\begin{aligned}
&\left\| c_{d,\varepsilon}\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} = \left\| c_{d,\varepsilon}\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi-S}\right\|_{\mathscr{A}} \\
&\le C_{H,d} \sup_{|\alpha|\le K_d} \|\partial^\alpha (c_{d,\varepsilon}\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi-S})\|_{\mathscr{L}^2([-\varepsilon,\varepsilon]^d)} \\
&\le C_{H,d} \sup_{|\alpha|\le K_d} \left(\sum_{\beta+\gamma+\delta=\alpha} \prod_{i=1}^d \binom{\alpha_i}{\beta_i,\gamma_i,\delta_i} \|\partial^\beta c_{d,\varepsilon}\,\partial^\gamma(\psi-\chi)\|_{\mathscr{L}^\infty([-\varepsilon,\varepsilon]^d)} \|\partial^\delta \mathrm{e}^{\lambda\phi-S}\|_{\mathscr{L}^2([-\varepsilon,\varepsilon]^d)} \right).
\end{aligned}$$

On the last line, we can bound $\|\partial^\beta c_{d,\varepsilon}\|_{\mathscr{L}^\infty([-\varepsilon,\varepsilon]^d)}$ by $C(d,c)\,\varepsilon^{-|\beta|}$, and $\|\partial^\gamma(\psi-\chi)\|_{\mathscr{L}^\infty([-\varepsilon,\varepsilon]^d)}$ by $\beta_{r+1}(\varepsilon)\,(d\varepsilon)^{r+1-|\gamma|}$. Thus,

$$\begin{aligned}
&\left\| c_{d,\varepsilon}\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}} \\
&\le C(d,c)\,\varepsilon^{r+1}\,\beta_{r+1}(\varepsilon) \left(\sup_{|\alpha|\le K_d} \sum_{\rho+\delta=\alpha} \prod_{i=1}^d \binom{\alpha_i}{\rho_i,\delta_i} \varepsilon^{-|\rho|}\,\|\partial^\delta \mathrm{e}^{\lambda\phi-S}\|_{\mathscr{L}^2([-\varepsilon,\varepsilon]^d)} \right)
\end{aligned}$$

where $C(d,c)$ is some positive constant depending only on d and the cut function c.

To bound the norms $\|\partial^\delta \mathrm{e}^{\lambda\phi-S}\|_{\mathscr{L}^2([-\varepsilon,\varepsilon]^d)}$, we use the same combinatorics of Bell polynomials as in the previous lemma. Specifically, we have

$$\left| \frac{\partial^n \left(\mathrm{e}^{\lambda\phi_i(\theta)-\mathrm{i}\lfloor\lambda m_i\rfloor\theta}\right)}{\partial\theta^n} \right| \le \left|\mathrm{e}^{\lambda\phi_i(\theta)}\right| \left| \sum_{k=1}^n B_{n,k}(\lambda\phi_i'(\theta) - \mathrm{i}\lfloor\lambda m_i\rfloor, \lambda\phi_i''(\theta), \dots, \lambda\phi_i^{(n-k+1)}(\theta)) \right|,$$

and the variables of the incomplete Bell polynomials that appear are bounded as follows:

$$\begin{aligned}
|\lambda\phi_i'(\theta) - \mathrm{i}\lfloor\lambda m_i\rfloor| &\le \lambda|\phi_i'(\theta)-\phi_i'(0)| + |\lambda m_i - \lfloor\lambda m_i\rfloor| \\
&\le \lambda\,\|\phi_i''\|_{\mathscr{L}^\infty([-\varepsilon,\varepsilon])}\,|\theta| + 1 \le 2\lambda\,\sigma_i^2\,|\theta| + 1; \\
|\lambda\phi_i^{(r\ge 2)}(\theta)| &\le \lambda\,\|\phi_i^{(r)}\|_\infty \le \lambda\,\|\phi\|_{\mathscr{C}^n(\mathbb{T}^d)},
\end{aligned}$$

Rewriting the complete Bell polynomial $B_n = \sum_{k=1}^n B_{n,k}$ as a sum over integer partitions of size n, we thus obtain

$$\begin{aligned}
&\left| \frac{\partial^n \left(\mathrm{e}^{\lambda\phi_i(\theta)-\mathrm{i}\lfloor\lambda m_i\rfloor\theta}\right)}{\partial\theta^n} \right| \\
&\le \mathrm{e}^{-\frac{\lambda\sigma_i^2\xi_i^2}{4}} \sum_{\substack{L \text{ partition} \\ \text{of size } n}} \frac{n!}{\prod_{i\ge1}(i!^{m_i(L)}\,m_i(L)!)} \left(2\lambda\,\sigma_i^2\,|\theta|+1\right)^{m_1(L)} \left(\lambda\,\|\phi\|_{\mathscr{C}^n(\mathbb{T}^d)}\right)^{\ell(L)-m_1(L)}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\sqrt{\int_{[-\varepsilon,\varepsilon]} \left| \frac{\partial^n \left(\mathrm{e}^{\lambda\phi_i(\theta)-\mathrm{i}\lfloor\lambda m_i\rfloor\theta}\right)}{\partial\theta^n} \right|^2 \frac{d\theta}{2\pi}} \\
&\le \sum_{\substack{L \text{ partition} \\ \text{of size } n}} \frac{n!\,\left(\lambda\,\|\phi\|_{\mathscr{C}^n(\mathbb{T}^d)}\right)^{\ell(L)-m_1(L)}}{\prod_{i\ge1}(i!^{m_i(L)}\,m_i(L)!)} \sqrt{\int_{[-\varepsilon,\varepsilon]} \mathrm{e}^{-\frac{\lambda\sigma_i^2\xi_i^2}{2}} \left(2\lambda\,\sigma_i^2\,|\theta|+1\right)^{2m_1(L)} \frac{d\theta}{2\pi}},
\end{aligned}$$

and each integral under the square root sign is equal to

$$\sqrt{\frac{1}{\sqrt{2\pi\lambda\,\sigma_i^2}}\sum_{k=0}^{2m_1(L)}\binom{2m_1(L)}{k}(2\sqrt{\lambda}\,\sigma_i)^k\int_{\mathbb{R}}\mathrm{e}^{-\frac{x^2}{2}}\,|x|^k\,\frac{dx}{\sqrt{2\pi}}}$$
$$=\frac{2^{m_1(L)}\sqrt{(2m_1(L)-1)!!}}{(2\pi)^{1/4}}\,(\lambda\,\sigma_i^2)^{\frac{m_1(L)}{2}-\frac{1}{4}}\left(1+O\left(\frac{1}{\sqrt{\lambda}}\right)\right).$$

Thus, the term corresponding to an integer partition L in the previous bound on the norm

$$\left\|\frac{\partial^n\left(\mathrm{e}^{\lambda\phi_i(\theta)-\mathrm{i}\lfloor\lambda m_i\rfloor\theta}\right)}{\partial\theta^n}\right\|_{\mathscr{L}^2([-\varepsilon,\varepsilon])}$$

is of order $O(\lambda^{\ell(L)-\frac{m_1(L)}{2}-\frac{1}{4}})$, where the constant in the $O(\cdot)$ can depend only on n and ϕ. The integer partitions that maximise the exponent $\ell(L)-\frac{m_1(L)}{2}$ among those of fixed size n are those with parts 1 or 2, of the form

$$L=(\underbrace{2,2,\ldots,2}_{j\text{ parts of size }2},\ \underbrace{1,1,\ldots,1}_{n-2j\text{ parts of size }1}).$$

Therefore,

$$\left\|\frac{\partial^n\left(\mathrm{e}^{\lambda\phi_i(\theta)-\mathrm{i}\lfloor\lambda m_i\rfloor\theta}\right)}{\partial\theta^n}\right\|_{\mathscr{L}^2([-\varepsilon,\varepsilon])}$$
$$\leq\frac{1}{(2\pi\lambda\,\sigma_i^2)^{\frac{1}{4}}}\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor}\frac{n!\,(\lambda\,\|\phi\|_{\mathscr{C}^n(\mathbb{T}^d)})^j}{2^j\,j!\,(n-2j)!}\,2^{n-2j}\,\sqrt{(2n-4j-1)!!}\,(\lambda\,\sigma_i^2)^{\frac{n}{2}-j}\left(1+O\left(\frac{1}{\sqrt{\lambda}}\right)\right)$$
$$\leq C(n,\phi)\,\lambda^{\frac{n}{2}-\frac{1}{4}}\left(1+O\left(\frac{1}{\sqrt{\lambda}}\right)\right),$$

where $C(n,\phi)$ and the constant hidden in the $O(\cdot)$ depend only on n and ϕ. Finally,

$$\left\|c_{d,\varepsilon}\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}}$$
$$\leq C(d,c,\phi)\,\frac{\varepsilon^{r+1}\,\beta_{r+1}(\varepsilon)}{\lambda^{\frac{d}{4}}}\left(1+O\left(\frac{1}{\sqrt{\lambda}}\right)\right)\left(\sup_{|\alpha|\leq K_d}\left(\sqrt{\lambda}+\frac{1}{\varepsilon}\right)^{|\alpha|}\right)$$
$$\leq C(d,c,\phi)\,\frac{\varepsilon^{r+1}\,\beta_{r+1}(\varepsilon)}{\lambda^{\frac{d}{4}}}\left(\sqrt{\lambda}+\frac{1}{\varepsilon}\right)^{K_d}\left(1+O\left(\frac{1}{\sqrt{\lambda}}\right)\right).$$

□

Proof of Theorem 5.3. As in the one dimensional case, it suffices now to write

$$\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\leq\|\psi-\chi\|_{\mathscr{A}}\left\|(1-c_{d,\varepsilon})\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}}+\left\|c_{d,\varepsilon}\,(\psi-\chi)\,\mathrm{e}^{\lambda\phi}\right\|_{\mathscr{A}},$$

and to use the two previous lemmas. □

Corollary 5.6. *With the assumptions of Theorem 5.3, for any $r\geq\lfloor\frac{d}{2}\rfloor$, there exists a positive constant $C_3(r,d,c,\phi)$ such that, if $\varepsilon>0$ is fixed and $\lambda\geq 2$ is sufficiently large (chosen according to ε), then*

$$\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\leq C_3(r,d,c,\phi)\,(\|\psi-\chi\|_{\mathscr{A}}+\beta_{r+1}(\varepsilon))\,\frac{(\log\lambda)^{r+1}}{\lambda^{\frac{r}{2}+\frac{1}{4}}}.$$

Proof. Set $\varepsilon = \frac{t\log\lambda}{\sqrt{\lambda}}$ with $t>0$. We then have, up to a multiplicative remainder $(1+o(1))$, a bound on the norm $\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}$ equal to the sum of the two terms:

$$C_1(d,c,\phi)\,\|\psi-\chi\|_{\mathscr{A}}\,\frac{\lambda^{K+\frac12-\frac{tM}{4}}}{t\log\lambda} \qquad\text{and}\qquad C_1(d,c,\phi)\,\beta_{r+1}(\varepsilon)\,(t\,\log\lambda)^{r+1}\,\lambda^{\frac{K}{2}-\frac{r+1}{2}-\frac{d}{4}}.$$

We set $t=\frac{1}{M}(d+2r+2K+4)$, so that the two powers of λ agree. Note that

$$\frac{K_d}{2}-\frac{r+1}{2}-\frac{d}{4}=\frac{\frac{d}{2}+1}{2}-\frac{r+1}{2}-\frac{d}{4}=\begin{cases}-\frac{r}{2} & \text{if } d \text{ is even,}\\ -\frac{r}{2}-\frac14 & \text{if } d \text{ is odd.}\end{cases}$$

So,

$$\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\le\begin{cases}C_3(r,d,c,\phi)\,(\|\psi-\chi\|_{\mathscr{A}}+\beta_{r+1}(\varepsilon))\,\frac{(\log\lambda)^{r+1}}{\lambda^{\frac{r}{2}}} & \text{if } d \text{ is even,}\\ C_3(r,d,c,\phi)\,(\|\psi-\chi\|_{\mathscr{A}}+\beta_{r+1}(\varepsilon))\,\frac{(\log\lambda)^{r+1}}{\lambda^{\frac{r}{2}+\frac14}} & \text{if } d \text{ is odd.}\end{cases}$$

If d is odd, the claim is proven. Otherwise, notice that if μ and ν are two measures satisfying the hypotheses of Theorem 5.3 in even dimension d, then

$$\mu^+(k_1,\dots,k_d,k_{d+1})=\mu(k_1,\dots,k_d)\,\frac{\mathrm{e}^{-\lambda}\,\lambda^{k_{d+1}}}{(k_{d+1})!};$$
$$\nu^+(k_1,\dots,k_d,k_{d+1})=\nu(k_1,\dots,k_d)\,\frac{\mathrm{e}^{-\lambda}\,\lambda^{k_{d+1}}}{(k_{d+1})!}$$

are signed measures on $\mathbb{Z}^{d+1}$ that again satisfy the hypotheses of Theorem 5.3:

$$\widehat{\mu^+}(\xi_1,\dots,\xi_{d+1})=\widehat{\mu}(\xi_1,\dots,\xi_{d+1})\,\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}\xi_{d+1}}-1)}=\mathrm{e}^{\lambda\Phi(x_1,\dots,x_{d+1})}\,\psi(\xi_1,\dots,\xi_d);$$
$$\widehat{\nu^+}(\xi_1,\dots,\xi_{d+1})=\widehat{\nu}(\xi_1,\dots,\xi_{d+1})\,\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}\xi_{d+1}}-1)}=\mathrm{e}^{\lambda\Phi(x_1,\dots,x_{d+1})}\,\chi(\xi_1,\dots,\xi_d),$$

where $\Phi(\xi_1,\dots,\xi_{d+1})=\phi(\xi_1,\dots,\xi_d)+\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}\xi_{d+1}}-1)}$. Since

$$\left\|\widehat{\mu^+}-\widehat{\nu^+}\right\|_{\mathscr{A}}=\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\,,$$

we can apply our result in dimension $d+1$, assuming $r\ge\lfloor\frac{d+1}{2}\rfloor$:

$$\|\widehat{\mu}-\widehat{\nu}\|_{\mathscr{A}}\le C_3(r,d+1,c,\phi)\,(\|\psi-\chi\|_{\mathscr{A}}+\beta_{r+1}(\varepsilon))\,\frac{(\log\lambda)^{r+1}}{\lambda^{\frac{r}{2}+\frac14}}.$$

However, for d even, $\lfloor\frac{d+1}{2}\rfloor=\lfloor\frac{d}{2}\rfloor$, so the minimum r that one can take is again $\lfloor\frac{d}{2}\rfloor$. □

In comparison to the one dimensional case, we lost a logarithmic factor $(\log\lambda)^{r+1}$ in our estimate of norms (*cf.* Theorem 2.3). This will not be a problem for the calculation of asymptotics of distances.

5.2. **Asymptotics of distances in the multi-dimensional setting.** As in the one-dimensional case, a combination of the Laplace method and of the norm estimates given by Corollary 5.6 yields under appropriate hypotheses the asymptotics of $d_{\mathrm{L}}(\mu_n,\nu_n)$ and of $d_{\mathrm{TV}}(\mu_n,\nu_n)$, where $(\nu_n)_{n\in\mathbb{N}}$ is a general scheme of approximation of the laws μ_n of the mod-ϕ convergent random variables X_n in $\mathbb{Z}^d$. The only difference is that, for the computation of the total variation distance, we shall need to assume the order of approximation r to be larger than some minimal value $\lfloor\frac{d}{2}\rfloor$.

Until the end of this paragraph, $(X_n)_{n\in\mathbb{N}}$ is a sequence of random variables with values in $\mathbb{Z}^d$, which converges mod-ϕ with parameters $\lambda_n\to\infty$ and limit residue $\psi(\xi_1,\dots,\xi_d)$. In

the sequel, we shall write $(\nu_n)_{n\in\mathbb{N}}$ a sequence of signed measures on $\mathbb{Z}^d$ such that $\widehat{\nu_n}(\xi) = \mathrm{e}^{\lambda_n\phi(\xi)}\,\chi_n(\xi_1,\dots,\xi_d)$, with $\lim_{n\to\infty}\chi_n(\xi_1,\dots,\xi_d) = \chi(\xi_1,\dots,\xi_d)$. The multi-dimensional analogue of the hypothesis (H1) is:

$$\forall n\in\mathbb{N},\ \psi_n(\xi) - \chi_n(\xi) = \sum_{\alpha_1+\cdots+\alpha_d=r+1} \beta_n^{(\alpha_1,\dots,\alpha_d)}\,(\mathrm{i}\xi_1)^{\alpha_1}\cdots(\mathrm{i}\xi_d)^{\alpha_d}\,(1+o_\xi(1));$$

$$\text{and}\ \ \psi(\xi) - \chi(\xi) = \sum_{\alpha_1+\cdots+\alpha_d=r+1} \beta^{(\alpha_1,\dots,\alpha_d)}\,(\mathrm{i}\xi_1)^{\alpha_1}\cdots(\mathrm{i}\xi_d)^{\alpha_d}\,(1+o_\xi(1)) \qquad \text{(H1d)}$$

with $\lim_{n\to\infty}\beta_n^\alpha = \beta^\alpha$ for any choice of index $\alpha = (\alpha_1,\dots,\alpha_d)$ with $|\alpha| = r+1$. On the other hand, we write in the following $\|x\|^2 = \sum_{i=1}^d (x_i)^2$, $H_\alpha(x) = \prod_{i=1}^d H_{\alpha_i}(x_i)$, $\sigma^\alpha = \prod_{i=1}^d(\sigma_i)^{\alpha_i}$, and finally

$$\frac{k-\lambda_n m}{\sqrt{\lambda_n}\,\sigma} = \left(\frac{k_i-\lambda_n m_i}{\sqrt{\lambda_n}\,\sigma_i}\right)_{i\in[\![1,d]\!]}$$

if $k\in\mathbb{Z}^d$.

Proposition 5.7. *Under the assumption* (H1d) *with any $r\geq 0$, one has*

$$\mu_n(k_1,\dots,k_d) - \nu_n(k_1,\dots,k_d) =$$
$$\left(\sum_{|\alpha|=r+1}\frac{\beta^\alpha}{\sigma^\alpha}\,\frac{1}{\prod_{i=1}^d(\sqrt{2\pi}\,\sigma_i)}\,\mathrm{e}^{-\frac{1}{2}\left\|\frac{k-\lambda_n m}{\sqrt{\lambda_n}\sigma}\right\|^2}\,H_\alpha\!\left(\frac{k-\lambda_n m}{\sqrt{\lambda_n}\,\sigma}\right)\right)\frac{1}{(\lambda_n)^{\frac{r+d+1}{2}}} + o\!\left(\frac{1}{(\lambda_n)^{\frac{r+d+1}{2}}}\right)$$

with a remainder that is uniform over $\mathbb{Z}^d$.

Proof. The same arguments as in dimension 1 (d-dimensional Fourier inversion formula and Laplace method) yield

$$\begin{aligned}
&\mu_n(k_1,\dots,k_d) - \nu_n(k_1,\dots,k_d)\\
&= \sum_{|\alpha|=r+1}\frac{\beta_n^\alpha}{(2\pi)^d}\int_{[-\varepsilon,\varepsilon]^d}\left(\prod_{i=1}^d(\mathrm{i}\xi_i)^{\alpha_i}\,\mathrm{e}^{\lambda_n(\phi_i(\mathrm{i}\xi_i)-\mathrm{i}m_i\xi_i)}\,\mathrm{e}^{-\mathrm{i}\,(k-\lambda_n m_i)\xi_i}\right)\frac{d\xi}{(2\pi)^d} + o\!\left(\frac{1}{(\lambda_n)^{\frac{r+d+1}{2}}}\right)\\
&= \sum_{|\alpha|=r+1}\frac{\beta_n^\alpha}{(2\pi)^{\frac{d}{2}}}\prod_{i=1}^d\left(\frac{1}{(\sqrt{\lambda_n}\,\sigma_i)^{\alpha_i+1}}\int_{\mathbb{R}}(\mathrm{i}x)^{\alpha_i}\,\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{e}^{-\mathrm{i}\left(\frac{k_i-\lambda_n m_i}{\sqrt{\lambda_n}\,\sigma_i}\right)x}\frac{dx}{\sqrt{2\pi}}\right) + o\!\left(\frac{1}{(\lambda_n)^{\frac{r+d+1}{2}}}\right),
\end{aligned}$$

and the same Hermite polynomials as in Proposition 3.2 appear. □

Using the same argument of density of the sequence of lattices $\{\frac{k-\lambda_n m}{\sqrt{\lambda_n}\,\sigma},\ k\in\mathbb{Z}^d\}$ as in the proof of Theorem 3.3, we conclude that:

Theorem 5.8. *Under the hypothesis* (H1d)*, one has*

$$\begin{aligned}
&d_{\mathrm{L}}(\mu_n,\nu_n)\\
&= \sup_{x\in\mathbb{R}^d}\left(\mathrm{e}^{-\frac{\|x\|^2}{2}}\left|\sum_{|\alpha|=r+1}\frac{\beta^\alpha\,H_\alpha(x)}{\sigma^\alpha}\right|\right)\times\frac{1}{\left(\prod_{i=1}^d\sqrt{2\pi}\,\sigma_i\right)(\lambda_n)^{\frac{r+d+1}{2}}} + o\!\left(\frac{1}{(\lambda_n)^{\frac{r+d+1}{2}}}\right).
\end{aligned}$$

The proof of the multi-dimensional analogue of Theorem 3.11 is a bit more involved, but the main difficulty was lying in the proof of the norm estimate (Corollary 5.6). In the following we fix $r\geq\lfloor\frac{d}{2}\rfloor$; then, $r+1\geq K_d$. The multi-dimensional version of the

hypothesis (H2) is the following. We assume the residues ψ, ψ_n, χ and χ_n to be of class $\mathscr{C}^{K_d}$ on $\mathbb{T}^d$, with $\partial^\delta(\psi_n-\chi_n)(0)=\partial^\delta(\psi-\chi)(0)=0$ for any $|\delta|<K_d$. We also assume that there exist families of coefficients $(\beta_n^\alpha)_{|\alpha|=r+1,n\in\mathbb{N}}$ and $(\gamma_n^\alpha)_{|\alpha|=r+2,n\in\mathbb{N}}$, such that for any multi-index δ with $|\delta|=K_d$,

$$\partial^\delta(\psi_n-\chi_n)(\xi)=\mathrm{i}^{|\delta|}\left(\sum_{|\rho|=r+1-|\delta|}\beta_n^{\rho+\delta}\,\frac{(\rho+\delta)!}{\rho!}\,(\mathrm{i}\xi)^\rho+\sum_{|\rho|=r+2-|\delta|}\gamma_n^{\rho+\delta}\,\frac{(\rho+\delta)!}{\rho!}\,(\mathrm{i}\xi)^\rho\,(1+o_\xi(1))\right)$$
$$\partial^\delta(\psi-\chi)(\xi)=\mathrm{i}^{|\delta|}\left(\sum_{|\rho|=r+1-|\delta|}\beta^{\rho+\delta}\,\frac{(\rho+\delta)!}{\rho!}\,(\mathrm{i}\xi)^\rho+\sum_{|\rho|=r+2-|\delta|}\gamma^{\rho+\delta}\,\frac{(\rho+\delta)!}{\rho!}\,(\mathrm{i}\xi)^\rho\,(1+o_\xi(1))\right) \tag{H2d}$$

with $\lim_{n\to\infty}\beta_n^\alpha=\beta^\alpha$ and $\lim_{n\to\infty}\gamma_n^\alpha=\gamma^\alpha$. Here, given a multi-index α, we write $\alpha!=\prod_{i=1}(\alpha_i)!$ and $(\mathrm{i}\xi)^\alpha=\prod_{i=1}^d(\mathrm{i}\xi_i)^{\alpha_i}$. These hypotheses are satisfied *e.g.* if the convergences $\psi_n\to\psi$ and $\chi_n\to\chi$ occur in $\mathscr{C}^{r+2}(\mathbb{T}^d)$, and if $(\nu_n)_{n\in\mathbb{N}}$ is the scheme of approximation of order r of the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ (its definition in dimension $d\geq 2$ is similar to the definition in dimension 1 given in Example 1.11, see also the two worked examples hereafter).

Under the hypothesis (H2d), if one wants to prove a bound on $d_{\mathrm{TV}}(\mu_n,\nu_n)$ of order $O((\lambda_n)^{-\frac{r+1}{2}})$, then one can replace ν_n with the measure ρ_n defined by the Fourier transform

$$\widehat{\rho_n}(\xi)=\mathrm{e}^{\lambda_n\phi(\xi)}\left(\psi_n(\xi)-\sum_{|\alpha|=r+1}\beta_n^\alpha\prod_{i=1}^d(\mathrm{e}^{\mathrm{i}\xi_i}-1)^{\alpha_i}\right).$$

Indeed, denoting $\widetilde{\psi_n}(\xi)$ the residue associated to ρ_n, the hypothesis (H2d) implies that, for any $|\delta|\leq K_d$,

$$\left|\partial^\delta(\widetilde{\psi_n}-\chi_n)(\xi)\right|\leq C\,|\xi|^{r+2-|\delta|}$$

with a constant C that is uniform, and also valid for the derivatives $|\partial^\delta(\widetilde{\psi}-\chi)(\xi)|$. This bound is the only one useful in the proof of Theorem 5.3 and its Corollary 5.6, and therefore, one can apply Corollary 5.6 with parameter $r+1$ instead of r. We thus get

$$d_{\mathrm{TV}}(\rho_n,\nu_n)=O\left(\frac{(\log\lambda_n)^{r+2}}{(\lambda_n)^{\frac{r+1}{2}+\frac14}}\right)=o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

From there, the proofs are essentially identical to the one dimensional case (Theorem 3.11), and one obtains:

Theorem 5.9. *Consider a general approximation scheme $(\nu_n)_{n\in\mathbb{N}}$ of a mod-ϕ convergent sequence $(X_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}^d$. We assume that $r\geq\lfloor\frac{d}{2}\rfloor$ and that the hypothesis (H2d) is satisfied. Then,*

$$d_{\mathrm{TV}}(\mu_n,\nu_n)=\frac{1}{(2\pi)^{\frac{d}{2}}\,(\lambda_n)^{\frac{r+1}{2}}}\left(\int_{\mathbb{R}^d}\mathrm{e}^{-\frac{\|x\|^2}{2}}\left|\sum_{|\alpha|=r+1}\frac{\beta^\alpha\,H_\alpha(x)}{\sigma^\alpha}\right|dx\right)+o\left(\frac{1}{(\lambda_n)^{\frac{r+1}{2}}}\right).$$

Last, the discussion of §3.4 holds *mutatis mutandis* in the multi-dimensional setting, and if $\|\chi_n-\chi\|_{\mathscr{C}^{K_d}}$ is asymptotically smaller than any negative power of λ_n, then one can replace $(\nu_n)_{n\in\mathbb{N}}$ by the derived scheme $(\sigma_n)_{n\in\mathbb{N}}$ with constant residue and keep the conclusions of Theorems 5.8 and 5.9.

5.3. **Cycle-type of random coloured permutations.** In this section, we study random elements of the group of d-coloured permutations:

$$G_n = \mathfrak{S}_n \wr (\mathbb{Z}/d\mathbb{Z}) = \mathfrak{S}_n \ltimes (\mathbb{Z}/d\mathbb{Z})^n .$$

It can be viewed as the group of complex matrices of size n times n, with one non-zero coefficient on each row and each column, and these non-zero coefficients belonging to the group of d-th roots of unity $\mathbb{Z}/d\mathbb{Z} = \{1, \omega, \omega^2, \dots, \omega^{d-1}\}$, with $\omega = \mathrm{e}^{\frac{2\mathrm{i}\pi}{d}}$. Another alternative definition is the following:

$$G_n = \{\sigma \in \mathfrak{S}_{nd} = \mathfrak{S}(\mathbb{Z}/n\mathbb{Z}\times\mathbb{Z}/d\mathbb{Z}) \mid \forall (a,b) \in \mathbb{Z}/n\mathbb{Z}\times\mathbb{Z}/d\mathbb{Z},\ \ \sigma(a,b+1) = \sigma(a,b)+(0,1)\}.$$

From this definition, it appears that a d-coloured permutation is entirely determined by the images of the elements $(a,0)$ with $a \in \mathbb{Z}/n\mathbb{Z}$:

$$\sigma(a,0) = (\rho(a), k(a)) \quad \text{with } \rho \in \mathfrak{S}_n = \mathfrak{S}(\mathbb{Z}/n\mathbb{Z}).$$

One then has $\sigma(a,b) = (\rho(a), k(a)+b)$. In particular, $\operatorname{card} G_n = n!\, d^n$, the factor $n!$ coming from the choice of $\rho \in \mathfrak{S}_n$, and the factor d^n from the choice of the elements $k(a)$.

Let us identify the conjugacy classes of elements of G_n (we refer to [CSST14, §2.3]). Suppose that $\sigma_2 = \tau\, \sigma_1\, \tau^{-1}$ in G_n, with σ_i associated to the pair $(\rho_i, k_i) \in \mathfrak{S}_n \times (\mathbb{Z}/d\mathbb{Z})^n$, and τ associated to the pair (ν, l). Note that $\tau^{-1} = (\nu^{-1}, -l \circ \nu^{-1})$, this identity coming from the product rule $(\rho_1, k_1)(\rho_2, k_2) = (\rho_1\rho_2, k_1 \circ \rho_2 + k_2)$. Then,

$$\begin{aligned}(\rho_2, k_2) &= (\nu, l)(\rho_1, k_1)(\nu^{-1}, -l \circ \nu^{-1}) \\ &= (\nu, l)(\rho_1\nu^{-1}, (k_1 - l) \circ \nu^{-1}) \\ &= (\nu\rho_1\nu^{-1}, (k_1 + l \circ \rho_1 - l) \circ \nu^{-1}),\end{aligned}$$

so in particular ρ_1 and ρ_2 are conjugated in $\mathfrak{S}_n$, so they have the same cycle-type (an integer partition L of size $|L| = n$). Then, to determine the conjugacy class of (ρ_1, k_1), taking the conjugate of (ρ_2, k_2) by $(\nu^{-1}, 0)$, we can assume that $\rho_1 = \rho_2 = \rho$, in which case

$$\sigma_1 = (\rho, k_1) \qquad ; \qquad \sigma_2 = (\rho, k_2 = k_1 + (l \circ \rho - l)),$$

that is to say that k_2 and k_1 differ by a *cocyle* $l \circ \rho - l$. This is possible if and only if, for every cycle $c = (a = \rho^r(a), \rho(a), \rho^2(a), \dots, \rho^{r-1}(a))$ of $\rho \in \mathfrak{S}_n$, one has

$$k_1(a) + k_1(\rho(a)) + \cdots + k_1(\rho^{r-1}(a)) = k_2(a) + k_2(\rho(a)) + \cdots + k_2(\rho^{r-1}(a)).$$

Indeed, one has

$$\sum_{j=0}^{r-1} (l \circ \rho - l)(\rho^j(a)) = l(\rho^r(a)) - l(a) = l(a) - l(a) = 0.$$

In the previous setting, denote $\Sigma(c,k) = \sum_{j=0}^{r-1} k(\rho^j(a))$, which is an element of $\mathbb{Z}/d\mathbb{Z}$. This quantity depends only on the cycle c, and not on the choice of a representative a of the corresponding orbit; it allows to one to "color" each cycle of ρ. Then (*cf.* [CSST14, Theorem 2.3.5], applied to the abelian case):

Theorem 5.10. *Two coloured permutations $\sigma_1 = (\rho_1, k_1)$ and $\sigma_2 = (\rho_2, k_2)$ are conjugated in $G_n = \mathfrak{S}_n \wr (\mathbb{Z}/d\mathbb{Z})$ if and only if:*

(1) the permutations ρ_1 and ρ_2 have the same lengths of cycles (encoded by an integer partition L of size n);

(2) and there is a size-preserving bijection $c \mapsto \psi(c)$ between the cycles of ρ_1 and ρ_2, such that the colors are also preserved:

$$\Sigma(c, k_1) = \Sigma(\psi(c), k_2).$$

Consequently, the conjugacy classes of G_n are labelled by the d-uples of integer partitions $L = (L^{(1)}, L^{(2)}, \dots, L^{(d)})$, such that $\sum_{i=1}^{d} |L^{(i)}| = n$. If $\sigma \in G_n$ is of type L, then the parts of $L^{(i)}$ are the sizes of the cycles of σ with color $i \in \mathbb{Z}/d\mathbb{Z}$.

Example 5.11. Consider the following 2-coloured permutation in $G_8 = \mathfrak{S}_8 \wr \mathbb{Z}/2\mathbb{Z}$:

$$\sigma = \big(38762415, 11010001\big),$$

where the first part is the word $\rho(1)\rho(2)\dots\rho(8)$ of the permutation ρ of size 8 underlying σ, and the second part is the word $k(1)k(2)\dots k(8)$. The three disjoint cycles of ρ are $(1,3,7)$, $(2,8,5)$ and $(4,6)$, and the associated colors are $1+0+0=1$, $1+1+0=0$ and $1+0=1$. Therefore, the cycle-type of σ is

$$L^{(1)} = (3,2), L^{(2)} = (3)$$

since the sizes of the cycles of color 1 are 3 and 2, and the size of the unique cycle of color 0 is 3. We leave the reader check that the other 2-coloured permutation

$$\sigma' = \big(73864512, 10011100\big)$$

has the same cycle-type $((3,2),(3))$. Therefore, σ and σ' are conjugated in G_n.

In the following, we denote σ_n a random uniform element of the wreath product G_n, and $\ell_n = (\ell_n^{(1)}, \ell_n^{(2)}, \dots, \ell_n^{(d)})$ the random d-uple of integers that consists in the lengths of the partitions $L^{(1)}, \dots, L^{(d)}$ of the cycle-type of σ_n. Our goal is to apply the results of the previous paragraph to these random vectors. To this purpose, it is convenient to construct a coupling of all the random permutations $(\sigma_n)_{n\in\mathbb{N}}$, which generalizes the Feller coupling in the case $d = 1$ of the Feller coupling. Algebraically, this amounts to the following:

Lemma 5.12. *We consider G_n as a subgroup of G_{n+1}, by sending a pair*

$$(\rho \in \mathfrak{S}_n, (k_1, \dots, k_n) \in (\mathbb{Z}/d\mathbb{Z})^n) \quad \text{to the pair} \quad (\rho \in \mathfrak{S}_{n+1}, (k_1, \dots, k_n, 0) \in (\mathbb{Z}/d\mathbb{Z})^{n+1}).$$

Let σ_{n+1} be an element of G_{n+1}. There exists a unique element $\sigma_n \in G_n$, and a unique coloured transposition

$$\tau_n = ((j, n+1), (0, \dots, 0, \alpha, 0, \dots, 0, -\alpha)) \quad \text{with } j \in [\![1, n+1]\!],$$

and the $\alpha \in \mathbb{Z}/d\mathbb{Z}$ at position j, such that $\sigma_{n+1} = \tau_n\sigma_n$. Here we agree that the trivial transposition $(n+1, n+1)$ is the identity permutation; in this case we put α in position $n+1$, and there is no term $-\alpha$.

Proof. Denote $\sigma_{n+1} = (\rho_{n+1}, (k_1, \dots, k_{n+1}))$. If $\rho_n = \rho_{n+1} \circ (j = \rho_{n+1}^{-1}(n+1), n+1)$, then ρ_n sends $n+1$ to $n+1$, so it can be considered as a permutation in $\mathfrak{S}_n$. We then have $\rho_{n+1} = \rho_n \circ (j, n+1)$. Set

$$\sigma_n = (\rho_n, (k_1, \dots, k_{j-1}, k_j + k_{n+1}, k_{j+1}, \dots, k_n)) \in G_n.$$

We then have

$$\begin{aligned}
\sigma_n\tau_n &= (\rho_n, (k_1, \dots, k_j + k_{n+1}, \dots, k_n, 0))((j, n+1), (0, \dots, k_j, \dots, -k_j)) \\
&= (\rho_n \circ (j, n+1), (k_1, \dots, 0, \dots, k_n, k_j + k_{n+1}) + (0, \dots, k_j, \dots, 0, -k_j)) \\
&= (\rho_{n+1}, (k_1, \dots, k_{n+1})) = \sigma_{n+1}.
\end{aligned}$$

The unicity of τ_n comes from a cardinality argument: $\frac{\operatorname{card} G_{n+1}}{\operatorname{card} G_n} = (n+1)d$, and this is the number of d-coloured transpositions $((j, n+1), (0, \dots, 0, \alpha, 0, \dots, 0, -\alpha))$. □

We denote $T(j, \alpha)$ the coloured transposition of the previous lemma. Then, in order to construct a random coloured permutation $\sigma_n \in G_n$, it suffices to take random independent positive integers $j_1 \leq 1, j_2 \leq 2, \dots, j_n \leq n$, and random independent elements $\alpha_1, \dots, \alpha_n \in \mathbb{Z}/d\mathbb{Z}$, and to consider the product

$$\sigma_n = T(j_1, \alpha_1) \circ T(j_2, \alpha_2) \circ \cdots \circ T(j_n, \alpha_n).$$

We also denote $(\mathcal{F}_n)_{n\in\mathbb{N}}$ the filtration of probability spaces associated to the sequence of random coloured permutations $(\sigma_n)_{n\in\mathbb{N}}$. The interest of our construction is that one can easily follow the evolution of the cycle type of σ_n. More precisely, if $j_n = n$, then to construct σ_n, one adds a cycle of length 1 to σ_{n-1}, with color α_n. Thus, for every color $i \in \mathbb{Z}/d\mathbb{Z}$, there is a probability $\frac{1}{nd}$ to increase the length $\ell^{(i)}$ by one unit:

$$\forall i \in [\![1, d]\!], \ \mathbb{P}[\ell_n = \ell_{n-1} + (0, \dots, 0, 1_i, 0, \dots, 0) | \mathcal{F}_{n-1}] = \frac{1}{nd}.$$

On the other hand, if $j_n \neq n$, then the multiplication by $T(j_n, \alpha_n)$ increases by 1 the size of the cycle containing j_n in σ_{n-1}, but it does not change its color, because of the terms α_n and $-\alpha_n$ that compensate one another. So,

$$\mathbb{P}[\ell_n = \ell_{n-1} | \mathcal{F}_{n-1}] = 1 - \frac{1}{n}.$$

We have therefore proven that ℓ_n is a sum of independent increments in $\mathbb{Z}^d$, with

$$\mathbb{P}[\ell_n - \ell_{n-1} = 0] = 1 - \frac{1}{n} \quad ; \quad \mathbb{P}[\ell_n - \ell_{n-1} = e_i] = \frac{1}{nd},$$

where $(e_i)_{i\in[\![1,d]\!]}$ is the canonical basis of the lattice $\mathbb{Z}^d$.

As an immediate consequence of the previous discussion, we have:

Theorem 5.13. *The sequence of random vectors $(\ell_n)_{n\in\mathbb{N}}$ converges modulo a d-dimensional Poisson law of exponent $\phi(\xi) = \frac{1}{d}\sum_{i=1}^{d}(\mathrm{e}^{\mathrm{i}\xi_i} - 1)$, with parameters H_n and limiting function*

$$\begin{aligned}\psi(\xi) &= \prod_{n=1}^{\infty}\left(1 + \frac{1}{nd}\sum_{i=1}^{d}(\mathrm{e}^{\mathrm{i}\xi_i} - 1)\right)\mathrm{e}^{-\frac{1}{nd}\sum_{i=1}^{d}(\mathrm{e}^{\mathrm{i}\xi_j} - 1)} \\ &= \mathrm{e}^{\gamma\,\phi(\xi)}\,\frac{1}{\Gamma\left(\frac{1}{d}\sum_{i=1}^{d}\mathrm{e}^{\mathrm{i}\xi_i}\right)}.\end{aligned}$$

Proof. Since the increments of ℓ_n are independent,

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\langle\xi\,|\,\ell_n\rangle}] = \prod_{j=1}^{n}\left(1 + \frac{1}{jd}\sum_{i=1}^{d}(\mathrm{e}^{\mathrm{i}\xi_i} - 1)\right) = \mathrm{e}^{H_n\,\phi(\xi)}\,\psi_n(\xi)$$

where $\psi_n(\xi)$ is the partial product over indices in $[\![1, n]\!]$ coming from the infinite product $\psi(\xi)$. □

In particular, since the Fourier transform of ℓ_n does not factorize over the coordinates ξ_i, the coordinates of ℓ_n are not independent, though in the asymptotics $n \to \infty$ and at first order, ℓ_n looks like a d-uple of independent Poisson variables of parameters $\frac{H_n}{d}$. In the following, in order to simplify a bit the discussion, we assume $d = 2$, which already contains all the subtleties of the general case. We denote (x, y) the coordinates in $\mathbb{R}^2$, and (ξ, ζ) the coordinates in the Fourier space. We can then write:

$$\begin{aligned}\psi_n(\xi,\zeta) &= \prod_{j=1}^{n}\left(1+\frac{1}{j}\left(\frac{\mathrm{e}^{\mathrm{i}\xi}+\mathrm{e}^{\mathrm{i}\zeta}}{2}-1\right)\right)\mathrm{e}^{-\frac{1}{j}\left(\frac{\mathrm{e}^{\mathrm{i}\xi}+\mathrm{e}^{\mathrm{i}\zeta}}{2}-1\right)}\\ &= \sum_{k=0}^{\infty}\mathfrak{e}_{k,n}\left(\frac{\mathrm{e}^{\mathrm{i}\xi}+\mathrm{e}^{\mathrm{i}\zeta}}{2}-1\right)^k = 1-\frac{\mathfrak{p}_{2,n}}{2}\left(\frac{\mathrm{e}^{\mathrm{i}\xi}+\mathrm{e}^{\mathrm{i}\zeta}}{2}-1\right)^2+o(\|(\xi,\zeta)\|^2)\end{aligned}$$

where the $\mathfrak{e}_{k,n}$'s are obtained from the parameters

$$\mathfrak{p}_{1,n}=0 \qquad ; \qquad \mathfrak{p}_{k\geq 2,n}=\sum_{j=1}^{n}\frac{1}{j^k}$$

by the same recipe as in Section 4. The scheme of approximation of $(\ell_n)_{n\in\mathbb{N}}$ of order $r = 1$ is the Poissonian approximation

$$\widehat{\nu_n}(\xi,\zeta)=\mathrm{e}^{H_n\left(\frac{\mathrm{e}^{\mathrm{i}\xi}+\mathrm{e}^{\mathrm{i}\zeta}}{2}-1\right)} \qquad ; \qquad \nu_n(k,l)=\mathrm{e}^{-H_n}\frac{(H_n)^{k+l}}{2^{k+l}\,k!\,l!}.$$

It satisfies the hypotheses (H1d) and (H2d), with $\beta_n^{(2,0)}=\beta_n^{(0,2)}=-\frac{\mathfrak{p}_{2,n}}{8}$ and $\beta_n^{(1,1)}=-\frac{\mathfrak{p}_{2,n}}{4}$. Since $H_{(2,0)}(x,y)=x^2-1$, $H_{(1,1)}(x,y)=xy$ and $H_{(0,2)}(x,y)=y^2-1$, Theorems 5.8 and 5.9 ensure that

$$\begin{aligned}d_{\mathrm{L}}(\mu_n,\nu_n) &= \frac{\pi}{6\,(\log n)^2}\sup_{(x,y)\in\mathbb{R}^2}\left|\mathrm{e}^{-\frac{x^2+y^2}{2}}\,(2-(x+y)^2)\right|+o\left(\frac{1}{(\log n)^2}\right);\\ d_{\mathrm{TV}}(\mu_n,\nu_n) &= \frac{\pi}{24\,\log n}\left(\int_{\mathbb{R}^2}\mathrm{e}^{-\frac{x^2+y^2}{2}}\,|2-(x+y)^2|\,dx\,dy\right)+o\left(\frac{1}{\log n}\right)\end{aligned}$$

since $\sum_{|\alpha|=2}\frac{\beta^\alpha H_\alpha(x)}{\sigma^\alpha}=-\frac{\mathfrak{p}_2}{2}((x+y)^2-2)=\frac{\pi^2}{12}\,(2-(x+y)^2)$.

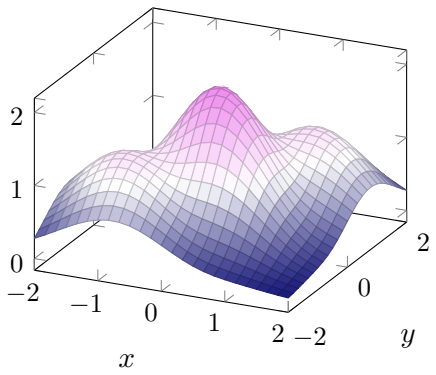

FIGURE 5. The function $\mathrm{e}^{-\frac{x^2+y^2}{2}}\,|2-(x+y)^2|$.

For the local distance, one checks at once that the maximum of the function $f(x,y) = \mathrm{e}^{-\frac{x^2+y^2}{2}}\,|2-(x+y)^2|$ is obtained when $x=y=0$, so

$$d_{\mathrm{L}}(\mu_n,\nu_n) = \frac{\pi}{3\,(\log n)^2}\,(1+o(1)),$$

see Figure 5. For the total variation distance, a computer algebra system (`Sage`) yields the approximate value $12.162\ldots$ for the integral. As in the one-dimensional case (Example 1.6 and §4.3), one can on the other hand construct better schemes of approximations, which yield distances smaller than any arbitrary negative power of $\log n$. It is important to notice that the dependence between the different coordinates of ℓ_n is directly involved in these asymptotics of distances.

5.4. **Distinct prime divisors counted according to their residue classes.** Similar to the generalisation of §4.3 by the study of random coloured permutations, there is a natural generalisation of the discussion of §4.5 on distinct prime divisors of a random integer, which involves *residue classes* of prime numbers.

Fix an integer $a \geq 2$, and denote $d = \varphi(a)$ the cardinality of the multiplicative group $(\mathbb{Z}/a\mathbb{Z})^*$. This quantity is the usual Euler φ-function, and it is equal to the number of integers in $[\![1,a]\!]$ that are coprime to a. We label the elements of $(\mathbb{Z}/a\mathbb{Z})^*$ as $b_1, b_2, \ldots, b_d$. If $n \in \mathbb{N}$ and $i \in [\![1,d]\!]$, we denote $\omega_i(n)$ the number of distinct prime divisors of n that have residue class b_i modulo a:

$$\omega_i(n) = \operatorname{card}\{p \in \mathbb{P},\ p \mid n \text{ and } p \equiv b_i \bmod a\}.$$

One then has

$$\sum_{i\in(\mathbb{Z}/a\mathbb{Z})^*} \omega_i(n) = \omega(n) - \operatorname{card}\{p \in \mathbb{P},\ p \mid a \text{ and } p \mid n\}.$$

We are interested in the random vector $\Omega(N_n) = (\omega_1(N_n), \omega_2(N_n), \ldots, \omega_d(N_n))$, where N_n is a random integer chosen uniformly in $[\![1,n]\!]$. In view of the theory developed in the previous sections, the asymptotics of these vectors are encoded in the multiple generating series

$$\frac{1}{n}\sum_{i=1}^{n} (z_1)^{\omega_1(i)}(z_2)^{\omega_2(i)}\cdots(z_d)^{\omega_d(i)}.$$

In number theory, the asymptotics of such quantities are classically related to the behavior of the Dirichlet series

$$F(z,s) = \sum_{n=1}^{\infty} \frac{(z_1)^{\omega_1(n)}(z_2)^{\omega_2(n)}\cdots(z_d)^{\omega_d(n)}}{n^s}.$$

In a moment we shall precise these relations, which amount to the so-called Selberg–Delange method, see [Ten95, Chapter II.5]. Note that, though we want to study a random vector in $\mathbb{Z}^{d\geq 2}$, the Dirichlet series written above is not a multiple Dirichlet series.

The main algebraic tool that is required in order to use Selberg–Delange method is the theory of Dirichlet characters and their L-series, see for instance Chapter II.8 in [Ten95]. In the following we recall the basics of this theory. The space of functions $(\mathbb{Z}/a\mathbb{Z})^* \to \mathbb{C}$ is endowed with a Hilbert structure

$$\langle f \mid g\rangle = \frac{1}{\varphi(a)} \sum_{i\in(\mathbb{Z}/a\mathbb{Z})^*} f(i)\overline{g(i)},$$

and a Hilbert basis consists of the so-called *Dirichlet characters* $\chi_{a,1}, \dots, \chi_{a,d}$, which are the morphisms of (multiplicative) groups $\chi : (\mathbb{Z}/a\mathbb{Z})^* \to \mathbb{C}^*$. In particular, there are as many Dirichlet characters as elements of the group $(\mathbb{Z}/a\mathbb{Z})^*$, *i.e.*, d distinct characters. In the following, we denote $\chi_{a,1}$ the trivial character: $\chi_{a,1}(i) = 1$ for all $i \in (\mathbb{Z}/a\mathbb{Z})^*$.

If χ is a character of $(\mathbb{Z}/a\mathbb{Z})^*$, introduce its L-function

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

where the function χ is extended to $\mathbb{Z}$ by

$$\chi(n) = \begin{cases} \chi(m) & \text{if } (n,a) = 1 \text{ and } m \equiv n \bmod a, \\ 0 & \text{if } (n,a) > 1. \end{cases}$$

These functions admit an Euler product representation:

$$L(\chi, s) = \prod_{p \in \mathbb{P}_a} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1},$$

where $\mathbb{P}_a$ is the set of prime numbers that do not divide a. In particular, $L(\chi_{a,1}, s)$ is almost the same as Riemann's ζ-function:

$$L(\chi_{a,1}, s) = \prod_{p \in \mathbb{P}_a} \left(1 - \frac{1}{p^s}\right)^{-1} = \zeta(s) \prod_{p \mid a} \left(1 - \frac{1}{p^s}\right).$$

Therefore, the L-function associated to the trivial character has abscissa of convergence 1, and can be extended to a meromorphic function on the half-plane $\mathrm{Re}(s) > 0$, with a single pole at $s = 1$. On the other hand, for the other characters $\chi_{a,2}, \dots, \chi_{a,d}$, the corresponding L-functions converge simply towards holomorphic functions on the same half-plane $\mathrm{Re}(s) > 0$.

We now form the Dirichlet series

$$F(z, s) = \sum_{n \geq 1} \frac{(z_1)^{\omega_1(n)} \cdots (z_d)^{\omega_d(n)}}{n^s}.$$

For any choice of parameters $z_1, \dots, z_d \in \mathbb{C}$, the series $F(z, s)$ converges absolutely on the half-plane $\mathrm{Re}(s) > 1$, as can be seen from the inequality

$$\sum_{n \geq 1} \left| \frac{(z_1)^{\omega_1(n)} \cdots (z_d)^{\omega_d(n)}}{n^s} \right| \leq \prod_{p \in \mathbb{P}_a} \left(1 + \frac{\max_{i \in [\![1,d]\!]} |z_i|}{p^s - 1}\right).$$

Given parameters $y_1, \dots, y_d$, we set

$$G_\chi(z, s) = F(z, s) \prod_{j=1}^{d} L(\chi_{a,j}, s)^{-y_i}.$$

Let us choose the parameters $y_1, \dots, y_d$ so that the series $G_\chi(z,s)$ is an holomorphic function on $\mathrm{Re}(s) > \frac{1}{2}$. If $\mathrm{Re}(s) > 1$, we can write without ambiguity

$$\begin{aligned} F(z,s) &= \prod_{i=1}^{d} \prod_{p \equiv b_i \bmod a} \left(1 + \frac{z_i}{p^s - 1}\right); \\ G_\chi(z,s) &= \prod_{i=1}^{d} \prod_{p \equiv b_i \bmod a} \left(1 + \frac{z_i}{p^s - 1}\right) \prod_{j=1}^{d} \left(1 - \frac{\chi_{a,j}(p)}{p^s}\right)^{y_j} \\ &= \prod_{p \in \mathbb{P}_a} \left(\left(1 + \frac{z_{i(p)}}{p^s - 1}\right) \prod_{j=1}^{d} \left(1 - \frac{\chi_{a,j}(b_{i(p)})}{p^s}\right)^{y_j} \right) \end{aligned}$$

where $i(p)$ is the unique $i \in [\![1,d]\!]$ such that $p \equiv b_i \bmod a$. The Taylor expansion of the term corresponding to $p \in \mathbb{P}_a$ is the previous product is

$$1 + \frac{1}{p^s} \left(z_{i(p)} - \sum_{j=1}^{d} y_j \, \chi_{a,j}(b_{i(p)}) \right) + O\left(\frac{1}{p^{2s}}\right).$$

Set $y_j = \frac{1}{d} \sum_{k=1}^{d} z_k \, \overline{\chi_{a,j}(b_k)}$. Then, by orthogonality of the Dirichlet characters, one has:

$$\begin{aligned} \sum_{j=1}^{d} y_j \, \chi_{a,j}(b_i) &= \frac{1}{d} \sum_{j,k=1}^{d} z_k \overline{\chi_{a,j}(b_k)} \, \chi_{a,j}(b_i) \\ &= d \sum_{j=1}^{d} \left\langle \sum_{k=1}^{d} z_k b_k \,\middle|\, \chi_{a,j} \right\rangle \langle \chi_{a,j} \mid b_i \rangle \\ &= d \left\langle \sum_{k=1}^{d} z_k b_k \,\middle|\, b_i \right\rangle = z_i. \end{aligned}$$

Hence, for this choice of parameters, $G_\chi(z,s) = \prod_{p \in \mathbb{P}_a} (1 + O(p^{-2s}))$, so $G_\chi(z,d)$ is an holomorphic function on $\mathrm{Re}(s) > \frac{1}{2}$. Note that $y_1 = \frac{z_1 + \dots + z_d}{d}$. Now, by the previous discussion on L-series, one can remultiply $G_\chi(z,s)$ by $\prod_{j \neq 1} L(\chi_{a,j}, s)^{y_j}$, hence:

Proposition 5.14. *For any choice of complex parameters $z_1, \dots, z_d$, the series*

$$G(z,s) = F(z,s) \, (\zeta(s))^{-\frac{z_1 + \dots + z_d}{d}}$$

has an holomorphic extension on the half-plane $\mathrm{Re}(s) > \frac{1}{2}$.

Proof. If $L(\chi_{a,1}, s) = \zeta_a(s)$ is the partial ζ-function associated to the integer a, then we shown that $F(z,s) \, (\zeta_a(s))^{-\frac{z_1 + \dots + z_d}{d}}$ is an holomorphic function on $\mathrm{Re}(s) > \frac{1}{2}$. It suffices then to multiply by the missing terms

$$\prod_{p \mid a} \left(1 - \frac{1}{p^s}\right)^{\frac{z_1 + \dots + z_d}{d}}. \qquad \square$$

We can now apply Theorem 3 in [Ten95, Chapter II.5]:

Theorem 5.15. *When n goes to infinity,*

$$\frac{1}{n} \sum_{i=1}^{n} (z_1)^{\omega_1(n)} \cdots (z_d)^{\omega_d(n)} = (\log n)^{\frac{z_1 + \dots + z_d}{d} - 1} \frac{G((z_1, \dots, z_d), 1)}{\Gamma\left(\frac{z_1 + \dots + z_d}{d}\right)} \left(1 + O\left(\frac{1}{\log n}\right)\right),$$

with a remainder that is locally uniform in the parameters $z_1, \dots, z_d$.

As a corollary, the random vectors $(\Omega(N_n))_{n\in\mathbb{N}}$ of numbers of distinct prime divisors in each residue class of $(\mathbb{Z}/a\mathbb{Z})^*$ converge mod-ϕ, where $\phi(\xi) = \frac{1}{d}\sum_{j=1}^{d}(\mathrm{e}^{\mathrm{i}\xi_j} - 1)$. The parameters of this mod-ϕ convergence are $\lambda_n = \log\log n$, and the limiting residue is

$$\psi(\xi) = \frac{G((\mathrm{e}^{\mathrm{i}\xi_1}, \dots, \mathrm{e}^{\mathrm{i}\xi_d}), 1)}{\Gamma\left(\frac{\mathrm{e}^{\mathrm{i}\xi_1}+\cdots+\mathrm{e}^{\mathrm{i}\xi_d}}{d}\right)}.$$

Moreover, the convergence happens at speed $O(\frac{1}{\log n})$. As a consequence, one can construct explicit schemes of approximations of the laws of the vectors $(\Omega(N_n))_{n\in\mathbb{N}}$, which yield distances that are $O((\log\log n)^{-\frac{p}{2}})$ with $p \geq 1$ arbitrary (see Theorems 5.8 and 5.9). The first of these schemes is the Poisson approximation:

$$\widehat{\nu_n}(\xi) = \mathrm{e}^{(\log\log n)\frac{1}{d}\sum_{i=1}^{d}(\mathrm{e}^{\mathrm{i}\xi_i}-1)}.$$

Unfortunately, it is then difficult to calculate the constants involved in these asymptotics of distances. Indeed, there are no simple expression for the values of $G((\mathrm{e}^{\mathrm{i}\xi_1}, \dots, \mathrm{e}^{\mathrm{i}\xi_d}), 1)$ and its partial derivatives around $\xi = 0$.

References

[ABT03] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic Combinatorial Structures: a Probabilistic Approach*. EMS Monographs in Mathematics. European Mathematical Society, 2003.

[AGG89] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for poisson approximations: the Chen–Stein method. *Ann. Probab.*, 17:9–25, 1989.

[Art64] E. Artin. *The gamma function*. Holt, Rinehart and Winston, 1964.

[BČ02] A. D. Barbour and V. Čekanavičius. Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.*, 30(2):509–545, 2002.

[Bel27] E. T. Bell. Partition polynomials. *Ann. Math., 2nd series*, 29(1/4):38–46, 1927.

[BHJ92] A. D. Barbour, L. Holst, and S. Janson. *Poisson Approximation*. Oxford University Press, 1992.

[BKN09] A. D. Barbour, E. Kowalski, and A. Nikeghbali. Mod-discrete expansions. *Probab. Th. Rel. Fields*, 158(3):859–893, 2009.

[BP96] K. A. Borovkov and D. Pfeifer. On improvements of the order of approximation in the Poisson limit theorem. *J. Appl. Probab.*, 33:146–155, 1996.

[BU09] V. Betz and D. Ueltschi. Spatial random permutations and infinite cycles. *Commun. Math. Phys.*, 285:469–501, 2009.

[BU11] V. Betz and D. Ueltschi. Spatial random permutations with small cycle weights. *Probab. Th. Rel. Fields*, 149:191–222, 2011.

[BUV11] V. Betz, D. Ueltschi, and Y. Velenik. Random permutations with cycle weights. *Ann. Appl. Probab.*, 21:312–331, 2011.

[BX99] A. D. Barbour and A. Xia. Poisson perturbations. *ESAIM Probab. Statist.*, 3:131–150, 1999.

[Cam60] L. Le Cam. An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.*, 10(4):1181–1197, 1960.

[Ček97] V. Čekanavičius. Asymptotic expansions in the exponent: a compound Poisson approach. *Adv. Appl. Probab.*, 29:374–387, 1997.

[Ček98] V. Čekanavičius. Poisson approximations for sequences of random variables. *Statist. Probab. Lett.*, 39:101–107, 1998.

[Che74] L. H. Y. Chen. On the convergence of Poisson binomial to Poisson distributions. *Ann. Probab.*, 2:178–180, 1974.

[Che75] L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probab.*, 3:534–545, 1975.

[ČM99] V. Čekanavičius and M. Mikalauskas. Signed Poisson approximations for Markov chains. *Stochastic Process. Appl.*, 82:205–227, 1999.

[CSST14] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. *Representation Theory and Harmonic Analysis of Wreath Products of Finite Groups*, volume 410 of *London Mathematical Society Lecture Note Series*. London Mathematical Society, 2014.

[DKN15] F. Delbaen, E. Kowalski, and A. Nikeghbali. Mod-ϕ convergence. *Int. Math. Res. Notices*, 2015:3445–3485, 2015.

[DP86] P. Deheuvels and D. Pfeifer. A semigroup approach to Poisson approximation. *Annals of Probability*, 14:663–676, 1986.

[DVJ02] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Volume 1.* Probability and Its Applications. Springer-Verlag, 2nd edition, 2002.

[EK40] P. Erdös and M. Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *American Journal of Mathematics*, 62:738–742, 1940.

[EU12] N. M. Ercolani and D. Ueltschi. Cycle structure of random permutations with cycle weights. *Random Structures & Algorithms*, 44(1):109–133, 2012.

[FMN16] V. Féray, P.-L. Méliot, and A. Nikeghbali. *Mod-ϕ convergence: Normality zones and precise deviations*. SpringerBriefs in Probability and Mathematical Statistics. Springer-Verlag, 2016.

[FMN17a] V. Féray, P.-L. Méliot, and A. Nikeghbali. Mod-ϕ convergence, II: Estimates on the speed of convergence. In preparation, 2017.

[FMN17b] V. Féray, P.-L. Méliot, and A. Nikeghbali. Mod-ϕ convergence, III: Multi-dimensional mod-Gaussian convergence and related estimates of probabilities. In preparation, 2017.

[FO90a] P. Flajolet and A. M. Odlyzko. Random mapping statistics. In *Advances in Cryptology - EUROCRYPT 1989*, volume 434 of *Lecture Notes in Computer Sciences*, pages 329–354. Springer-Verlag, 1990.

[FO90b] P. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3:216–240, 1990.

[FS09] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[Har09] A. J. Harper. Two new proofs of the Erdös–Kac theorem, with bound on the rate of convergence, by Stein's method for distributional approximations. *Math. Proc. Cam. Phil. Soc.*, 147:95–114, 2009.

[Hwa96] H.-K. Hwang. Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.*, 6(1):297–319, 1996.

[Hwa98] H.-K. Hwang. On convergence rates in the central limit theorems for combinatorial structures. *Europ. J. Combin.*, 19(3):329–343, 1998.

[Hwa99] H.-K. Hwang. Asymptotics of Poisson approximation to random discrete distributions: an analytic approach. *Advances in Applied Probability*, 31(2):448–491, 1999.

[JKK08] Y. Jiao, N. El Karoui, and D. Kurtz. Gauss and Poisson approximations: application to CDOs pricing. *Journal of Computational Finance*, 12(2):31–58, 2008.

[JKN11] J. Jacod, E. Kowalski, and A. Nikeghbali. Mod-Gaussian convergence: new limit theorems in probability and number theory. *Forum Mathematicum*, 23(4):835–873, 2011.

[Kah70] J.-P. Kahane. *Séries de Fourier absolument convergentes*, volume 50 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1970.

[Kat67] S. K. Katti. Infinite divisibility of integer-valued random variables. *Ann. Math. Statistic.*, 38:1306–1308, 1967.

[Kat04] Y. Katznelson. *An introduction to Harmonic Analysis*. Cambridge Mathematical Library, 2004.

[Ker64] J. Kerstan. Verallgemeinerung eines Satzes von Prochorow und Le Cam. *Z. Wahrsch. Verw. Gebiete*, 2:173–179, 1964.

[KN10] E. Kowalski and A. Nikeghbali. Mod-Poisson convergence in probability and number theory. *Intern. Math. Res. Not.*, 18:3549–3587, 2010.

[Kru86] J. Kruopis. Precision of approximations of the generalized binomial distribution by convolutions of Poisson measures. *Lithuanian Math. J.*, 26:37–49, 1986.

[Mac95] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. Clarendon Press, 1995.

[Mél17] P.-L. Méliot. *Representation Theory of Symmetric Groups*. CRC Press, 2017.

[New75] D. J. Newman. A simple proof of Wiener's $1/f$ theorem. *Proceedings of the American Mathematical Society*, 48(1):264–265, 1975.

[NZ13] A. Nikeghbali and D. Zeindler. The generalized weighted probability measure on the symmetric group and the asymptotic behavior of the cycles. *Annales de l'Institut Henri Poincaré*, 49(4):961–981, 2013.

[Pet75] V. V. Petrov. *Sums of independent random variables*, volume 82 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1975.

[Pre83] E. L. Presman. Approximation of binomial distributions by infinitely divisible ones. *Theory Probab. Appl.*, 28:393–403, 1983.

[Pro53] J. V. Prohorov. Asymptotic behavior of the binomial distribution. *Uspehi Mat. Nauk*, 8(3):135–142, 1953.

[Rho09] R. C. Rhoades. Statistics of prime divisors in function fields. *International Journal of Number Theory*, 5:141–152, 2009.

[RT58] A. Rényi and P. Turán. On a theorem of Erdös–Kac. *Acta Arithmetica*, 4:71–84, 1958.

[Sam75] S. M. Samuels. Positive integer-valued infinitely divisible distributions. Purdue University Technical Report, Department of Statistics, Division of Mathematical Sciences, Mimeograph Series 406, 1975.

[Sat99] K.-I. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1999.

[SH04] F. W. Steutel and K. Van Harn. *Infinite divisibility of probability distributions on the real line*, volume 259 of *Monographs and textbooks in pure and applied mathematics*. Marcel Dekker, 2004.

[Ste94] J. M. Steele. Le Cam's inequality and Poisson approximations. *The American Mathematical Monthly*, 101(1):48–54, 1994.

[Sze39] G. Szegö. *Orthogonal Polynomials*, volume 23 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, 1939.

[Ten95] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 46 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 1995.

Institut de Mathématiques de Toulouse – Université Paul Sabatier – 118, route de Narbonne, 31062 Toulouse cedex 9, France

E-mail address: reda.chhaibi@math.univ-toulouse.fr

ETH Zürich, Department of Mathematics and Universität Zürich, Institut für Mathematik – Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

E-mail address: delbaen@math.ethz.ch

Laboratoire de Mathématiques, Bâtiment 425 – Faculté des sciences d'Orsay – Université Paris-Sud, 91400 Orsay, France

E-mail address: pierre-loic.meliot@math.u-psud.fr

Universität Zürich – Institut für Mathematik – Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

E-mail address: ashkan.nikeghbali@math.uzh.ch