

Uniform asymptotics of Poisson approximation to the Poisson-binomial distribution

Hsien-Kuei Hwang, Vytas Zacharovas
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

February 12, 2008

Abstract

New uniform asymptotic approximations with error bounds are derived for a generalized total variation distance of Poisson approximations to the Poisson-binomial distribution. The method of proof is also applicable to other Poisson approximation problems.

MSC 2000 Subject Classifications: Primary 62E17; secondary 60C05.

1 Introduction

We study in this paper a fundamental problem in probability theory: how good is Poisson approximation to the Poisson-binomial distribution (of which binomial is a special case)? We re-examine this old and extensively studied problem and derive new estimates from a different viewpoint.

1.1 Poisson approximation to binomial distribution

We start with the simplest case of a binomial distribution. Let $\text{Bi}(n; p)$ denote a binomial distribution of parameters n and p , $0 < p < 1$. Let $\text{Po}(\lambda)$ denote a Poisson distribution with mean λ . It is well known since Siméon-Denis Poisson [23] that if $np \rightarrow c < \infty$ then the binomial distribution $\text{Bi}(n; p)$ converges to the Poisson distribution with mean c , as $n \rightarrow \infty$ (he proved indeed an equivalent version for negative binomial distribution). More than a century later, Prokhorov [24] extended Poisson's limit theorem to an approximation theorem and was the first to show that the total variation distance

$$d_{\text{TV}}(\text{Bi}(n; p), \text{Po}(np)) := \frac{1}{2} \sum_{j \geq 0} \left| \binom{n}{j} p^j (1-p)^{n-j} - e^{-np} \frac{(np)^j}{j!} \right|$$

satisfies (see [5] for a minor correction)

$$d_{\text{TV}}(\text{Bi}(n; p), \text{Po}(np)) = \frac{p}{\sqrt{2\pi e}} \left(1 + O\left(\min(1, p + (np)^{-1/2})\right) \right). \quad (1.1)$$

This says that *Poisson approximation to binomial distribution is good as long as p is small*. Observe that when $np \rightarrow \infty$ the O -term in (1.1) is asymptotically smaller than the dominant term, while in the case of bounded np , the right-hand side becomes an upper bound. Thus (1.1) extends Poisson's original *limit theorem* in two directions: in *degree of precision* and in *range of variations* of p . Such an extension has since attracted the interest of many probabilists and many powerful tools have been developed; see, in particular, [4, 5, 8, 9, 16, 21, 27] and the references therein for more information.

Second-order refinement. A natural question regarding the approximation (1.1) is that how good the leading term is. Or, is the error term optimal? This raises the question of second-order expansion in Poisson approximation problems, which, as far as we were aware, was first addressed by Kerstan in [18]. His results implicitly imply that if $p = o(1)$, then ($\lambda := np$)

$$d_{\text{TV}}(\text{Bi}(n; p), \text{Po}(\lambda)) = pT + O(p^2), \quad (1.2)$$

where

$$T := \frac{1}{4} \sum_{j \geq 0} \frac{\lambda^j}{j!} e^{-\lambda} \left| \frac{(j - \lambda)^2 - j}{\lambda} \right|.$$

Note that $T = (2\pi e)^{-1/2}(1 + \Theta(\lambda^{-1}))$; see also [9, 15]. In particular, for $p = o(1)$ and $\lambda \rightarrow \infty$,

$$d_{\text{TV}}(\text{Bi}(n; p), \text{Po}(\lambda)) = \frac{p}{\sqrt{2\pi e}} (1 + O(p + \lambda^{-1})); \quad (1.3)$$

compare (1.1). The asymptotic nature of Prokhorov's result (1.1) is now clearer. Note that the minimum error (of the O -term) is reached when $p = \lambda^{-1}$, or when $p = n^{-1/2}$. Kerstan's result was later extended in [2, 4, 8, 11, 27] by different approaches; see also the book [5] for more details.

New uniform approximations. Instead of the total variation distance, we consider in this paper the distance

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(X), \mathcal{L}(Y)) := \frac{1}{2} \sum_m |\mathbb{P}(X = m) - \mathbb{P}(Y = m)|^\alpha, \quad (1.4)$$

between the distributions of the two discrete random variables X and Y , where, *here and throughout this paper*, $\alpha > 0$ is a fixed constant. The interest of considering $d_{\text{TV}}^{(\alpha)}$ is multifold. First, such a quantity is, modulo power, the natural analogue of the usual ℓ_p -norm; see [25, Ch. 2] and [27]. When $\alpha = 1$ the above quantity coincides with the usual total variation distance d_{TV} . Second, it can be regarded as an effective measure of robustness of Poisson approximation; see Corollary 1.4 and its discussions. Third, several approaches to Poisson approximation apply well only for special values of α but not for all. Thus, its consideration also introduces more methodological interests. Indeed, most of our proofs will largely simplify if we consider only the case when $\alpha \geq 1$.

Throughout this paper, $q := 1 - p$.

Theorem 1.1. *If $p \leq 1/2$ and $\lambda := np \rightarrow \infty$, then*

$$d_{\text{TV}}^{(\alpha)}(\text{Bi}(n; p), \text{Po}(\lambda)) = p^\alpha \lambda^{(1-\alpha)/2} J_\alpha(p) \left(1 + O\left(\lambda^{-(\alpha+1)/2} + \lambda^{-1}\right) \right), \quad (1.5)$$

where $J_\alpha(p)$ is bounded for $p \in (0, 1)$ and $\alpha > 0$, and defined by

$$J_\alpha(p) = \frac{1}{2(2\pi)^{\alpha/2} p^\alpha} \int_{-\infty}^{\infty} e^{-\alpha t^2/2} \left| \frac{e^{-pt^2/(2q)}}{\sqrt{q}} - 1 \right|^\alpha dt. \quad (1.6)$$

Note that if $p > 1/2$, we can interchange the role of p and q and the same approximation holds with appropriate changes of p and q .

While less explicit, the approximation (1.9), especially the function $J_\alpha(p)$, is seen to contain much information. Roughly, we may say that most dependence on p of $d_{\text{TV}}^{(\alpha)}$ is encapsulated into $J_\alpha(p)$ (although λ itself depends on p). Also the smooth transition of $d_{\text{TV}}^{(\alpha)}$ from $p = o(1)$ to $p \in [\varepsilon, 1/2]$ is visible from (1.5). On the other hand, the uniformity provided by such an approximation is of practical value since most practical parameters are finite and it is often not easy to tell if a given small p is $o(1)$ or $O(1)$.

For integer values of α , we have, by splitting the integral according to the sign of $1 - e^{-px^2/(2q)}/\sqrt{q}$ and then evaluating the corresponding integrals,

$$J_{2m}(p) = \frac{1}{2(2\pi)^{m-1/2} p^{2m}} \sum_{0 \leq j \leq 2m} \binom{2m}{j} (-1)^j q^{-(j-1)/2} (2mq + pj)^{-1/2} \quad (m \geq 1),$$

$$J_{2m+1}(p) = \frac{1}{2(2\pi)^m p^{2m+1}} \sum_{0 \leq j \leq 2m+1} \binom{2m+1}{j} (-1)^j q^{-(j-1)/2} ((2m+1)q + pj)^{-1/2}$$

$$\times \left(3 - 4\Phi \left(\sqrt{((2m+1)q + pj) \frac{1}{p} \log \frac{1}{q}} \right) \right), \quad (m \geq 0),$$

where $\Phi(x) := (2\pi)^{-1/2} \int_{-\infty}^x e^{-t^2/2} dt$ denotes the standard normal distribution function. In particular,

$$J_1(p) = \frac{2}{p} \left(\Phi \left(\sqrt{\frac{1}{p} \log \frac{1}{1-p}} \right) - \Phi \left(\sqrt{\frac{1-p}{p} \log \frac{1}{1-p}} \right) \right), \quad (1.7)$$

$$J_2(p) = \frac{\sqrt{q(1+q)} - 2\sqrt{2q} + \sqrt{1+q}}{4\sqrt{\pi} p^2 \sqrt{q(1+q)}}.$$

For non-integral values of α , no simple explicit expressions are known for J_α . However, numerical calculation does not pose any problem since the integrand decays exponentially fast for large parameter and p bounded away from 1; see Figure 1 for a plot of $J_1(p)$ and $J_{1/2}(p)$. Note that $J_1(0) := \lim_{t \rightarrow 0^+} J_1(t) = 1/\sqrt{2\pi e}$ and $J_1(1) := \lim_{t \rightarrow 1^-} J_1(t) = 1$.

In particular, taking $\alpha = 1$, we obtain the following refinement to (1.3).

Corollary 1.2. *If $p \leq 1/2$ and $\lambda \rightarrow \infty$, then*

$$d_{\text{TV}}(\text{Bi}(n; p), \text{Po}(\lambda)) = pJ_1(p) (1 + O(\lambda^{-1})). \quad (1.8)$$

The result (1.8) is to be compared with (1.3).

On the other hand, $J_1'(0) = J_1(0)/2 = 1/(2\sqrt{2\pi e})$; we thus have

$$d_{\text{TV}}(\text{Bi}(n; p), \text{Po}(\lambda)) = \frac{p}{\sqrt{2\pi e}} \left(1 + \frac{p}{2} + O(p^2 + (np)^{-1}) \right),$$

as $\lambda \rightarrow \infty$ and $p = o(1)$; compare (1.3). We see that the remainder term in the above formula is of order r_n^2 , where $r_n = p + (np)^{-1/2}$, while the error term in the original Prokhorov's result (1.1) is of order r_n .

The second uniform asymptotic approximation we derive to $d_{\text{TV}}^{(\alpha)}(\text{Bi}(n; p), \text{Po}(np))$ for small p is as follows, which extends Kerstan's (1.2); see also [27].

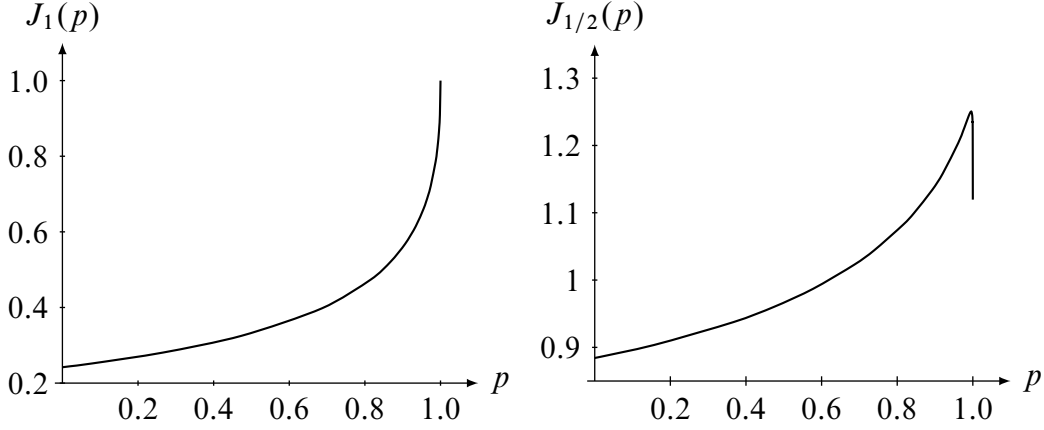


Figure 1: The two functions $J_1(p)$ and $J_{1/2}(p)$ plotted against p . Note that as p grows, $J_{1/2}(p)$ first increases and then suddenly decreases near unity. Such a behavior is typical for $J_\alpha(p)$ when $\alpha \in (0, 1)$. The reason is that when $\alpha < 1$ and $p \rightarrow 1^-$ the integral in (1.6) over $e^{-pt^2/(2\alpha)}/\sqrt{q} \leq 1$ is asymptotic to $\sqrt{\pi/(2\alpha)}(1-p)^{(1-\alpha)/2}$ while the remaining integral is asymptotic to $\sqrt{\pi/(2\alpha)}$. On the other hand, $J_\alpha(p) \rightarrow \infty$ when $\alpha > 1$ and $p \rightarrow 1^-$.

Theorem 1.3. If $p = o(1)$ and $\lambda \geq K > 0$, then

$$d_{\text{TV}}^{(\alpha)}(\text{Bi}(n; p), \text{Po}(np)) = p^\alpha \lambda^{(1-\alpha)/2} T_\alpha(\lambda) \left(1 + O_K\left(p + p^\alpha \lambda^{-1/2}\right)\right), \quad (1.9)$$

where T_α is bounded for all λ and $\alpha > 0$ and given by

$$T_\alpha(\lambda) = \frac{1}{(2\sqrt{\lambda})^{1+\alpha}} \sum_{j \geq 0} \left(e^{-\lambda} \frac{\lambda^j}{j!}\right)^\alpha |(j-\lambda)^2 - j|^\alpha. \quad (1.10)$$

Like J_α , the series T_α also contains much information. From a computational point of view, the function T_α looks complicated and is not much simpler than directly computing the sum-definition of $d_{\text{TV}}^{(\alpha)}(\text{Bi}(n; p), \text{Po}(np))$. However, in the case of general Poisson-binomial and other distributions, a direct calculation of $d_{\text{TV}}^{(\alpha)}$ is very messy and less tractable for large n . In such a case, it is useful to use (1.9) since T_α is distribution-independent. Moreover, the calculation of T_α does not introduce any problem in practice because the terms in (1.10) converge factorially fast for small λ , while for large λ , one can use straightforward Gaussian approximation with the desired error bounds.

Note that the only case not covered by Theorems 1.1 and 1.3 is when $p = o(n^{-1})$ for which both distributions (binomial and Poisson) degenerate and the corresponding approximation problem has been well studied in the literature; see [9, 17].

Robustness of Poisson approximation. From an approximation point of view, an immediate consequence of the two uniform estimates (1.5) and (1.9) is as follows.

Corollary 1.4. The distance $d_{\text{TV}}^{(\alpha)}(\text{Bi}(n; p), \text{Po}(np))$ tends to zero as $n \rightarrow \infty$ if (i) $\alpha > 1$ or (ii) $\alpha \leq 1$ and $p = o(n^{-(1-\alpha)/(1+\alpha)})$.

Thus Poisson approximation to binomial distribution is always good if $\alpha > 1$ (even when p is not small), or if p is suitably small when $\alpha \leq 1$.

On the other hand, another natural question is: which of the two approximations (1.5) and (1.9) is better in the overlapping range (for numerical purposes) when all parameters are given? Roughly, we see that (1.5) is preferable for large λ while (1.9) is better for small λ . More precisely, if $\alpha \geq 1$, then $p = n^{-1/2}$ is the threshold of choosing (1.5) or (1.9), namely, if $p \geq n^{-1/2}$ then use (1.5), otherwise use (1.9). Note that in either case, the error term is of order $n^{-1/2}$ in the worst case, in contrast to (1.3) for which the error term is $n^{-1/2}$ in the best case.

When $\alpha \in (0, 1)$, then by comparing the two error terms $n^{-(\alpha+1)/2}$ and $p + p^\alpha \lambda^{-1/2}$, we see that the threshold becomes $p = n^{-(\alpha+1)/(\alpha+3)}$, namely, if $p \geq n^{-(\alpha+1)/(\alpha+3)}$, then the error term $\lambda^{-(\alpha+1)/2}$ is smaller than that in (1.9), while if $p \leq n^{-(\alpha+1)/(\alpha+3)}$, then the use of (1.9) yields a smaller error than that of (1.5). In either case, the worst error is order $n^{-(\alpha+1)/(\alpha+3)}$.

1.2 Poisson approximation to the Poisson-binomial distribution

Similar results as above can be obtained for the Poisson-binomial distribution. In this case, we have $S_n := X_1 + X_2 + \dots + X_n$, where the X_j 's are independent Bernoulli random variables with

$$\mathbb{P}(X_j = 1) = 1 - \mathbb{P}(X_j = 0) = p_j.$$

Let $\lambda_k = \sum_{1 \leq j \leq n} p_j^k$ for $k \geq 1$. With this notation, we have $\mathbb{E}(S_n) = \lambda_1$ and $\mathbb{V}(S_n) = \lambda_1 - \lambda_2$. For convenience, we write $\lambda = \lambda_1$. Let $\theta := \lambda_2/\lambda$.

Theorem 1.5. *If $\theta \leq 1/2$ and $\lambda \rightarrow \infty$, then*

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda)) = \theta^\alpha \lambda^{(1-\alpha)/2} J_\alpha(\theta) \left(1 + O\left(\lambda^{-(\alpha+1)/2} + \lambda^{-1}\right) \right),$$

where J_α is defined in (1.6).

Note that Theorem 1.1 is a special case of Theorem 1.5, and θ here plays the role of p .

Theorem 1.6. *If $\theta = o(1)$ and $\lambda \geq K > 0$, then*

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda)) = \theta^\alpha \lambda^{(1-\alpha)/2} T_\alpha(\lambda) \left(1 + O_K\left(\zeta + \zeta^\alpha \lambda^{-1/2}\right) \right).$$

where T_α is defined in (1.10) and $\zeta := \theta + \lambda_3/(\lambda_2 \sqrt{\lambda})$.

Note that the quantity ζ becomes $p + p\lambda^{-1/2}$ when $p_j = p$ for all $j = 1, \dots, n$. Also the O -term tends to zero when $\theta \rightarrow 0$ and $\lambda \rightarrow \infty$; it is also infinitesimally small when $\theta \rightarrow 0$ and $\lambda \geq K > 0$ because $\lambda_3/(\lambda_2 \sqrt{\lambda}) \leq \sqrt{\lambda_2/\lambda} = \sqrt{\theta}$.

The results of both theorems are new except for Theorem 1.6 in the special case when $\alpha \geq 1$, which is covered by Theorem 3 in Roos (1999). For earlier results for $\alpha = 1$ (with weaker error bounds), see [8, 12, 13, 18]; the optimal error term was first derived in [4] by improving the Stein-Chen method. Effectiveness of both results can be addressed as in the binomial case and is not given here (indeed all results there hold by replacing p by θ).

1.3 Methods of proof and organization of the paper

The tools we developed are based on Fourier analysis with several new ingredients. They are general and applicable to many other metrics. A brief discussion of the Kolmogorov distance and the point metric is given in Section 2.10. The main ideas of our proofs consist in deriving a few precise local limit theorems

(LLTs) for S_n , with a more careful control of the error bounds (LLTs in the usual form or with large deviations being insufficient for our uses). Then we decompose the sum in (1.4) into major and minor parts, and evaluate the contribution of each. The idea is straightforward, but the technicalities, especially the error analysis, are more delicate and highly nontrivial.

The approach we used is also applicable to many other Poisson and non-Poisson approximation problems. A large number of examples and extensions can be found in [15]; for many others, see for example [5].

Poisson approximation has received extensive attention in recent probability and applied probability literature, and several different approaches have been proposed; see, for example, [1, 3, 5, 15, 20, 22, 28, 30] and the references therein.

This paper is organized as follows. To prove Theorems 1.5 and 1.6, we first establish another (more) uniform approximation to $d_{\text{TV}}^{(\alpha)}$ from which both theorems will follow. Although Theorems 1.1 and 1.3 are special cases of Theorems 1.5 and 1.6, respectively, we will sketch an elementary approach in Appendix for more methodological interests.

Notation. Except for the parameter λ whose use is clear in each occurrence, the use of all other symbols will be kept consistent.

2 A more uniform approximation to $d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda))$

Although the ranges specified by the two main approximation theorems 1.5 and 1.6 differ, their forms suggest the possibility of an asymptotic approximation holding uniformly in an even wider range than that covered by both theorems, which is the aim of this section. Indeed, the proof of the two main approximation theorems 1.5 and 1.6 relies on the following result, which is more uniform but at the price of a more complicated sum-approximant.

Throughout this section, $\sigma := \sqrt{\lambda - \lambda_2}$.

Theorem 2.1. *Let $x = x_m := (m - \lambda)/\sigma$. If $\theta \leq 1/2$ and $\lambda \geq K > 0$, then*

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda)) = W_\alpha(\lambda) \left(1 + O_K \left(\zeta_1^\alpha \lambda^{-\alpha-1/2} + \zeta_1 \lambda^{-1} \right) \right),$$

where $\zeta_1 := \theta + \lambda_3/\lambda_2$,

$$W_\alpha(\lambda) := \frac{1}{2} \sum_{m \geq 0} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| \frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}} \left(1 + \frac{C_1 x + C_2 x^3}{\sigma} \right) - 1 \right|^\alpha, \quad (2.1)$$

the two constants C_1, C_2 being bounded for all p_j 's defined by

$$C_1 := \frac{\lambda(\lambda_2 - 2\lambda_3) + \lambda_2^2}{2\lambda\sigma^2} \quad \text{and} \quad C_2 := \frac{2\lambda_3\lambda^2 - 3\lambda_2^2\lambda + \lambda_2^3}{6\lambda^2\sigma^2}. \quad (2.2)$$

Since $W_\alpha(\lambda)$ is a function of λ and θ , we can derive more precise expansions when λ is large or when θ is small. More precisely, if λ is large, then the sum W_α can be approximated by an integral with the use of the LLT for Poisson distribution. This will prove Theorem 1.5. On the other hand, if θ is small, then we can expand the terms inside the absolute-value sign in (2.1) with respect to θ and then deduce the estimate of Theorem 1.6.

To prove Theorem 2.1, let $\delta_{n,m} := \mathbb{P}(S_n = m) - e^{-\lambda} \lambda^m / m!$, and

$$F(z) := \sum_{0 \leq m \leq n} \mathbb{P}(S_n = m) z^m = \prod_{1 \leq j \leq n} (q_j + p_j z) = \prod_{1 \leq j \leq n} (1 + p_j(z - 1)),$$

where $q_j := 1 - p_j$. We first derive estimates for $|\delta_{n,m}|$, λ_k and $|F(z)|$. Then we prove several different versions of LLTs for S_n , from which Theorem 2.1 will be deduced.

2.1 An estimate for Poisson distribution

We start with an inequality for the Poisson distribution that will be used later. It is taken from [5, p. 259].

Lemma 2.2. For $m \geq 1$

$$e^{-\lambda} \frac{\lambda^m}{m!} \leq \frac{e^{-(m-\lambda)^2/(2(m+\lambda))}}{\sqrt{2\pi m}}. \quad (2.3)$$

Proof. A direct proof is as follows. By the inequality $m! \geq \sqrt{2\pi m} (m/e)^m$, we have

$$e^{-\lambda} \frac{\lambda^m}{m!} \leq \frac{1}{\sqrt{2\pi m}} \exp\left(-\lambda \left(1 - \frac{m}{\lambda} + \frac{m}{\lambda} \log \frac{m}{\lambda}\right)\right).$$

Then the upper bound (2.3) will follow from the elementary inequality $1 - x + x \log x \geq (1-x)^2/(2(1+x))$ for $x > 0$, or, equivalently,

$$\int_0^x \log(1+t) dt \geq \frac{x^2}{2(2+x)} \quad (x > -1). \quad (2.4)$$

To prove (2.4), observe first that $\log(1+t) \geq t/(1+t)$ for $t > -1$ since $\int_0^t \log(1+v) dv \geq 0$. Then $\int_0^x \log(1+t) dt \geq \int_0^x t/(1+t) dt$, and the right-hand side is bounded below by $x^2/(2(2+x))$ by considering the two cases $x \geq 0$ and $x \in (-1, 0]$. \square

2.2 A crude estimate for $|\delta_{n,m}|$

Lemma 2.3. For $m \geq 1$, we have

$$|\delta_{n,m}| \leq c_1 \lambda_2 \left(1 + (m/\lambda)^2\right) e^{-(m-\lambda)^2/(2(m+\lambda))}, \quad (2.5)$$

where $c_1 := e^2/2$.

Proof. First, by partial summation,

$$\begin{aligned} e^{\lambda(z-1)} - F(z) &= \sum_{1 \leq j \leq n} \left(e^{(p_j + \dots + p_n)(z-1)} \prod_{1 \leq \ell < j} (1 + p_\ell(z-1)) \right. \\ &\quad \left. - e^{(p_{j+1} + \dots + p_n)(z-1)} \prod_{1 \leq \ell \leq j} (1 + p_\ell(z-1)) \right) \\ &= \sum_{1 \leq j \leq n} e^{(p_{j+1} + \dots + p_n)(z-1)} (e^{p_j(z-1)} - 1 - p_j(z-1)) \prod_{1 \leq \ell < j} (1 + p_\ell(z-1)); \end{aligned}$$

from this and the elementary inequalities

$$|1 + p(z-1)| \leq 1 - p + p|z| = 1 + p(|z| - 1) \leq e^{p(|z|-1)} \quad (p \in [0, 1]),$$

$$|e^z - 1 - z| = \left| z^2 \int_0^1 e^{tz}(1-t)dt \right| \leq \frac{|z|^2}{2} e^{|z|},$$

we obtain

$$\begin{aligned} |F(z) - e^{\lambda(z-1)}| &\leq \sum_{1 \leq j \leq n} e^{(|z|-1) \sum_{\ell \neq j} p_\ell} |e^{p_j(z-1)} - 1 - p_j(z-1)| \\ &\leq c_1 \lambda_2 e^{\lambda(|z|-1)} |z-1|^2. \end{aligned}$$

Substituting this bound in the Cauchy integral formula, we get

$$|\delta_{n,m}| \leq c_1 \lambda_2 \min_{r>0} r^{-m} e^{\lambda(r-1)} (1+r^2),$$

for any $r > 0$. Taking $r = m/\lambda$ gives

$$\begin{aligned} |\delta_{n,m}| &\leq c_1 \lambda_2 \left(1 + (m/\lambda)^2\right) e^{m-\lambda-m \log(m/\lambda)} \\ &\leq c_1 \lambda_2 \left(1 + (m/\lambda)^2\right) e^{-(m-\lambda)^2/(2(m+\lambda))}, \end{aligned}$$

where the last estimate is obtained by the same proof of Lemma 2.2. This proves the Lemma. \square

2.3 Estimates for quantities involving λ_k

We prove in this subsection a few estimates for λ_k and $\lambda_k(r)$ defined as follows. Let $r := 1 + x/\sigma$, where $x := (m - \lambda)/\sigma$. (Recall that $\sigma = \sqrt{\lambda - \lambda_2}$.) Then

$$\lambda_k(r) := \sum_{1 \leq j \leq n} p_j(r)^k, \quad p_j(r) := \frac{p_j r}{q_j + p_j r}.$$

The reason of introducing the $\lambda_k(r)$'s is because the usual LLT is not sufficient for our purpose and we will need LLTs for moderate and large deviations. More precisely, define the Bernoulli random variables $X_j(r)$ such that

$$\mathbb{P}(X_j(r) = 1) = 1 - \mathbb{P}(X_j(r) = 0) = p_j(r).$$

Let $S_n(r) := X_1(r) + \dots + X_n(r)$. Then S_n and $S_n(r)$ are connected by the relation

$$\mathbb{P}(S_n = m) = r^{-m} F(r) \mathbb{P}(S_n(r) = m), \quad (2.6)$$

because $(q_j(r) = 1 - p_j(r))$

$$F(r; z) := \prod_{1 \leq j \leq n} (q_j(r) + p_j(r)z) = \sum_{0 \leq m \leq n} \mathbb{P}(S_n(r) = m) z^m = \frac{F(rz)}{F(r)}.$$

Lemma 2.4. For any $\rho > 0$,

$$0 \leq \lambda_1(\rho) - \lambda_k(\rho) \leq \max\{\rho, \rho^{-1}\} (k-1) \sigma^2 \quad (k \geq 1), \quad (2.7)$$

and

$$\lambda_1(\rho) \geq \frac{\rho}{2} (\lambda - \rho \lambda_2). \quad (2.8)$$

Proof. The upper bound of (2.7) follows immediately from

$$\begin{aligned} 0 \leq \lambda_1(\rho) - \lambda_k(\rho) &= \rho \sum_{1 \leq j \leq n} \left(\frac{p_j}{1 + (\rho - 1)p_j} \right) \left[1 - \left(1 - \frac{1 - p_j}{1 + (\rho - 1)p_j} \right)^{k-1} \right] \\ &\leq \rho(k-1) \sum_{1 \leq j \leq n} \frac{p_j - p_j^2}{(1 + (\rho - 1)p_j)^2} \leq \frac{\rho}{(\min\{1, \rho\})^2} (k-1)\sigma^2, \end{aligned}$$

since

$$1 + (\rho - 1)p_j = q_j + p_j\rho \geq \min\{1, \rho\}. \quad (2.9)$$

In particular, when $\rho = 1$, (2.7) has the form

$$0 \leq \lambda - \lambda_k \leq (k-1)\sigma^2. \quad (2.10)$$

For (2.8), we have

$$\begin{aligned} \lambda_1(\rho) &= \sum_{1 \leq j \leq n} \frac{p_j\rho}{1 + (\rho - 1)p_j} \geq \sum_{\substack{1 \leq j \leq n \\ p_j\rho \leq 1}} \frac{p_j\rho}{1 + (\rho - 1)p_j} \\ &\geq \frac{1}{2} \sum_{\substack{1 \leq j \leq n \\ p_j\rho \leq 1}} p_j\rho \geq \frac{1}{2} \sum_{\substack{1 \leq j \leq n \\ p_j\rho \leq 1}} p_j\rho(1 - p_j\rho) = \frac{\rho}{2}(\lambda - \lambda_2\rho). \end{aligned}$$

□

Lemma 2.5. For $r = 1 + x/\sigma$,

$$0 \leq m - \lambda_1(r) \leq (\lambda_2 - \lambda_3) \frac{(r-1)^2}{\min\{1, r\}}, \quad (2.11)$$

and, if $x = o(\sigma)$ as $\sigma \rightarrow \infty$, then

$$\frac{1}{\sqrt{\lambda_1(r) - \lambda_2(r)}} = \frac{1}{\sigma} \left(1 - \frac{\lambda - 3\lambda_2 + 2\lambda_3}{2\sigma^3} x + O\left(\frac{x^2}{\sigma^2}\right) \right). \quad (2.12)$$

Proof. We have

$$\begin{aligned} \lambda_1(r) - \lambda_2(r) &= r \sum_{1 \leq j \leq n} \frac{p_j - p_j^2}{(1 + (r-1)p_j)^2} \\ &= r\sigma^2 \left(1 - 2\frac{\lambda_2 - \lambda_3}{\sigma^3} x + O\left(\frac{\lambda_3 - \lambda_4}{\sigma^4} x^2\right) \right), \end{aligned}$$

which, together with the estimate $r^{-1/2} = 1 - x/(2\sigma) + O(x^2/\sigma^2)$, yields (2.12). Note that the factor $(\lambda - 3\lambda_2 + 2\lambda_3)/\sigma^3$ is of order σ^{-1} .

Similarly, since $m = \lambda + (r-1)\sigma^2$, we have $m - \lambda_1(r) = (r-1)^2 \sum_{1 \leq j \leq n} p_j^2(1 - p_j)/(1 + (r-1)p_j)$, from which (2.11) follows. □

2.4 Estimates for $|F(z)|$ and $|F(r; z)|$

Since our approach is based on Cauchy's integral formula, we also need some estimates for the probability generating functions $F(z)$ and $F(r; z)$.

Let $\sigma(r) := \sqrt{\lambda_1(r) - \lambda_2(r)}$.

Lemma 2.6. (i) For $|x| \leq \sigma^{1/3}$ with $\sigma \geq 2$,

$$\frac{F(r)}{r^m} = e^{-x^2/2} \left(1 + \frac{\lambda - 3\lambda_2 + 2\lambda_3}{6\sigma^3} x^3 + O\left(\frac{1+x^6}{\sigma^2}\right) \right). \quad (2.13)$$

(ii) For $|t| \leq \pi$,

$$|F(r; e^{it})| \leq e^{-c_2\sigma(r)^2 t^2}, \quad (2.14)$$

where $c_2 := 2/\pi^2$.

(iii) For $|t| \leq \sigma^{-2/3}$ and $|x| \leq \sigma^{1/3}$ with $\sigma \geq 2$,

$$\begin{aligned} F(r; e^{it})e^{-mit} &= e^{-\sigma(r)^2 t^2/2} \left(1 + (\lambda_1(r) - m)it \right. \\ &\quad \left. + \frac{\lambda_1(r) - 3\lambda_2(r) + 2\lambda_3(r)}{6} (it)^3 + O(x^4 t^2 + \sigma^4 t^6 + \sigma^2 t^4) \right). \end{aligned} \quad (2.15)$$

Proof. Since $r = 1 + x/\sigma$, $\sigma \geq 2$ and $|x| \leq \sigma^{1/3}$, we see that $|r - 1| \leq 2^{-1/6} < 1$, and thus

$$\begin{aligned} \log \frac{F(r)}{r^m} &= \sum_{j \geq 1} \frac{(-1)^{j-1} (\lambda_j - m)}{j} (r - 1)^j \\ &= -\frac{x^2}{2} + \frac{\lambda - 3\lambda_2 + 2\lambda_3}{6\sigma^3} x^3 + O\left(\frac{x^4}{\sigma^2}\right), \end{aligned}$$

because, by (2.10), $\lambda_j - m = \lambda_j - \lambda - x\sigma = O(j\sigma^2)$ for $j \geq 1$. This proves (2.13).

For (ii), we start from the relation

$$|q + pe^{it}|^2 = q^2 + p^2 + 2pq \cos t = 1 - 2pq(1 - \cos t) \quad (t \in \mathbb{R}),$$

which yields the inequality $|q + pe^{it}|^2 \leq e^{-2pq(1 - \cos t)}$. This and the inequality $1 - \cos t \geq c_2 t^2$ for $|t| \leq \pi$ give

$$|F(r; e^{it})| = \prod_{1 \leq j \leq n} |q_j(r) + p_j(r)e^{it}| \leq e^{-\sigma(r)^2(1 - \cos t)} \leq e^{-c_2\sigma(r)^2 t^2},$$

from which (2.14) follows.

Finally, for (iii), by Taylor expansion and (2.10), we have

$$\begin{aligned} \log(F(r; e^{it})e^{-mit}) &= (\lambda_1(r) - m)it + \sum_{k \geq 2} \frac{(-1)^{k-1}}{k} (\lambda_k(r) - \lambda_1(r))(e^{it} - 1)^k \\ &= (\lambda_1(r) - m)it - \frac{\sigma(r)^2}{2} t^2 + \frac{\lambda_1(r) - 3\lambda_2(r) + 2\lambda_3(r)}{6} (it)^3 + O(\sigma^2 t^4). \end{aligned}$$

This proves (2.15) and the lemma. \square

2.5 The usual LLT for S_n

With the above estimates available, we can now prove four LLTs, starting from the one in the usual form.

Proposition 2.7. *If $\sigma \rightarrow \infty$ as $n \rightarrow \infty$, then*

$$\mathbb{P}(S_n = m) = \frac{e^{-x^2/2}}{\sqrt{2\pi}\sigma} \left(1 + \frac{\lambda - 3\lambda_2 + 2\lambda_3}{6\sigma^3} (x^3 - 3x) + O\left(\frac{x^6 + 1}{\sigma^2}\right) \right). \quad (2.16)$$

uniformly for $m = \lambda + x\sigma$ with $|x| \leq \sigma^{1/3}$.

Proof. By Cauchy's (or Fourier's) integral formula

$$\begin{aligned} \mathbb{P}(S_n(r) = m) &= \frac{1}{2\pi} \int_{|t| \leq t_0} e^{-mit} F(r; e^{it}) dt + \frac{1}{2\pi} \int_{t_0 < |t| \leq \pi} e^{-mit} F(r; e^{it}) dt \\ &=: I_1 + I_2, \end{aligned} \quad (2.17)$$

where $t_0 := \sigma(r)^{-4/5}$. Consider first I_2 . By (2.14) of lemma 2.6,

$$I_2 = O\left(\int_{t_0}^{\pi} e^{-c_2\sigma(r)^2 t^2} dt\right) = O\left(\sigma(r)^{-6/5} e^{-c_2\sigma(r)^{2/5}}\right).$$

For I_1 , we use the estimate (2.15) of Lemma 2.6 and obtain

$$\begin{aligned} I_1 &= \frac{1}{2\pi\sigma(r)} \left(\int_{-\sigma(r)^{1/5}}^{\sigma(r)^{1/5}} e^{-t^2/2} dt + O\left(\frac{1+x^4}{\sigma^2}\right) \right) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \left(1 - \frac{\lambda - 3\lambda_2 + 2\lambda_3}{2\sigma^3} x + O\left(\frac{1+x^6}{\sigma^2}\right) \right), \end{aligned}$$

where the last estimate follows from (2.12).

Substituting the estimates for I_1 and I_2 in (2.17), we obtain an LLT for $\mathbb{P}(S_n(r) = m)$, which in turn, by (2.6), yields

$$\mathbb{P}(S_n = m) = \frac{F(r)}{\sqrt{2\pi} r^m \sigma} \left(1 - \frac{\lambda - 3\lambda_2 + 2\lambda_3}{2\sigma^3} x + O\left(\frac{1+x^6}{\sigma^2}\right) \right);$$

this implies (2.16) by the expansion for $F(r)/r^m$ given in (2.13). □

Remark. By a similar and simpler proof, we obtain the usual LLT for Poisson distribution

$$e^{-\lambda} \frac{\lambda^m}{m!} = \frac{e^{-x^2/2}}{\sqrt{2\pi\lambda}} \left(1 + \frac{x(x^2 - 3)}{6\sqrt{\lambda}} + O\left(\frac{1+x^6}{\lambda}\right) \right), \quad (2.18)$$

uniformly for $m = \lambda + x\sqrt{\lambda}$, where $x = o(\lambda^{1/6})$. This is often derived by Stirling's formula and will be used later.

2.6 A refined LLT for S_n

The error term of the LLT (2.16) is insufficient for our purpose. We derive a crucial LLT with a better error term in this subsection. Let $\Pi_m(\lambda) := e^{-\lambda} \lambda^m / m!$ and define the k -th difference operator

$$\begin{aligned} \Delta^k \Pi_m(\lambda) &:= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-mit + \lambda(e^{it} - 1)} (1 - e^{it})^k dt \\ &= \Pi_m(\lambda) \sum_{0 \leq j \leq k} \binom{k}{j} (-1)^j \frac{m(m-1) \cdots (m-j+1)}{\lambda^j} \quad (k = 0, 1, \dots). \end{aligned} \quad (2.19)$$

Proposition 2.8. *If the p_j 's satisfy $\lambda_2/\lambda \leq 1/2$ and $\lambda \geq c$, where $c > 0$ is an arbitrary fixed constant, then uniformly for $m \geq 0$*

$$\mathbb{P}(S_n = m) - \Pi_m(\lambda) = -\frac{\lambda_2}{2} \Delta^2 \Pi_m(\lambda) + \frac{\lambda_3}{3} \Delta^3 \Pi_m(\lambda) + I_\chi + O\left(\lambda_4 \lambda^{-5/2} + \theta^2 \lambda^{-3/2}\right), \quad (2.20)$$

where the implied constant in the O -symbol depends only on c and

$$I_\chi := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(\lambda-m)it - \lambda t^2/2} \left(1 + \frac{\lambda}{6} (it)^3\right) \chi(t) dt,$$

with

$$\chi(t) := e^{\lambda_2 t^2/2} - 1 - \frac{\lambda_2}{2} t^2 - \left(\frac{\lambda_2}{2} - \frac{\lambda_3}{3}\right) (it)^3 \left(e^{\lambda_2 t^2/2} - 1\right). \quad (2.21)$$

For results similar to (2.20), see [7, 19].

Proof. Assume first that $\max_{1 \leq j \leq n} p_j \leq 1/10$. Then $\lambda_k \leq 10^{1-k} \lambda$ for $k \geq 2$. By the expansion

$$\left| \log(1+z) - z + \frac{z^2}{2} - \frac{z^3}{3} \right| \leq \frac{|z|^4}{4(1-|z|)} \quad (|z| < 1),$$

we have, with some function $c(t)$ satisfying $|c(t)| \leq 1/2$ for $|t| \leq \pi$,

$$\begin{aligned} F(e^{it}) - e^{\lambda(e^{it}-1)} &= e^{\lambda(e^{it}-1)} \left(e^{-\lambda_2(e^{it}-1)^2/2 + \lambda_3(e^{it}-1)^3/3 + c(t)\lambda_4|e^{it}-1|^4} - 1 \right) \\ &= e^{\lambda(e^{it}-1)} \left(E(t) + O((\lambda_4 t^4 + \lambda_3^2 t^6) e^{c_3 \lambda_2 t^2}) \right), \end{aligned} \quad (2.22)$$

where we can take $c_3 = 3/5$ and $E(t) := e^{-\lambda_2(e^{it}-1)^2/2} (1 + \lambda_3(e^{it}-1)^3/3) - 1$. Now

$$\begin{aligned} E(t) + \frac{\lambda_2}{2} (e^{it} - 1)^2 - \frac{\lambda_3}{3} (e^{it} - 1)^3 \\ &= e^{-\lambda_2(e^{it}-1)^2/2} - 1 + \frac{\lambda_2}{2} (e^{it} - 1)^2 \\ &\quad + \frac{\lambda_3}{3} (e^{it} - 1)^3 \left(e^{-\lambda_2(e^{it}-1)^2/2} - 1 + \frac{\lambda_2}{2} (e^{it} - 1)^2 \right) - \frac{\lambda_2 \lambda_3}{6} (e^{it} - 1)^5. \end{aligned}$$

By the inequality $|(e^{it} - 1)^2 - (it)^2| \leq |t|^3$, we have the expansion

$$e^{-\lambda_2(e^{it}-1)^2/2} = e^{\lambda_2 t^2/2} \left(1 - \frac{\lambda_2}{2} ((e^{it} - 1)^2 - (it)^2) \right) + O\left(\lambda_2 t^6 e^{\lambda_2 t^2/2}\right).$$

It follows that

$$\begin{aligned}
E(t) &+ \frac{\lambda_2}{2}(e^{it} - 1)^2 - \frac{\lambda_3}{3}(e^{it} - 1)^3 \\
&= e^{\lambda_2 t^2/2} \left(1 - \frac{\lambda_2}{2}[(e^{it} - 1)^2 - (it)^2] \right) - 1 + \frac{\lambda_2}{2}(e^{it} - 1)^2 \\
&\quad + \frac{\lambda_3}{3}(it)^3 \left(e^{-\lambda_2(it)^2/2} - 1 + \frac{\lambda_2}{2}(it)^2 \right) - \frac{\lambda_2 \lambda_3}{6}(it)^5 \\
&\quad + O\left(\lambda_2 \lambda_3^2 t^8 e^{\lambda_2 t^2/2} + \lambda_2^2 t^6 e^{\lambda_2 t^2/2}\right) \\
&= e^{\lambda_2 t^2/2} \left(1 - \frac{\lambda_2}{2}t^2 \right) - 1 - \frac{\lambda_2}{2}((it)^2 + (it)^3) \left(e^{\lambda_2 t^2/2} - 1 \right) \\
&\quad + \frac{\lambda_3}{3}(it)^3 \left(e^{\lambda_2 t^2/2} - 1 \right) + O\left(\lambda_2^2 \lambda_3 t^8 e^{t^2 \lambda_2/2} + \lambda_2^2 t^6 e^{t^2 \lambda_2/2}\right).
\end{aligned}$$

Thus

$$E(t) = -\frac{\lambda_2}{2}(e^{it} - 1)^2 + \frac{\lambda_3}{3}(e^{it} - 1)^3 + \chi(t) + O\left(\lambda_2^2 \lambda_3 t^8 e^{t^2 \lambda_2/2} + \lambda_2^2 t^6 e^{t^2 \lambda_2/2}\right).$$

Substituting this expression into (2.22), we then obtain, by Cauchy's integral representation,

$$\begin{aligned}
\mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} &= -\frac{\lambda_2}{2} \Delta^2 \Pi_m(\lambda) + \frac{\lambda_3}{3} \Delta^3 \Pi_m(\lambda) \\
&\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-mit + \lambda(e^{it} - 1)} \chi(t) dt + O\left(\lambda_4 \lambda^{-5/2} + \theta^2 \lambda^{-3/2}\right).
\end{aligned}$$

We now simplify the integral involving $\chi(t)$. Since

$$\chi(t) = O\left(\lambda_2^2 t^4 e^{\lambda_2 t^2/2}\right), \quad (2.23)$$

we have, for a suitably small $\varepsilon > 0$,

$$\begin{aligned}
\int_{-\pi}^{\pi} e^{-mit + \lambda(e^{it} - 1)} \chi(t) dt &= \int_{-\varepsilon}^{\varepsilon} e^{(\lambda - m)it - \lambda t^2/2} \left(1 + \frac{\lambda}{6}(it)^3 + O(\lambda t^4 + \lambda^2 t^6) \right) \chi(t) dt + O\left(\theta^2 \lambda^{-3/2}\right) \\
&= 2\pi I_\chi + O\left(\theta^2 \lambda^{-3/2}\right).
\end{aligned}$$

To complete the proof, we consider now the case when $\max_{1 \leq j \leq n} \{p_j\} \geq 1/10$. In this case, we split the integral into two parts

$$\mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} = \frac{1}{2\pi} \left(\int_{-1/10}^{1/10} + \int_{1/10 < |t| \leq \pi} \right) e^{-mit} \left(F(e^{it}) - e^{\lambda(e^{it} - 1)} \right) dt.$$

The first integral can be estimated exactly as in the case when $\max_{1 \leq j \leq n} \{p_j\} \leq 1/10$, and the second integral is bounded above by $O\left(e^{-c_2(\lambda - \lambda_2)/100} + e^{-\lambda(1 - \cos(1/10))}\right) = O\left(\lambda_2^2 \lambda^{-5}\right)$, by using (2.14) and the inequality $\lambda_2 \geq 1/100$.

□

2.7 A simpler version of the refined LLT

Although the integral I_χ in (2.20) can be computed explicitly, the resulting expression is rather complicated and not needed in this paper. We derive instead a simpler LLT at the price of weaker error terms; this LLT is also crucial in the development of our argument.

Proposition 2.9. *If $\lambda_2/\lambda = \theta \leq 1/2$ and $\lambda \geq c > 0$, then*

$$\begin{aligned} \mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} &= -\frac{\lambda_2}{2} \Delta^2 \Pi_m(\lambda) + \frac{1}{\sqrt{2\pi\lambda}} \left(\frac{1}{\sqrt{1-\theta}} - 1 - \frac{\theta}{2} \right) \\ &\quad + O\left(\frac{\theta^2(|m-\lambda|^2+1)}{\lambda^{3/2}} + \frac{\lambda_3(|m-\lambda|+1)}{\lambda^{5/2}} \right), \end{aligned} \quad (2.24)$$

uniformly for $m \geq 0$. The constant in the O -symbol depends only on c .

This expansion is useful when $|m - \lambda|$ is not too large.

Proof. First, by (2.19),

$$\Delta^3 \Pi_m(\lambda) = O\left(e^{-\lambda} \frac{\lambda^m}{m!} \left(\frac{1+|m-\lambda|}{\lambda^2} + \frac{|m-\lambda|^3}{\lambda^3} \right) \right).$$

It remains, by (2.20), to simplify the integral I_χ .

$$\begin{aligned} I_\chi &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\lambda t^2/2} \left\{ (1 + (\lambda - m)it + O((\lambda - m)^2 t^2)) \left(1 + \lambda \frac{(it)^3}{3!} \right) \right. \\ &\quad \left. \times \left(e^{\lambda_2 t^2/2} - 1 - \frac{\lambda_2}{2} t^2 - \left(\frac{\lambda_2}{2} - \frac{\lambda_3}{3} \right) (it)^3 (e^{\lambda_2 t^2/2} - 1) \right) \right\} dt. \end{aligned}$$

It follows, by expanding the factors inside the curly braces and by estimating term by term, that

$$\begin{aligned} I_\chi &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\lambda t^2/2} \left(e^{\lambda_2 t^2/2} - 1 - \frac{\lambda_2}{2} t^2 \right) dt \\ &\quad + O\left(\lambda |\lambda - m| \int_{-\infty}^{\infty} e^{-\lambda t^2/2} t^4 \left(e^{\lambda_2 t^2/2} - 1 - \frac{\lambda_2}{2} t^2 \right) dt \right. \\ &\quad \left. + \lambda_2 |\lambda - m| \int_{-\infty}^{\infty} e^{-\lambda t^2/2} t^4 \left(e^{\lambda_2 t^2/2} - 1 \right) dt + \frac{\lambda_2^2 (|\lambda - m|^2 + 1)}{\lambda^{7/2}} \right) \\ &= \frac{1}{\sqrt{2\pi\lambda}} \left(\frac{1}{\sqrt{1-\lambda_2/\lambda}} - 1 - \frac{\lambda_2}{2\lambda} \right) + O\left(\frac{\theta^2(|m-\lambda|^2+1)}{\lambda^{3/2}} \right), \end{aligned}$$

since the integrals involving odd integrands are equal to zero. This proves (2.24). \square

2.8 Yet another LLT when $\theta \leq 1/2$

We now derive another LLT for S_n when $\theta = \lambda_2/\lambda \leq 1/2$. This LLT is based on Proposition 2.8, and can be regarded as a hybrid of Propositions 2.7 and 2.8.

Proposition 2.10. Assume $\lambda \geq K > 0$ and that $m = \lambda + x\sigma$ satisfies

$$|x| + \frac{|x|^3}{m} = O\left(\theta^{-1/4}\lambda^{1/6}\right) \quad \text{and} \quad m \geq 2\lambda_2. \quad (2.25)$$

Then for $\theta \leq 1/2$

$$\mathbb{P}(S_n = m) = e^{-\lambda} \frac{\lambda^m e^{-\theta x^2/2}}{m! \sqrt{1-\theta}} \left(1 + \frac{C_1 x + C_2 x^3}{\sigma} + O_K \left(\left(\frac{\theta^2}{\lambda} + \frac{\lambda_3}{\lambda^2} \right) \left(1 + \frac{x^2}{m} \right)^4 (1+x^6) \right) \right), \quad (2.26)$$

where C_1, C_2 are given in (2.2).

Propositions 2.7 and 2.10 provide asymptotic approximations to the probability $\mathbb{P}(S_n = m)$ when m lies in the range

$$-\sigma^{1/3} < \frac{m - \lambda}{\sigma} < \sigma^{1/3},$$

as $\sigma \rightarrow \infty$. Proposition 2.7 can be used to estimate the closeness between $\mathbb{P}(S_n = m)$ and the density of the normal distribution when $\theta \geq \varepsilon > 0$, while proposition 2.10 yields a more accurate estimate when $\theta = o(1)$.

Proof. Since the range (2.25) is wider than that specified by the usual LLT ($x = O(1)$ or $x = o(\lambda^{1/6})$; see (2.16)) when $\theta \rightarrow 0$, our method of proof here is to apply Proposition 2.8 to $p_j(r)$ instead of to p_j , which gives an LLT for $S_n(r)$, and then to use the relationship (2.6) between S_n and $S_n(r)$ to deduce (2.26). As the proof of Proposition 2.8, the error analysis constitutes the hard part of the proof.

In what follows, we assume that $\theta \rightarrow 0$, for otherwise (2.26) follows directly from Proposition 2.7.

Smallness of $\lambda_2(r)/\lambda_1(r) \leq 1/2$ when $\theta \rightarrow 0$. We first show that the probabilities $p_j(r)$ satisfy the condition of Proposition 2.8, namely, $\lambda_2(r)/\lambda_1(r) \leq 1/2$ when $\theta \rightarrow 0$. To this purpose, we observe, by (2.9), that, for any $\rho > 0$,

$$\lambda_k(\rho) \leq \max\{1, \rho^k\} \lambda_k \quad (k \geq 1). \quad (2.27)$$

This (with $k = 2$) and the lower bound for $\lambda_1(r)$ given in (2.8) yield

$$\frac{\lambda_2(r)}{\lambda_1(r)} \leq 2 \max\left\{r, \frac{1}{r}\right\} \frac{\lambda_2}{\lambda - r\lambda_2} \leq 2 \left(r + \frac{1}{r}\right) \frac{\lambda_2}{\lambda - r\lambda_2}. \quad (2.28)$$

Note that r can be expressed in terms of m as $r = (m - \lambda_2)/\sigma^2$. From this, the condition $m \geq 2\lambda_2$, and $\lambda_2/\lambda \rightarrow 0$, it follows that $r + 1/r = m/\lambda + \lambda/m = O(1 + x^2/m)$. Consequently, by condition (2.25), we obtain

$$|r - 1| \left(r + \frac{1}{r}\right) = O\left(\sigma^{-1} \left(|x| + \frac{|x|^3}{m}\right)\right) = O\left(\theta^{-1/4}\lambda^{-1/3}\right); \quad (2.29)$$

this estimate will be used several times below. In particular, this implies, by (2.28), that

$$\frac{\lambda_2(r)}{\lambda_1(r)} = O\left(\frac{\lambda_2}{\lambda} \left(r + \frac{1}{r}\right)\right) = O(\theta^{3/4}) = o(1), \quad (2.30)$$

(by considering the two cases $|x| \leq \lambda^{1/6}$ and $|x| > \lambda^{1/6}$).

Application of Proposition 2.8 and simplification. From (2.30), we see that the probabilities $p_j(r)$ satisfy the condition of the Proposition 2.8, which, together with the two estimates,

$$\begin{aligned} e^{-\lambda_1(r)} \frac{\lambda_1(r)^m}{m!} &= O(\lambda_1(r)^{-1/2}), \\ 1 - 2 \frac{m}{\lambda_1(r)} + \frac{m(m-1)}{\lambda_1(r)^2} &= -\frac{1}{\lambda_1(r)} + O\left(\frac{|m - \lambda_1(r)|^2 + |m - \lambda_1(r)|}{\lambda_1(r)^2}\right), \end{aligned} \quad (2.31)$$

(see (2.19)) for $m \geq 0$, implies that $\mathbb{P}(S_n(r) = m) = Q_1 + O(E_2)$, where

$$\begin{aligned} Q_1 &:= e^{-\lambda_1(r)} \frac{\lambda_1(r)^m}{m!} \left(1 + \frac{\lambda_2(r)}{2\lambda_1(r)}\right) + \frac{1}{\sqrt{2\pi\lambda_1(r)}} \left(\frac{1}{\sqrt{1 - \lambda_2(r)/\lambda_1(r)}} - 1 - \frac{\lambda_2(r)}{2\lambda_1(r)}\right), \\ E_2 &:= \left(\frac{\lambda_2(r)}{\lambda_1(r)}\right)^2 \frac{|m - \lambda_1(r)|^2 + 1}{\lambda_1(r)^{3/2}} + \frac{\lambda_2(r)(|m - \lambda_1(r)| + |m - \lambda_1(r)|^2) + \lambda_3(r)}{\lambda_1(r)^{5/2}}. \end{aligned} \quad (2.32)$$

Now from (2.11) and (2.29) it follows that

$$|m - \lambda_1(r)| = O\left(\theta x^2 \left(1 + \frac{x^2}{m}\right)\right) \quad (m \geq 0). \quad (2.33)$$

By substituting this, (2.27), and (2.30) into (2.32), we deduce that $E_2 = O(E_3)$, where

$$E_3 := \lambda^{-3/2} \left(\theta^2 + \frac{\lambda_3}{\lambda}\right) \left(1 + \frac{x^2}{m}\right)^4 (1 + x^4).$$

On the other hand, by the estimate

$$e^{-\lambda_1(r)} \frac{\lambda_1(r)^m}{m!} - \frac{1}{\sqrt{2\pi\lambda_1(r)}} = O\left(\frac{|m - \lambda_1(r)|^2 + 1}{\lambda_1(r)^{3/2}}\right) \quad (m \geq 0),$$

we have

$$Q_1 = e^{-\lambda_1(r)} \frac{\lambda_1(r)^m}{m!} \cdot \frac{1}{\sqrt{1 - \lambda_2(r)/\lambda_1(r)}} + O\left(\frac{|m - \lambda_1(r)|^2 + 1}{\lambda_1(r)^{3/2}}\right),$$

where the O -term is also bounded above by E_3 . Thus we obtain

$$\mathbb{P}(S_n(r) = m) = e^{-\lambda_1(r)} \frac{\lambda_1(r)^m}{m!} \cdot \frac{1}{\sqrt{1 - \lambda_2(r)/\lambda_1(r)}} + O(E_3). \quad (2.34)$$

Further simplification. Consider first the dominant term on the right-hand side of (2.34). By (2.33), we have

$$\frac{(m - \lambda_1(r))^2}{m} = O\left(\frac{\theta^2}{\lambda} x^4 \left(1 + \frac{x^2}{m}\right)^2\right) = O(\theta \lambda^{-1/3}) = o(1). \quad (2.35)$$

Thus

$$\begin{aligned} e^{-\lambda_1(r)} \frac{\lambda_1(r)^m}{m!} &= \frac{1}{m!} \left(\frac{m}{e}\right)^m \exp\left\{m \left(\log\left(1 - \frac{m - \lambda_1(r)}{m}\right) + \frac{m - \lambda_1(r)}{m}\right)\right\} \\ &= \frac{1}{m!} \left(\frac{m}{e}\right)^m \left(1 + O\left(\frac{\theta^2}{\lambda} x^4 \left(1 + \frac{x^2}{m}\right)^2\right)\right). \end{aligned} \quad (2.36)$$

Note that by (2.35) the terms inside the curly braces are bounded.

Next we express the factor $(1 - \lambda_2(r)/\lambda_1(r))^{-1/2}$ in terms of the λ_j 's. By straightforward expansion

$$\frac{1}{\sqrt{1 - \lambda_2(r)/\lambda_1(r)}} = \frac{1}{\sqrt{1 - \theta}} \left(1 + C_1 \frac{x}{\sigma} + O \left(\frac{x^2}{\lambda} \left(\theta^2 + \frac{\lambda_3}{\lambda} \right) \left(1 + \frac{x^2}{m} \right)^3 \right) \right),$$

where C_1 is given in (2.2). Substituting this and (2.36) into (2.34), we obtain

$$\mathbb{P}(S_n(r) = m) = \frac{(m/e)^m}{m! \sqrt{1 - \theta}} \left(1 + C_1 \frac{x}{\sigma} \right) + O(E_3), \quad (2.37)$$

where the new error terms introduced for the dominant term in (2.34) are absorbed in E_3 .

Connection between the distribution of $S_n(r)$ and that of S_n . We now use (2.6) to derive (2.26). Consider $F(r)$. By Taylor expansion

$$\log F(r) = \lambda(r - 1) - \frac{\lambda_2}{2}(r - 1)^2 + \frac{\lambda_3}{3}(r - 1)^3 + O \left(\frac{(r - 1)^4 \lambda_4}{\min\{1, r\}} \right).$$

The O -term is, by (2.25) and (2.29), of order $O(\theta^{3/4} \lambda^{-1/3}) = o(1)$. Thus, similar to (2.36),

$$\begin{aligned} \frac{F(r)}{r^m} \cdot \frac{(m/e)^m}{m!} &= e^{-\lambda \frac{\lambda^m}{m!}} F(r) e^{-\lambda(r-1)} \left(\frac{\lambda r}{m} \right)^{-m} e^{\lambda r - m} \\ &= e^{-\lambda \frac{\lambda^m}{m!}} \exp \left\{ -\frac{\lambda_2}{2}(r - 1)^2 + \frac{\lambda_3}{3}(r - 1)^3 + \frac{(m - \lambda r)^2}{2m} + \frac{(m - \lambda r)^3}{3m^2} \right\} \\ &\quad \times \left(1 + O \left(\frac{(r - 1)^4 \lambda_4}{\min\{1, r\}} + \frac{(m - \lambda r)^4}{m^3} \right) \right), \end{aligned}$$

because

$$m = \lambda r - \lambda_2(r - 1) \quad (2.38)$$

and $(m - \lambda r)^4/m^3 = O(\lambda_2^4(r - 1)^4 \lambda^{-3}(r + 1/r)^4) = o(1)$. Hence by (2.38) and the conditions (2.25) on m , we can simplify the above estimate and obtain

$$\frac{F(r)}{r^m} \cdot \frac{(m/e)^m}{m!} = e^{-\lambda \frac{\lambda^m}{m!}} e^{-\theta x^2/2} \left(1 + C_2 \frac{x^3}{\sigma} + O(E_4) \right),$$

where C_2 is given in (2.2) and $E_4 = x^4 \lambda^{-1} (\theta^2 + \lambda_3/\lambda) (1 + x^2)$. This also implies in particular that $F(r)/(r^m \sqrt{m}) = O(e^{-\lambda \lambda^m} e^{-\theta x^2/2}/m!)$. Combining this with (2.6) and (2.37), we deduce that

$$\mathbb{P}(S_n = m) = e^{-\lambda \frac{\lambda^m e^{-\theta x^2/2}}{m! \sqrt{1 - \theta}}} \left(1 + \frac{C_1 x + C_2 x^3}{\sigma} + O(E_5) \right),$$

where $E_5 := \lambda^{1/2} E_3 + E_4$. It is easily checked term by term that E_5 is bounded by the O -term on the right-hand side of (2.26). This completes the proof of the theorem. \square

2.9 Proof of Theorem 2.1

We are now ready to prove Theorem 2.1. The idea of proof is to split the sum $\sum_{m \geq 0} |\delta_{n,m}|^\alpha$ into two parts according to $m < 2\lambda_2$ and $m \geq 2\lambda_2$. The former is easily estimated by applying Lemma 2.3 and is asymptotically negligible, while the latter is more involved and requires the more precise expansion (2.26).

An estimate for $\sum_{m < 2\lambda_2} |\delta_{n,m}|^\alpha$. Consider first the sum of $|\delta_{n,m}|^\alpha$ over those m satisfying $m < 2\lambda_2$. We further distinguish between two cases: $\lambda_2 \geq 1/4$ and $\lambda_2 < 1/4$.

– If $\lambda_2 \geq 1/4$, then $\theta \geq 1/(4\lambda)$. By applying the estimate (2.5) of $\delta_{n,m}$, we obtain

$$\sum_{m \leq 2\lambda_2} \left| |\delta_{n,m}|^\alpha - \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^\alpha \right| = O\left(\lambda_2 e^{-\alpha\lambda/4}\right) = O\left(\theta^\alpha \zeta_1^\alpha \lambda^{-3\alpha/2}\right), \quad (2.39)$$

(recalling that $\zeta_1 := \theta + \lambda_3/\lambda_2$).

– If $\lambda_2 < 1/4$, then the above sum contains only one term corresponding to $m = 0$ and we have

$$\begin{aligned} |\delta_{n,0}| &= |\mathbb{P}(S_n = 0) - e^{-\lambda}| = e^{-\lambda} \frac{\lambda_2}{2} \left(1 + O\left(\frac{\lambda_3}{\lambda_2} + \lambda_2\right) \right), \\ e^{-\lambda} |e^{-\theta\varphi(-\sigma)} - 1| &= e^{-\lambda} \frac{\lambda_2}{2} \left(1 + O\left(\frac{\lambda_3}{\lambda_2} + \lambda_2\right) \right). \end{aligned}$$

Thus the estimate (2.39) also holds in the case $\lambda_2 \leq 1/4$.

An asymptotic approximation to $\sum_{m \geq 2\lambda_2} |\delta_{n,m}|^\alpha$. We now assume that $m \geq 2\lambda_2$. Since $\zeta_1 \geq \theta$, we see that the condition (2.25) of Proposition 2.10 is satisfied whenever

$$m \geq 2\lambda_2 \quad \text{and} \quad |x| \leq (\lambda/\zeta_1)^{1/18}. \quad (2.40)$$

Thus, for m lying in this region, we have, by (2.26),

$$\delta_{n,m} = e^{-\lambda} \frac{\lambda^m}{m!} \left(e^{\theta\varphi(x)/2} - 1 + O\left(\theta \frac{\zeta_1}{\lambda} (1 + x^{10})\right) \right),$$

where

$$\varphi(x) = -x^2 - \frac{\log(1-\theta)}{\theta} + \frac{2}{\theta} \log\left(1 + \frac{C_1 x + C_2 x^3}{\sigma}\right). \quad (2.41)$$

We now discuss when the dominant term $e^{\theta\varphi(x)/2} - 1$ is not smaller than the O -term. Observe that

$$C_1 = \frac{\theta}{2} + O(\theta\zeta_1) \quad \text{and} \quad C_2 = O(\theta\zeta_1). \quad (2.42)$$

This implies that

$$\varphi(x) = -x^2 + 1 + \frac{x}{\sqrt{\lambda}} + O(\zeta_1 (1 + x^6) \lambda^{-1/2}).$$

Since $\zeta_1 = O(\min\{\theta, \sqrt{\lambda_2}\})$, we have, for x satisfying $4\sqrt{1 + 1/(4K)} \leq |x| \leq (\lambda/\zeta_1)^{1/18}$,

$$-c_4 x^2 < \varphi(x) < -c_5 x^2, \quad (2.43)$$

for sufficiently small θ , where c_4 and c_5 are positive absolute constants.

Now for $|x| \leq 4\sqrt{1 + 1/(4K)}$, we have $\varphi'(x) = -2(x - x_0) + O(\zeta_1 \lambda^{-1/2})$, where $x_0 := 1/(2\sqrt{\lambda})$; integrating this estimate, we obtain

$$\varphi(x) = \varphi(x_0) - (x - x_0)^2 + O(\zeta_1 \lambda^{-1/2}) \geq \frac{3}{4} + O(\zeta_1 \lambda^{-1/2})$$

for $|x - x_0| \leq 1/2$, and

$$1 + O(\zeta_1 \lambda^{-1/2}) \leq |\varphi'(x)| = O(1), \quad (2.44)$$

for $|x - x_0| \leq |x| \leq 4\sqrt{1 + 1/(4K)}$. Furthermore, it is easily checked that

$$\varphi\left(x_0 \pm 2\sqrt{1 + \frac{1}{4K}}\right) \leq -1 + O(\zeta_1 \lambda^{-1/2}).$$

From these estimates, we conclude that $\varphi(x)$ has two zeroes $x_+ = x_+(\lambda, \lambda_2, \lambda_3)$ and $x_- = x_-(\lambda, \lambda_2, \lambda_3)$ in the interval $|x| \leq 4\sqrt{1 + 1/(4K)}$, such that

$$x_{\pm} \rightarrow \xi_{\pm} := x_0 \pm \sqrt{1 + \frac{1}{4\lambda}}, \quad (2.45)$$

uniformly for all $\lambda \geq K > 0$ as $\theta \rightarrow 0$.

Since $\varphi(x_+) = 0$, we can integrate the estimate (2.44) and obtain

$$|\varphi(x)| = \left| \int_{x_+}^x \varphi'(t) dt \right| \geq |x_+ - x|(1 + O(\zeta_1 \lambda^{-1/2})). \quad (2.46)$$

Consequently, there can be at most a finite number of m (indeed two if θ is sufficiently small) such that $|\varphi(x)| \leq \zeta_1/\lambda$ and in that case

$$|\delta_{n,m}|^\alpha - \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^\alpha = O\left(\theta^\alpha \zeta_1^\alpha \lambda^{-3\alpha/2} \right). \quad (2.47)$$

On the other hand, when $|\varphi(x)| \geq \zeta_1/\lambda$ and x satisfies (2.40), we have

$$\begin{aligned} & |\delta_{n,m}|^\alpha - \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^\alpha \\ &= \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^\alpha \left| \left(1 + O\left(\frac{\theta\zeta_1(1+x^{10})}{\lambda |e^{\theta\varphi(x)/2} - 1|} \right) \right)^\alpha - 1 \right| \\ &= O\left(\left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^{\alpha-1} \theta \frac{\zeta_1}{\lambda} (1+x^{10}) \right). \end{aligned} \quad (2.48)$$

Combining (2.47) and (2.48), we obtain

$$\begin{aligned} \sum_{|x| \leq (\lambda/\zeta_1)^{1/18}} |\delta_{n,m}|^\alpha &= \sum_{|x| \leq (\lambda/\zeta_1)^{1/18}} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^\alpha \\ &+ O\left(\frac{\theta\zeta_1}{\lambda} \sum_{|\varphi(x)| \geq \zeta_1/\lambda} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^\alpha \left| e^{\theta\varphi(x)/2} - 1 \right|^{\alpha-1} \right) + O\left(\theta^\alpha \zeta_1^\alpha \lambda^{-3\alpha/2} \right) \\ &= 2W_\alpha(\lambda) + O\left(\theta^\alpha \zeta_1 \lambda^{-(\alpha+1)/2} + \theta^\alpha \zeta_1^\alpha \lambda^{-3\alpha/2} \right), \end{aligned}$$

where $W_\alpha(\lambda)$ is defined in (2.1).

The sum of $|\delta_{n,m}|^\alpha$ over the remaining range of m , namely, $|x| \leq (\lambda/\xi_1)^{1/18}$ can be estimated by means of Lemma 2.2 and we obtain

$$\sum_{|x| > (\lambda/\xi_1)^{1/18}} |\delta_{n,m}|^\alpha = O\left(\theta^\alpha \xi_1^\alpha \lambda^{-3\alpha/2}\right).$$

Lower and upper bounds for $W_\alpha(\lambda)$. To finish the proof, we need to estimate the order of $W_\alpha(\lambda)$. First, since $C_1 = O(\theta)$ and $C_2 = O(\theta)$, we have the upper bound

$$W_\alpha(\lambda) = O\left(\theta^\alpha \sum_{m \geq 0} \left(e^{-\lambda \frac{\lambda^m}{m!}}\right)^\alpha (1 + |x|^3)^\alpha\right) = O\left(\theta^\alpha \lambda^{(1-\alpha)/2}\right).$$

Next, we see that the same estimate holds from below since

$$\begin{aligned} W_\alpha(\lambda) &\geq \frac{1}{2} \sum_{|x-x_0| \leq 1/2} \left(e^{-\lambda \frac{\lambda^m}{m!}}\right)^\alpha \left|\theta \frac{\varphi(x)}{2}\right|^\alpha \\ &\geq \left|\theta \frac{3}{8}\right|^\alpha \left(\frac{1}{2} + o(1)\right) \sum_{|x-x_0| \leq 1/2} \left(e^{-\lambda \frac{\lambda^m}{m!}}\right)^\alpha \\ &\geq O\left(\theta^\alpha \lambda^{-\alpha/2} \sum_{|x-x_0| \leq 1/2} e^{-\alpha x^2(1-\theta)/2}\right) \geq O\left(\theta^\alpha \lambda^{(1-\alpha)/2}\right), \end{aligned}$$

as $\theta \rightarrow 0$ and $\lambda \rightarrow \infty$. This completes the proof of Theorem 2.1.

2.10 Kolmogorov distance and the point metric

The methods of proof we used above can be readily amended for the consideration of other distances. We briefly discuss the Kolmogorov distance and the point metric and start with the following lemma.

Lemma 2.11. For $m = \lambda + x\sigma \geq 0$

$$\begin{aligned} \mathbb{P}(S_n \leq m) - \sum_{j \leq m} e^{-\lambda} \frac{\lambda^j}{j!} &= -\frac{\lambda_2}{2} \Delta^1 \Pi_m(\lambda) + \Phi\left(\sqrt{1-\theta}x\right) - \Phi(x) - \frac{\theta\sqrt{1-\theta}}{2\sqrt{2\pi}} x e^{-(1-\theta)x^2/2} \\ &\quad + O\left(\lambda_3 \lambda^{-3/2} + \theta^2 \lambda^{-1/2}\right), \end{aligned} \tag{2.49}$$

$$\begin{aligned} \mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} &= -\frac{\lambda_2}{2} \Delta^2 \Pi_m(\lambda) + \frac{e^{-(1-\theta)x^2/2}}{\sqrt{2\pi\lambda}} \left(\frac{\theta}{2}((1-\theta)x^2 - 1) - 1 + \frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}}\right) \\ &\quad + O\left(\lambda_3 \lambda^{-2} + \theta^2 \lambda^{-1}\right). \end{aligned} \tag{2.50}$$

Proof. First of all, since

$$\mathbb{P}(S_n \leq m) - \sum_{j \leq m} e^{-\lambda} \frac{\lambda^j}{j!} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-mit} \frac{F(e^{it}) - e^{\lambda(e^{it}-1)}}{1 - e^{it}} dt,$$

we can apply the same proof used for (2.20) (Proposition 2.8), together with the expansions

$$\frac{1}{1 - e^{it}} = -\frac{1}{it} + \frac{1}{2} + O(|t|),$$

and (2.22), and deduce that

$$\mathbb{P}(S_n \leq m) - \sum_{j \leq m} e^{-\lambda} \frac{\lambda^j}{j!} = -\frac{\lambda_2}{2} \Delta^1 \Pi_m(\lambda) + \frac{\lambda_3}{3} \Delta^2 \Pi_m(\lambda) + I_3 + O(\lambda_4 \lambda^{-2} + \theta^2 \lambda^{-1}),$$

where $\chi(t)$ is given in (2.21))

$$I_3 := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(\lambda-m)it - \lambda t^2/2} \left(1 + \frac{\lambda}{6}(it)^3\right) \left(\frac{i}{t} + \frac{1}{2}\right) \chi(t) dt;$$

compare Theorem 1 in [29]. By the O -estimate (2.23) of $\chi(t)$, we can simplify I_3 as follows.

$$\begin{aligned} I_3 &= \frac{i}{2\pi} \int_{-\infty}^{\infty} e^{(\lambda-m)it - \lambda t^2/2} \frac{\chi(t)}{t} dt + O(\theta^2 \lambda^{-1/2}) \\ &= \frac{i}{2\pi} \int_{-\infty}^{\infty} e^{(\lambda-m)it - \lambda t^2/2} \left(e^{\lambda_2 t^2/2} - 1 - \frac{\lambda_2}{2} t^2\right) \frac{dt}{t} + O(\theta^2 \lambda^{-1/2}) \\ &= \Phi\left(\frac{m-\lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{m-\lambda}{\sqrt{\lambda-\lambda_2}}\right) + \frac{\lambda_2(m-\lambda)}{2\lambda^{3/2}\sqrt{2\pi}} e^{-(m-\lambda)^2/(2\lambda)} + O(\theta^2 \lambda^{-1/2}). \end{aligned}$$

On the other hand, $\Delta^2 \Pi_m(\lambda) = O(\lambda^{-3/2})$. Collecting these estimates, we obtain (2.49).

Similarly, we can show that $\Delta^2 \pi(\lambda, m)$ and $\Delta^3 \pi(\lambda, m)$ are smaller in order than the error terms of the estimates (2.49) and (2.50), respectively. Then (2.50) follows by the same line of arguments using the expression

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-yit - \lambda t^2/2} \left(e^{\lambda_2 t^2/2} - 1 - \frac{\lambda_2}{2} t^2\right) dt = \frac{e^{-y^2/(2\lambda-2\lambda_2)}}{\sqrt{\lambda-\lambda_2}} + \frac{e^{-y^2/(2\lambda)}}{2\lambda^{5/2}} (-2\lambda^2 - \lambda\lambda_2 + \lambda_2 y^2).$$

□

We now derive asymptotic approximations to the Kolmogorov distance and the point metric between the distribution of S_n and $\text{Po}(\lambda)$.

Theorem 2.12. *If $\lambda \rightarrow \infty$ and $\theta \leq 1/2$, then*

$$\sup_{m \geq 0} \left| \mathbb{P}(S_n \leq m) - \sum_{j \leq m} e^{-\lambda} \frac{\lambda^j}{j!} \right| = \frac{\theta}{2} J_1(\theta) \left(1 + O(\lambda^{-1/2})\right), \quad (2.51)$$

$$\sup_{m \geq 0} \left| \mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} \right| = \frac{1}{\sqrt{2\pi\lambda}} \left(\frac{1}{\sqrt{1-\theta}} - 1\right) \left(1 + O(\lambda^{-1/2})\right), \quad (2.52)$$

where J_1 is defined in (1.7).

Proof. We first prove (2.51). By the usual LLT for Poisson distribution (2.18), we have, for $m = \lambda + x\sigma$, $|x| \leq \lambda^{1/6}$,

$$\Delta^1 \Pi_m(\lambda) = e^{-\lambda} \frac{\lambda^m}{m!} \left(1 - \frac{m}{\lambda}\right) = -\frac{\sqrt{1-\theta}}{\sqrt{2\pi\lambda}} x e^{-(1-\theta)x^2/2} \left(1 + O\left(\frac{1+|x|^3}{\sqrt{\lambda}}\right)\right).$$

Substituting this estimate into (2.49), we obtain

$$\mathbb{P}(S_n \leq m) - \sum_{j \leq m} e^{-\lambda} \frac{\lambda^j}{j!} = \Phi(\sqrt{1-\theta}x) - \Phi(x) + O(\theta\lambda^{-1/2}),$$

for $|x| \leq \lambda^{1/6}$. Thus

$$\sup_{m \geq 0} \left| \mathbb{P}(S_n \leq m) - \sum_{j \leq m} e^{-\lambda} \frac{\lambda^j}{j!} \right| = \left(1 + O(\lambda^{-1/2})\right) \sup_{x \in \mathbb{R}} \left| \Phi(\sqrt{1-\theta}x) - \Phi(x) \right|.$$

Now it is straightforward to show that the maximum of the function $|\Phi(\sqrt{1-\theta}x) - \Phi(x)|$ is reached at $x = \pm \sqrt{-\theta^{-1} \log(1-\theta)}$ (see also [31]). This proves (2.51).

We now prove (2.52). First, by the usual LLT (2.18) and (2.19) with $k = 2$, we have, for $m = \lambda + x\sigma$, $|x| \leq \lambda^{1/6}$,

$$\Delta^2 \Pi_m(\lambda) = \frac{e^{-(1-\theta)x^2/2}}{\sqrt{2\pi\lambda}} \left((1-\theta)^2 x^2 - 1 + O\left(\frac{1+|x|^3}{\sqrt{\lambda}}\right) \right).$$

This, together with the estimate (2.50), gives

$$\mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} = \frac{e^{-(1-\theta)x^2/2}}{\sqrt{2\pi\lambda}} \left(\frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}} - 1 \right) + O(\theta\lambda^{-1}).$$

Then we deduce that

$$\sup_{m \geq 0} \left| \mathbb{P}(S_n = m) - e^{-\lambda} \frac{\lambda^m}{m!} \right| = \frac{1}{\sqrt{2\pi\lambda}} \sup_{x \in \mathbb{R}} e^{-(1-\theta)x^2/2} \left| \frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}} - 1 \right| + O(\theta\lambda^{-1}).$$

Since the function $x \mapsto e^{-(1-\theta)x^2/2} \left(-1 + e^{-\theta x^2/2} / \sqrt{1-\theta} \right)$ reaches extrema at the points $x = 0$ and $x = \pm \sqrt{-(3/\theta) \log(1-\theta)}$. Thus

$$\sup_{x \in \mathbb{R}} e^{-(1-\theta)x^2/2} \left| \frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}} - 1 \right| = \max \left\{ \frac{1}{\sqrt{1-\theta}} - 1, \frac{\theta}{(1-\theta)^{3(1-\theta)/(2\theta)}} \right\} = \frac{1}{\sqrt{1-\theta}} - 1.$$

This proves (2.52) and completes the proof of the theorem. \square

For more results on both distances, see [6, 10, 26, 27, 32] and the references cited there.

3 Proof of Theorems 1.5 and 1.6

The proof of Theorems 1.5 and 1.6 relies on Theorem 2.1 and Proposition 2.7. We prove first Theorem 1.6 since its proof is shorter.

Proof of Theorem 1.6. We start with the asymptotic estimate of $d_{TV}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda))$ provided by Theorem 2.1. First of all, since $\theta = o(1)$, by (2.42),

$$\begin{aligned} \frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}} \left(1 + \frac{C_1 x + C_2 x^3}{\sigma} \right) - 1 &= -\frac{\theta}{2} \left(x^2 - \frac{x}{\sqrt{\lambda}} - 1 \right) + O(\theta\zeta(1+x^4)) \\ &= -\frac{\theta}{2} (x - \xi_+)(x - \xi_-) \left(1 + O\left(\frac{\zeta(1+x^4)}{|x - \xi_+||x - \xi_-|} \right) \right), \end{aligned}$$

uniformly for x bounded away from the two zeros $\xi_{\pm} := (1 \pm \sqrt{1+4\lambda})/(2\sqrt{\lambda})$ of the polynomial $x^2 - x/\sqrt{\lambda} - 1$; see (2.45). Obviously, the two zeros are bounded whenever $\lambda \geq \varepsilon$. Also, $\xi_+ > 0$, $\xi_- < 0$, and $\xi_+ - \xi_- > 2$. Thus, by applying the estimate $|1 + O(|t|)|^{\alpha} - 1 = O(|t|)$ for $|t| \leq 1$, we obtain

$$\begin{aligned} E_6 &:= \sum_{m \geq 0} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^{\alpha} \left(\left| \frac{e^{-\theta x^2/2}}{\sqrt{1-\theta}} \left(1 + \frac{C_1 x + C_2 x^3}{\sigma} \right) - 1 \right|^{\alpha} - \left| \frac{\theta}{2} \left(x^2 - \frac{x}{\sqrt{\lambda}} - 1 \right) \right|^{\alpha} \right) \\ &= O \left(\theta^{\alpha} \zeta \sum_{\substack{|x| \leq \zeta^{-1/2} \\ |x - \xi_{\pm}| > \zeta}} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^{\alpha} |(x - \xi_+)(x - \xi_-)|^{\alpha-1} (1 + x^4) + \theta^{\alpha} \zeta^{\alpha} \sum_{|x - \xi_{\pm}| \leq \zeta} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^{\alpha} \right), \end{aligned} \quad (3.1)$$

where $|x - \xi_{\pm}| > \zeta$ represents the two inequalities $|x - \xi_+| > \zeta$ and $|x - \xi_-| > \zeta$. For the first sum in the O -symbol, we use (2.3) and deduce that

$$\sum_{\substack{|x| \leq \zeta^{-1/2} \\ |x - \xi_{\pm}| > \zeta}} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^{\alpha} |(x - \xi_+)(x - \xi_-)|^{\alpha-1} (1 + x^4) = O \left(\lambda^{-\alpha/2} (\zeta^{\alpha-1} + \lambda^{1/2}) \right);$$

similarly, by the crude bound (2.31), we see that the second sum in the O -symbol of (3.1) satisfies

$$\sum_{|x - \xi_{\pm}| \leq \zeta} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right)^{\alpha} = O \left(\lambda^{-\alpha/2} (1 + \zeta \lambda^{1/2}) \right).$$

Thus

$$\begin{aligned} E_6 &= O \left(\theta^{\alpha} \zeta \lambda^{-\alpha/2} (\zeta^{\alpha-1} + \lambda^{1/2}) + \theta^{\alpha} \zeta^{\alpha} \lambda^{-\alpha/2} (1 + \zeta \lambda^{1/2}) \right) \\ &= O \left(\theta^{\alpha} \lambda^{(1-\alpha)/2} (\zeta + \zeta^{\alpha} \lambda^{-1/2}) \right). \end{aligned}$$

This completes the proof. \square

Proof of Theorem 1.5. Our method of proof is straightforward: we start from Theorem 2.1, approximate the sum in (2.1) over the central range $|m - \lambda| \leq \lambda^{3/5}$ by means of the LLT (2.18) of Poisson distribution and then apply the Euler-Maclaurin summation formula, the sum of terms over the remaining range of m being negligible. As most proofs we have seen so far, a more delicate error analysis is needed.

Dominant part. Estimating the tail of $W_{\alpha}(\lambda)$ by means of (2.3) and replacing the factor $e^{-\lambda} \lambda^m / m!$ by (2.18), we obtain,

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda)) = \frac{1}{2} \sum_{|m-\lambda| \leq \lambda^{3/5}} \left(\frac{\Psi(x)}{\sqrt{2\pi\lambda}} \right)^{\alpha} \left| e^{\theta\varphi(x)/2} - 1 \right|^{\alpha} + O \left(\theta^{\alpha} \lambda^{-(\alpha+1)/2} \right),$$

where φ is defined in (2.41) and $\Psi(x) := e^{-x^2(1-\theta)/2} (1 + \omega_1(x)/\sqrt{\lambda})$, with $\omega_1(x) := -\sqrt{1-\theta}(3x - (1-\theta)x^3)/6$.

Let m_+ and m_- denote the nearest integers to $\lambda + x_+ \sigma$ and to $\lambda + x_- \sigma$, respectively, where $\varphi(x_{\pm}) = 0$; see (2.41) and (2.45). Then for $m \neq m_{\pm}$, we have $|x - x_{\pm}| \geq 1/(2\sigma)$. Thus, letting $L_{\pm} := \lfloor \lambda \pm \lambda^{3/5} \rfloor$,

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda)) = \frac{1}{2(2\pi\lambda)^{\alpha/2}} \sum_{\substack{L_- < m \leq L_+ \\ m \neq m_{\pm}}} \Psi(x)^{\alpha} \left| e^{\theta\varphi(x)/2} - 1 \right|^{\alpha} + O \left(\theta^{\alpha} \lambda^{-\alpha} + \theta^{\alpha} \lambda^{-(\alpha+1)/2} \right),$$

because, by the relations $|x_{\pm} - (m_{\pm} - \lambda)/\sigma| \leq 1/(2\sigma)$ and $\varphi(x_{\pm}) = 0$,

$$\begin{aligned} \left| \exp\left(\frac{\theta}{2}\varphi\left(\frac{m_{\pm} - \lambda}{\sigma}\right)\right) - 1 \right|^{\alpha} &= O\left(\theta^{\alpha} \left|\varphi\left(\frac{m_{\pm} - \lambda}{\sigma}\right)\right|^{\alpha}\right) \\ &= O\left(\theta^{\alpha} \left|\varphi\left(\frac{m_{\pm} - \lambda}{\sigma}\right) - \varphi(x_{\pm})\right|^{\alpha}\right) \\ &= O\left(\theta^{\alpha} \lambda^{-\alpha/2}\right). \end{aligned}$$

Euler-Maclaurin formula. Let $s(t) := t - [t] - 1/2$ denote the sawtooth function, and $s_1(t) := \int_0^{\{t\}} s(v) dv$. We will use the Euler-Maclaurin formula

$$\sum_{a \leq m < b} f(m) = \int_a^b f(t) dt - \frac{f(b) - f(a)}{2} - \int_a^b f''(t) s_1(t) dt,$$

for any $f \in C^2[0, 1]$ and $a, b \in \mathbb{Z}$. Applying this formula to the function $f(t) = h((t - \lambda)/\sigma)$ with $a = L_- + 1$ and $b = L_+$, where $h(t) := \Psi(t)^{\alpha} |g(t)|^{\alpha}$ and $g(t) := e^{\theta\varphi(t)/2} - 1$, we have

$$\sum_{L_- < m \leq L_+} h\left(\frac{m - \lambda}{\sigma}\right) = \sigma \int_{(L_- + 1 - \lambda)/\sigma}^{(L_+ - \lambda)/\sigma} h(t) dt + O\left(e^{-\alpha\lambda^{1/5}/4} + \lambda^{-1/2} \int_{(L_- + 1 - \lambda)/\sigma}^{(L_+ - \lambda)/\sigma} |h''(t)| dt\right).$$

Error analysis. To estimate the integral in the O -term, we observe first that for t in the range $t \in (x_+, \lambda^{1/10})$, g satisfies, by (2.42), $g(t) = O(1)$, $g'(t) = O(\theta)$, $g''(t) = O(\theta^2)$. This implies that

$$h''(t) = O\left((1 + t^2)e^{-\alpha t^2/2} (|g(t)|^{\alpha} + |g(t)|^{\alpha-1}\theta + |\alpha - 1||g(t)|^{\alpha-2}\theta^2)\right),$$

for $t \in (x_+, \lambda^{1/10})$, which can be split into two ranges: $t \in (x_+, 2]$ and $t \in (2, \lambda^{1/10})$. In view of the estimate (2.43), we have $\theta t^2/(1 + \theta t^2) = O(|g(t)|) = O(\theta t^2)$ for $t \in (2, \lambda^{1/10})$. It follows that

$$\lambda^{-1/2} \int_2^{(L_+ - \lambda)/\sigma} |h''(t)| dt = O\left(\theta^{\alpha} \lambda^{-1/2}\right).$$

On the other hand, by (2.46), g satisfies $|g(t)| \asymp \theta|t - x_+|$, for $t \in (x_+, 2]$. Thus

$$\begin{aligned} &\lambda^{-1/2} \int_{(m_+ + 1 - \lambda)/\sigma}^2 |h''(t)| dt \\ &= O\left(\theta^{\alpha} \lambda^{-1/2} \int_{(m_+ + 1 - \lambda)/\sigma}^2 (|t - x_+|^{\alpha} + |t - x_+|^{\alpha-1} + |\alpha - 1||t - x_+|^{\alpha-2}) dt\right) \\ &= O\left(\theta^{\alpha} \lambda^{-1/2} + \theta^{\alpha} \lambda^{-1/2} |\alpha - 1| \int_{(m_+ + 1 - \lambda)/\sigma}^2 |t - x_+|^{\alpha-2} dt\right) \\ &= O\left(\theta^{\alpha} \lambda^{-1/2} + \theta^{\alpha} \lambda^{-\alpha/2}\right). \end{aligned}$$

The integral of $|h''(t)|$ over the remaining ranges $(L_- + 1 - \lambda)/\sigma < t \leq x_-$ and $x_- \leq t < x_+$ is estimated similarly and the same error bound holds. Collecting these estimates, we get

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(S_n), \text{Po}(\lambda)) = \frac{1 - \theta}{2(2\pi)^{\alpha/2}} \lambda^{(1-\alpha)/2} \int_{-\lambda^{1/10}}^{\lambda^{1/10}} h(t) dt + O\left(\theta^{\alpha} \lambda^{-(\alpha+1)/2} + \theta^{\alpha} \lambda^{-\alpha}\right).$$

The dominant integral. We now evaluate the integral in the above expression. For convenience, we write $e^{\theta\varphi(x)/2} = e^{\theta u(x)/2}(1 + \omega_2(x)/\sigma)$, where

$$u(x) = -x^2 - \frac{\log(1 - \theta)}{\theta} \quad \text{and} \quad \omega_2(x) = C_1x + C_2x^3.$$

Obviously, $u(-x) = u(x)$ and $\omega_2(-x) = -\omega_2(x)$. Thus, for x satisfying $|u(x)| \geq \lambda^{-1/2}$,

$$h(x) + h(-x) = \Psi(x)^\alpha \left| e^{\theta u(x)/2} - 1 \right|^\alpha \left(2 + O \left(|\alpha - 1| \left| \frac{\omega_2(x)}{\sigma(e^{\theta u(x)/2} - 1)} \right|^2 + \frac{1 + x^6}{\lambda} \right) \right).$$

Consequently,

$$\begin{aligned} \int_{-\lambda^{1/10}}^{\lambda^{1/10}} h(t) dt &= \int_{\substack{0 < t \leq \lambda^{1/10} \\ |u(t)| \geq \lambda^{-1/2}}} (h(t) + h(-t)) dt + O \left(\theta^\alpha \lambda^{-(\alpha+1)/2} \right) \\ &= 2 \int_{\substack{0 < t \leq \lambda^{1/10} \\ |u(t)| \geq \lambda^{-1/2}}} e^{-\alpha(1-\theta)x^2/2} |e^{\theta u(x)/2} - 1|^\alpha dt + O \left(\theta^\alpha \lambda^{-(\alpha+1)/2} + \theta^\alpha \lambda^{-1} \right) \\ &= \int_{-\infty}^{\infty} e^{-\alpha(1-\theta)x^2/2} |e^{\theta u(x)/2} - 1|^\alpha dt + O \left(\theta^\alpha \lambda^{-(\alpha+1)/2} + \theta^\alpha \lambda^{-1} \right), \end{aligned}$$

because the integrals involving odd integrands cancel out. □

Appendix. Elementary proofs for the binomial case

In the special case of binomial distribution, most proofs become much simpler. We sketch here the major steps.

Let as above $\delta_{n,m} := \binom{n}{m} p^m q^{n-m} - e^{-\lambda} \lambda^m / m!$ and $\sigma := \sqrt{pq\lambda}$.

Lemma 3.1. *Assume $p \leq 1/2$. Then*

$$\mathbb{P}(S_n = m) = e^{-\lambda} \frac{\lambda^m}{m!} \cdot \frac{e^{-px^2/2}}{\sqrt{q}} \left(1 + \frac{p(3x - px^3)}{6\sigma} + O_K \left(\frac{p^2}{\lambda} (1 + x^2 + p^2 x^6) \right) \right),$$

uniformly for $x := (m - \lambda)/\sigma = O(\lambda^{1/6} p^{-2/3})$ and $\lambda \geq K > 0$.

Proof. We may, without loss of generality, assume that $p = o(1)$, for otherwise, the local limit theorem follows from (2.7). Write first $\mathbb{P}(S_n = m) = \binom{n}{m} p^m q^{n-m} = e^{-\lambda} \lambda^m e^{G_{n,m}} / m!$, where

$$\begin{aligned} G_{n,m} &:= \frac{n! p^m q^{n-m}}{(n-m)! \lambda^m} e^\lambda \\ &= (n-m) \left(\log \left(1 + \frac{m-\lambda}{n-m} \right) - \frac{m-\lambda}{n-m} \right) - \frac{1}{2} \log \left(1 - \frac{m}{n} \right) + O \left(\frac{m}{n(n-m)} \right), \end{aligned} \quad (3.2)$$

by an application of Stirling's formula $N! = \sqrt{2\pi N} (N/e)^N e^{1/(12N) + O(1/N^2)}$. Since $m = \lambda + x\sigma$, we have

$$\frac{m-\lambda}{n-m} = \frac{px}{\sigma} \cdot \frac{1}{1-px/\sigma}, \quad n-m = \frac{\sigma^2}{p} \left(1 - \frac{px}{\sigma} \right)$$

and

$$\log\left(1 - \frac{m}{n}\right) = \log(1 - p) + \log\left(1 - \frac{px}{\sigma}\right).$$

Substituting these expressions into (3.2), we obtain

$$\left|G_{n,m} - \frac{\sigma^2}{p} f\left(\frac{px}{\sigma}\right) + \frac{1}{2} \log(1 - p) + \frac{1}{2} \log\left(1 - \frac{px}{\sigma}\right)\right| = O\left(\frac{p^2(1 + x/\sigma)}{\sigma^2(1 - px/\sigma)}\right),$$

for $m \leq n - 1$, where $f(t) := -t - (1 - t) \log(1 - t)$. This together with the inequality

$$\left|f(t) + \frac{t^2}{2} + \frac{t^3}{6}\right| \leq \frac{t^4}{12(1 - t)}$$

yields

$$\left|G_{n,m} - \frac{px}{2q\sqrt{\lambda}} + \frac{px^2}{2} + \frac{p^2x^3}{6q^2\sqrt{\lambda}} + \frac{1}{2} \log(1 - p)\right| = O\left(\frac{p^2}{\lambda} \left(1 + x^2 + \frac{|x|}{\sqrt{\lambda}} + px^4\right)\right).$$

This completes the proof. □

Proof of theorems 1.1 and 1.3. Assume $p \leq 1/2$ and take $R = m/\lambda$. Then

$$|\delta_{n,m}| \leq \frac{(1 + p(R - 1))^n + e^{\lambda(R-1)}}{R^m} \leq 2 \frac{e^{\lambda(R-1)}}{R^m} = 2e^{-\lambda \int_1^{m/\lambda} \log u \, du}. \quad (3.3)$$

Now we split the sum of $d_{\text{TV}}^{(\alpha)}(\mathcal{L}(X), \mathcal{L}(Y))$ into two parts

$$d_{\text{TV}}^{(\alpha)}(\mathcal{L}(X), \mathcal{L}(Y)) = \sum_{|x| \leq \lambda^{1/6} p^{-2/3}} |\delta_{n,m}|^\alpha + \sum_{|x| > \lambda^{1/6} p^{-2/3}} |\delta_{n,m}|^\alpha.$$

The first sum is evaluated by applying Lemma 3.1 and the second sum is estimated by inequality (3.3). These two estimates are enough to prove Theorem 2.1 when all p_j 's are equal. Once Theorem 2.1 is established, Theorems 1.1 and 1.3 will follow by exactly the same argument as we used above for Poisson-binomial distribution. □

References

- [1] D. ALDOUS, *Probability Approximations via the Poisson Clumping Heuristic*, Springer-Verlag, New York, 1989.
- [2] A. D. BARBOUR, *Asymptotic expansions in the Poisson limit theorem*, Ann. Probab., 15 (1987), pp. 748–766.
- [3] A. D. BARBOUR AND L. H.-Y. CHEN, *Stein's Method and Applications*, Singapore University Press and World Scientific Publishing Co., Singapore, 2005.
- [4] A. D. BARBOUR AND P. HALL, *On the rate of Poisson convergence*, Math. Proc. Cambridge Philos. Soc., 95 (1984), pp. 473–480.

- [5] A. D. BARBOUR, L. HOLST AND S. JANSON, *Poisson Approximation*, Oxford Science Publications, Clarendon Press, Oxford, 1992.
- [6] A. D. BARBOUR AND J. L. JENSEN, *Local and tail approximations near the Poisson limit*, Scand. J. Statist., 16 (1989), pp. 75–87.
- [7] V. ČEKANA VIČIUS AND J. KRUIOPIS, *Signed Poisson approximation: a possible alternative to normal and Poisson laws*, Bernoulli, 6 (2000), 591–606.
- [8] L. H. Y. CHEN, *Poisson approximation for dependent trials*, Ann. Probab., 3 (1975), pp. 534–545.
- [9] P. DEHEUVELS AND D. PFEIFER, *Operator semigroups and Poisson convergence in selected metrics*, Semigroup Forum, 34 (1986), pp. 203–224.
- [10] P. DEHEUVELS AND D. PFEIFER, *On a relationship between Uspensky’s theorem and Poisson approximation*, Ann. Inst. Statist. Math., 40 (1988), pp. 671–681.
- [11] P. DEHEUVELS, D. PFEIFER AND M. L. PURI, *A new semigroup technique in Poisson approximation. Semigroups and differential operators*, Semigroup Forum, 38 (1989), pp. 189–201.
- [12] P. FRANKEN, *Approximation des Verteilungen von Summen unabhängiger nichtnegativer ganzzahler Zufallsgrößen durch Poissonsche verteilungen*, Math. Nachr., 23 (1964), pp. 303–340.
- [13] H. HERRMANN, *Variationsabstand zwischen der Verteilung einer Summe unabhängiger nichtnegativer ganzzahliger Zufallsgrößen und Poissonschen Verteilungen*, Math. Nachr., 29 (1965), pp. 265–289.
- [14] H.-K. HWANG, *Asymptotic estimates of elementary probability distributions*, Stud. Appl. Math., 99 (1997), pp. 393–417.
- [15] H.-K. HWANG, *Asymptotics of Poisson approximation to random discrete distributions: an analytic approach*, Adv. in Appl. Probab., 31 (1999), pp. 448–491.
- [16] S. JANSON, *Coupling and Poisson approximation*, Acta Appl. Math., 34 (1994), pp. 7–15.
- [17] J. E. KENNEDY AND M. P. QUINE, *The total variation distance between the binomial and Poisson distributions*, Ann. Probab., 17 (1989), pp. 396–400.
- [18] J. KERSTAN, *Verallgemeinerung eines Satzes von Prochorow und Le Cam*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 2 (1964), pp. 173–179.
- [19] Y. KRUIOPIS, *The accuracy of approximation of the generalized binomial distribution by convolutions of Poisson measures*, Lithuanian Math. J., 26 (1986), pp. 37–49.
- [20] K. LANGE, *Applied Probability*, Springer Texts in Statistics. Springer-Verlag, New York, 2003.
- [21] L. LE CAM, *An approximation theorem for the Poisson binomial distribution*, Pacific J. Math., 10 (1960), pp. 1181–1197.
- [22] M. PENROSE, *Random Geometric Graphs*, Oxford University Press, Oxford, 2003.
- [23] S. D. POISSON, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile: précédés des règles générales du calcul des probabilités*, Bachelier, Paris, 1837.

- [24] YU. V. PROKHOROV, *Asymptotic behavior of the binomial distribution*, Uspekhi Matematicheskikh Nauk, 8 (1953), pp. 135–142. Also in *Selected Translations in Mathematical Statistics and Probability*, Volume 1, pp. 87–95.
- [25] S. T. RACHEV, *Probability Metrics and the Stability of Stochastic Models*, John Wiley & Sons, Chichester, 1991.
- [26] B. ROOS, *A semigroup approach to Poisson approximation with respect to the point metric*, Statistics and Probability Letters, 24 (1995), pp. 305–314.
- [27] B. ROOS, *Asymptotics and sharp bounds in the Poisson approximation to the Poisson-binomial distribution*, Bernoulli, 5 (1999), pp. 1021–1034.
- [28] B. ROOS, *On variational bounds in the compound Poisson approximation of the individual risk model*, Insurance Math. Econom., 40 (2007), pp. 403–414.
- [29] S. Y. SHORGIN, *Approximation of a generalized Binomial distribution*, Theory Probab. Appl., 22 (1977), pp. 846–850.
- [30] M. S. WATERMAN, *Introduction to Computational Biology: Maps, Sequences, and Genomes*, Chapman and Hill, London, 1995.
- [31] M. WEBA, *Bounds for the total variation distance between the binomial and the Poisson distribution in case of medium-sized success probabilities*, J. Appl. Probab., 36 (1999), pp. 97–104.
- [32] H.-J. WITTE, *A unification of some approaches to Poisson approximation*, J. Appl. Probab., 27 (1990), pp. 611–621.