Poisson type approximations for the Markov binomial distribution

Vydas Čekanavičius* and Bero Roos†
Vilnius University and University of Leicester

Abstract

The Markov binomial distribution is approximated by the Poisson distribution with the same mean, by a translated Poisson distribution and by two-parametric Poisson type signed measures. Using an adaptation of Le Cam's operator technique, estimates of accuracy are proved for the total variation, local, and Wasserstein norms. In a special case, asymptotically sharp constants are obtained. For some auxiliary results, we used Stein's method.

Keywords: Markov binomial distribution, Poisson approximation, translated Poisson distribution, signed compound Poisson measure, total variation norm, local norm, Wasserstein norm.

MSC 2000 Subject Classification: Primary 60J10; Secondary 60F05.

1 Introduction

The Markov binomial distribution plays an important role in probability theory. Frequently, it is approximated by the compound Poisson distribution, see, for example, [9, 10, 14, 18, 21, 30, 32, 33] and the references therein. For papers dealing with related problems, see, for example, [7, 12, 17, 31, 35]. On the other hand, Poisson approximation of the non-stationary Markov binomial distribution was not thoroughly investigated, see [29] and [9].

The purpose of this paper is the estimation of accuracy of Poisson approximation and various two-parametric Poisson type approximations to the Markov binomial distribution. In particular, we consider a second-order Poisson asymptotic expansion, a translated Poisson distribution and a signed compound Poisson measure. Note that, to some extent, the last two can be viewed as lattice counterparts of the normal distribution. Though we usually assume that certain transition probabilities are small, we allow them to be constants, thus

*Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 03225, Lithuania. E-mail: vydas.cekanavicius@mif.vu.lt

†Department of Mathematics, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom. E-mail: b.roos@leicester.ac.uk

including the case which is usually associated with the normal approximation. The estimates are obtained in the total variation, local, and Wasserstein norms. In a special case, we derive asymptotically sharp constants. The proofs are based on an adaptation of Le Cam's operator technique. For some auxiliary results, we used Stein's method.

We need the following notation. Let I_k denote the distribution concentrated at an integer $k \in \mathbb{Z}$ and set $I = I_0$. Throughout this paper, we use the abbreviation

$$U = I_1 - I. \tag{1}$$

In what follows, let V and W be two finite signed measures on $\mathbb{Z}$. Products and powers of V, W are understood in the convolution sense, i.e. $VW\{A\} = \sum_{k=-\infty}^{\infty} V\{A-k\}\, W\{k\}$ for a set $A \subseteq \mathbb{Z}$; further $W^0 = I$. Here and henceforth, we write $W\{k\}$ for $W\{\{k\}\}$, $(k \in \mathbb{Z})$. The total variation norm, the local norm, and the Wasserstein norm of W are denoted by

$$\|W\| = \sum_{k=-\infty}^{\infty} |W\{k\}|, \qquad \|W\|_\infty = \sup_{k\in\mathbb{Z}} |W\{k\}|, \qquad \|W\|_{\mathrm{Wass}} = \sum_{k=-\infty}^{\infty} |W\{(-\infty, k]\}|,$$

respectively. Using the simple equality

$$\|UW\|_{\mathrm{Wass}} = \|W\|, \tag{2}$$

it is possible to switch from the Wasserstein norm to the total variation norm. The logarithm and exponential of W are given by

$$\ln(I+W) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\, W^k \quad (\text{if } \|W\| < 1), \qquad \mathrm{e}^W = \exp\{W\} = \sum_{k=0}^{\infty} \frac{1}{k!}\, W^k.$$

In particular, $\mathrm{Pois}(\lambda) = \mathrm{e}^{\lambda U}$ is the Poisson distribution with parameter $\lambda \in [0,\infty)$. Note that

$$\|VW\|_\infty \leqslant \|V\|\|W\|_\infty, \qquad \|VW\| \leqslant \|V\|\|W\|, \qquad \|\mathrm{e}^W\| \leqslant \mathrm{e}^{\|W\|}.$$

Let $\widehat{W}(t)$ $(t \in \mathbb{R})$ be the Fourier transform of W. We denote by C positive absolute constants. The letter Θ stands for any finite signed measure on $\mathbb{Z}$ satisfying $\|\Theta\| \leqslant 1$. The values of C and Θ can vary from line to line, or even within the same line. For $x \in \mathbb{R}$ and $k \in \mathbb{N} = \{1, 2, 3, \dots\}$, we set

$$\binom{x}{k} = \frac{1}{k!}\, x(x-1)\dots(x-k+1), \qquad \binom{x}{0} = 1.$$

Let $\xi_0, \xi_1, \dots, \xi_n, \dots$ be a Markov chain with the initial distribution

$$\mathrm{P}(\xi_0 = 1) = p_0, \qquad \mathrm{P}(\xi_0 = 0) = 1 - p_0, \qquad p_0 \in [0,1]$$

and transition probabilities

$$\mathrm{P}(\xi_i = 1 \,|\, \xi_{i-1} = 1) = p, \qquad \mathrm{P}(\xi_i = 0 \,|\, \xi_{i-1} = 1) = q,$$
$$\mathrm{P}(\xi_i = 1 \,|\, \xi_{i-1} = 0) = \overline{q}, \qquad \mathrm{P}(\xi_i = 0 \,|\, \xi_{i-1} = 0) = \overline{p},$$
$$p + q = \overline{q} + \overline{p} = 1, \qquad p, \overline{q} \in (0,1), \qquad i \in \mathbb{N}.$$

The distribution of $S_n = \xi_1 + \cdots + \xi_n$ ($n \in \mathbb{N}$) is called the Markov binomial distribution. We denote it by F_n, that is $\mathrm{P}(S_n = m) = F_n\{m\}$ for $m \in \mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. We should note that the definition of the Markov binomial distribution slightly varies from paper to paper, see [14, 29, 32]. Sometimes ξ_0 is added to S_n or stationarity of the chain is assumed. For example, Dobrushin [14] assumed that $p_0 = 1$ and considered $S_{n-1} + 1$. However, if $p = \overline{q}$, then Dobrushin's Markov binomial distribution becomes a binomial distribution *shifted by unity*. This is not very natural, since we want the Markov binomial distribution to be a generalization of the binomial one. Therefore, we use the definition above which contains the binomial distribution as a special case. Moreover, it obviously allows the rewriting of our results for $S_{n-1} + 1$.

Further on, we need various characteristics of S_n. Let

$$\begin{aligned}
\nu_1 &= \frac{\overline{q}}{q+\overline{q}}, \qquad \nu_2 = \frac{2q\overline{q}\,(p-\overline{q})}{(q+\overline{q})^3}, \qquad \nu_3 = 6(\overline{q}-p)\,\frac{q\overline{q}\,(\overline{q}+q(\overline{q}-p))}{(q+\overline{q})^5},\\
A_0 &= \frac{|2qp - 3q\overline{q} - \overline{q}^2|}{2(q+\overline{q})^2}, \qquad A_1 = \frac{(\overline{q}-p)(\nu_1 - p_0)}{(q+\overline{q})},\\
A_2 &= \frac{\overline{q}-p}{(q+\overline{q})^2}\Big(\frac{\nu_1 q}{q+\overline{q}} + (\nu_1 - p_0)(p - 2\nu_1)\Big).
\end{aligned}$$

Note that $q + \overline{q} > 0$. From Lemma 4.4 below, it follows that

$$\begin{aligned}
\mathrm{E}S_n &= n\nu_1 + A_1 - A_1(p-\overline{q})^n,\\
\mathrm{Var}S_n &= n(\nu_2 + \nu_1 - \nu_1^2) + A_1 - A_1^2 + 2A_2\\
&\quad + (p-\overline{q})^n\Big[2nA_1\,\frac{\overline{q}-q}{q+\overline{q}} + A_1^2(2-(p-\overline{q})^n) - A_1 - 2A_2\Big].
\end{aligned}$$

2 Known results

It is known that a suitably normalized binomial distribution can have only two non-degenerate limit laws – the normal and the Poisson one. In contrast, S_n has seven different limit laws, see [14, Table 1]. However, as already noted above, the compound Poisson approximation dominates the field of research. Such a limit occurs, for example, when $n\overline{q} \to \lambda \in (0, \infty)$ and $p \to b \in (0, 1)$. Consequently, we cannot expect that Poisson approximation is good for $\overline{q} \ll p$. However, if p and $\overline{q}$ are of similar magnitude, we show that Poisson approximation can be sufficiently accurate, see Theorem 3.1 below.

The Markov binomial distribution is a generalization of the binomial one. Let us therefore recall the classical Poisson approximation bound for the binomial distribution. Let $\widetilde{p} \in (0, 1]$. Then

$$\|((1-\widetilde{p})I + \widetilde{p}I_1)^n - \mathrm{Pois}(n\widetilde{p})\| \leqslant 2\widetilde{p}\min(1, n\widetilde{p}) = 2n\widetilde{p}^2 \min\Big(1, \frac{1}{n\widetilde{p}}\Big), \tag{3}$$

see [4, formula (1.23), p. 8]. Speaking in terms of Barbour et al. [4, Introduction], the factor $(n\widetilde{p})^{-1}$ is the 'magic factor'. In fact, it often implies satisfactory accuracy but is difficult to obtain. It should be mentioned that the estimate $2n\widetilde{p}^2$ is principally due to Khintchine [20] and Doeblin [15] (see also [23, p. 1183]), whereas the bound $2\widetilde{p}$ can be called the Prokhorov type bound, since Prokhorov [24] was the first to get the estimate $C\widetilde{p}$.

Estimates similar to (3) also hold for the Markov binomial distribution. Let us consider the stationary Markov chain, that is, let $p_0 = \nu_1$. Then, as follows from the more general result [4, Theorem 8.H and Example 8.5.4], we have

$$\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\| \leqslant \frac{2}{q+\overline{q}}(1-\mathrm{e}^{-n\nu_1})\Big(\overline{q} + \frac{2q|p-\overline{q}|}{1-|p-\overline{q}|}\Big). \tag{4}$$

The right-hand side of (4) is of the order $O((p+\overline{q})\min(1,n\overline{q}))$, if p and $\overline{q}$ are bounded away from unity. Consequently, if p is close to $\overline{q}$, then the estimate (4) is a direct analogue of (3). However, the stationarity of the chain means that $p_0 = \nu_1$. In this paper, we investigate the case when p_0 is arbitrary.

Čekanavičius and Mikalauskas [9, Theorem 3.2] obtained an estimate which holds for any $p_0 \in [0,1]$ and which contains a magic factor. If the condition

$$p \leqslant \frac{1}{20}, \qquad \nu_1 \leqslant \frac{1}{30} \tag{5}$$

is satisfied, then

$$\|F_n - \mathrm{Pois}(n\nu_1)\| \leqslant C\,(p+\overline{q})\min(1,n\overline{q}) + C\,|p-\overline{q}|\min\Big(1,\frac{1}{\sqrt{n\overline{q}}}\Big). \tag{6}$$

In principle, here, the condition (5) can be dropped, since otherwise the right-hand side of (6) becomes greater than some absolute constant, whereas the left-hand side is in any case bounded by 2. On the other hand, for the results below, (5) is also assumed, where it is unclear whether it is superfluous. Note that, though condition (5) requires the smallness of p and $\overline{q}$, it allows for both parameters to be constants. In [8, Corollary 3] it was shown that, for $p_0 = 1$, $n\overline{q} \geqslant 1$ and $(p+\overline{q})^2 \leqslant |p-\overline{q}|$ the estimate (6) is of the right order. Estimate (6) is uniform over p_0 in a sense, that the right-hand side of (6) does not depend on p_0. However, this means that, for some values of p_0, the estimate (6) can be too rough. Indeed, let us consider the stationary case and $p = O(n^{-1/2})$ and $\overline{q} = O(n^{-2})$. Then (4) and (6) are of the order $O(n^{-3/2})$ and $O(n^{-1/2})$, respectively.

To the best of our knowledge, the second-order Poisson approximation to the Markov binomial distribution was not considered previously. In contrast, two-parametric signed compound Poisson measures were used. It should be mentioned that, for sums of independent random variables, such approximations of general order were investigated in numerous papers, see, for example, [2, 22, 28], and the references therein. In the present context, there is a result of Čekanavičius and Mikalauskas [9, formula (3.7)], which tells us that, if (5) is satisfied, then

$$\Big\|F_n - \exp\Big\{n\nu_1 U + n\frac{\nu_2-\nu_1^2}{2}U^2\Big\}\Big\| \leqslant C\,(p+\overline{q})^2\min\Big(n\overline{q},\frac{1}{\sqrt{n\overline{q}}}\Big) + C\,|p-\overline{q}|\min\Big(1,\frac{1}{\sqrt{n\overline{q}}}\Big). \tag{7}$$

In contrast to (6), the bound in (7) can be small if n is large without supposing the smallness of p and $\overline{q}$. On the other hand, the approximation in (7) is not a distribution but a signed measure, which might be less preferable in applications, see, for example, the discussion in [6, p. 1375].

There are other two-parametric Poisson type approximations, which differ from the second order asymptotic expansion and signed compound Poisson measure. Kruopis [22] proposed to use a suitably translated Poisson distribution. Such translated approximations are comparable to the normal distribution and can be accurate, when the standard Poisson approximation fails, see [2, 11, 25].

The choice of parameters for the translated Poisson approximating distribution is determined by the following considerations. Let us take the normal characteristic function $\exp\{\mu it - \sigma^2 t^2/2\}$ and replace $-t^2/2$ by $\mathrm{e}^{\mathrm{i}t} - 1 - \mathrm{i}t$. We get the characteristic function $\exp\{\mu it + \sigma^2(\mathrm{e}^{\mathrm{i}t} - 1 - \mathrm{i}t)\} = \exp\{(\mu - \sigma^2)\mathrm{i}t + \sigma^2(\mathrm{e}^{\mathrm{i}t} - 1)\}$ of a translated Poisson distribution. However, in view of the norms used in this paper, we need approximations concentrated on integers. Therefore, we translate the Poisson distribution by an integer quantity and add some fractional part to the Poisson parameter for compensation. For $\mu \in \mathbb{R}$ and $\sigma \in [0, \infty)$, set

$$\mathrm{TPois}(\mu, \sigma) = I_{\lfloor \mu - \sigma^2 \rfloor} \,\mathrm{Pois}(\sigma^2 + \delta).$$

Here $\lfloor \mu - \sigma^2 \rfloor$ and δ denote the integer and fractional parts of $\mu - \sigma^2$, respectively, i.e.

$$\mu - \sigma^2 = \lfloor \mu - \sigma^2 \rfloor + \delta, \qquad \delta \in [0, 1), \qquad \lfloor \mu - \sigma^2 \rfloor \in \mathbb{Z}. \tag{8}$$

As an example of translated Poisson approximation to the binomial distribution, we formulate an analogue of Theorem 2 from [22] for the total variation norm. Let $\widetilde{p} \in (0, 1/2]$ and $n\widetilde{p} \geqslant 1$. Then

$$\|((1 - \widetilde{p})I + \widetilde{p}I_1)^n - \mathrm{TPois}(n\widetilde{p}, \sqrt{n\widetilde{p}(1 - \widetilde{p})})\| \leqslant C\Big(\frac{\widetilde{p}}{\sqrt{n\widetilde{p}}} + \frac{1}{n\widetilde{p}}\Big), \tag{9}$$

see [2, Corollary 3.2 and discussion thereafter]. Note that recently Barbour and Lindvall [5] applied the translated Poisson approximation to Markov chains. However, for the Markov binomial distribution, their results apparently do not allow explicit estimates in terms of p and $\overline{q}$. It should be noted that Goldstein and Xia [19] introduced a new family of discrete distributions which includes translated Poisson distribution as a special case. It was shown that the members of the family can be used for approximation of the distribution of the sum of independent integer-valued random variables in total variation.

3 Results

The main goal of this paper is to investigate various second-order Poisson type approximations to the Markov binomial law containing magic factors. For this, we make use of the explicit structure of F_n and assume (5). For completeness of investigation we begin from a slight improvement of (6). In (6), the parameter of the approximating Poisson distribution was chosen as one of the parts of $\mathrm{E}S_n$, which grows when $n \to \infty$ and the remaining parameters are some absolute constants. The next result shows that Poisson approximation with exactly the same mean can improve the accuracy.

Theorem 3.1 *If* (5) *is satisfied, then*

$$\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\| \leqslant C\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1, \frac{1}{n\overline{q}}\Big), \tag{10}$$

$$\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\|_\infty \leqslant C\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1, \frac{1}{n\overline{q}\sqrt{n\overline{q}}}\Big), \tag{11}$$

$$\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\|_{\mathrm{Wass}} \leqslant C\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1, \frac{1}{\sqrt{n\overline{q}}}\Big). \tag{12}$$

Note that Theorem 3.1 remains valid if $\mathrm{Pois}(\mathrm{E}S_n)$ is replaced by $\mathrm{Pois}(n\nu_1 + A_1)$, see (47) below. For the stationary case, estimate (10) is of the order $O((p+\overline{q})\min(1, n\overline{q}))$. This is the same order of accuracy as in (4). It can be seen that, in view of the bounds, stationary and non-stationary cases can be different for small values of $\overline{q}$ only. Meanwhile, for the case $n\overline{q} \geqslant 1$, both estimates are of the order $O(p+\overline{q})$. We note that, if $n\overline{q} \geqslant 1$ then in (10) the assumption (5) can be dropped. Indeed, if (5) is not valid, then $p+\overline{q}$ is greater than some absolute constant, meanwhile the left-hand side of (10) is always less or equal to 2.

Due to the method of proof, the absolute constants in Theorem 3.1 are not given explicitly. However, in a special case, we can calculate asymptotically sharp constants.

Theorem 3.2 *Let condition* (5) *be satisfied and let* $n\overline{q} \geqslant 1$. *Then*

$$\Big|\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\| - \frac{4A_0}{\sqrt{2\pi \mathrm{e}}}\Big| \leqslant C\,(p+\overline{q})\Big(p+\overline{q}+\frac{1}{\sqrt{n\overline{q}}}\Big), \tag{13}$$

$$\Big|\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\|_\infty - \frac{A_0}{\sqrt{2\pi n\nu_1}}\Big| \leqslant C\,\frac{p+\overline{q}}{\sqrt{n\overline{q}}}\Big(p+\overline{q}+\frac{1}{\sqrt{n\overline{q}}}\Big), \tag{14}$$

$$\Big|\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\|_{\mathrm{Wass}} - \frac{2A_0\sqrt{n\nu_1}}{\sqrt{2\pi}}\Big| \leqslant C\,(p+\overline{q})\sqrt{n\overline{q}}\Big(p+\overline{q}+\frac{1}{\sqrt{n\overline{q}}}\Big). \tag{15}$$

As a consequence of (13), we note that, if $p+\overline{q} = O(|2p-3\overline{q}|)$, $p \to 0$, $\overline{q} \to 0$, and $n\overline{q} \to \infty$, then $\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\| \sim 2|2p-3\overline{q}|/\sqrt{2\pi\mathrm{e}}$. If, in addition $p = \overline{q}$, we have a Poisson approximation of the binomial distribution and (13) principally coincides with a result obtained by Prokhorov [24, Theorem 2]. The same applies for the local and Wasserstein norms, see, for example, [26, formula (32) and discussion thereafter].

The remaining results are devoted to two-parametric approximations. Here we expect better upper bounds, since, in contrast to the simpler Poisson approximation, we can match mean and variance of F_n. We begin with the second-order Poisson approximation. Recall that U is defined in (1). Let $M_0 = 2^{-1}(\mathrm{Var}S_n - \mathrm{E}S_n)U^2$.

Theorem 3.3 *If* (5) *is satisfied, then*

$$\begin{aligned}
&\|F_n - \mathrm{Pois}(\mathrm{E}S_n)(I+M_0)\| \\
&\qquad\leqslant C\,(p+\overline{q})\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1,\frac{1}{n\overline{q}}\Big), &(16)\\
&\|F_n - \mathrm{Pois}(\mathrm{E}S_n)(I+M_0)\|_\infty \\
&\qquad\leqslant C\,(p+\overline{q})\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1,\frac{1}{n\overline{q}\sqrt{n\overline{q}}}\Big), &(17)\\
&\|F_n - \mathrm{Pois}(\mathrm{E}S_n)(I+M_0)\|_{\mathrm{Wass}} \\
&\qquad\leqslant C\,(p+\overline{q})\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1,\frac{1}{\sqrt{n\overline{q}}}\Big). &(18)
\end{aligned}$$

Note that, in the case $n \geqslant 2$, Theorem 3.3 also holds, if we replace $\mathrm{Pois}(\mathrm{E}S_n)$ by $\mathrm{Pois}(n\nu_1 + A_1)$ and M_0 by $\widetilde{M_0} = 2^{-1}\left(n(\nu_2-\nu_1^2) - A_1^2 + 2A_2\right)U^2$, see (47) below.

Now, let us consider the translated Poisson approximation. Though it is possible to use $\mathrm{E}S_n$ and $\mathrm{Var}S_n$ as parameters, for simplicity, we shall drop the parts of the moments which are, at least, exponentially vanishing. Therefore, let

$$\mu = n\nu_1 + A_1, \qquad \sigma^2 = n(\nu_2 + \nu_1 - \nu_1^2) + A_1 - A_1^2 + 2A_2. \tag{19}$$

As shown in the following theorem, the translated Poisson approximation gives a bound similar to that of (9).

Theorem 3.4 *Let condition* (5) *be satisfied and let* $n\overline{q} \geqslant 1$. *Then*

$$\begin{aligned}
\|F_n - \mathrm{TPois}(\mu,\sigma)\| &\leqslant \frac{C}{\sqrt{n\overline{q}}}\Big(p+\overline{q}+\frac{1}{\sqrt{n\overline{q}}}\Big), &(20)\\
\|F_n - \mathrm{TPois}(\mu,\sigma)\|_\infty &\leqslant \frac{C}{n\overline{q}}\Big(p+\overline{q}+\frac{1}{\sqrt{n\overline{q}}}\Big), &(21)\\
\|F_n - \mathrm{TPois}(\mu,\sigma)\|_{\mathrm{Wass}} &\leqslant C\Big(p+\overline{q}+\frac{1}{\sqrt{n\overline{q}}}\Big). &(22)
\end{aligned}$$

If $\overline{q}$ is an absolute constant, then the estimate (20) is of order $O(n^{-1/2})$, which, in this case, is impossible for non-shifted Poisson approximation, see (10) and (16). If $p = \overline{q}$, then, up to constants, (20) coincides with (9).

Finally, we formulate a result for the signed compound Poisson measure. Let μ and σ^2 be defined as in (19) and set

$$\mathrm{SPois}(\mu,\sigma) = \exp\Big\{\mu U + \frac{\sigma^2-\mu}{2}\,U^2\Big\}.$$

Theorem 3.5 *Let condition* (5) *be satisfied and let* $n \geqslant 2$. *Then*

$$\begin{aligned}
\|F_n - \mathrm{SPois}(\mu,\sigma)\| &\leqslant C\,(p+\overline{q})\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1,\frac{1}{n\overline{q}\sqrt{n\overline{q}}}\Big), &(23)\\
\|F_n - \mathrm{SPois}(\mu,\sigma)\|_\infty &\leqslant C\,(p+\overline{q})\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1,\frac{1}{(n\overline{q})^2}\Big), &(24)\\
\|F_n - \mathrm{SPois}(\mu,\sigma)\|_{\mathrm{Wass}} &\leqslant C\,(p+\overline{q})\big(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q})\big)\min\Big(1,\frac{1}{n\overline{q}}\Big). &(25)
\end{aligned}$$

We note that direct calculations show that, for $n = 1$, (23)–(25) remain valid, if the right-hand sides are replaced by the larger value $C((p+\overline{q})^3 + (p-\overline{q})^2)$. It can be seen that, for the case $n\overline{q} \geqslant 1$, the upper bound in (23) has an additional multiplier $(n\overline{q})^{-1/2}$ in comparison with (16), which means an essential improvement in accuracy of approximation. Moreover, $\mathrm{SPois}(\mu, \sigma)$ is more accurate than the signed approximation used in (7). Indeed, in the stationary case $p_0 = \nu_1$, if $p = O(n^{-1/2})$ and $\overline{q} = O(n^{-2})$, then the upper bounds in (23) and (7) are of order $O(n^{-2})$ and $O(n^{-1/2})$, respectively. Finally, it is easy to check that, if $p = o(1)$, $\overline{q} = o(1)$, and $n\overline{q} \geqslant 1$, then $\mathrm{SPois}(\mu, \sigma)$ is more accurate than $\mathrm{TPois}(\mu, \sigma)$. The main advantage of $\mathrm{TPois}(\mu, \sigma)$ over $\mathrm{SPois}(\mu, \sigma)$ is the fact that it is a distribution and is simpler structured.

4 Auxiliary results

In the following two lemmas, $C(k)$ denotes an absolute positive constant depending on k.

Lemma 4.1 *Let $t \in (0, \infty)$ and $k \in \mathbb{Z}_+$. Then we have*

$$\|U^2 \mathrm{e}^{tU}\| \leqslant \frac{3}{t\mathrm{e}}, \qquad \|U^k \mathrm{e}^{tU}\| \leqslant \Big(\frac{2k}{t\mathrm{e}}\Big)^{k/2}, \qquad \|U^k \mathrm{e}^{tU}\|_\infty \leqslant \frac{C(k)}{t^{(k+1)/2}}.$$

The first inequality was proved in [27, Lemma 3]. The second bound follows from formula (3.8) in [13] and the properties of the total variation norm. Here and throughout this paper, we set $0^0 = 1$. The third relation is a simple consequence of the formula of inversion. Our asymptotically sharp results require the following lemma. Set

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}, \qquad \varphi_k(x) = \frac{\mathrm{d}^k}{\mathrm{d}x^k}\,\varphi_0(x) \qquad (k \in \mathbb{N},\ x \in \mathbb{R}),$$
$$\|\varphi_k\|_1 = \int_{\mathbb{R}} |\varphi_k(x)|\,\mathrm{d}x, \qquad \|\varphi_k\|_\infty = \sup_{x\in\mathbb{R}} |\varphi_k(x)| \qquad (k \in \mathbb{Z}_+).$$

Lemma 4.2 *Let $t \in (0, \infty)$ and $k \in \mathbb{Z}_+$. Then we have*

$$\begin{aligned}
\Big|\|U^k \mathrm{e}^{tU}\| - \frac{\|\varphi_k\|_1}{t^{k/2}}\Big| &\leqslant \frac{C(k)}{t^{(k+1)/2}},\\
\Big|\|U^k \mathrm{e}^{tU}\|_\infty - \frac{\|\varphi_k\|_\infty}{t^{(k+1)/2}}\Big| &\leqslant \frac{C(k)}{t^{k/2+1}},\\
\Big|\|U^k \mathrm{e}^{tU}\|_{\mathrm{Wass}} - \frac{\|\varphi_{k-1}\|_1}{t^{(k-1)/2}}\Big| &\leqslant \frac{C(k)}{t^{k/2}} \qquad (k \neq 0).
\end{aligned}$$

The proof trivially follows from the more general Proposition 4 of [26] together with (2). The next lemma is devoted to some properties of the characteristic function of F_n.

Lemma 4.3 *Let* (5) *be satisfied. Then*

$$\widehat{F}_n(t) = \widehat{\Lambda}_1^n(t)\,\widehat{W}_1(t) + \widehat{\Lambda}_2^n(t)\,\widehat{W}_2(t), \tag{26}$$

where

$$
\begin{aligned}
\widehat{\Lambda}_{1,2}(t) &= \frac{p\mathrm{e}^{\mathrm{i}t}+\overline{p}\pm\widehat{D}^{1/2}(t)}{2}, & (27)\\
\widehat{W}_{1,2}(t) &= \frac{p_0}{2}\Big(1\pm\frac{q+\overline{q}+p(\mathrm{e}^{\mathrm{i}t}-1)}{\widehat{D}^{1/2}(t)}\Big)+\frac{1-p_0}{2}\Big(1\pm\frac{q+\overline{q}+(2\overline{q}-p)(\mathrm{e}^{\mathrm{i}t}-1)}{\widehat{D}^{1/2}(t)}\Big), & (28)\\
\widehat{D}(t) &= (p\mathrm{e}^{\mathrm{i}t}+\overline{p})^2+4\mathrm{e}^{\mathrm{i}t}(\overline{q}-p)\\
&= (1+\overline{q}+2\nu_1(\mathrm{e}^{\mathrm{i}t}-1)-p\mathrm{e}^{\mathrm{i}t})^2\Big(1+\frac{4\nu_1(p-\nu_1)(\mathrm{e}^{\mathrm{i}t}-1)^2}{(1+\overline{q}+2\nu_1(\mathrm{e}^{\mathrm{i}t}-1)-p\mathrm{e}^{\mathrm{i}t})^2}\Big). & (29)
\end{aligned}
$$

Here, for $\widehat{\Lambda}_1$ and $\widehat{W}_1$, we use the sign '+', and, for $\widehat{\Lambda}_2$ and $\widehat{W}_2$, the sign '−'.

Proof. Expression (26) was already used in [9]. However, the comment on its derivation was very short. Therefore, for the sake of completeness, we give a more detailed explanation on how (26)–(29) are obtained. Using the standard matrix product, we obtain

$$
\widehat{F}_n(t)=(p_0,1-p_0)\,\widetilde{P}^n(t)\begin{pmatrix}1\\1\end{pmatrix},\qquad \widetilde{P}(t)=\begin{pmatrix}p\mathrm{e}^{\mathrm{i}t} & q\\ \overline{q}\mathrm{e}^{\mathrm{i}t} & \overline{p}\end{pmatrix},
$$

see [16] or, for example, [34]. Now, we can apply the standard spectral decomposition of matrices. Under condition (5), we have two different eigenvalues $\widehat{\Lambda}_{1,2}(t)$ of $\widetilde{P}(t)$. In fact, for $j=1,2$,

$$
\widetilde{P}(t)\,\vec{x}_j=\widehat{\Lambda}_j(t)\,\vec{x}_j,\qquad \vec{y}_j^{\,T}\,\widetilde{P}(t)=\widehat{\Lambda}_j(t)\,\vec{y}_j^{\,T},\qquad \vec{y}_j^{\,T}\,\vec{x}_j=1,
$$

where T stands for transposition and

$$
\begin{aligned}
\vec{x}_j^{\,T} &= (x_{j1},x_{j2})=(q(\widehat{\Lambda}_j(t)-\overline{p}),\,q\overline{q}\mathrm{e}^{\mathrm{i}t}),\\
\vec{y}_j^{\,T} &= (y_{j1},y_{j2})=\Big(\frac{\widehat{\Lambda}_j(t)-\overline{p}}{q^2\overline{q}\mathrm{e}^{\mathrm{i}t}+q(\widehat{\Lambda}_j(t)-\overline{p})^2},\,\frac{1}{q\overline{q}\mathrm{e}^{\mathrm{i}t}+(\widehat{\Lambda}_j(t)-\overline{p})^2}\Big).
\end{aligned}
$$

It is now easy to check that $\widetilde{P}^n(t)=\widehat{\Lambda}_1^n(t)\,\vec{x}_1\,\vec{y}_1^{\,T}+\widehat{\Lambda}_2^n(t)\,\vec{x}_2\,\vec{y}_2^{\,T}$ and

$$
\begin{aligned}
\widehat{W}_j(t) &= p_0x_{j1}(y_{j1}+y_{j2})+(1-p_0)x_{j2}(y_{j1}+y_{j2})\\
&= p_0\,\frac{(\widehat{\Lambda}_j(t)-\overline{p})(\widehat{\Lambda}_j(t)-\overline{p}+q)}{q\overline{q}\mathrm{e}^{\mathrm{i}t}+(\widehat{\Lambda}_j(t)-\overline{p})^2}+(1-p_0)\frac{\overline{q}\mathrm{e}^{\mathrm{i}t}(\widehat{\Lambda}_j(t)-\overline{p}+q)}{q\overline{q}\mathrm{e}^{\mathrm{i}t}+(\widehat{\Lambda}_j(t)-\overline{p})^2}.
\end{aligned}
$$

The proof is easily completed. □

Due to Lemma 4.3, F_n can be decomposed into several signed measures. Let condition (5) be satisfied. Then $F_n = \Lambda_1^n W_1 + \Lambda_2^n W_2$, where

$$\begin{aligned}
W_{1,2} &= \frac{1}{2}\Big\{I \pm [(q+\overline{q})I + pU]H\sum_{j=0}^{\infty}\binom{-1/2}{j}L^j\Big\} \pm (1-p_0)(\overline{q}-p)UH\sum_{j=0}^{\infty}\binom{-1/2}{j}L^j,\\
\Lambda_1 &= I + \nu_1 U + \nu_1(p-\nu_1)U^2 H + 8\nu_1^2(p-\nu_1)^2 U^4 B H^4 \sum_{j=2}^{\infty}\binom{1/2}{j}L^{j-2},\\
\Lambda_2 &= (\nu_1-\overline{q})I + (p-\nu_1)I_1 - 2\nu_1(p-\nu_1)U^2 H\sum_{j=1}^{\infty}\binom{1/2}{j}L^{j-1},\\
H &= \frac{1}{q+\overline{q}}\sum_{j=0}^{\infty}\Big(\frac{p-2\nu_1}{q+\overline{q}}\Big)^j U^j = \frac{1}{1+\overline{q}-2\nu_1}\sum_{j=0}^{\infty}\Big(\frac{p-2\nu_1}{1+\overline{q}-2\nu_1}\Big)^j I_j,\\
B &= (1+\overline{q})I + 2\nu_1 U - pI_1, \quad L = 4\nu_1(p-\nu_1)U^2H^2.
\end{aligned}$$

The following lemma is the main tool in the proofs.

Lemma 4.4 *Let condition* (5) *be satisfied. Then*

$$\Lambda_1 = I + \nu_1 U + \nu_2 U^2\Theta, \tag{30}$$

$$\Lambda_1 = I + \nu_1 U + \frac{\nu_2}{2}U^2 + \frac{\nu_3}{6}U^3 + C\,\overline{q}|p-\overline{q}|(p+\overline{q})^2U^4\Theta, \tag{31}$$

$$\ln\Lambda_1 = \nu_1 U + \frac{\nu_2-\nu_1^2}{2}U^2 + \frac{\nu_3 - 3\nu_1\nu_2 + 2\nu_1^3}{6}U^3 + C\,\overline{q}(p+\overline{q})^3U^4\Theta, \tag{32}$$

$$\Lambda_2 = 2|p-\overline{q}|\Theta, \qquad \Lambda_2^n = C(b)\,|p-\overline{q}|^b\mathrm{e}^{-C(b)n}\Theta \quad (\textit{if } n\geqslant b\geqslant 0), \tag{33}$$

$$W_1 = I + A_1U + A_2U^2 + C\,(p+\overline{q})|p-\overline{q}|(p_0p+\overline{q})U^3\Theta, \quad W_1 = I + \frac{1}{2}\Theta, \tag{34}$$

$$\ln W_1 = A_1U + \frac{2A_2 - A_1^2}{2}U^2 + C\,(p+\overline{q})|p-\overline{q}|(p_0p+\overline{q})U^3\Theta, \tag{35}$$

$$W_2 = C\,|p-\overline{q}|(\overline{q}+|\nu_1-p_0|)U\Theta. \tag{36}$$

For any finite signed measure W on $\mathbb{Z}$ and any $t\in(0,\infty)$, we have

$$\|W\mathrm{e}^{t\ln\Lambda_1}\| \leqslant C\,\|W\mathrm{e}^{0.1t\nu_1U}\|. \tag{37}$$

Estimate (37) *also holds if the total variation norm on both sides is replaced by the local one.*

Proof. Some of the estimates improve the ones obtained in [9]. Condition (5) implies that

$$|p-\nu_1|\leqslant\frac{1}{20}, \quad |p-2\nu_1|\leqslant\frac{1}{15}, \quad \frac{1}{q+\overline{q}}\leqslant\frac{20}{19}, \qquad \overline{q}\leqslant\frac{1}{29}. \tag{38}$$

Using straight forward calculus, it is shown that

$$\|H\|\leqslant\frac{1}{1+\overline{q}-2\nu_1}\sum_{j=0}^{\infty}\Big(\frac{|p-2\nu_1|}{1+\overline{q}-2\nu_1}\Big)^j\leqslant\frac{15}{13}, \quad H = \frac{1}{q+\overline{q}}I + 0.172\Theta, \tag{39}$$

$$\|U\| = 2, \quad \|L\|\leqslant 0.04, \quad \|B\| = 1+\overline{q}-2\nu_1+|2\nu_1-p|\leqslant 1+|2\nu_1-p|\leqslant\frac{16}{15} \tag{40}$$

and

$$H=\frac{1}{q+\overline{q}}I+\frac{p-2\nu_1}{(q+\overline{q})^2}U+\frac{(p-2\nu_1)^2}{(q+\overline{q})^3}U^2+C(p+\overline{q})^3U^3\Theta. \tag{41}$$

Taking into account (38), (39), (40) and (41) it is easy to obtain (30), (31), (33) and the second equality of (34). For example, the proof of (30) follows from

$$\begin{aligned}\Lambda_1 &= I+\nu_1U+\nu_1U^2\Big((p-\nu_1)H+8\nu_1(p-\nu_1)^2U^2BH^4\sum_{j=2}^{\infty}\binom{1/2}{j}L^{j-2}\Big)\\ &= I+\nu_1U+\nu_1U^2\bigg(\frac{1}{20}\cdot\frac{15}{13}+\frac{8}{30}\cdot\frac{4}{20^2}\cdot\frac{16}{15}\cdot\Big(\frac{15}{13}\Big)^4\cdot\frac{1}{2}\sum_{j=2}^{\infty}0.04^{j-2}\bigg)\Theta.\end{aligned}$$

For the first expansion in (34) note that

$$\sum_{j=0}^{\infty}\binom{-1/2}{j}L^j=I-\frac{1}{2}L+C\overline{q}^2|p-\overline{q}|^2U^4\Theta=I+C\overline{q}|p-\overline{q}|U^2\Theta.$$

Consequently,

$$\begin{aligned}W_1 &= \frac{1}{2}I+\frac{1}{2}(q+\overline{q})H-\frac{1}{4}(q+\overline{q})HL+\frac{1}{2}pUH\\ &\quad+(1-p_0)(\overline{q}-p)UH+C\overline{q}|p-\overline{q}|(p+\overline{q})U^3\Theta.\end{aligned}$$

Moreover,

$$\begin{aligned}&\frac{1}{2}(q+\overline{q})H+\frac{1}{2}[p+2(1-p_0)(\overline{q}-p)]UH\\ &\quad= \frac{1}{2}I+\frac{1}{2}\Big(I+\frac{p+2(1-p_0)(\overline{q}-p)}{p-2\nu_1}\Big)\sum_{j=1}^{\infty}\Big(\frac{p-2\nu_1}{q+\overline{q}}\Big)^jU^j\\ &\quad= \frac{1}{2}I+\frac{(p_0-\nu_1)(p-\overline{q})}{p-2\nu_1}\sum_{j=1}^{\infty}\Big(\frac{p-2\nu_1}{q+\overline{q}}\Big)^jU^j=\frac{1}{2}I+(p_0-\nu_1)(p-\overline{q})UH.\end{aligned}$$

Now for (34) it suffices to use (41). The estimate (36) follows from (34) and relation $W_1+W_2=I$. Taking into account (30), we get

$$\Lambda_1-I=\frac{3}{2}\nu_1U\Theta=3\nu_1\Theta=\frac{1}{10}\Theta$$

and hence

$$\sum_{j=4}^{\infty}\frac{(-1)^{j+1}}{j}(\Lambda_1-I)^j=\Big(\frac{3}{2}\Big)^4\sum_{j=4}^{\infty}\Big(\frac{1}{10}\Big)^{j-4}\nu_1^4U^4\Theta=C\,\overline{q}^4U^4\Theta.$$

The proof of (32) now follows from the definition of $\ln\Lambda_1$ and (31). The proof of (35), is very similar to the proof of (32) and is, therefore, omitted. Estimate (37) is shown by applying

$$\begin{aligned}\ln\Lambda_1 &= \Lambda_1-I+(\Lambda_1-I)^2\sum_{j=2}^{\infty}\frac{(-1)^{j+1}}{j}(\Lambda_1-I)^{j-2}\\ &= \nu_1U+\frac{\nu_1}{4}U^2\Theta+\frac{3^2}{2^2}\sum_{j=2}^{\infty}\frac{1}{j}\Big(\frac{1}{10}\Big)^{j-2}\nu_1^2U^2\Theta=\nu_1U+\frac{7}{24}\nu_1U^2\Theta.\end{aligned}$$

In fact, this gives

$$\|W\mathrm{e}^{t\ln\Lambda_1}\| \leqslant \Big\|\exp\Big\{0.9t\nu_1U+\frac{7}{24}t\nu_1U^2\Theta\Big\}\Big\|\|W\mathrm{e}^{0.1t\nu_1U}\|.$$

In view of Lemma 4.1, we see that

$$\begin{aligned}\Big\|\exp\Big\{0.9t\nu_1U+\frac{7}{24}t\nu_1U^2\Theta\Big\}\Big\| &\leqslant 1+\sum_{r=1}^{\infty}\frac{1}{r!}\Big\|\frac{7}{24}t\nu_1U^2\exp\Big\{\frac{0.9t\nu_1}{r}U\Big\}\Big\|^r\\ &\leqslant 1+\sum_{r=1}^{\infty}\frac{\mathrm{e}^r}{r^r\sqrt{2\pi r}}\Big(\frac{21r}{24\cdot 0.9\mathrm{e}}\Big)^r\leqslant C, \qquad (42)\end{aligned}$$

which implies (37). The proof is completed. □

Note that Lemma 4.4 immediately leads to shorter expressions for Λ_1 and W_1, e.g.

$$\Lambda_1 = I+\nu_1U+\frac{\nu_2}{2}U^2+C\,(p+\overline{q})\overline{q}|p-\overline{q}|U^3\Theta.$$

Further, in the presence of condition (5), we can use Lemma 4.4 to expand $\widehat{F}_n(t)$ in powers of $(\mathrm{i}t)$, which leads to the exact expressions of $\mathrm{E}S_n$ and $\mathrm{Var}S_n$ as given in the introduction. However, it is easily seen that these formulas remain valid, if condition (5) is dropped.

The next results are needed for the estimation of the closeness of $\mathrm{TPois}(\mu,\sigma)$ and $\mathrm{SPois}(\mu,\sigma)$. We use Stein's method for the proof. In the remaining part of this section, we assume that condition (5) is satisfied and $n\overline{q}\geqslant 1$. Further, let μ and σ be defined as in (19), $a=\lfloor\mu-\sigma^2\rfloor$, and let δ be the fractional part of $\mu-\sigma^2$, see (8). It is not difficult to check that

$$n\nu_1\geqslant\frac{1}{2},\quad |A_1|\leqslant\frac{1}{19},\quad |A_2|\leqslant\frac{|p-\overline{q}|}{(q+\overline{q})^3}(3\overline{q}+2q|p-\overline{q}|)\leqslant\frac{20}{19^2}(3\nu_1+2|p-\overline{q}|)\leqslant\frac{4}{19^2},$$

and hence

$$\mu\geqslant n\nu_1-|A_1|\geqslant\frac{n\overline{q}}{4},\qquad \sigma^2\geqslant\mu-|\sigma^2-\mu|\geqslant\frac{3}{4}\,\mu\geqslant\frac{n\overline{q}}{6},$$

since

$$|\sigma^2-\mu|\leqslant n|\nu_2-\nu_1^2|+A_1^2+2|A_2|\leqslant\mu\,\frac{n\nu_1}{\mu}\Big(\frac{|\nu_2|}{\nu_1}+\nu_1+\frac{A_1^2+2|A_2|}{n\nu_1}\Big)\leqslant\frac{\mu}{4}. \qquad (43)$$

Let $g:\mathbb{Z}\to\mathbb{R}$ be a bounded function and set

$$\Delta g(j)=g(j+1)-g(j)\quad(j\in\mathbb{Z}),\qquad \|g\|_\infty=\sup_{j\in\mathbb{Z}}|g(j)|,\qquad \|g\|=\sum_{j\in\mathbb{Z}}|g(j)|.$$

Further, let

$$\lambda_1=2\mu-\sigma^2,\qquad \lambda_2=\frac{\sigma^2-\mu}{2},\qquad \mathrm{SP}(g)=\sum_{j\in\mathbb{Z}}g(j)\mathrm{SPois}(\mu,\sigma)\{j\},$$
$$(\mathcal{A}g)(j)=2\lambda_2g(j+2)+\lambda_1g(j+1)-jg(j),\quad j\in\mathbb{Z}.$$

Lemma 4.5 *Let $f:\mathbb{Z}\to\mathbb{R}$ satisfy one of the following conditions: $\|f\|_\infty\leqslant 1$, $\|f\|\leqslant 1$, or $\|\Delta f\|_\infty\leqslant 1$. Then there exists a bounded function $g:\mathbb{Z}\to\mathbb{R}$ such that, for $j\in\mathbb{Z}$,*

$$g(j)=0\quad (j\leqslant 0),\qquad (\mathcal{A}g)(j)=f(j)-\mathrm{SP}(f)\quad (j\geqslant 0).$$

Moreover, if $\|f\|_\infty\leqslant 1$, then $\|\Delta g\|_\infty\leqslant C\mu^{-1}$. If $\|f\|\leqslant 1$, then $\|\Delta g\|\leqslant C\mu^{-1}$. If $\|\Delta f\|_\infty\leqslant 1$, then $\|\Delta g\|_\infty\leqslant C\mu^{-1/2}$.

Proof. Due to (43), we have

$$\gamma:=\frac{4|\lambda_2|}{2\lambda_2+\lambda_1}=\frac{2|\sigma^2-\mu|}{\mu}\leqslant\frac{1}{2}.$$

Now, the statement of the lemma follows from Theorem 2.1 and Example 3.3 in [3]. □

Let X be a $\mathrm{Pois}(\sigma^2+\delta)$ distributed random variable and set $Z=a+X$. Then Z has distribution $\mathrm{TPois}(\mu,\sigma)$. Further, we have $\mathrm{SPois}(\mu,\sigma)=\exp\{\lambda_1U+\lambda_2(I_2-I)\}$.

Lemma 4.6 *Let f and g be defined as in Lemma 4.5. If, for some $\varepsilon_1=\varepsilon_1(n,p,\overline{q},p_0)>0$, $|\mathrm{E}(\mathcal{A}g)(Z)|\leqslant\varepsilon_1\|\Delta g\|_\infty$, then*

$$\begin{aligned}\|\mathrm{TPois}(\mu,\sigma)-\mathrm{SPois}(\mu,\sigma)\| &\leqslant C\,(\varepsilon_1\mu^{-1}+\mathrm{e}^{-Cn\overline{q}}),\\ \|\mathrm{TPois}(\mu,\sigma)-\mathrm{SPois}(\mu,\sigma)\|_{\mathrm{Wass}} &\leqslant C\,(\varepsilon_1\mu^{-1/2}+\mathrm{e}^{-Cn\overline{q}}).\end{aligned}$$

If, for some $\varepsilon_2=\varepsilon_2(n,p,\overline{q},p_0)>0$, $|\mathrm{E}(\mathcal{A}g)(Z)|\leqslant\varepsilon_2\|\Delta g\|$, then

$$\|\mathrm{TPois}(\mu,\sigma)-\mathrm{SPois}(\mu,\sigma)\|_\infty\leqslant C\,(\varepsilon_2\mu^{-1}+\mathrm{e}^{-Cn\overline{q}}).$$

Proof. If $a\geqslant 0$, then we have $|\mathrm{E}f(Z)-\mathrm{SP}(f)|=|\mathrm{E}(\mathcal{A}g)(Z)|$, and the statement of Lemma 4.6 follows directly from the definition of the norms, see, for example, [4, Appendix A1]. Now, let $a<0$. Applying (43) and Bernstein's inequality (see [1, Theorem 1.4.1 and comment on p. 37]), we obtain

$$\mathrm{P}(Z\leqslant 0)=\mathrm{P}(X-\mathrm{E}X\leqslant-\mu)\leqslant\exp\Big\{-\frac{\mu^2}{4(\sigma^2+\delta)}\Big\}\leqslant\mathrm{e}^{-\mu/14}\leqslant\mathrm{e}^{-Cn\overline{q}}.\tag{44}$$

Let $\widetilde{Z}$ be a random variable on $\mathbb{Z}_+$ with $\mathrm{P}(\widetilde{Z}=0)=\mathrm{P}(Z\leqslant 0)$ and $\mathrm{P}(\widetilde{Z}=j)=\mathrm{P}(Z=j)$ for $j\in\mathbb{N}$. Without loss of generality, we can assume that $f(j)=0$ for $j<a$. Due to the assumption made on f, for $j\geqslant a$, we then have $|f(j)|\leqslant j+|a|+1$. Now, we obtain

$$\begin{aligned}|\mathrm{E}f(Z)-\mathrm{E}f(\widetilde{Z})| &\leqslant \sum_{j=a}^{-1}|f(j)|\mathrm{P}(Z=j)+|f(0)|\mathrm{P}(Z<0)\\ &\leqslant C\,|a|\mathrm{P}(Z\leqslant 0)\leqslant C\,n\overline{q}\mathrm{e}^{-Cn\overline{q}}\leqslant C\,\mathrm{e}^{-Cn\overline{q}}\end{aligned}\tag{45}$$

and

$$\begin{aligned}|\mathrm{E}f(\widetilde{Z})-\mathrm{SP}(f)| &= \Big|\sum_{j=0}^{\infty}\big(f(j)-\mathrm{SP}(f)\big)\mathrm{P}(\widetilde{Z}=j)\Big|=\Big|\sum_{j=0}^{\infty}(\mathcal{A}g)(j)\mathrm{P}(\widetilde{Z}=j)\Big|\\ &\leqslant |\mathrm{E}(\mathcal{A}g)(Z)|+|(\mathcal{A}g)(0)|\mathrm{P}(Z\leqslant 0)+\Big|\sum_{j=a}^{0}(\mathcal{A}g)(j)\mathrm{P}(Z=j)\Big|\\ &\leqslant |\mathrm{E}(\mathcal{A}g)(Z)|+C\,(|(\mathcal{A}g)(0)|+|(\mathcal{A}g)(-1)|)\mathrm{P}(Z\leqslant 0).\end{aligned}\tag{46}$$

Taking into account Lemma 4.5, we obtain

$$|(\mathcal{A}g)(0)| \leqslant (\lambda_1 + 2|\lambda_2|)|\Delta g(0)| + 2|\lambda_2||\Delta g(1)| \leqslant C\,\mu\|\Delta g\|_\infty \leqslant C\,n\overline{q}.$$

Similarly, $|(\mathcal{A}g)(-1)| \leqslant C\,n\overline{q}$. Combining the last two estimates with (45), (46), and (44), the asserted inequalities are easily proved. □

Lemma 4.7 *Let μ and σ be given by* (19)*, δ be as in* (8)*, $n\overline{q} \geqslant 1$, and let condition* (5) *be satisfied. Then*

$$\begin{aligned}
\|\mathrm{TPois}(\mu,\sigma) - \mathrm{SPois}(\mu,\sigma)\| &\leqslant \frac{C}{\sqrt{n\overline{q}}}\Big(p + \overline{q} + \frac{1}{\sqrt{n\overline{q}}}\Big),\\
\|\mathrm{TPois}(\mu,\sigma) - \mathrm{SPois}(\mu,\sigma)\|_\infty &\leqslant \frac{C}{n\overline{q}}\Big(p + \overline{q} + \frac{1}{\sqrt{n\overline{q}}}\Big),\\
\|\mathrm{TPois}(\mu,\sigma) - \mathrm{SPois}(\mu,\sigma)\|_{\mathrm{Wass}} &\leqslant C\Big(p + \overline{q} + \frac{1}{\sqrt{n\overline{q}}}\Big).
\end{aligned}$$

Proof. Taking into account that $\mathrm{E}(Z-a)g(Z) = (\sigma^2+\delta)\mathrm{E}g(Z+1)$ and applying Lemma 4.1, we obtain

$$\begin{aligned}
|\mathrm{E}(\mathcal{A}g)(Z)| &= |(\sigma^2 - \mu)\mathrm{E}\Delta^2 g(Z) - \delta\mathrm{E}\Delta g(Z)|\\
&\leqslant |\sigma^2 - \mu|\|\Delta g\|_\infty\|U\mathrm{e}^{(\sigma^2+\delta)U}\| + \delta\|\Delta g\|_\infty\\
&\leqslant C\,\|\Delta g\|_\infty\Big(\frac{|\sigma^2 - \mu|}{\sqrt{\sigma^2+\delta}} + \delta\Big) \leqslant C\,\|\Delta g\|_\infty(\sqrt{n\overline{q}}(p+\overline{q}) + \delta),\\
|\mathrm{E}(\mathcal{A}g)(Z)| &\leqslant \|\Delta g\|\Big(|\sigma^2 - \mu|\|U\mathrm{e}^{(\sigma^2+\delta)U}\|_\infty + \delta\|\mathrm{e}^{(\sigma^2+\delta)U}\|_\infty\Big)\\
&\leqslant C\,\|\Delta g\|\Big(\frac{|\sigma^2-\mu|}{\sigma^2+\delta} + \frac{\delta}{\sqrt{\sigma^2+\delta}}\Big) \leqslant C\,\|\Delta g\|\Big(p + \overline{q} + \frac{\delta}{\sqrt{n\overline{q}}}\Big).
\end{aligned}$$

Now it suffices to use Lemma 4.6. □

5 Proofs of the theorems

For the proofs of the theorems, we adapt Le Cam's [23] operator technique, which is mainly based on signed measures and their convolutions. Though this approach is natural for distributions of sums of independent random variables, we nevertheless show that it can also be applied to the Markov binomial distribution. The idea of the proofs of Theorems 3.1, 3.3, 3.5 is the following. The assumptions of this paper allow to write $F_n = W_1\Lambda_1^n + W_2\Lambda_2^n$. As a rule, $\|W_2\Lambda_2^n\|$ is sufficiently small (for Theorem 3.3 this is true if $n \geqslant 2$). It remains to approximate $W_1\Lambda_1^n$, which we write as an exponential measure $\exp\{\ln W_1 + n\ln\Lambda_1\}$. Then, taking into account the properties of exponential measures and applying Lemma 4.4, we obtain expressions of the form $\|U^k\exp\{n\nu_1 U\}\|$. Application of Lemma 4.1 completes the proofs. For the Wasserstein metric we use (2) whenever possible and further on work with the total variation norm. In general, the method of this paper might be applied when all but one eigenvalues of the transition matrix of the characteristic function (see $\widetilde{P}(t)$ of proof of Lemma 4.3) are very small. Otherwise, the main problem would be to get the analogue of Lemma 4.4. In our case, we take advantage of the explicit dependence of Λ_1 on the transition probabilities.

Proof of Theorem 3.1. The estimates are proved similarly, so we give the details of (10) only. We have

$$\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\| \leqslant \|\Lambda_1^n W_1 - \mathrm{Pois}(\mathrm{E}S_n)\| + \|\Lambda_2^n\|\|W_2\|.$$

Let $M_1 = n\ln\Lambda_1 + \ln W_1$ and $M_2 = \mathrm{E}S_n U$. Direct calculations show that

$$|A_1| \leqslant C|p-\overline{q}||p_0-\nu_1|, \quad |A_1|^2 \leqslant C|p-\overline{q}|(p_0p+\overline{q}), \quad |A_2| \leqslant C|p-\overline{q}|(p_0p+\overline{q}).$$

Applying Lemma 4.4 and the properties of the total variation norm (see Introduction), we get

$$\begin{aligned}
\|\Lambda_1^n W_1 - \mathrm{Pois}(\mathrm{E}S_n)\| &= \|\mathrm{e}^{M_1} - \mathrm{e}^{M_2}\| = \Big\|\int_0^1 \big(\mathrm{e}^{tM_1+(1-t)M_2}\big)'\,\mathrm{d}t\Big\| \\
&\leqslant \int_0^1 \|(M_1-M_2)\mathrm{e}^{tM_1+(1-t)M_2}\|\,\mathrm{d}t \leqslant \int_0^1 \|(M_1-M_2)\mathrm{e}^{tn\ln\Lambda_1+(1-t)M_2}\|\;\mathrm{e}^{t\|\ln W_1\|}\mathrm{d}t \\
&\leqslant C\int_0^1 \|(M_1-M_2)\mathrm{e}^{0.1tn\nu_1 U+(1-t)M_2}\|\,\mathrm{e}^{t\|\ln W_1\|}\,\mathrm{d}t \\
&\leqslant C\,\|(M_1-M_2)\mathrm{e}^{0.1n\nu_1 U}\|\int_0^1 \|\mathrm{e}^{0.9(1-t)M_2}\|\exp\{0.1(1-t)|A_1-A_1(p-\overline{q})^n|\|U\|\}\,\mathrm{d}t \\
&\leqslant C\,\|(n\ln\Lambda_1 - n\nu_1U + \ln W_1 - A_1U)\mathrm{e}^{0.1n\nu_1U}\| + C\,|A_1(p-\overline{q})^n|\|U\mathrm{e}^{0.1n\nu_1U}\| \\
&\leqslant C\,(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q}))\|U^2\mathrm{e}^{0.1n\nu_1U}\| + C\,|p-\overline{q}|^2|\nu_1-p_0|\mathrm{e}^{-Cn}.
\end{aligned}$$

We used the fact that $\exp\{0.9(1-t)M_2\}$ is a distribution. Consequently $\|\exp\{0.9(1-t)M_2\}\| = 1$. Similarly, $\|U\mathrm{e}^{0.1n\nu_1U}\| \leqslant \|U\| \leqslant 2$. Moreover, $\|\ln W_1\| \leqslant C$. Applying Lemma 4.1, (33), and (36), we obtain

$$\|F_n - \mathrm{Pois}(\mathrm{E}S_n)\| \leqslant C\,(n\overline{q}(p+\overline{q}) + |p-\overline{q}|(p_0p+\overline{q}))\min\Big(1,\frac{1}{n\overline{q}}\Big) + C\,|p-\overline{q}|^2(\overline{q}+|\nu_1-p_0|)\mathrm{e}^{-Cn},$$

which leads us to (10). □

Above, we mentioned that Theorem 3.1 remains valid, when $\mathrm{Pois}(\mathrm{E}S_n)$ is replaced by $\mathrm{Pois}(n\nu_1 + A_1)$. This follows from the simple inequalities

$$\begin{aligned}
\|\mathrm{Pois}(\mathrm{E}S_n) - \mathrm{Pois}(n\nu_1+A_1)\| &\leqslant \|U\||\mathrm{E}S_n - (n\nu_1+A_1)| \\
&\leqslant C|p-\overline{q}|^2|p_0-\nu_1|\mathrm{e}^{-Cn} \leqslant C|p-\overline{q}|(p_0p+\overline{q})\mathrm{e}^{-Cn}. \quad (47)
\end{aligned}$$

Proof of Theorem 3.3. Let $n=1$ and $\omega = p_0p + (1-p_0)\overline{q}$. Then $F_n = I + \omega U$, $\mathrm{E}S_1 = \omega$, $2M_0 = -\omega^2U^2$ and the proof follows from the expansion

$$\mathrm{Pois}(\mathrm{E}S_1)(I+M_0) = I + \omega U + C\omega^3U^3\Theta.$$

If $n \geqslant 2$, the proof is similar to the previous one. Applying (36) and (33) we obtain

$$\|\Lambda_2^n W_2\| \leqslant C(p-\overline{q})^2(p_0 p+\overline{q})\mathrm{e}^{-Cn}.$$

Similarly to the proof of (42), we obtain

$$\|\Lambda_1^n W_1 - \mathrm{e}^{M_2+M_0}\| \leqslant C\,\|(M_1-M_2-M_0)\mathrm{e}^{0.1n\nu_1 U}\|$$

For the proof we used the fact that

$$\begin{aligned}
&\|\exp\{0.9(1-t)M_2+(1-t)M_0\}\|\\
&\qquad\leqslant\ C\|\exp\{(1-t)0.9n\nu_1 U+0.5(1-t)n(\nu_2-\nu_1^2)U^2\}\|\\
&\qquad\leqslant\ C\Big\|\exp\Big\{0.9(1-t)n\nu_1 U+\frac{7}{24}0.9(1-t)n\nu_1 U^2\Theta\Big\}\Big\|\leqslant C. \qquad (48)
\end{aligned}$$

The last inequality is a consequence of (42). Similarly,

$$\begin{aligned}
\Big\|\mathrm{e}^{M_2}\big(\mathrm{e}^{M_0}-I-M_0\big)\Big\| &= \Big\|M_0^2\int_0^1 \mathrm{e}^{M_2+tM_0}(1-t)\,\mathrm{d}t\Big\| \leqslant C\int_0^1\Big\|M_0^2\mathrm{e}^{n\nu_1 U+tM_0}\Big\|(1-t)\mathrm{d}t\\
&\leqslant C\Big\|M_0^2\mathrm{e}^{0.1n\nu_1 U}\Big\|\int_0^1\Big\|\exp\{0.9n\nu_1 U+tM_0\}\Big\|\mathrm{d}t \leqslant C\,\|\mathrm{e}^{0.1n\nu_1 U}M_0^2\|.
\end{aligned}$$

The proof is completed by applying Lemmas 4.4, 4.1, and (2). The estimates for local and Wasserstein norms are proved in the same way. □

If $n \geqslant 2$, Theorem 3.3 remains valid, when $\mathrm{Pois}(\mathrm{E}S_n)$ and M_0 are replaced by $\mathrm{Pois}(n\nu_1 + A_1)$ and $\widetilde{M}_0$, respectively. Indeed, then

$$\begin{aligned}
\|M_0-\widetilde{M}_0\| &\leqslant Cn|p-\overline{q}|^n(|A_1|+|A_2|)\|U^2\| \leqslant Cn|p-\overline{q}|^{n+1}(p_0+\overline{q})\\
&\leqslant Cn\mathrm{e}^{-Cn}(p-\overline{q})^2(p+\overline{q})(p_0+\overline{q}) \leqslant Cn\mathrm{e}^{-Cn}(p-\overline{q})^2(p_0p+\overline{q}),\\
\|M_0\| &\leqslant Cn, \quad |\mathrm{E}S_n-n\nu_1-A_1| \leqslant C|p-\overline{q}|^{n+1}(p_0+\overline{q}) \leqslant C(p-\overline{q})^2(p_0p+\overline{q})\mathrm{e}^{-Cn}.
\end{aligned}$$

Now by the properties of the total variation norm

$$\begin{aligned}
&\|\mathrm{Pois}(\mathrm{E}S_n)(I+M_0)-\mathrm{Pois}(n\nu_1+A_1)(I+\widetilde{M}_0)\|\\
&\qquad\leqslant \|[\mathrm{Pois}(\mathrm{E}S_n)-\mathrm{Pois}(n\nu_1+A_1)](I+M_0)\| + \|\mathrm{Pois}(n\nu_1+A_1)(M_0-\widetilde{M}_0)\|\\
&\qquad\leqslant C(1+\|M_0\|)|\mathrm{E}S_n-n\nu_1-A_1| + C\|M_0-\widetilde{M}_0\|
\end{aligned}$$

and it suffices to use previous estimates.

Proof of Theorem 3.2. Note that

$$\frac{4A_0}{\sqrt{2\pi\mathrm{e}}} = \frac{n}{2}|\nu_2-\nu_1^2|\frac{\|\varphi_2\|_1}{n\nu_1}.$$

Therefore,

$$\begin{aligned}
\Big|\|F_n-\mathrm{Pois}(\mathrm{E}S_n)\| - \frac{4A_0}{\sqrt{2\pi\mathrm{e}}}\Big| &\leqslant \Big\|F_n-\mathrm{Pois}(\mathrm{E}S_n)\Big(I+\frac{n(\nu_2-\nu_1^2)-A_1^2+2A_2}{2}U^2\Big)\Big\|\\
&+\Big\|\frac{2A_2-A_1^2}{2}U^2\mathrm{Pois}(\mathrm{E}S_n)\Big\| + \Big\|\frac{n}{2}(\nu_2-\nu_1^2)U^2(\mathrm{Pois}(\mathrm{E}S_n)-\mathrm{Pois}(n\nu_1))\Big\|\\
&+\frac{n}{2}|\nu_2-\nu_1^2|\Big|\|U^2\mathrm{Pois}(n\nu_1)\| - \frac{\|\varphi_2\|_1}{n\nu_1}\Big|.
\end{aligned}$$

Now it suffices to use Lemmas 4.2 and 4.1 and Theorem 3.3. The difference between two Poisson distributions can be estimated via the same approach as used in the proof of Theorem 3.1:

$$
\begin{aligned}
&\Big\| \frac{n}{2}(\nu_2-\nu_1^2)U^2(\mathrm{Pois}(\mathrm{E}S_n)-\mathrm{Pois}(n\nu_1))\Big\| \\
&\qquad \leqslant \; Cn\overline{q}(p+\overline{q})\Big\|U^2\mathrm{e}^{n\nu_1U}\int_0^1(\exp\{(\mathrm{E}S_nU-n\nu_1U)t\})'\mathrm{d}t\Big\| \\
&\qquad \leqslant \; Cn\overline{q}(p+\overline{q})|p-\overline{q}|\Big\|U^3\int_0^1\exp\{t\mathrm{E}S_nU+(1-t)n\nu_1U\}\mathrm{d}t\Big\| \\
&\qquad \leqslant \; Cn\overline{q}(p+\overline{q})|p-\overline{q}|\Big\|U^3\mathrm{e}^{n\nu_1U}\Big\| \leqslant C(p+\overline{q})|p-\overline{q}|(n\overline{q})^{-1/2}.
\end{aligned}
$$

Estimates for other norms are obtained similarly. □

Proof of Theorem 3.5. Let $M_3=\mu U+(\sigma^2-\mu)U^2/2$. Taking into account the last inequality in (48) and arguing similarly to the proof of Theorem 3.1, we obtain

$$
\begin{aligned}
&\|\Lambda_1^nW_1-\mathrm{SPois}(\mu,\sigma)\| \\
&\quad\leqslant \; C\int_0^1\Big\|\mathrm{e}^{tM_1+(1-t)M_3}\Big\|\mathrm{d}t\leqslant C\int_0^1\Big\|(M_1-M_3)\mathrm{e}^{0.1tn\nu_1U+(1-t)M_3}\Big\|\mathrm{d}t \\
&\quad\leqslant \; C\int_0^1\Big\|(M_1-M_3)\exp\{0.1tn\nu_1U+(1-t)n\nu_1U+0.5(1-t)n(\nu_2-\nu_1^2)U^2\}\Big\|\mathrm{d}t \\
&\quad\leqslant \; C\Big\|(M_1-M_3)\mathrm{e}^{0.1n\nu_1U}\Big\|\int_0^1\|\exp\{0.9(1-t)n\nu_1U+0.5(1-t)n(\nu_2-\nu_1^2)U^2\}\|\mathrm{d}t \\
&\quad\leqslant \; C\Big\|(M_1-M_3)\mathrm{e}^{0.1n\nu_1U}\Big\|.
\end{aligned}
$$

Similar estimates hold for local and Wasserstein norms. Moreover,

$$M_1-M_3=C\,(p+\overline{q})^2(n\overline{q}+|p-\overline{q}|)U^3\Theta.$$

Now the proof of Theorem 3.5 can be completed applying Lemma 4.1. □

The proof of Theorem 3.4 follows from Lemma 4.7 and Theorem 3.5.

Acknowledgment

We are grateful to V. Kazakevičius for valuable remarks, which helped to improve the results of the paper. This work was finished during the second author's stay at the Institute of Mathematics at Carl von Ossietzky University of Oldenburg in Summer term 2006. The second author would like to thank Professor Dietmar Pfeifer for the invitation to Oldenburg and the Institute for its hospitality. We are also grateful to the referees for useful remarks.

References

[1] T.V. Arak, A.Yu. Zaĭtsev, Uniform Limit Theorems for Sums of Independent Random Variables. Transl. from Russian. (English). Proc. Steklov Inst. Math., **174**(1), (1988) 222 p., AMS, Providence, RI.

[2] A.D. Barbour, V. Čekanavičius, Total variation asymptotics for sums of independent integer random variables, Ann. Probab. **30** (2002) 509–545.

[3] A.D. Barbour, V. Čekanavičius, A. Xia, On Stein's method and perturbations. ALEA **3** (2007) 31–53.

[4] A.D. Barbour, L. Holst, S. Janson, Poisson Approximation, (1992) Clarendon Press, Oxford.

[5] A.D. Barbour, T. Lindvall, Translated Poisson approximation for Markov chains. J. Theor. Probab. **19** (2006) 609–630.

[6] T.C. Brown, A. Xia, Stein's method and birth-death processes, Ann. Probab. **29** (2001) 1373–1403.

[7] S.A. Campbell, A.P. Godbole, S. Schaller, Discriminating between sequences of Bernoulli and Markov-Bernoulli trials, Commun. Statist. Theory Meth. **23** (1994) 2787–2814.

[8] V. Čekanavičius, A. Elijio, Lower-bound estimates for Poisson-type approximations, Lithuanian Math. J. **45** (2005) 405–423.

[9] V. Čekanavičius, M. Mikalauskas, Signed Poisson approximations for Markov chains, Stochastic. Proc. Appl. **82** (1999) 205–227.

[10] V. Čekanavičius, M. Mikalauskas, Local theorems for the Markov binomial distribution, Lithuanian Math. J. **41** (2001) 219–231.

[11] V. Čekanavičius, P. Vaitkus, Centered Poisson approximation via Stein's method, Lithuanian Math. J. **41** (2001) 319–329.

[12] O. Chryssaphinou, E. Vaggelatou, Compound Poisson approximation for multiple runs in a Markov chain, Ann. Inst. Statist. Math. **54** (2002) 411–424.

[13] P. Deheuvels, D. Pfeifer, On a relationship between Uspensky's theorem and Poisson approximations, Ann. Inst. Statist. Math. **40** (1988) 671–681.

[14] R.L. Dobrushin, Limit theorems for a Markov chain of two states, Izv. Akad. Nauk. USSR Ser. mat. **17** (1953) 291–330. (Russian). English translation in Select. Transl. Math. Stat. and Probab. **1**, Inst. Math. Stat. and Amer. Math. Soc. (1961) 97–134.

[15] W. Doeblin, Sur les sommes d'un grand nombre de variables aléatoires indépendantes, Bull. Sci. Math., II. Ser. **63** (1939) 23–32 and 35–64.

[16] A.W.F. Edwards, The meaning of the binomial distribution, Nature, London **186** (1960) 1074.

[17] T. Erhardsson, Compound Poisson approximation for Markov chains using Stein's method, Ann. Probab. **27** (1999) 565–596.

[18] J. Gani, On the probability generating function of the sum of Markov-Bernoulli random

variables, J. Appl. Probab. (Special vol.) **19A** (1982) 321–326.

[19] L. Goldstein, A. Xia, Zero biasing and a discrete central limit theorem, Ann. Probab. **34** (2006) 1782–1806.

[20] A.Ya. Khintchine, Asymptotische Gesetze der Wahrscheinlichkeitsrechnung, (1933) Springer–Verlag, Berlin.

[21] B.O. Koopman, A generalization of Poisson's distribution for Markoff chains, Proc. Nat. Acad. Sci. USA **36** (1950) 202–207.

[22] J. Kruopis, Precision of approximation of the generalized binomial distribution by convolutions of Poisson measures, Lithuanian Math. J. **26** (1986) 37–49.

[23] L. Le Cam, An approximation theorem for the Poisson binomial distribution, Pacific J. Math. **10** (1960) 1181–1197.

[24] Yu.V. Prokhorov, Asymptotic behavior of the binomial distribution, Uspekhi Mat. Nauk. **8** (1953) 135–142. (Russian). English translation in Select. Transl. Math. Stat. and Probab. **1**, Inst. Math. Stat. and Amer. Math. Soc. (1961) 87–95 .

[25] A. Röllin, Approximation of sums of conditionally independent variables by the translated Poisson distribution, Bernoulli **11** (2005) 1115–1128.

[26] B. Roos, Asymptotics and sharp bounds in the Poisson approximation to the Poisson-binomial distribution, Bernoulli **5** (1999) 1021–1034.

[27] B. Roos, Sharp constants in the Poisson approximation, Statist. Probab. Lett. **52** (2001) 155–168.

[28] B. Roos, On Hipp's compound Poisson approximations via concentration functions, Bernoulli **11** (2005) 533–557.

[29] R.J. Serfling, A general Poisson approximation theorem, Ann. Probab. **3** (1975) 726–731.

[30] R.F. Serfozo, Compound Poisson approximations for sums of random variables, Ann. Probab. **14** (1986) 1391–1398. Correction. Ann. Probab. **16** (1988) 429–430.

[31] P. Vellaisamy, Poisson approximation for (k_1, k_2)-events via the Stein-Chen method, J. Appl. Probab. **41** (2004) 1081–1092.

[32] Y.H. Wang, On the limit of the Markov binomial distribution, J. Appl. Probab. **18** (1981) 937–942.

[33] Y.H. Wang, Approximating kth-order two-state Markov chains, J. Appl. Probab. **29** (1992) 861–868.

[34] Y.H. Wang, L. Tang, Poisson style convergence theorems for additive processes defined on Markov chains, Statistica Sinica **13** (2003) 227–242.

[35] Y.H. Wang, Hsiao-Fen Chang, Shun-Yi Chen, Convergence theorems for the lengths of consecutive successes of Markov Bernoulli sequences, J. Appl. Probab. **40** (2003) 741–749.