MOD–DISCRETE EXPANSIONS

A. D. BARBOUR, E. KOWALSKI, AND A. NIKEGHBALI

ABSTRACT. In this paper, we consider approximating expansions for the distribution of integer valued random variables, in circumstances in which convergence in law cannot be expected. The setting is one in which the simplest approximation to the n'th random variable X_n is by a particular member R_n of a given family of distributions, whose variance increases with n. The basic assumption is that the ratio of the characteristic function of X_n and that of R_n converges to a limit in a prescribed fashion. Our results cover a number of classical examples in probability theory, combinatorics and number theory.

1. INTRODUCTION

The topic of this paper is the explicit approximation, in various metrics, of random variables which, in terms of characteristic functions, behave like a sum

$$X_n = Z_n + Y_n \tag{1.1}$$

of a "model" variable Z_n (for instance, a Poisson random variable) and an independent perturbation Y_n, when the model variable has "large" parameter. Our interest is in discrete random variables, and in cases where this simple-minded decomposition does not in fact exist. We have two motivations:

(1) In probabilistic number theory, it has been known since the proof by Rényi and Turán of the Erdős-Kac theorem that the random variable $\omega(N_n)$ given by the number of prime divisors (without multiplicity, for definiteness) of an integer N_n uniformly chosen in the interval $\{1, 2, \ldots, n\}$ has characteristic function given by

$$\mathbb{E}\{e^{i\theta\omega(N_n)}\} = \mathbb{E}\{e^{i\theta Z_n}\}\Phi(\theta)(1 + o(1))$$

as $n \to \infty$, where $Z_n \sim \mathrm{Po}\,(\log\log n)$ is a Poisson variable with mean $\log\log n$ and $\Phi(\theta)$ is defined by

$$\Phi(\theta) = \frac{1}{\Gamma(e^{i\theta})} \prod_{p \text{ prime}} \left(1 + \frac{e^{i\theta} - 1}{p}\right)\left(1 - \frac{1}{p}\right)^{e^{i\theta} - 1},$$

the product being absolutely convergent for all θ real. This $\Phi(\theta)$ is not the characteristic function of a probability distribution, and hence formula (1.1) with $Z_n \sim \mathrm{Po}\,(\log\log n)$ cannot be true. However, we are nonetheless able to obtain explicit approximation statements for the law of $\omega(N_n)$ from the characteristic function identity.

Theorem 1.1. *For every integer $r \geq 0$, there exist explicitly computable signed measures $\nu_{r,n}$ on the positive integers such that the total variation distance between the law of $\omega(N_n)$ and $\nu_{r,n}$ is of order $O\{(\log\log n)^{-(r+1)/2}\}$ for $n \geq 3$.*

This is proved, and formulas for the measures $\nu_{1,n}$ and $\nu_{2,n}$ are given, in Section 8.3. For more on the significance of the Rényi-Turán formula, comparison with the Keating-Snaith conjectures for the Riemann zeta function, and finite-field analogues, see Kowalski & Nikeghbali (2009).

(2) In a beautiful paper, Hwang (1999) considered sequences of non-negative integer valued random variables X_n, whose probability generating functions f_{X_n} satisfy

$$e^{\lambda_n(1-z)} f_{X_n}(z) \ \rightarrow \ g(z),$$

for all $z \in \mathbb{C}$ with $|z| \leq \eta$, for some $\eta > 1$, where the function g is analytic, and $\lim_{n\to\infty} \lambda_n = \infty$. This assumption is also intuitively related to a model (1.1). Under some extra conditions, Hwang exhibits bounds of order $O(\lambda_n^{-1})$ on the accuracy of the approximation of the distribution of X_n by a Poisson distribution with carefully chosen mean, close to λ_n. Hwang (1999) also notes that his methods can be applied to families of distributions other than the Poisson family, and gives examples using the Bessel family.

In this paper, we systematically consider sequences of integer valued random variables X_n, whose characteristic functions ϕ_{X_n} satisfy a condition which, in the Poisson context, is some strengthening of the convergence

$$\exp\{\lambda_n(1 - e^{i\theta})\}\phi_{X_n}(\theta) \to \psi(\theta), \quad 0 < |\theta| \leq \pi. \tag{1.2}$$

Under suitable conditions, we derive explicit approximations to the distribution of X_n, in various metrics, by measures related to the Poisson model. The approximations can be made close to any given polynomial order in $\lambda_n^{-1/2}$, if the conditions are sharp enough and the measure is correspondingly chosen. The conditions that we require for these expansions are much weaker than those of Hwang (1999). For instance, his conditions require the X_n to take only non-negative values, and to have exponential tails, neither of which conditions we need to impose.

Our basic result, Proposition 2.1, is very simple and explicit. It enables us to dispense with asymptotic settings, and to prove concrete error bounds. It also allows us to consider approximation by quite general families of distributions on the integers, instead of just the Poisson family, requiring only the replacement of the Poisson characteristic function in (1.2) by the characteristic function corresponding to the family chosen. This enables us to deduce expansions based on any discrete family of distributions, as shown in Section 4, without any extra effort. Indeed, the main problem would seem to be to identify the higher order terms in the expansions, but these turn out simply to be linear combinations of the higher order differences of the basic distribution: see (2.6).

In the Poisson context, the measures that result are the Poisson–Charlier measures. Our general results enable us to deduce a Poisson–Charlier approximation with error of order $O(\lambda_n^{-t/2})$, for any prescribed t, assuming that Hwang's conditions hold. We also show that the Poisson–Charlier expansions are valid under more general conditions, in which the X_n may have only a few finite moments. The expansions are established in Section 5, and simpler, translated Poisson approximations are considered in Section 6. We discuss some examples, to sums of independent integer valued random variables, to Hwang's setting and to our first motivation, proving Theorem 1.1, in Section 8.

2. The basic estimate

The essence of our argument is the following elementary result, linking the closeness of finite signed measures μ and ν to the closeness of their characteristic functions, when these have a common factor involving a 'large' parameter ρ; for a finite signed measure ζ on $\mathbb{Z}$, the characteristic function ϕ_ζ is defined by $\phi_\zeta(\theta) := \sum_{j \in \mathbb{Z}} e^{ij\theta} \zeta\{j\}$, for $|\theta| \leq \pi$.

Proposition 2.1. *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions ϕ_μ and ϕ_ν respectively. Suppose that $\phi_\mu = \psi_\mu \chi$ and $\phi_\nu = \psi_\nu \chi$, where, for some $\gamma, \rho, t > 0$,*

$$|\psi_\mu(\theta) - \psi_\nu(\theta)| \leq \gamma|\theta|^t \quad and \quad |\chi(\theta)| \leq e^{-\rho\theta^2} \quad for\ all \quad |\theta| \leq \pi. \qquad (2.1)$$

Then there are explicit constants α_{1t} and α_{2t} such that

$$1. \quad \sup_{j \in \mathbb{Z}} |\mu\{j\} - \nu\{j\}| \leq \alpha_{1t} \gamma (\rho \vee 1)^{-(t+1)/2};$$

$$2. \quad \sup_{a \leq b \in \mathbb{Z}} |\mu\{[a,b]\} - \nu\{[a,b]\}| \leq \alpha_{2t} \gamma (\rho \vee 1)^{-t/2}.$$

Proof. For any $j \in \mathbb{Z}$, the Fourier inversion formula gives

$$\mu\{j\} - \nu\{j\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ij\theta} (\psi_\mu(\theta) - \psi_\nu(\theta)) \chi(\theta) \, d\theta, \qquad (2.2)$$

from which our assumptions imply directly that

$$|\mu\{j\} - \nu\{j\}| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma|\theta|^t \exp\{-\rho\theta^2\} \, d\theta.$$

For $\rho \leq 1$, we thus have

$$|\mu\{j\} - \nu\{j\}| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma|\theta|^t \, d\theta \leq \frac{\pi^t \gamma}{t+1} =: \beta_{1t}\gamma.$$

For $\rho \geq 1$, it is immediate that

$$|\mu\{j\} - \nu\{j\}| \leq \frac{\gamma}{2\pi} \left(\frac{1}{\sqrt{2\rho}}\right)^{t+1} \int_{-\infty}^{\infty} |y|^t e^{-y^2/2} \, dy \leq \beta'_{1t} \gamma \rho^{-(t+1)/2},$$

with $\beta'_{1t} := 2^{-(t+1)/2} m_t / \sqrt{2\pi}$; here, m_t denotes the t-th absolute moment of the standard normal distribution. Setting

$$\alpha_{1t} := \max\{\beta_{1t}, \beta'_{1t}\} = \max\left\{2^{-(t+1)/2} m_t/\sqrt{2\pi}, \, \pi^t/(t+1)\right\},$$

this proves part 1. The second part is similar, adding (2.2) over $a \leq j \leq b$, and estimating

$$\frac{\left|e^{-ia\theta} - e^{-i(b+1)\theta}\right|}{|1 - e^{-i\theta}|} \leq \frac{\pi}{|\theta|}, \qquad |\theta| \leq \pi.$$

This gives part 2, with

$$\alpha_{2t} := \max\{2^{-t/2} m_{t-1} \sqrt{\pi/2}, \, \pi^t/t\}. \qquad \blacksquare$$

We shall principally be concerned with taking μ to be the distribution of a random variable X. We allow ν to be a signed measure, because in many cases, such as in the following canonical example and in the Poisson–Charlier expansions of Section 5, signed measures appear as the natural approximations.

Let X be an integer valued random variable with characteristic function $\phi_X := \psi\chi$, where χ is the characteristic function of a (well known) probability distribution R on $\mathbb{Z}$. Suppose that χ satisfies

$$|\chi(\theta)| \leq e^{-\rho\theta^2}, \tag{2.3}$$

as for Proposition 2.1, and that ψ can be approximated by a polynomial expansion around $\theta = 0$ of the form

$$\tilde{\psi}_r(\theta) := \sum_{l=0}^{r} \tilde{a}_l(e^{i\theta} - 1)^l, \tag{2.4}$$

for real coefficients $\tilde{a}_l$ (and with $\tilde{a}_0 = 1$) and some $r \in \mathbb{N}_0$, in that

$$|\psi(\theta) - \tilde{\psi}_r(\theta)| \leq K_{r\delta}|\theta|^{r+\delta}, \qquad |\theta| \leq \pi, \tag{2.5}$$

for some $0 < \delta \leq 1$. Clearly, this is much like supposing that ψ has a Taylor expansion around zero (in powers of $i\theta$), and hence that X has a corresponding number of finite moments. However, the particular form chosen for the expansion has the advantage that the signed measure $\nu_r = \nu_r(R; \tilde{a}_1, \ldots, \tilde{a}_r)$ having $\tilde{\psi}_r\chi$ as characteristic function can immediately be identified:

$$\nu_r = \sum_{l=0}^{r} (-1)^l \tilde{a}_l D^l R, \tag{2.6}$$

where the differences $D^l R$ of the probability measure R are determined by iterating the relation $DR\{j\} := R\{j\} - R\{j-1\}$. The following theorem now follows directly from Proposition 2.1.

Theorem 2.2. *Let X be a random variable on $\mathbb{Z}$ with distribution P_X. Suppose that its characteristic function ϕ_X is of the form $\psi\chi$, where χ is the characteristic function of a probability distribution R and satisfies (2.3) above. Suppose also that (2.5) is satisfied, for some $r \in \mathbb{N}_0$, $\tilde{a}_1, \ldots, \tilde{a}_r \in \mathbb{R}$ and $\delta \geq 0$. Then, writing $t = r + \delta$, we have*

1. $$\sup_{j \in \mathbb{Z}} |P_X\{j\} - \nu_r\{j\}| \leq \alpha_{1t} K_{r\delta}(\rho \vee 1)^{-(t+1)/2};$$

2. $$\sup_{a \leq b \in \mathbb{Z}} |P_X\{[a,b]\} - \nu_r\{[a,b]\}| \leq \alpha_{2t} K_{r\delta}(\rho \vee 1)^{-t/2},$$

with α_{1t} and α_{2t} as in Proposition 2.1, and with $\nu_r = \nu_r(R; \tilde{a}_1, \ldots, \tilde{a}_r)$ as defined in (2.6).

Remark. Note that Proposition 2.1 can be applied with $\psi_\mu = 0$, corresponding to μ the zero measure, and $\psi_\nu(\theta) = \tilde{a}_l(e^{i\theta} - 1)^l$, for any $1 \leq l \leq r$, showing that the contribution from the l-th term in the expansion to $\nu_r\{j\}$ is at most $|\tilde{a}_l|\alpha_{1l}(\rho \vee 1)^{-(l+1)/2}$, and that to $\nu_r\{[a,b]\}$ at most $|\tilde{a}_l|\alpha_{2l}(\rho \vee 1)^{-l/2}$. Thus, if ρ is large and the coefficients $\tilde{a}_l$ moderate, the contributions decrease in powers of $\rho^{-1/2}$ as l increases. In such circumstances, the signed measure ν_r can be seen as a perturbation of the underlying distribution R.

The simplest application of the above results arises when $\phi_X = \phi_Y p_\lambda$, where $p_\lambda(\theta) = e^{\lambda(e^{i\theta} - 1)}$ is the characteristic function of the Poisson distribution $\text{Po}(\lambda)$ with mean λ, which satisfies (2.3) with $\rho = 2\pi^{-2}\lambda$, and ϕ_Y is the characteristic function associated with a random variable Y on the integers. In this case, $X = Z + Y$ is the sum of two independent random variables, as in (1.1), with $Z \sim \text{Po}(\lambda)$, and the situation is probabilistically very clear. For $w = w_\theta = e^{i\theta} - 1$, we have $\phi_Y(\theta) = \mathbb{E}\{(1 + w)^Y\}$. The latter expression has an expansion in powers of w up to the term in w^r if the r-th moment

4

of Y exists, with coefficients $\tilde{a}_k := F_k(Y)/k!$, $1 \le k \le r$, where $F_k(Y)$ denotes the k-th factorial moment of Y:

$$F_k(Y) := \sum_{l \ge k} \frac{l!}{(l-k)!} \mathbb{P}[Y = l] + \sum_{l \ge 1} (-1)^k \frac{(l+k-1)!}{(l-1)!} \mathbb{P}[Y = -l].$$

Thus the asymptotic expansion of X around $\mathrm{Po}\,(\lambda)$ is simply derived from the factorial moments of the perturbing random variable Y, if they exist.

For example, we could take ϕ_Y to be the characteristic function of a random variable Y_s with distribution

$$\mathbb{P}[Y_s = -l] = s! \frac{s}{l(l+1)\dots(l+s)}, \quad l \ge 1,$$

for some integer $s \ge 1$; the random variable has only $s-1$ moments, and takes negative values, so that Hwang's (1999) theorems cannot be applied. However, Y_s has factorial moments

$$F_k(Y_s) = (-1)^k s! \sum_{l \ge 1} \frac{s}{(l+k)\dots(l+s)} = (-1)^k k! \frac{s}{s-k}, \quad 1 \le k \le s-1,$$

and characteristic function

$$\psi_{Y_s}(\theta) = 1 + \sum_{k=1}^{s-1} (-1)^k \frac{s}{s-k} (e^{i\theta} - 1)^k - s(1 - e^{i\theta})^s \log(1 - e^{-i\theta}),$$

and (2.5) holds for $\tilde{\psi}_r$ as in (2.4), with $r = s-1$ and any $\delta < 1$, for $\tilde{a}_k = F_k(Y)/k! = (-1)^k s/(s-k)$. Hence, if $X = Z + Y_s$, where $Z \sim \mathrm{Po}\,(\lambda)$ is independent of Y_s, then Theorem 2.2 can be applied, approximating the distribution of X by the signed measure $\nu_{s-1}(\mathrm{Po}\,(\lambda); \tilde{a}_1, \dots, \tilde{a}_{s-1})$.

3. Refinements

We frame our approximations in terms of three distances between (signed) measures μ and ν on the integers: the point metric

$$d_{\mathrm{loc}}(\mu, \nu) := \sup_{j \in \mathbb{Z}} |\mu\{j\} - \nu\{j\}|,$$

the Kolmogorov distance

$$d_{\mathrm{K}}(\mu, \nu) := \sup_{j \in \mathbb{Z}} |\mu\{(-\infty, j]\} - \nu\{(-\infty, j]\}|,$$

and the total variation norm

$$\|\mu - \nu\| := \sum_{j \in \mathbb{Z}} |\mu\{j\} - \nu\{j\}|.$$

Other metrics could also be treated using our results. Proposition 2.1 immediately yields explicit bounds on $d_{\mathrm{loc}}(\mu, \nu)$ and $d_{\mathrm{K}}(\mu, \nu)$ in terms of the quantities specified in (2.1). However, for many applications, a slight weakening of its conditions is useful. The following result is proved in exactly the same way as Proposition 2.1.

Proposition 3.1. *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions ϕ_μ and ϕ_ν respectively. Suppose that $\phi_\mu = \psi_\mu \chi$ and $\phi_\nu = \psi_\nu \chi$, where, for some*

$\theta_0, \varepsilon, \eta, \rho > 0$ *and for positive pairs* γ_m, t_m, $1 \le m \le M$, *we have*

$$|\psi_\mu(\theta) - \psi_\nu(\theta)| \le \sum_{m=1}^{M} \gamma_m |\theta|^{t_m} + \varepsilon \quad and \quad |\chi(\theta)| \le e^{-\rho\theta^2}, \quad 0 \le |\theta| \le \theta_0;$$

$$|\phi_\mu(\theta) - \phi_\nu(\theta)| \le \eta, \qquad \theta_0 < |\theta| \le \pi.$$

Then, with notation as for Proposition 2.1, we have

1. $$\sup_{j \in \mathbb{Z}} |\mu\{j\} - \nu\{j\}| \le \sum_{m=1}^{M} \gamma_m \alpha_{1t_m} (\rho \vee 1)^{-(t_m+1)/2} + \tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta;$$

2. $$\sup_{a_0 \le a \le b \le b_0} |\mu\{[a,b]\} - \nu\{[a,b]\}|$$
$$\le \sum_{m=1}^{M} \gamma_m \alpha_{2t_m} (\rho \vee 1)^{-t_m/2} + (b_0 - a_0 + 1)(\tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta),$$

where

$$\tilde{\alpha}_1 := \left(\frac{\theta_0}{\pi} \wedge \frac{1}{2\sqrt{\pi\rho}} \right); \qquad \tilde{\alpha}_2 := \left(1 - \frac{\theta_0}{\pi} \right).$$

The presence of the factor $(b_0 - a_0 + 1)$ in the second bound means that a direct bound on the Kolmogorov distance between the signed measures μ and ν is not immediately visible. The following corollary is however easily deduced; for a signed measure μ, $|\mu|$ as usual denotes its variation.

Corollary 3.2. *Under the conditions of Proposition 3.1,*

$$d_{\mathrm{K}}(\mu, \nu) \le \inf_{a \le b} \left(\varepsilon_{ab}^{(\mathrm{K})} + (|\mu| + |\nu|)\{[a,b]^c\} \right);$$

$$\|\mu - \nu\| \le \inf_{a \le b} \left(\varepsilon_{ab}^{(1)} + (|\mu| + |\nu|)\{[a,b]^c\} \right),$$

where

$$\varepsilon_{ab}^{(\mathrm{K})} := \sum_{m=1}^{M} \gamma_m \alpha_{2t_m} (\rho \vee 1)^{-t_m/2} + (b - a + 1)(\tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta);$$

$$\varepsilon_{ab}^{(1)} := (b - a + 1) \left\{ \sum_{m=1}^{M} \gamma_m \alpha_{1t_m} (\rho \vee 1)^{-(t_m+1)/2} + (\tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta) \right\}.$$

If also μ is a probability measure and $\nu(\mathbb{Z}) = 1$, then

$$d_{\mathrm{K}}(\mu, \nu) \le 2 \inf_{a \le b} \left(\varepsilon_{ab}^{(\mathrm{K})} + |\nu|\{[a,b]^c\} \right);$$

$$\|\mu - \nu\| \le \inf_{a \le b} \left(\varepsilon_{ab}^{(1)} + \varepsilon_{ab}^{(\mathrm{K})} + 2|\nu|\{[a,b]^c\} \right).$$

Proof. The inequality for the total variation norm is immediate. For the Kolmogorov distance, by considering the possible positions of x in relation to $a < b$, we have

$$|\mu\{(-\infty, x]\} - \nu\{(-\infty, x]\}|$$
$$\leq \sup_{y < a} |\mu\{(-\infty, y]\} - \nu\{(-\infty, y]\}| + \sup_{a \leq y \leq b} |\mu\{[a, y]\} - \nu\{[a, y]\}|$$
$$+ \sup_{y > b} |\mu\{(b, y]\} - \nu\{(b, y]\}|$$
$$\leq (|\mu| + |\nu|)\{(-\infty, a) \cup (b, \infty)\} + \varepsilon_{ab}^{(\mathrm{K})}.$$

∎

Remark. If μ is a probability measure and $\nu(\mathbb{Z}) = 1$, we have

$$|\mu|\{[a, b]^c\} = 1 - \mu\{[a, b]\} \leq |1 - \nu\{[a, b]\}| + \varepsilon_{ab}^{(\mathrm{K})} \leq |\nu|\{[a, b]^c\} + \varepsilon_{ab}^{(\mathrm{K})}.$$

Under slightly stronger conditions than those of Proposition 2.1, a much neater total variation bound can be deduced; the argument is reminiscent of that in Presman (1983).

Proposition 3.3. *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions $\phi_\mu = \psi_\mu \chi$ and $\phi_\nu = \psi_\nu \chi$ respectively, where $\chi(\theta) := e^{i\zeta\theta - u(\theta)}$ for some $\zeta \in \mathbb{R}$, and $u(0) = u'(0) = 0$. Suppose now that u and the difference $d_{\mu\nu} := \psi_\mu - \psi_\nu$ are both twice differentiable, that $d'_{\mu\nu}(0) = 0$ and that, for some $\gamma, \gamma' > 0$, $\rho \geq 1$ and $t \geq 2$,*

$$|d''_{\mu\nu}(\theta)| \leq \gamma|\theta|^{t-2}, \quad |u''(\theta)| \leq \gamma'\rho \quad and \quad \Re\{u(\theta)\} \geq \rho\theta^2, \quad for \ all \ |\theta| \leq \pi.$$

Then there is a constant $\alpha' := \alpha'(t, \gamma')$ such that

$$\|\mu - \nu\| \leq \gamma\alpha'\rho^{-t/2}.$$

Proof. First, the assumptions on $d_{\mu\nu}$ and u give

$$|d'_{\mu\nu}(\theta)| \leq \frac{\gamma}{t-1}|\theta|^{t-1}; \quad |d_{\mu\nu}(\theta)| \leq \frac{\gamma}{t(t-1)}|\theta|^t; \tag{3.1}$$
$$|u'(\theta)| \leq \gamma'\rho|\theta|.$$

In particular, for $|j - \zeta| < 1 + \lceil\sqrt{\rho}\rceil$, we can apply part 1 of Proposition 2.1, which gives

$$|\mu\{j\} - \nu\{j\}| \leq \frac{\alpha_{1t}\gamma}{t(t-1)}\rho^{-(t+1)/2}. \tag{3.2}$$

For the remaining j, integrating the Fourier inversion formula (2.2) twice by parts gives

$$\mu\{j\} - \nu\{j\} = -\frac{1}{2\pi(j-\zeta)^2}\int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta}\Big(d''_{\mu\nu}(\theta) - 2d'_{\mu\nu}(\theta)u'(\theta)$$
$$+ d_{\mu\nu}(\theta)\{(u'(\theta))^2 - u''(\theta)\}\Big)e^{-u(\theta)}\, d\theta. \tag{3.3}$$

Substituting the bounds from (3.1) into (3.3) gives

$$|\mu\{j\} - \nu\{j\}|$$
$$\leq \frac{1}{2\pi(j-\zeta)^2}\int_{-\pi}^{\pi}\gamma\left\{|\theta|^{t-2} + \frac{2\gamma'\rho}{t-1}|\theta|^t + \frac{\gamma'\rho}{t(t-1)}|\theta|^t(1 + \gamma'\rho\theta^2)\right\}e^{-\rho\theta^2}\, d\theta$$
$$\leq \frac{1}{(j-\zeta)^2}\gamma\,\beta'(t, \gamma')\rho^{-(t-1)/2}, \tag{3.4}$$

after some calculation, where, with m_t as in Proposition 2.1,

$$\beta'(t, \gamma') := \frac{m_{t-2}}{4t \, 2^{t/2} \sqrt{\pi}} \{4t + 2(2t+1)\gamma' + (t+1)(\gamma')^2\}.$$

Hence

$$\sum_{|j-\zeta| \geq 1 + \lceil \sqrt{\rho} \rceil} |\mu\{j\} - \nu\{j\}| \leq 2\gamma \, \beta'(t, \gamma') \rho^{-t/2},$$

and the proposition follows directly, with $\alpha'(t, \gamma') := 2\beta'(t, \gamma') + \frac{6\alpha_{1t}}{t(t-1)}$. ∎

Note that, in the Poisson context, with $u(\theta) = \lambda(1 - e^{i\theta} + i\theta)$, we take $\rho = 2\lambda/\pi^2$ as before, and then $\gamma' = \pi^2/2$. Note also that, if $d_{\mu\nu}(\theta) = \hat{d}_{\mu\nu}(e^{i\theta} - 1)$, with $\hat{d}_{\mu\nu}(0) = \hat{d}'_{\mu\nu}(0) = 0$ and $|\hat{d}''_{\mu\nu}(w)| \leq \hat{\gamma}|w|^{t-2}$ in $|w| \leq 2$, then $d_{\mu\nu}(0) = d'_{\mu\nu}(0) = 0$ and

$$|d''_{\mu\nu}(\theta)| \leq \left(1 + \frac{2}{t-1}\right)\hat{\gamma}|\theta|^{t-2} \tag{3.5}$$

in $|\theta| \leq \pi$.

The above proposition, like Proposition 2.1, has conditions that are often satisfied almost, but not quite, in applications. The next result allows the sort of freedom envisaged in Proposition 3.1. It is useful when ρ is large, the quantities ε and η are, for instance, exponentially small with ρ, and the same is true for $|\nu|\{[-\rho, \rho]^c\}$. In such circumstances, it yields a bound of (slightly) sharper order than can be deduced from Corollary 3.2, since there it is necessary to choose a and b in such a way that $\rho^{-1/2}(b-a)$ is large, if $|\nu|\{[a, b]^c\}$ is to be made small, and the factor $\rho^{-1/2}(b-a)$ then multiplies the main element of the bound thus obtained.

Proposition 3.4. *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions $\phi_\mu = \psi_\mu \chi$ and $\phi_\nu = \psi_\nu \chi$ respectively, where $\chi(\theta) := e^{i\zeta\theta - u(\theta)}$ for some $\zeta \in \mathbb{R}$, and for some function u satisfying the same conditions as in Proposition 3.3. Suppose that $\tilde{d}_{\mu\nu}$ is a twice differentiable function such that $\tilde{d}_{\mu\nu}(0) = \tilde{d}'_{\mu\nu}(0) = 0$ and that*

$$|\tilde{d}''_{\mu\nu}(\theta)| \leq \sum_{m=1}^{M} \gamma_m |\theta|^{t_m - 2} \tag{3.6}$$

for some $t_m \geq 2$ and $\gamma_m > 0$, $1 \leq m \leq M$; assume that

$$e^{-\rho\theta^2}|\psi_\mu(\theta) - \psi_\nu(\theta) - \tilde{d}_{\mu\nu}(\theta)| \leq \varepsilon, \qquad |\theta| \leq \theta_0;$$
$$|\phi_\mu(\theta) - \phi_\nu(\theta)| \leq \eta, \qquad \theta_0 < |\theta| \leq \pi,$$

for some $\varepsilon, \eta > 0$ and for some θ_0 with $\theta_0 \leq \pi/4$ and $\rho\theta_0^2 \geq \log \rho \geq 0$. Then there is a function $\alpha' := \alpha'(t, \gamma)$ such that, with γ' as in Proposition 3.3,

$$\|\mu - \nu\| \leq \sum_{m=1}^{M} \gamma_m \alpha'(t_m, \gamma') \rho^{-t_m/2} + 3\rho \max\{\varepsilon, \eta\} + (|\mu| + |\nu|)\{[-\rho, \rho]^c\}.$$

If μ is a probability measure and $\nu(\mathbb{Z}) = 1$, then

$$\|\mu - \nu\| \leq 2\sum_{m=1}^{M} \gamma_m \alpha'(t_m, \gamma') \rho^{-t_m/2} + 6\rho \max\{\varepsilon, \eta\} + 2|\nu|\{[-\rho, \rho]^c\}.$$

8

Proof. We begin by observing that a function $w \colon (-\infty, 0] \to \mathbb{R}$ with $w(0) = a$ and $w'(0) = b$ can be continued differentiably on $[0, \infty)$ by a piecewise quadratic function such that $|w''(x)| \leq c$ for all $x > 0$ for which $w''(x)$ is defined, and such that $w(x) = 0$ for all

$$x \geq \frac{1}{c}\left\{ |b| + 2\sqrt{|ac + \tfrac{1}{2}\mathrm{sgn}(b)b^2|} \right\},$$

with $\max_{x \geq 0} |w(x)| \leq |a| + b^2/2c$. We use this to continue the real and imaginary parts of $\tilde{w}(\theta) := e^{-u(\theta)}\tilde{d}_{\mu\nu}(\theta)$ into $\theta_0 \leq |\theta| \leq \pi$ in such a way that $\tilde{w}$ is piecewise twice differentiable and satisfies

$$\tilde{w}(-\pi) = \tilde{w}(\pi) = \tilde{w}'(-\pi) = \tilde{w}'(\pi) = 0,$$

with the second derivatives of the real and imaginary parts suitably bounded. We do so by treating $\tilde{w}$ as a sum of M pieces, reflecting the condition (3.6); for the m'th piece $\tilde{w}_m$, defining $\kappa_1(t, \gamma) := (t + \gamma)/\{t(t-1)\}$, we have

$$
\begin{aligned}
|\tilde{w}_m(\theta_0)| &\leq \frac{\gamma_m}{t_m(t_m - 1)} \theta_0^{t_m} e^{-\rho\theta_0^2} \\
&\leq |a_m| := t_m^{-1}\gamma_m\kappa_1(t_m, \gamma')\theta_0^{t_m}e^{-\rho\theta_0^2}; \\
|\tilde{w}_m'(\theta_0)| &\leq \frac{\gamma_m}{t_m(t_m - 1)} \theta_0^{t_m - 1} e^{-\rho\theta_0^2}\{t_m + \gamma'\rho\theta_0^2\} \\
&\leq |b_m| := \gamma_m\kappa_1(t_m, \gamma')\rho\theta_0^{t_m+1}e^{-\rho\theta_0^2},
\end{aligned}
$$

with the same bounds holding also at $-\theta_0$, and we restrict the second derivatives of its continuation by taking

$$c_m := 4\gamma_m\kappa_1(t_m, \gamma')\rho^2\theta_0^{t_m+2}e^{-\rho\theta_0^2} \qquad (3.9)$$

in $\theta_0 \leq |\theta| \leq \pi$. The length of the θ-interval beyond θ_0 on which $\tilde{w}_m$ is not identically zero is thus bounded by

$$\frac{1}{c_m}\left\{ |b_m|(1 + \sqrt{2}) + 2\sqrt{|a_m c_m|} \right\} \leq \frac{1 + 3\sqrt{2}}{4\rho\theta_0} \leq h := \frac{2}{\rho\theta_0}, \qquad (3.10)$$

from (3.7) and (3.8), the bound being the same for all m; note that

$$h \leq \frac{2\theta_0}{\rho\theta_0^2} \leq \frac{\pi}{2},$$

since $\theta_0 \leq \pi/4$ and $\rho\theta_0^2 \geq 1$. From this and (3.9), it follows also that

$$\int_{\theta_0 \leq |\theta| \leq \pi} |\tilde{w}_m''(\theta)|\, d\theta \leq 4hc_m \leq 32\gamma_m\kappa_1(t_m, \gamma')\rho\theta_0^{t_m+1}e^{-\rho\theta_0^2}, \qquad (3.11)$$

and, using (3.7) and (3.8), that

$$\rho\int_{\theta_0 \leq |\theta| \leq \pi} |\tilde{w}_m(\theta)|\, d\theta \leq 4h\rho\{|a_m| + b_m^2/2c_m\} \leq 5\gamma_m\kappa_1(t_m, \gamma')\theta_0^{t_m-1}e^{-\rho\theta_0^2}. \qquad (3.12)$$

The proof now follows the lines of those of Propositions 3.1 and 3.3. We begin by writing

$$
\begin{aligned}
\mu\{j\} - \nu\{j\} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ij\theta}(\phi_\mu(\theta) - \phi_\nu(\theta))\, d\theta \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\varsigma)\theta}\left\{ e^{-u(\theta)}[\psi_\mu(\theta) - \psi_\nu(\theta) - \tilde{d}_{\mu\nu}(\theta)] + \tilde{w}(\theta) \right\} d\theta. \qquad (3.13)
\end{aligned}
$$

Now, by assumption,

$$\left| e^{-u(\theta)}[\psi_\mu(\theta) - \psi_\nu(\theta) - \tilde{d}_{\mu\nu}(\theta)] \right| \leq \varepsilon$$

9

in $|\theta| \leq \theta_0$, whereas, in $\theta_0 < |\theta| \leq \pi$, it is bounded by $\eta + |\tilde{w}(\theta)|$. Hence, for any j, we use (3.12) to give

$$\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta} e^{-u(\theta)} [\psi_\mu(\theta) - \psi_\nu(\theta) - \tilde{d}_{\mu\nu}(\theta)] \, d\theta \right|$$

$$\leq \max\{\varepsilon, \eta\} + \frac{5}{2\pi\rho} \sum_{m=1}^{M} \gamma_m \kappa_1(t_m, \gamma') \theta_0^{t_m - 1} e^{-\rho\theta_0^2}. \tag{3.14}$$

Noting also that, if $\rho\theta^2 \geq \log \rho \geq 0$ and $\theta > 0$, then

$$\rho^{t/2} \theta^{t-1} e^{-\rho\theta^2} = \rho^{1/2} (\rho\theta^2)^{(t-1)/2} e^{-\rho\theta^2} \leq k_2(t),$$

for $k_2(t) = \{(t-1)/e\}^{(t-1)/2} \max\{1, 2^{-(t-1)/2}\sqrt{t-1}\}$, it follows that

$$\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta} e^{-u(\theta)} [\psi_\mu(\theta) - \psi_\nu(\theta) - \tilde{d}_{\mu\nu}(\theta)] \, d\theta \right|$$

$$\leq \max\{\varepsilon, \eta\} + \frac{5}{2\pi\rho} \sum_{m=1}^{M} \gamma_m \kappa_1(t_m, \gamma') k_2(t_m) \rho^{-t_m/2}. \tag{3.15}$$

This bounds the first element of (3.13).

For the second element of (3.13), we begin by considering values of j such that $|j - \zeta| < 1 + \lceil \sqrt{\rho} \rceil$. Here, we write

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |\tilde{w}(\theta)| \, d\theta \leq \frac{1}{2\pi} \int_{|\theta| \leq \theta_0} e^{-\rho\theta^2} |\tilde{d}_{\mu\nu}(\theta)| \, d\theta + \frac{1}{2\pi} \sum_{m=1}^{M} \int_{\theta_0 < |\theta| \leq \pi} |\tilde{w}_m(\theta)| \, d\theta.$$

The first integral is bounded, as in the proof of Propositions 2.1 and 3.1, by

$$\sum_{m=1}^{M} \gamma_m \frac{\alpha_{1t_m}}{t_m(t_m - 1)} \rho^{-(t_m+1)/2},$$

and the second, as above, by

$$\frac{5}{2\pi\rho} \sum_{m=1}^{M} \gamma_m \kappa_1(t_m, \gamma') k_2(t_m) \rho^{-t_m/2},$$

giving the bound

$$\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta} \tilde{w}(\theta) \, d\theta \right|$$

$$\leq \sum_{m=1}^{M} \gamma_m \left\{ \frac{\alpha_{1t_m}}{t_m(t_m - 1)} + \frac{5}{2\pi} \kappa_1(t_m, \gamma') k_2(t_m) \right\} \rho^{-(t_m+1)/2}, \tag{3.16}$$

since $\rho \geq 1$. The interval containing such j values contains at most $5 + 2\sqrt{\rho}$ integers.

For $|j - \zeta| \geq 1 + \lceil \sqrt{\rho} \rceil$, integrating twice by parts, it follows that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta} \tilde{w}(\theta) \, d\theta = -\frac{1}{2\pi(j-\zeta)^2} \int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta} \tilde{w}''(\theta) \, d\theta.$$

10

The right hand side can be bounded as in (3.4), but with an additional element of

$$\frac{1}{2\pi(j-\zeta)^2} \sum_{m=1}^{M} \int_{\theta_0<|\theta|\leq\pi} |\tilde{w}_m''(\theta)|\, d\theta$$

$$\leq \frac{16}{\pi(j-\zeta)^2} \sum_{m=1}^{M} \gamma_m \kappa_1(t_m, \gamma') \rho \theta_0^{t_m+1} e^{-\rho\theta_0^2}$$

$$\leq \frac{16}{\pi(j-\zeta)^2} \sum_{m=1}^{M} \gamma_m \kappa_1(t_m, \gamma')\, k_3(t_m) \rho^{-(t_m-1)/2}, \qquad (3.17)$$

from (3.11), with $k_3(t) := \{(t+1)/2e\}^{(t+1)/2}$. As a result, we find that, for $|j-\zeta| \geq 1 + \lceil\sqrt{\rho}\rceil$, the second element of (3.13) can be bounded by

$$\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta)\theta} \tilde{w}(\theta)\, d\theta \right|$$

$$\leq \frac{1}{(j-\zeta)^2} \sum_{m=1}^{M} \gamma_m \left\{ \beta'(t_m, \gamma') + \frac{16}{\pi} \kappa_1(t_m, \gamma')\, k_3(t_m) \right\} \rho^{-(t_m-1)/2}. \qquad (3.18)$$

The final step is to write

$$\|\mu - \nu\| \leq \sum_{|j|\leq\rho} |\mu\{j\} - \nu\{j\}| + (|\mu| + |\nu|)\{[-\rho, \rho]^c\},$$

and to note that, if μ is a probability measure and $\nu(\mathbb{Z}) = 1$, then

$$|\mu|\{[a,b]^c\} = 1 - \mu\{[a,b]\} \leq |1 - \nu\{[a,b]\}| + |\nu\{[a,b]\} - \mu\{[a,b]\}|$$

$$\leq |\nu|\{[a,b]^c\} + \sum_{j=a}^{b} |\mu\{j\} - \nu\{j\}|. \qquad \blacksquare$$

Remark. The interval $[-\rho, \rho]$ could be replaced by an interval $[-k\rho, k\rho]$, for any fixed $k \geq 1$, in the statement of the theorem, at the expense of multiplying the remaining bounds by the factor k.

4. Applying the bounds

Let R_λ, $\lambda \geq 0$, be a family of probability distributions on the integers. Assume that, for $Z_\lambda \sim R_\lambda$, $\mu(\lambda) := \mathbb{E}Z_\lambda$ and $\sigma^2(\lambda) := \operatorname{Var} Z_\lambda$ exist, and are both uniformly continuous functions of λ, with $\sigma^2(\lambda)$ strictly increasing from zero to infinity with λ. Suppose also that the characteristic function r_λ of R_λ can be written in the form $r_\lambda(\theta) = \exp\{i\zeta_\lambda\theta - u_\lambda(\theta)\}$, with $u_\lambda'(0) = 0$, and that

$$h(\lambda) := \inf_{0<|\theta|\leq\pi} \left(-\theta^{-2}\Re\{u_\lambda(\theta)\} \right) > 0. \qquad (4.19)$$

Then $|r_\lambda(\theta)| \leq e^{-h(\lambda)\theta^2}$ for $0 \leq |\theta| \leq \pi$, and so R_λ can play the role of R in Theorem 2.2, with $h(\lambda)$ for ρ. This leads to the following result.

Corollary 4.1. *Let X be an integer valued random variable with distribution P_X and characteristic function $\phi_X := \psi r_\lambda$, where r_λ is as above and $h(\lambda) \geq 1$. Then, if (2.5) is satisfied for $\tilde{\psi}_r$, as defined in (2.4), for some $r \in \mathbb{N}_0$, $\tilde{a}_1, \ldots, \tilde{a}_r \in \mathbb{R}$ and $\delta \geq 0$, then,*

writing $t = r + \delta$, we have

$$1. \quad \sup_{j \in \mathbb{Z}} |P_X\{j\} - \nu_r\{j\}| \;\leq\; \alpha_{1t} K_{r\delta} h(\lambda)^{-(t+1)/2};$$

$$2. \quad \sup_{a \leq b \in \mathbb{Z}} |P_X\{[a,b]\} - \nu_r\{[a,b]\}| \;\leq\; \alpha_{2t} K_{r\delta} h(\lambda)^{-t/2},$$

with α_{1t} and α_{2t} as in Proposition 2.1, and with $\nu_r = \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)$ as defined in (2.6). Furthermore, if $r \geq 2$ and $\psi - \tilde{\psi}_r$ and u_λ are both twice differentiable, and if

$$|\psi''(\theta) - \tilde{\psi}''_r(\theta)| \;\leq\; \widetilde{K}_{r\delta} |\theta|^{t-2} \quad and \quad |u''_\lambda(\theta)| \;\leq\; \gamma' h(\lambda) \quad for\ all\ |\theta| \leq \pi, \qquad (4.20)$$

for constants $\widetilde{K}_{r\delta}$ and γ', then

$$3. \quad \|P_X - \nu_r\| \;\leq\; \alpha' \widetilde{K}_{r\delta} h(\lambda)^{-t/2},$$

for $\alpha' := \alpha'(t, \gamma')$ as in Proposition 3.3.

Remark. Taking $\psi_\mu = 0$ and $\psi_\nu = (e^{i\theta} - 1)^l$ in Proposition 3.3 for $l \geq 2$ gives $|d''_{\mu\nu}(\theta)| \leq l(l+1)|\theta|^{l-2}$ and $d'_{\mu\nu}(0) = 0$. Hence the contribution from the l-th term in the signed measure ν_r of (2.6) has total variation norm at most $\alpha' l(l+1)|\tilde{a}_l| h(\lambda)^{-l/2}$, for $2 \leq l \leq r$.

The use of signed measures to approximate probability distributions is convenient, but not very natural. However, the signed measures $\nu_1(R_\lambda; \tilde{a}_1)$ and $\nu_2(R_\lambda; \tilde{a}_1, \tilde{a}_2)$ can be replaced by suitably translated members of the family $\{R_\lambda, \lambda > 0\}$, with the same *asymptotic* rate of approximation, by fitting the first two moments, a procedure analogous to that used in the Berry–Esseen theorem. The signed measure ν_2 has mean and variance given by

$$\mu_* \;=\; \mu(\lambda) + \tilde{a}_1; \qquad \sigma_*^2 \;=\; \sigma^2(\lambda) + 2\tilde{a}_2 + \tilde{a}_1(1 - \tilde{a}_1), \qquad (4.21)$$

and the corresponding equations for ν_1 just have $\tilde{a}_2 = 0$. However, when matching these moments, only integer translations m of R_λ can be allowed, since the distributions must remain on the integers, and so it is not possible to match both moments exactly within the family. To circumvent this, we extend to approximation by a member of the family of probability distributions $Q_{mp}(R_{\lambda'})$, for $\lambda' > 0$, $m \in \mathbb{Z}$ and $0 \leq p < 1$, where

$$Q_{mp}(R_{\lambda'})\{j\} \;:=\; p R_{\lambda'}\{j - m - 1\} + (1 - p)R_{\lambda'}\{j - m\}. \qquad (4.22)$$

If $Z \sim R_{\lambda'}$, then $Q_{mp}(R_{\lambda'})$ is the distribution of $Z + m + I$, where $I \sim \text{Be}(p)$ is independent of Z. $Q_{mp}(R_{\lambda'})$ has characteristic function $q_{mp}(R_{\lambda'})$ given by

$$q_{mp}(R_{\lambda'})(\theta) \;:=\; e^{im\theta}(1 + p(e^{i\theta} - 1))r_{\lambda'}(\theta), \qquad (4.23)$$

similar to the measure $\nu_2\{R_{\lambda'}; m+p, \binom{m}{2}+mp\}$, but with terms of higher order as powers of $(e^{i\theta} - 1)$ as well.

Within the family $\{Q_{mp}(R_{\lambda'}); \lambda' > 0, m \in \mathbb{Z}, 0 \leq p < 1\}$, we can always find a member having a given mean μ_* and variance σ_*^2, provided that $\sigma_*^2 \geq 1/4$, by solving the equations

$$\mu_* \;=\; \mu(\lambda') + m + p; \qquad \sigma_*^2 \;=\; \sigma^2(\lambda') + p(1 - p). \qquad (4.24)$$

To do so, let λ_p solve $\sigma^2(\lambda_p) = \sigma_*^2 - p(1-p)$, possible for $0 \leq p \leq 1$, since $\sigma_*^2 \geq 1/4$ and the function σ^2 has an inverse; note also that $\lambda_0 = \lambda_1$. Define $m_p := \mu_* - \mu(\lambda_p) - p$, continuous under the assumptions on μ, and observe that $m_0 = m_1 + 1$. Hence the value $m = \lfloor m_0 \rfloor$ is realized in the form m_p for some $0 \leq p < 1$, and then λ_p, m_p and p satisfy (4.24).

We now show that $Q_{mp}(R_{\lambda'})$ and $\nu_r = \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)$ are close to order $O\{(h(\lambda) \vee 1)^{-3/2}\}$, under some assumptions on the family R_λ. We assume that μ and the inverse function $(\sigma^2)^{-1}$ are both uniformly continuous, and that, for any $\lambda, \lambda' > 0$,

$$d_{\lambda,\lambda'}(\theta) \;:=\; \frac{r_{\lambda'}(\theta)}{r_\lambda(\theta)} - \left\{1 + b_1(e^{i\theta} - 1) + b_2(e^{i\theta} - 1)^2\right\}$$

satisfies

$$d_{\lambda,\lambda'}(0) \;=\; d'_{\lambda,\lambda'}(0) \;=\; 0; \qquad |d''_{\lambda,\lambda'}(\theta)| \;\leq\; \gamma(|\lambda - \lambda'|)|\theta| \text{ in } |\theta| \leq \pi,$$

where, from moment considerations,

$$\begin{aligned} b_1 &\;:=\; b_1(\lambda, \lambda') \;:=\; \mu(\lambda') - \mu(\lambda); \\ b_2 &\;:=\; b_2(\lambda, \lambda') \;:=\; \tfrac{1}{2}\{\sigma^2(\lambda') - \sigma^2(\lambda) - b_1(1 - b_1)\}. \end{aligned}$$

If this is the case, then the choices of λ', m, p in (4.24) when μ_* and σ_*^2 are given by (4.21) are such as to give

$$\tilde{a}_1 \;=\; b_1 + m + p; \qquad \tilde{a}_2 \;=\; b_2 + (m + p)b_1 + mp + \binom{m}{2},$$

so that $Q_{mp}(R_{\lambda'})$ and $\nu_r = \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)$ have the same mean and variance; what is more, m, $|\lambda' - \lambda|$, $|\mu(\lambda') - \mu(\lambda)|$ and $|\sigma^2(\lambda') - \sigma^2(\lambda)|$ are uniformly bounded for $(\tilde{a}_1, \tilde{a}_2)$ in any compact set. It thus follows, with this choice of λ', m, p, that $q_{mp}(R_{\lambda'})(\theta)$ can be written as $r_\lambda(\theta)\psi_{\lambda,\lambda'}(\theta)$, with

$$\psi_{\lambda,\lambda'}(\theta) \;=\; \{d_{\lambda,\lambda'}(\theta) + [1 + b_1(e^{i\theta} - 1) + b_2(e^{i\theta} - 1)^2]\}e^{i\theta m}(1 + p(e^{i\theta} - 1)), \qquad (4.25)$$

and with

$$\hat{d}(\theta) \;:=\; [1 + b_1(e^{i\theta} - 1) + b_2(e^{i\theta} - 1)^2]e^{i\theta m}(1 + p(e^{i\theta} - 1)) - \tilde{\psi}_r(\theta) \qquad (4.26)$$

satisfying

$$\hat{d}(0) \;=\; \hat{d}'(0) \;=\; 0; \qquad |\hat{d}''(\theta)| \leq \hat{\gamma}|\theta|, \text{ in } |\theta| \leq \pi, \qquad (4.27)$$

where $\hat{\gamma} = \hat{\gamma}(\tilde{a}_1, \ldots, \tilde{a}_r)$ remains bounded if $\tilde{a}_1, \ldots, \tilde{a}_r$ do. In view of (4.25) – (4.27), the difference $\psi_{\lambda,\lambda'} - \tilde{\psi}_r$ satisfies (4.20) for a constant $\widetilde{K}_{r\delta} = \widetilde{K}_{r\delta}(\tilde{a}_1, \ldots, \tilde{a}_r)$, whose definition depends on the family R_λ, and hence, from Corollary 4.1, we have

$$\|Q_{mp}(R_{\lambda'}) - \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)\| \;\leq\; \alpha' \widetilde{K}_{r\delta} h(\lambda)^{-3/2}. \qquad (4.28)$$

This implies the following result.

Corollary 4.2. *Under the above conditions on the family $\{R_\lambda, \lambda > 0\}$, if X has characteristic function $\phi_X = \psi(\theta)r_\lambda$ and if (2.5) is satisfied with $r \geq 2$ and $\delta > 0$, then, for λ', m and p solving (4.24) and for $t' := \min\{3, r + \delta\}$, we have translated R_λ-approximation of the form*

$$\begin{aligned} &1. \quad d_{\mathrm{loc}}(P_X, Q_{mp}(R_{\lambda'})) \;\leq\; \gamma'_1 h(\lambda)^{-(t'+1)/2}; \\ &2. \quad d_K(P_X, Q_{mp}(R_{\lambda'})) \;\leq\; \gamma'_2 h(\lambda)^{-t'/2}, \end{aligned}$$

for suitable choice of γ'_1, γ'_2. Furthermore, if (4.20) holds, then

$$3. \quad \|P_X - Q_{mp}(R_{\lambda'})\| \;\leq\; \gamma'_3 h(\lambda)^{-t'/2},$$

for suitable choice of γ'_3. For given family $\{R_\lambda, \lambda > 0\}$, the constants γ'_l, $l = 1, 2, 3$, are bounded for $(\tilde{a}_1, \ldots, \tilde{a}_r)$ in compact sets.

Suppose that, instead of (2.5), we are given an approximation to ψ by a Taylor expansion ψ_r around $\theta = 0$:

$$|\psi(\theta) - \psi_r(\theta)| \leq K_{r\delta}|\theta|^{r+\delta}, \qquad |\theta| \leq \pi, \tag{4.29}$$

for some $0 < \delta \leq 1$, where

$$\psi_r(\theta) := \sum_{l=0}^{r} a_l(i\theta)^l, \tag{4.30}$$

for real coefficients a_l (and with $a_0 = 1$) and some $r \in \mathbb{N}_0$. Then, equating coefficients of $i\theta$, it follows that

$$|\psi_r(\theta) - \tilde{\psi}_r(\theta)| \leq G_r|\theta|^{r+1}, \qquad |\theta| \leq \pi, \tag{4.31}$$

for $G_r := G_r(a_1, \ldots, a_r)$, if $\tilde{a}_1, \ldots, \tilde{a}_r$ are defined implicitly by

$$a_j := \sum_{l=1}^{j} \tilde{a}_l \sum_{(s_1,\ldots,s_l) \in S_{j-l}} \prod_{t=1}^{l} \frac{1}{(s_t+1)!}, \tag{4.32}$$

where $S_m := \{(s_1, \ldots, s_l): \sum_{t=1}^{l} s_t = m\}$. Hence we can replace (2.5) by (4.29) in the assumptions of the theorems, if, in the conclusions, $K_{r\delta}$ is replaced by $K_{r\delta} + \pi^{1-\delta}G_r$.

5. Poisson–Charlier expansions

The most common family of distributions in applications is the Poisson family, with $R_\lambda = \mathrm{Po}(\lambda)$ and $r_\lambda = p_\lambda$. Here, as observed in Section 2, the bound (2.3) is satisfied with $\rho = 2\pi^{-2}\lambda$, so that Theorem 2.2 can be applied to approximate the distribution P_X of a random variable X whose characteristic function ϕ_X has the form ψp_λ, if (2.5) holds for some r and $\tilde{a}_1, \ldots, \tilde{a}_r$. The signed measures ν_r on $\mathbb{N}_0$ have the explicit representation

$$\nu_r\{j\} := \nu_r(\mathrm{Po}(\lambda); \tilde{a}_1, \ldots, \tilde{a}_r)\{j\} := \mathrm{Po}(\lambda)\{j\}\Big\{1 + \sum_{l=1}^{r}(-1)^l \tilde{a}_l C_l(j; \lambda)\Big\}, \tag{5.1}$$

where

$$C_l(j; \lambda) := \sum_{k=0}^{l}(-1)^k \binom{l}{k}\binom{j}{k} k! \, \lambda^{-k} \tag{5.2}$$

denotes the l-th Charlier polynomial (Chihara 1978, (1.9), p. 171).

Note that, if $\binom{j}{k}$ is replaced by $j^k/k!$ in (5.2), one obtains the binomial expansion of $(1 - j/\lambda)^l$. As this suggests, the values of $C_l(j; \lambda)$ are in fact small for j near λ if λ is large:

$$|C_l(j; \lambda)| \leq 2^{l-1}\{|1 - j/\lambda|^l + (l/\sqrt{\lambda})^l\} \tag{5.3}$$

(Barbour & Čekanavičius 2002, Lemma 6.1). (5.3) thus implies that, in any interval of the form $|j - \lambda| \leq c\sqrt{\lambda}$, which is where the probability mass of $\mathrm{Po}(\lambda)$ is mostly to be found, the correction to the Poisson measure $\mathrm{Po}(\lambda)$ is of uniform relative order $O(\lambda^{-l/2})$. Indeed, the Chernoff inequalities for $Z \sim \mathrm{Po}(\lambda)$ can be expressed in the form

$$\max\{\mathbb{P}[Z > \lambda(1+\delta)], \mathbb{P}[Z < \lambda(1-\delta)]\}$$
$$\leq \exp\{-\lambda\delta^2/2(1+\delta/3)\} \leq \exp\{-\lambda\delta^2/3(\delta \vee 1)\}, \tag{5.4}$$

(Chung & Lu 2006, Theorem 3.2). Since also, from (5.2),

$$|C_l(j; \lambda)| \leq (1 + j/\lambda)^l \leq 2^l \quad \text{if} \quad 0 \leq j \leq \lambda,$$

14

and since

$$\binom{j}{k} k! \, \lambda^{-k} \frac{e^{-\lambda}\lambda^j}{j!} \;=\; \frac{e^{-\lambda}\lambda^{j-k}}{(j-k)!} \;\leq\; \frac{e^{-\lambda}\lambda^{j-l}}{(j-l)!}$$

if $0 \leq k \leq l$ and $j \geq l + \lambda$, it follows that, for any $l \geq 0$, we have

$$\sum_{j=0}^{m} |C_l(j;\lambda)| \mathrm{Po}\,(\lambda)\{j\} \;\leq\; 2^l \mathbb{P}[Z \leq m] \;\leq\; 2^l \exp\{-(\lambda-m)^2/3\lambda\}$$

for $m \leq \lambda$, and, for $l \leq r$ and $m \geq \lambda + r$,

$$\sum_{j\geq m} |C_l(j;\lambda)| \mathrm{Po}\,(\lambda)\{j\} \;\leq\; 2^l \mathbb{P}[Z \geq m-l] \;\leq\; 2^l \mathbb{P}[Z \geq m-r]$$

$$\leq\; 2^l\big\{\exp\{-(m-r-\lambda)^2/3\lambda\} \vee \exp\{-(m-r-\lambda)/3\}\big\}.$$

It thus follows that

$$\begin{aligned}
|\nu_r|\{[0,m]\} &\;\leq\; \bar{A}_r \, e^{-(\lambda-m)^2/3\lambda}, && 0 \leq m \leq \lambda; \\
|\nu_r|\{[m,\infty)\} &\;\leq\; \bar{A}_r \,\big\{ e^{-(m-r-\lambda)^2/3\lambda} \vee e^{-(m-r-\lambda)/3} \big\}, && m \geq \lambda + r,
\end{aligned} \tag{5.5}$$

where $\bar{A}_r := 1 + \sum_{l=1}^{r} 2^l |\tilde{a}_l|$, demonstrating concentration of measure for ν_r on a scale of $\sqrt{\lambda}$ around λ. Moreover, it can be deduced from (5.3) that there exists a positive constant $d = d(\tilde{a}_1, \ldots, \tilde{a}_r)$ such that $\nu_r\{j\} \geq 0$ for $|j - \lambda| \leq d\lambda$, and it follows from (5.5) that $|\nu_r|\{j \colon |j - \lambda| > d\lambda\} = O(e^{-\alpha\lambda})$ for some $\alpha > 0$. Since also $\nu_r\{\mathbb{N}_0\} = 1$, it thus follows that, even if ν_r is formally a signed measure, it differs from a probability only on a set of measure exponentially small with λ.

If (2.5) holds, Theorem 2.2 enables one to deduce simple bounds for $d_{\mathrm{loc}}(P_X, \nu_r)$ and $d_{\mathrm{K}}(P_X, \nu_r)$. For the total variation norm, with judicious choice of m_1 and m_2, we can use part 1 of Theorem 2.2 to bound

$$\sum_{j=m_1+1}^{m_2-1} |P_X\{j\} - \nu_r\{j\}| \;\leq\; (m_2 - m_1 - 1) \sup_{j\in\mathbb{N}_0} |P_X\{j\} - \nu_r\{j\}|, \tag{5.6}$$

and then (5.5) and part 2 to take care of the remaining tail probabilities:

$$\begin{aligned}
\sum_{j\leq m_1} |P_X\{j\} - \nu_r\{j\}| &\leq P_X\{(-\infty, m_1]\} + |\nu_r|\{[0, m_1]\} \\
&\leq \sup_{l\in\mathbb{N}_0} |P_X\{(-\infty, l]\} - \nu_r\{[0, l]\}| + 2|\nu_r|\{[0, m_1]\},
\end{aligned} \tag{5.7}$$

and

$$\begin{aligned}
\sum_{j\geq m_2} |P_X\{j\} - \nu_r\{j\}| &\leq P_X\{[m_2, \infty)\} + |\nu_r|\{[m_2, \infty)\} \\
&\leq \sup_{l\in\mathbb{N}_0} |P_X\{(-\infty, l]\} - \nu_r\{[0, l]\}| + 2|\nu_r|\{[m_2, \infty)\}.
\end{aligned} \tag{5.8}$$

This gives the following theorem.

Theorem 5.1. *Suppose that X is as above, having characteristic function $\phi_X := \psi p_\lambda$, and that (2.5) holds; write $t = r + \delta$ as before. If $\lambda \geq 1$, there is a constant $\alpha_{4t} = \alpha_{4t}(\tilde{a}_0, \ldots, \tilde{a}_r)$ such that*

$$\|P_X \;-\; \nu_r\| \;\leq\; \alpha_{4t} K_{r\delta} (\lambda \;\vee\; 1)^{-t/2} \max\{1, \sqrt{|\log K_{r\delta}|}, \sqrt{\log(\lambda+1)}\}; \tag{5.9}$$

15

if $\lambda < 1$, then there is a constant $\alpha_{5t} = \alpha_{5t}(\tilde{a}_0, \ldots, \tilde{a}_r)$ such that

$$\|P_X - \nu_r\| \ \leq \ \alpha_{5t} K_{r\delta} \lambda^{-t/2} \max\{1, |\log K_{r\delta}|\}. \tag{5.10}$$

Proof. For $\lambda \geq 1$, we use both parts of (5.5), with

$$m_1 \ := \ \lfloor \lambda - c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} \rfloor \quad \text{and} \quad m_2 \ := \ \lceil \lambda + r + c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} \rceil,$$

where $\lfloor x \rfloor \leq x \leq \lceil x \rceil$ denote the integers closest to x, and with

$$c_{r\lambda} \ := \ 3\{(r+1)/2 + |\log K_{r\delta}|/\log(\lambda+1)\}.$$

If $r + c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} \leq \lambda$, we obtain

$$|\nu_r|\{[m_1, m_2]^c\} \ \leq \ 2\bar{A}_r(\lambda+1)^{-c_{r\lambda}^2/3} \ \leq \ 2\bar{A}_r(\lambda+1)^{-c_{r\lambda}/3},$$

since $c_{r\lambda} \geq 1$, and, if $r + c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} > \lambda$, we get

$$|\nu_r|\{[m_1, m_2]^c\} \ \leq \ 2\bar{A}_r \exp\{-c_{r\lambda}\sqrt{\lambda \log(\lambda+1)}/3\} \ \leq \ 2\bar{A}_r(\lambda+1)^{-c_{r\lambda}/3},$$

since $\lambda \geq \log(\lambda+1)$ in $\lambda \geq 0$. Hence, in either case, from the definition of $c_{r\lambda}$, we have

$$|\nu_r|\{[m_1, m_2]^c\} \ \leq \ 2\bar{A}_r K_{r\delta}(\lambda+1)^{-(r+1)/2}. \tag{5.11}$$

Hence, from Theorem 2.2 and (5.6)–(5.8), it follows that

$$\begin{aligned}\|P_X - \nu_r\| \ \leq \ & \{2c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} + r + 2\}\alpha'_{1t} K_{r\delta}\lambda^{-(t+1)/2} \\ & + 2\alpha'_{2t} K_{r\delta}\lambda^{-t/2} + 4\bar{A}_r K_{r\delta}\lambda^{-(r+1)/2},\end{aligned}$$

with

$$\alpha'_{1t} \ := \ \alpha_{1t}(\pi^2/2)^{(t+1)/2}; \qquad \alpha'_{2t} \ := \ \alpha_{2t}(\pi^2/2)^{t/2}, \tag{5.12}$$

so that

$$\|P_X - \nu_r\| \ \leq \ \beta_{3t} K_{r\delta}\lambda^{-t/2}\sqrt{\log(\lambda+1)} \max\left\{1, \frac{|\log K_{r\delta}|}{\log(\lambda+1)}\right\},$$

with $\beta_{3t} := \alpha'_{1t}\{4r+11\} + 2\alpha'_{2t} + 4\bar{A}_r$.

For $\lambda < 1$, we take $m_2 := \lceil 2 + r + 3|\log K_{r\delta}| \rceil$ in (5.5), giving

$$|\nu_r|\{[m_2, \infty)\} \ \leq \ \bar{A}_r K_{r\delta},$$

and then, from Theorem 2.2, (5.6) and (5.8), it follows that

$$\|P_X - \nu_r\| \ \leq \ (r + 3 + 3|\log K_{r\delta}|)\alpha'_{1t} K_{r\delta} + 2\alpha'_{2t} K_{r\delta} + 2\bar{A}_r K_{r\delta},$$

so that

$$\|P_X - \nu_r\| \ \leq \ \beta'_{3t} K_{r\delta}\max\{1, |\log(K_{r\delta})|\},$$

with $\beta'_{3t} := \alpha'_{1t}\{r+6\} + 2\alpha'_{2t} + 2\bar{A}_r$. ∎

Sometimes it is convenient, for simplicity, to use parameters in the expansions that are not those emerging naturally from the proofs. The following theorem shows that such alterations can easily be allowed for.

Theorem 5.2. *Suppose that*

$$\phi_\mu := p_\lambda A; \qquad \phi_{\nu^{(1)}} := p_\lambda A'; \qquad \phi_{\nu^{(2)}} := p_{\lambda'} A,$$

with $A(\theta) := 1 + \sum_{l=1}^r a_l \theta^l$, $A'(\theta) := 1 + \sum_{l=1}^r a_l' \theta^l$ *and with* $\lambda > \lambda'$. *Then, with* $\rho := 2\pi^{-2}\lambda$, $\rho' := 2\pi^{-2}\lambda'$ *and* $a_0 := 1$,

$$d_{\mathrm{loc}}(\mu, \nu^{(1)}) \leq \sum_{l=1}^r \alpha_{1l} |a_l - a_l'| (\rho \vee 1)^{-(l+1)/2};$$

$$d_{\mathrm{K}}(\mu, \nu^{(1)}) \leq \sum_{l=1}^r \alpha_{2l} |a_l - a_l'| (\rho \vee 1)^{-l/2};$$

$$d_{\mathrm{loc}}(\mu, \nu^{(2)}) \leq (\lambda - \lambda') \sum_{l=1}^{r+1} \alpha_{1l} |a_{l-1}| (\rho' \vee 1)^{-(l+1)/2};$$

$$d_{\mathrm{K}}(\mu, \nu^{(2)}) \leq (\lambda - \lambda') \sum_{l=1}^{r+1} \alpha_{2l} |a_{l-1}| (\rho' \vee 1)^{-l/2}.$$

Proof. For the comparison between μ and $\nu^{(1)}$, we have

$$|A(\theta) - A'(\theta)| \leq \sum_{l=1}^r |a_l - a_l'| \, |\theta|^l, \quad 0 < |\theta| \leq \pi,$$

and Proposition 2.1 completes the proof. For that between μ and $\nu^{(2)}$, note that $p_\lambda = p_{\lambda-\lambda'} p_{\lambda'}$, and that, for $\lambda > \lambda'$ and $0 < |\theta| \leq \pi$,

$$|p_{\lambda-\lambda'}(\theta) - 1| \, |A(\theta)| \leq (\lambda - \lambda')|\theta| \left\{ 1 + \sum_{l=1}^r |a_l| |\theta|^l \right\},$$

from which and Proposition 2.1 the remaining results follow. ∎

6. Poisson approximation

The measures ν_r used in the approximations of the previous section are very explicit. However, approximation in terms of a Poisson distribution is more convenient, if possible. Clearly, if (2.5) holds for any $r = r^*$, $\delta = \delta^*$, then it holds with $r = 0$ and $\psi_0(\theta) = 1$ for all θ, with the exponent $r + \delta$ replaced by $\delta_0 = \delta^*$ if $r^* = 0$ and by $\delta_0 = 1$ if $r^* \geq 1$, with $K_{0\delta_0}$ depending on r^*, $K_{r^*\delta^*}$ and $\tilde{a}_1, \ldots, \tilde{a}_{r^*}$. Theorem 2.2 then gives approximation by $\mathrm{Po}(\lambda)$ with accuracy in Kolmogorov distance of order $O(\lambda^{-t_0/2})$, for $t_0 = \min\{1, r^* + \delta^*\}$.

However, if $r^* \geq 1$, one can also write

$$\psi(\theta) p_\lambda(\theta) = \hat{\psi}(\theta) p_{\lambda'}(\theta),$$

for any $\lambda' > 0$, where

$$\hat{\psi}(\theta) := \psi(\theta) \exp\{(\lambda - \lambda')(e^{i\theta} - 1)\}.$$

Taking $\lambda' - \lambda = \tilde{a}_1$ now gives a bound

$$|\hat{\psi}(\theta) - 1| \leq K_{1\delta_1} |\theta|^{t_1}, \quad 0 \leq |\theta| \leq \pi,$$

of the form (2.5), with $\delta_1 = \min\{r^* + \delta^* - 1, 1\}$, $t_1 = 1 + \delta_1$ and with $K_{1\delta_1}$ depending on r^*, K_{r^*} and $\tilde{a}_1, \ldots, \tilde{a}_{r^*}$. Hence, Theorem 2.2 implies the following approximation.

Corollary 6.1. *If X has characteristic function $\phi_X = \psi p_\lambda$ such that (2.5) is satisfied with $r \geq 1$ and $0 < \delta \leq 1$, then we have*

$$\text{1.} \qquad d_{\text{loc}}(P_X, \text{Po}(\lambda')) \leq \alpha_{1t'} K_{1,t'-1}(\rho' \vee 1)^{-(t'+1)/2};$$

$$\text{2.} \qquad d_{\text{K}}(P_X, \text{Po}(\lambda')) \leq \alpha_{2t'} K_{1,t'-1}(\rho' \vee 1)^{-t'/2},$$

where $\lambda' = \lambda + \tilde{a}_1$, $t' := \min\{2, r + \delta\}$ and $\rho' = 2\pi^{-2}\lambda'$, for a suitable constant $K_{1,t'-1}$ depending on r, $K_{r\delta}$ and $\tilde{a}_1, \ldots, \tilde{a}_r$.

The parameter λ' is chosen to make the Poisson mean λ' equal to the mean $\lambda + a_1$ of X. This choice of the Poisson parameter improves the rate, in the asymptotic sense that, if $a_1, \ldots, a_r$ and $K_{r\delta}$ remain bounded but $\lambda \to \infty$, and if $r + \delta \geq 2$, then the approximation error for Kolmogorov distance is of order $O(\lambda^{-1})$, as opposed to the rate of order $O(\lambda^{-1/2})$ in general obtained when approximating by $\text{Po}(\lambda)$.

Analogously, fitting the second moment as well (if it is finite) can lead to asymptotically sharper approximation within the translated Poisson family. The corresponding result has already been proved in Corollary 4.2 under the more general assumptions of Section 3, which the Poisson family satisfies. If X has characteristic function $\phi_X = \psi p_\lambda$ such that (2.5) is satisfied with $r \geq 2$ and $0 < \delta \leq 1$, then the first two moments of $Q_{mp}(\text{Po}(\lambda'))$ match those of X if

$$\lambda' + m + p = \lambda + \tilde{a}_1; \qquad \lambda' + p(1 - p) = \lambda + 2\tilde{a}_2 + \tilde{a}_1(1 - \tilde{a}_1).$$

λ', m and p are found as prescribed following (4.24), giving

$$\begin{aligned} m &= \lfloor \tilde{a}_1^2 - 2\tilde{a}_2 \rfloor; \qquad p^2 = \langle \tilde{a}_1^2 - 2\tilde{a}_2 \rangle; \\ \lambda' &= \lambda + 2\tilde{a}_2 + \tilde{a}_1(1 - \tilde{a}_1) - p(1 - p), \end{aligned} \tag{6.1}$$

where $\langle x \rangle$ denotes the fractional part of x. Corollary 4.2 now yields the following result.

Corollary 6.2. *If X has characteristic function $\phi_X = \psi p_\lambda$ such that (2.5) is satisfied with $r \geq 2$ and $0 < \delta \leq 1$, then for λ', m amd p given in (6.1) and for $t' := \min\{3, r + \delta\}$, we have translated Poisson approximation:*

$$\text{1.} \qquad d_{\text{loc}}(P_X, Q_{mp}(\text{Po}(\lambda'))) \leq \gamma_1'(\lambda \vee 1)^{-(t'+1)/2};$$

$$\text{2.} \qquad d_{\text{K}}(P_X, Q_{mp}(\text{Po}(\lambda'))) \leq \gamma_2'(\lambda \vee 1)^{-t'/2},$$

and, if (4.20) holds, then

$$\text{3.} \qquad \|P_X - Q_{mp}(\text{Po}(\lambda'))\| \leq \gamma_3'(\lambda \vee 1)^{-t'/2},$$

for constants $\gamma_l' = \gamma_l'(\tilde{a}_1, \ldots, \tilde{a}_r)$, $l = 1, 2, 3$, that are bounded for $(\tilde{a}_1, \ldots, \tilde{a}_r)$ in compact sets.

In particular, if $a_1, \ldots, a_r$ and $K_{r\delta}$ remain bounded but $\lambda \to \infty$, and if $r + \delta \geq 3$, then $t' = 3$ and the order of approximation in Kolmogorov distance is of order $O(\lambda^{-3/2})$.

7. Compound Poisson approximation

The theory of Section 3 can also be applied when the distributions R_λ come from a compound Poisson family. For $\lambda > 0$ and for μ a probability distribution on $\mathbb{Z}$, let $\text{CP}(\lambda, \mu)$ denote the distribution of the sum $Y := \sum_{j \in \mathbb{Z} \setminus \{0\}} jZ_j$, where Z_j, $j \neq 0$, are independent, and $Z_j \sim \text{Po}(\lambda \mu_j)$. Then, if $\mu_1 > 0$, the characteristic function of Y is of the form $R_\lambda := \zeta_\lambda p_{\lambda_1}$, where ζ_λ is the characteristic function of $\sum_{j \in \mathbb{Z} \setminus \{0,1\}} jZ_j$ and $\lambda_1 = \lambda \mu_1$. Thus, for the purposes of applying Corollaries 4.1 and 4.2, ρ can be taken

18

to be $2\pi^{-2}\mu_1\lambda$. Corollary 4.2, for instance, then gives conditions under which translated compound Poisson distribution can be achieved, with approximation at rate $O(\lambda^{-3/2})$.

These considerations apply as long as $\mu_1 > 0$, and could also be invoked if $\mu_{-1} > 0$. If $\mu_1 = \mu_{-1} = 0$, there is then no factor of the form p_λ to guarantee that, for some $\rho > 0$, the characteristic function ϕ_Y of Y (corresponding to the characteristic function χ of Proposition 2.1) satisfies $|\phi_Y(\theta)| \leq \exp\{-\rho\theta^2\}$ for all $|\theta| \leq \pi$. Some additional aperiodicity condition needs to be satisfied, if the family $\{\mathrm{CP}(\lambda,\mu), \lambda \geq 1\}$ is to satisfy (4.19). Indeed, if $Y = 2Z$ where $Z \sim \mathrm{Po}(\lambda)$, and if $W \sim \mathrm{Be}(1/2)$ is independent of Y, it is not true that the distribution of $Y + W$ is close to that of Y in total variation, even though $|\phi_{Y+W}(\theta) - \phi_Y(\theta)| \leq K_0|\theta|\,|\phi_Y(\theta)|$.

8. Applications

8.1. Sums of independent random variables.
Let $X_1, \ldots, X_n$ be independent integer valued random variables, and let S_n denote their sum. In contexts in which a central limit approximation to the distribution of S_n would be appropriate, the classical Edgeworth expansion (see, e.g., Petrov 1975, Chapter 5) is unwieldy, because S_n is confined to the integers. As an alternative, Barbour and Čekanavičius (2002, Theorem 5.1) give a Poisson–Charlier expansion, for S_n 'centred' so that its mean and variance are almost equal, with an error bound expressed in the total variation norm. Here, we show that an error bound for their expansion of the same order can be established by the techniques of this paper, but with respect to the less stringent Kolmogorov distance.

Assume that each of the X_j has finite $(r + 1 + \delta)$'th moment, with $r \geq 1$, and define

$$A^{(r)}(w) \ := \ 1 + \sum_{l \geq 2} \tilde{a}_l^{(r)} w^l \ = \ \exp\left\{\sum_{l=2}^{r+1} \frac{\kappa_l w^l}{l!}\right\}, \tag{8.1}$$

where $\kappa_l := \kappa_l(S_n)$ and $\kappa_l(X)$ denotes the l'th factorial cumulant of the random variable X. Then the approximation that we establish is to the Poisson–Charlier signed measure ν_r with

$$\nu_r\{j\} \ := \ \mathrm{Po}(\lambda)\{j\}\left\{1 + \sum_{l=2}^{L_r} (-1)^l \tilde{a}_l^{(r)} C_l(j; \lambda)\right\}, \tag{8.2}$$

where $L_r := \max\{1, 3(r-1)\}$, and where $\lambda := \mathbb{E}S_n$; ν_r has characteristic function

$$\phi_{\nu_r} \ := \ p_\lambda(\theta)\,\widetilde{A}^{(r)}(\theta), \tag{8.3}$$

where

$$\widetilde{A}^{(r)}(\theta) \ := \ 1 + \sum_{l=2}^{L_r} \tilde{a}_l^{(r)} (e^{i\theta} - 1)^l. \tag{8.4}$$

We need two further quantities involving the X_j:

$$K^{(n)} \ := \ \left|\sum_{j=1}^{n} \kappa_2(X_j)\right| \ = \ |\mathrm{Var}\,S_n - \mathbb{E}S_n|, \tag{8.5}$$

and

$$p_j \ := \ 1 - \tfrac{1}{2}\|\mathcal{L}(X_j) - \mathcal{L}(X_j + 1)\|. \tag{8.6}$$

Theorem 8.1. *Suppose that there are constants K_l, $1 \le l \le r + 1$, such that, for each j,*

$$|\kappa_l(X_j)| \le K_l, \quad 2 \le l \le r + 1; \qquad \mathbb{E}|X_j|^{r+1+\delta} \le K_1^{r+1+\delta}.$$

Suppose also that $p_j \ge p_0 > 0$ for all j, and that $\lambda \ge n\lambda_0$. Then

$$d_{\mathrm{K}}(\mathcal{L}(S_n), \nu_r) \le G(K_1, \ldots, K_{r+1}, K^{(n)}, p_0^{-1}, \lambda_0^{-1}) n^{-(r-1+\delta)/2},$$

for a function G that is bounded on compact sets.

Remark. For asymptotics in n, with triangular arrays of variables, the error is of order $O(n^{-(r-1+\delta)/2})$ when λ_0 and p_0 are bounded away from zero, and $K_1, \ldots, K_{r+1}$ and $K^{(n)}$ remain bounded. The requirements on λ_0 and p_0 can often be achieved by grouping the random variables appropriately, though attention then has to be paid to the consequent changes in the K_l. The final condition can always be satisfied with $K^{(n)} \le 1$, by replacing the X_j by translates, where necessary. For more discussion, we refer to Barbour and Čekanavičius (2002). The above conditions are designed to cover sums of independent random variables, each of which has non-trivial variance, has uniformly bounded $(r + 1 + \delta)$'th moment, and whose distribution overlaps with its unit translate.

Proof. We check the conditions of Proposition 3.1. First, in view of (8.6), we can write

$$\mathbb{E}\big(e^{i\theta X_j}\big) = \tfrac{1}{2} p_j (e^{i\theta} + 1) \phi_{1j}(\theta) + (1 - p_j) \phi_{2j}(\theta),$$

where both ϕ_{1j} and ϕ_{2j} are characteristic functions. Hence we have

$$\big|\mathbb{E}\big(e^{i\theta X_j}\big)\big| \le 1 - p_j + p_j \cos(\theta/2) \le 1 - p_j \theta^2/4\pi, \quad 0 \le |\theta| \le \pi.$$

Hence $\phi_\mu(\theta) := \mathbb{E}\big(e^{i\theta S_n}\big)$ satisfies

$$|\phi_\mu(\theta)| \le \exp\{-np_0 \theta^2/4\pi\}, \quad 0 \le |\theta| \le \pi. \tag{8.7}$$

On the other hand, from the additivity of the factorial cumulants, we have

$$|\kappa_l(S_n)| \le nK_l, \quad 3 \le l \le r + 1,$$

with $|\kappa_2(S_n)| \le K^{(n)}$ from (8.5). From (8.1), we thus deduce the bound $|\tilde{a}_l^{(r)}| \le c_l n^{\lfloor l/3 \rfloor}$, for $c_l = c_l(K^{(n)}, K_3, \ldots, K_{r+1})$, $l \ge 1$. Hence

$$|\phi_{\nu_r}(\theta)| \le \exp\{-2n\lambda_0\theta^2/\pi^2\} c' n^{\lfloor L_r/3 \rfloor} \le \exp\{-n\lambda_0\theta^2/\pi^2\} c'', \tag{8.8}$$

for $c'' = c''(K^{(n)}, K_3, \ldots, K_{r+1})$. Combining (8.7) and (8.8), we can thus take $\eta := Ce^{-n\rho'\theta_0^2}$ in Proposition 3.1, for

$$\rho' = \min\{\lambda_0/\pi^2, p_0/4\pi\}$$

and a suitable $C = C(K^{(n)}, K_3, \ldots, K_{r+1})$. The choice of θ_0 we postpone for now.

For $|\theta| \le \theta_0$, we take $\chi(\theta) := p_\lambda(\theta)$, and check the approximation of

$$\phi_\mu(\theta) \exp\{-\lambda(e^{i\theta} - 1)\} = \mathbb{E}\left\{(1 + w)^{S_n}\right\} e^{-w\mathbb{E}S_n}$$

by $\widetilde{A}^{(r)}(\theta)$ as a polynomial in $w := e^{i\theta} - 1$. We begin with the inequality

$$\left| (1 + w)^s - \sum_{l=0}^{r+1} \frac{w^l}{l!} s_{(l)} \right| \le \frac{|s_{(r+2)}|}{(r+2)!} |w|^{r+2} \wedge 2 \frac{|s_{(r+1)}|}{(r+1)!} |w|^{r+1}$$

$$\le \frac{|s_{(r+1)}|}{(r+2)!} |w|^{r+1+\delta} \{|s| + r + 1\}^\delta \{2(r+2)\}^{1-\delta},$$

20

derived using Taylor's expansion, true for any $s \in \mathbb{Z}$ and $0 < \delta \le 1$, where $s_{(l)} := s(s-1)\ldots(s-l+1)$. Hence, for each j, we have

$$\left| \mathbb{E}\left\{(1+w)^{X_j}\right\} - \sum_{l=0}^{r+1} \frac{\mathbb{E}\{(X_j)_{(l)}\}}{l!} w^l \right| \le c_{r,\delta} |\theta|^{r+1+\delta}(K_1 + K_1^{r+1+\delta}), \qquad (8.9)$$

for a universal constant $c_{r,\delta}$. Then, writing

$$Q_{r+1}^{(s)}(w;X) := \exp\left\{ \sum_{l=s}^{r+1} \kappa_l(X)w^l/l! \right\},$$

and using the differentiation formula in Petrov (1975, p. 170), we have

$$\left| Q_{r+1}^{(1)}(w;X_j) - \sum_{l=0}^{r+1} \frac{\mathbb{E}\{(X_j)_{(l)}\}}{l!} w^l \right|$$
$$\le \frac{|\theta|^{r+2}}{(r+2)!} \sup_{|\theta'| \le \theta_0} \left| \frac{d^{r+2}}{dz^{r+2}} Q_{r+1}^{(1)}(z;X_j) \right|_{z=e^{i\theta'}-1}$$
$$\le |\theta|^{r+2} c(K_1, \ldots, K_{r+1}), \qquad (8.10)$$

for a suitable function c and for all $|\theta| \le \pi$. Combining these estimates, we deduce that, for $w = e^{i\theta} - 1$ and for all $|\theta| \le \pi$,

$$\left| \mathbb{E}\left\{(1+w)^{X_j}\right\} e^{-\mathbb{E}X_j w} - Q_{r+1}^{(2)}(w;X_j) \right| \le k_1 |\theta|^{r+1+\delta}, \qquad (8.11)$$

where $k_1 = k_1(K_1, \ldots, K_{r+1})$.

Now a standard inequality shows that, for $u_j := \prod_{l=1}^{j} x_l \prod_{l=j+1}^{n} y_l$, for complex x_l, y_l with $y_l \ne 0$ and $|x_l/y_l - 1| \le \varepsilon_l$, then

$$|u_n - u_0| \le |u_0| \left\{ \prod_{s=1}^{n-1}(1+\varepsilon_s) \right\} \sum_{l=1}^{n} \varepsilon_l. \qquad (8.12)$$

Taking $x_j := \mathbb{E}\left\{(1+w)^{X_j}\right\} e^{-\mathbb{E}X_j w}$ and $y_j := Q_{r+1}^{(2)}(w;X_j)$, (8.11) shows that we can take $\varepsilon_l := \varepsilon := k_1 |\theta|^{r+1+\delta} e^E$ for each l, with

$$E := \exp\{\sum_{l=2}^{r+1} K_l/l!\},$$

provided that $|\theta| \le \theta_0 \le 1$. For $r \ge 2$, choosing $\theta_0 := n^{-1/3}$ then ensures that $(1+\varepsilon)^n$ is suitably bounded, and (8.12) yields

$$\left| \mathbb{E}\left\{(1+w)^{S_n}\right\} e^{-w\mathbb{E}S_n} - Q_{r+1}^{(2)}(w;S_n) \right| \le k_2 n |\theta|^{r+1+\delta}, \qquad (8.13)$$

for $k_2 = k_2(K^{(n)}, K_1, \ldots, K_{r+1})$, since

$$|u_0| := |Q_{r+1}^{(2)}(w;S_n)| \le \exp\{|\kappa_2(S_n)|\theta_0^2/2\} \exp\left\{ \sum_{l=3}^{r+1} nK_l\theta_0^l/l! \right\}$$

is bounded for $\theta_0 = n^{-1/3}$, in view of (8.5). For $r = 1$, $|u_0|$ is uniformly bounded if $\theta_0 \le 1$, and the choice $\theta_0 = n^{-1/(2+\delta)}$ ensures that $(1+\varepsilon)^n$ remains bounded.

The remaining step is to note that, for $w := e^{i\theta} - 1$, $\widetilde{A}^{(r)}(\theta)$ contains all terms up to the power w^{L_r} in the power series expansion of $Q_{r+1}^{(2)}(w; S_n)$, giving

$$\left| Q_{r+1}^{(2)}(w; S_n) - \widetilde{A}^{(r)}(\theta) \right| \leq \frac{|\theta|^{L_r+1}}{(L_r+1)!} \sup_{|\theta'| \leq |\theta|} \left| \frac{d^{L_r+1}}{dz^{L_r+1}} Q_{r+1}^{(2)}(z; S_n) \right|_{z=e^{i\theta'}-1}. \tag{8.14}$$

Now $|\kappa_2(S_n)|$ is bounded by $K^{(n)}$, and, for $l \geq 3$, each $\kappa_l(S_n)$, for which we have only the weak bound nK_l, occurs associated with the power w^l in the exponent of $Q_{r+1}^{(2)}(w; S_n)$. Writing

$$\frac{d^s}{dz^s} Q_{r+1}^{(2)}(z; S_n) = P_s(n, z) Q_{r+1}^{(2)}(z; S_n),$$

the monomials that make up $P_s(n, z)$ thus have coefficients of magnitude n^l associated with powers z^m with $m \geq (2l - (s - l))_+ = (3l - s)_+$, so that they are themselves of magnitude at most $O(n^{l-(3l-s)_+/3}) = O(n^{s/3})$ in $|\theta'| \leq n^{-1/3}$. Taking $s = L_r + 1$ and $r \geq 2$, $m = 0$ requires that $l \leq r - 1$, and $l \geq r$ entails $m \geq 2$, so that, for $r \geq 2$ and $|\theta| \leq \theta_0$,

$$\sup_{|\theta'| \leq |\theta|} \left| \frac{d^{L_r+1}}{dz^{L_r+1}} Q_{r+1}^{(2)}(z; S_n) \right|_{z=e^{i\theta'}-1} \leq k_3 n^{r-1}(1 + n|\theta|^2),$$

with $k_3 = k_3(K^{(n)}, K_1, \ldots, K_{r+1})$. If $|\theta| \geq n^{-1/2}$, it follows that the bound in (8.14) is at most $2k_3\{(L_r+1)!\}^{-1} n^r |\theta|^{3r}$; if $|\theta| \leq n^{-1/2}$, the bound is at most $2k_3\{(L_r+1)!\}^{-1} n|\theta|^{r+2}$. Combining this with (8.13), we have established that for $|\theta| \leq n^{-1/3}$ and $r \geq 2$, we have

$$|\phi_\mu(\theta) \exp\{-\lambda(e^{i\theta} - 1)\} - \widetilde{A}^{(r)}(\theta)| \leq k_4 n |\theta|^{r+1+\delta}(1 + (n|\theta|^2)^{r-1}), \tag{8.15}$$

where $k_4 = k_4(K^{(n)}, K_1, \ldots, K_{r+1})$. This gives

$$\gamma_1 = nk_4, \quad t_1 = r + 1 + \delta, \quad \gamma_2 = n^r k_4, \quad t_2 = 3r - 1 + \delta$$
$$M = 2, \quad \rho = 2\lambda/\pi^2, \quad \varepsilon = 0 \quad \text{and} \quad \theta_0 = n^{-1/3}$$

in Proposition 3.1, together with $\eta = Ce^{-n^{1/3}\rho'}$ from the bounds (8.7) and (8.8). Applying Corollary 3.2, and using the tail properties of the Poisson–Charlier measures (5.5), the theorem follows for $r \geq 2$.

For $r = 1$, the bound in (8.14) is easily of order $|\theta|^2$, giving a bound in (8.15) of $k_4'(n|\theta|^{2+\delta} + |\theta|^2)$. This leads to the choices

$$\gamma_1 = nk_4', \quad t_1 = 2 + \delta, \quad \gamma_2 = k_4', \quad t_2 = 2d$$
$$M = 2, \quad \rho = 2\lambda/\pi^2, \quad \varepsilon = 0 \quad \text{and} \quad \theta_0 = n^{-1/(2+\delta)}$$

in Proposition 3.1, together with $\eta = Ce^{-n^{\delta/(2+\delta)}\rho'}$, and the remainder of the proof is as before. ∎

Remark. A total variation bound for the error, of the slightly larger order $O\big(n^{-(r-1+\delta)/2}\sqrt{\log n}\big)$, can also be deduced from Corollary 3.2, by taking $a = \lfloor \lambda - k\sqrt{\lambda \log \lambda} \rfloor$ and $b = \lceil \lambda + k\sqrt{\lambda \log \lambda} \rceil$, for suitable choice of $k = k_r$.

8.2. **Analytic combinatorial schemes.** An extremely interesting range of applications is to be found in the paper of Hwang (1999). His conditions are motivated by examples from combinatorics, in which generating functions are natural tools. He works in an asymptotic setting, assuming that X_n is a random variable whose probability generating function R_n is of the form

$$R_n(z) = z^h(g(z) + \varepsilon_n(z))e^{\lambda(z-1)},$$

where h is a non-negative integer, and both g and ε_n are analytic in a closed disc of radius $\eta > 1$. As $n \to \infty$, he assumes that $\lambda \to \infty$ and that $\sup_{z:|z|\le\eta} |\varepsilon_n(z)| \le K\lambda^{-1}$, uniformly in n. He then proves a number of results describing the accuracy of the approximation of P_{X_n-h} by $\text{Po}\,(\lambda + g'(1))$.

Under his conditions, it is immediate that we can write

$$g(z) \;=\; \sum_{j\ge 0} g_j(z-1)^j \quad \text{and} \quad \varepsilon_n(z) \;=\; \sum_{j\ge 0} \varepsilon_{nj}(z-1)^j \tag{8.16}$$

for $|z| < \eta - 1$, with

$$|g_j| \;\le\; k_g(\eta-1)^{-j} \quad \text{and} \quad |\varepsilon_{nj}| \;\le\; \lambda^{-1}k_\varepsilon(\eta-1)^{-j} \tag{8.17}$$

for all $j \ge 0$. Hence $X := X_n - h$ has characteristic function of the form $\psi^{(n)}p_\lambda$, where

$$\psi^{(n)}(\theta) \;=\; g(e^{i\theta}) + \varepsilon_n(e^{i\theta}),$$

and hence, for any $r \in \mathbb{N}_0$,

$$|\psi^{(n)}(\theta) - \tilde\psi_r^{(n)}(\theta)| \;\le\; K_{r1}|\theta|^{r+1}, \qquad |\theta| \le (\eta-1)/2, \tag{8.18}$$

with $\tilde\psi_r^{(n)}$ defined as in (2.4), taking $\tilde a_j^{(n)} = g_j + \varepsilon_{nj}$; note that the constant K_{r1} can indeed be taken to be uniform for all n. Since also g and ε_n are both uniformly bounded on the unit circle, and since $\tilde\psi_r^{(n)}$ is bounded (uniformly in n) for $|\theta| \le \pi$, it is clear that (8.18) can be extended to all $|\theta| \le \pi$, albeit with a different uniform constant K_{r1}', so that (2.5) holds with $\delta = 1$ for any $r \in \mathbb{N}_0$. Thus Corollary 4.1 (with $R_\lambda = \text{Po}\,(\lambda)$ and $h(\lambda) = 2\lambda/\pi^2$) and Theorem 5.1 can be applied with any choice of r, giving progressively more accurate approximations to P_{X_n-h}, as far as the λ-order is concerned, in terms of progressively more complicated perturbations of the Poisson distribution. These theorems are thus applicable to all the examples that Hwang considers, including the numbers of components (counted in various ways) in a wide class of logarithmic assemblies, multisets and selections.

For instance, Corollary 6.2 gives an approximation to P_{X_n-h} by the mixture $Q_{mp}(\text{Po}\,(\lambda'))$ with

$$m \;:=\; \lfloor m_n - v_n\rfloor; \qquad p^2 \;:=\; \langle m_n - v_n\rangle; \qquad \lambda' \;:=\; \lambda + v_n - p(1-p),$$

where $m_n := g_n'(1)$, $v_n := g_n''(1) + g_n'(1) - \{g_n'(1)\}^2$ and $g_n := g + \varepsilon_n$. Hwang's approximation by $\text{Po}\,(\lambda + g'(1))$ has asymptotically the same mean as ours (and as that of $X_n - h$), but a variance asymptotically differing by $\kappa := g''(1) - \{g'(1)\}^2$. As a consequence, Hwang's approximation has an error of larger asymptotic order, in which the quantity κ appears; for instance, for Kolmogorov distance, his Theorem 1 gives an error of order $O(\lambda^{-1})$, whereas that from Corollary 6.2 is of order $O(\lambda^{-3/2})$.

Although our Poisson expansion theorems are automatically applicable under Hwang's conditions, they also apply to examples that do not satisfy his conditions: the simple example at the end of Section 2 is one such. Conversely, Hwang's Theorem 2, which establishes Poisson approximation in the lower tail with good *relative* accuracy, cannot be proved using only our conditions; the conclusion would not be true, for instance, in the example just mentioned.

Note also that Hwang examines problems from combinatorial settings in which approximation is not by Poisson distributions: he has examples concerning the Bessel family of distributions,

$$B(\lambda)\{j\} \;:=\; L(\lambda)^{-1}\frac{\lambda^j}{j!(j-1)!}, \quad j \in \mathbb{N},$$

for the appropriate choice of normalizing constant $L(\lambda)$. This family also satisfies the conditions for R_λ of Section 3, and we could apply Corollary 4.2 to obtain slightly sharper approximations than his within the translated Bessel family, or Corollary 4.1 to obtain asymptotically more accurate expansions.

8.3. Prime divisors.

The numbers of prime divisors of a positive integer n, counted either with $(\Omega(n))$ or without $(\omega(n))$ multiplicity, can also be treated by these methods, since excellent information is available about their generating functions. For our purposes, we use only the shortest expansion, taken from Tenenbaum (1995, Theorems II.6.1 and 6.2). One finds that, for N_n uniformly distributed on $\{1, 2, \ldots, n\}$, the characteristic functions of the we have

$$
\begin{aligned}
\mathbb{E}\{e^{i\theta\omega(N_n)}\} &= p_{\log\log n}(\theta)\left\{\Phi_1(e^{i\theta}-1)+\eta_1(\theta)\right\}; \\
\mathbb{E}\{e^{i\theta\Omega(N_n)}\} &= p_{\log\log n}(\theta)\left\{\Phi_2(e^{i\theta}-1)+\eta_2(\theta)\right\},
\end{aligned}
$$

where $|\eta_s(\theta)| \leq C_s/\log n$, $s = 1, 2$, for some constants C_1 and C_2, and

$$
\Phi_1(w) := \frac{1}{\Gamma(1+w)}\prod_q\left(1+\frac{w}{q}\right)\left(1-\frac{1}{q}\right)^w;
$$

$$
\Phi_2(w) := \frac{1}{\Gamma(1+w)}\prod_q\left(1-\frac{w}{q-1}\right)^{-1}\left(1-\frac{1}{q}\right)^w,
$$

q running here over prime numbers. These expansions were established and used by Rényi and Turán (1958) in their proof of the Erdős–Kac Theorem, but they are also sketched by Selberg (1954). We refer to Kowalski and Nikeghbali (2009) for the structural interpretation of the two factors in these functions, with $1/\Gamma(1+w)$ being related to the number of cycles of large random permutations.

Let $\tilde{a}_{ls}$, $s = 1, 2$, denote the Taylor coefficients of the functions $\Phi_s(w)$ as power series in w (around $w = 0$, which corresponds to $\theta = 0$). By analyticity near 0, it follows that, for any r, we have

$$
\left|\Phi_s(w)-1-\sum_{l=1}^r\tilde{a}_{ls}w^l\right| \leq C_{rs}|w|^{r+1}; \qquad \left|\Phi_s''(w)-\sum_{l=2}^r\tilde{a}_{ls}l(l-1)w^{l-2}\right| \leq C_{rs}'|w|^{r-1},
$$

for suitable constants C_{rs}, C_{rs}' and for $|w| \leq 2$. In order to approximate the distributions $P_{\omega(N_n)}$ and $P_{\Omega(N_n)}$, we define the measures $\nu_r^{(s)}$ by

$$
\nu_r^{(s)}\{j\} := \mathrm{Po}\,(\log\log n)\{j\}\Big(1+\sum_{l=1}^r(-1)^l\tilde{a}_{ls}C_l(j;\log\log n)\Big),
$$

and invoke Propositions 3.1 and 3.4 together with (3.5); this leads to the following conclusion, which refines the Erdős–Kac theorem.

Theorem 8.2. *For the measures $\nu_r^{(s)}$ defined above, we have*

$$d_{\mathrm{loc}}(P_{\omega(N_n)}, \nu_r^{(1)}) \;\leq\; \alpha'_{1,r+1} C_{r1} (\log\log n)^{-1-r/2} + \tilde{\alpha}_1 C_1 / \log n;$$

$$\|P_{\omega(N_n)} - \nu_r^{(1)}\| \;\leq\; 2\alpha'(r+1, \pi^2/2) C'_{r1}\Big(1 + \frac{2}{r}\Big)(\log\log n)^{-(r+1)/2}$$
$$+ \widetilde{C}_1 \log\log n / \log n;$$

$$d_{\mathrm{loc}}(P_{\Omega(N_n)}, \nu_r^{(2)}) \;\leq\; \alpha'_{1,r+1} C_{r2} (\log\log n)^{-1-r/2} + \tilde{\alpha}_1 C_2 / \log n;$$

$$\|P_{\Omega(N_n)} - \nu_r^{(2)}\| \;\leq\; 2\alpha'(r+1, \pi^2/2) C'_{r2}\Big(1 + \frac{2}{r}\Big)(\log\log n)^{-(r+1)/2}$$
$$+ \widetilde{C}_2 \log\log n / \log n,$$

for suitable constants $\widetilde{C}_1$ and $\widetilde{C}_2$, and with α'_{1l} as defined in (5.12).

Remark. As far as we know, total variation approximation was first considered in this context by Harper (2009), who proved a bound with error of size $1/(\log\log n)$ (for a truncated version of $\omega(n)$, counting only prime divisors of size up to $n^{1/(3(\log\log n)^2)}$), and deduced explicit bounds in Kolmogorov distance.

To indicate what this means in concrete terms for number theory readers, consider the case of $\omega(n)$ for $r = 1$. Taylor expansion gives

$$\Phi_1(w) \;=\; 1 + B_1 w + O(w^2)$$

as $w \to 0$, where $B_1 \approx 0.26149721$ is the Mertens constant, i.e., the real number such that

$$\sum_{\substack{q \leq x \\ q \text{ prime}}} \frac{1}{q} = \log\log x + B_1 + o(1),$$

as $x \to +\infty$. An application of Theorem 8.2 gives

$$\left| \frac{1}{n} |\{k \leq n \mid \omega(n) \in A\}| - \nu_1^{(1)}\{A\} \right| \;\leq\; \tfrac{1}{2}\|P_{\omega(N_n)} - \nu_1^{(1)}\|$$
$$=\; O\Big(\frac{1}{\log\log n}\Big),$$

for any set A of positive integers, where

$$\nu_1^{(1)}\{j\} \;=\; \mathrm{Po}\,(\log\log n)\{j\}\Big(1 - B_1\Big\{1 - \frac{j}{\log\log n}\Big\}\Big).$$

Higher expansions could be computed in much the same way.

Alternatively, a more accurate approximation is available from Corollary 6.2, while staying within the realm of (translated) Poisson distributions. For this, we compute the expansion of Φ_1 to order 2, obtaining (after some calculations) that

$$\Phi_1(w) \;=\; 1 + \tilde{a}_1 w + \tilde{a}_2 w^2 + O(w^3), \quad \text{as } w \to 0,$$

where

$$\tilde{a}_1 \;:=\; B_1; \qquad \tilde{a}_2 \;:=\; \frac{B_1^2}{2} - \frac{\pi^2}{12} - \frac{1}{2} \sum_{q \text{ prime}} \frac{1}{q^2}$$

(use $1/\Gamma(1+w) = 1 + \gamma w + (\gamma^2/2 - \pi^2/12)w^2 + O(w^3)$, as well as the Mertens identity

$$\gamma + \sum_{q \text{ prime}} \Big(\frac{1}{q} + \log\Big(1 - \frac{1}{q}\Big)\Big) = B_1,$$

and expand every term in the Euler product). This corresponds to (2.5), since $w = e^{i\theta} - 1$.

25

We can then apply Corollary 6.2 to get the translated Poisson approximation $Q_{mp}(\mathrm{Po}\,(\lambda'))$, with parameters calculated using (6.1). With

$$x := \tilde{a}_1^2 - 2\tilde{a}_2 = \frac{\pi^2}{6} + \sum_{q\ \mathrm{prime}} \frac{1}{q^2} \approx 2.0971815,$$

this gives

$$p = \sqrt{\langle x \rangle} \approx 0.31173945; \qquad m = 2;$$
$$\lambda' = \log\log n + B_1 - x - p(1-p) \approx \log\log n - 2.0502422.$$

Thus, for any positive integer n and any set A of positive integers, we have

$$\left| \frac{1}{n} |\{k \le n \mid \omega(n) \in A\}| - \{p\mathrm{Po}\,(\lambda')\{A-3\} + (1-p)\mathrm{Po}\,(\lambda')\{A-2\}\} \right|$$
$$= O\Big(\frac{1}{(\log\log n)^{3/2}} \Big).$$

Similar results hold for $\Omega(n)$, where one obtains the following approximate values for the quantities p, m, λ':

$$p \approx 0.5195; \qquad m = 0; \qquad \lambda' \approx \log\log n + 0.5152.$$

References

[1] A. D. Barbour & V. Čekanavičius (2002) Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.* **30**, 509–545.
[2] T. S. Chihara (1978) *An introduction to orthogonal polynomials.* Gordon and Breach, New York.
[3] F. Chung & L. Lu (2006) Concentration inequalities and martingale inequalities: a survey. *Internet Math.* **3**, 79–127.
[4] A. J. Harper (2009) Two new proofs of the Erdős–Kac Theorem, with bound on the rate of convergence, by Stein's method for distributional approximations. *Math. Proc. Cam. Phil. Soc.* **147**, 95–114.
[5] H.-K. Hwang (1999) Asymptotics of Poisson approximation to random discrete distributions: an analytic approach. *Adv. Appl. Prob.* **31**, 448–491.
[6] J. Jacod, E. Kowalski & A. Nikeghbali (2008) Mod–Gaussian convergence: new limit theorems in probability and number theory, to appear in *Forum Math.*; see also `arXiv:0807.4739`.
[7] E. Kowalski & A. Nikeghbali (2009) Mod–Poisson convergence in probability and number theory. `arXiv:0905.0318`.
[8] V. V. Petrov (1975) *Limit theorems of probability theory.* Oxford University Press, Oxford.
[9] E. L. Presman (1983) Approximation of binomial distributions by infinitely divisible ones. *Theor. Prob. Appl.* **28**, 393–403.
[10] A. Rényi & P. Turán (1958) On a theorem of Erdős-Kac. *Acta Arith.* **4**, 71–84.
[11] A. Selberg (1954) Note on the paper by L.G. Sathe. *J. Indian Math. Soc.* **18** 83–87.
[12] G. Tenenbaum (1995) *Introduction à la théorie analytique et probabiliste des nombres.* Société Mathématique de France.

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
  E-mail address: `a.d.barbour@math.uzh.ch`

ETH Zurich, D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland
  E-mail address: `kowalski@math.ethz.ch`

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
  E-mail address: `ashkan.nikeghbali@math.uzh.ch`