Bernoulli **23**(4A), 2017, 2828–2859
DOI: 10.3150/16-BEJ829

On Stein operators for discrete approximations

NEELESH S. UPADHYE[1], VYDAS ČEKANAVIČIUS[2] and P. VELLAISAMY[3]

[1]*Department of Mathematics, Indian Institute of Technology Madras, Chennai 600036, India.*
E-mail: neelesh@iitm.ac.in
[2]*Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 03225, Lithuania.*
E-mail: vydas.cekanavicius@mif.vu.lt
[3]*Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India.*
E-mail: pv@math.iitb.ac.in

In this paper, a new method based on probability generating functions is used to obtain multiple Stein operators for various random variables closely related to Poisson, binomial and negative binomial distributions. Also, the Stein operators for certain compound distributions, where the random summand satisfies Panjer's recurrence relation, are derived. A well-known perturbation approach for Stein's method is used to obtain total variation bounds for the distributions mentioned above. The importance of such approximations is illustrated, for example, by the binomial convoluted with Poisson approximation to sums of independent and dependent indicator random variables.

Keywords: binomial distribution; compound Poisson distribution; Panjer's recursion; perturbation; Stein's method; total variation norm

1. Introduction

Stein's method is known to be one of the powerful techniques for probability approximations and there is a vast literature available on this topic. For details and applications of Stein's method, see [7,12,35] and [29]. For some recent developments, see [12,16,19,27–29] and the references therein. The method is based on the construction of a characteristic operator for an approximation problem. Different approaches are used for deriving Stein operators (see, [32]). For instance, a Stein operator can be treated in the framework of birth–death processes [10]. Stein's method for discrete distributions has been independently and simultaneously developed by [23,27,28]. More recently, [26] has proposed a canonical operator, for both continuous and discrete distributions, and a general approach to obtain bounds on approximation problems.

In this paper, we consider the random variables (r.v.s) concentrated on $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ with distributions having the form of convoluted measures or random sums. Using their probability generating functions (*PGF*'s), we derive Stein operators for discrete probability approximations. In particular, the existence of multiple Stein operators (in the case of convoluted measures) for an approximation problem is shown and the corresponding bounds are derived, using perturbation technique, and compared for the case of indicator r.v.s. Although the existence of infinite families of Stein operators for many common distributions is already well known (see, [22] and [26]), this comparison may benefit the readers, as it is illustrated for the first time (in case of convoluted measures) to the best of our knowledge.

Next, we describe a typical procedure for Stein's method on $\mathbb{Z}_+$-valued r.v.s. Let Y be a $\mathbb{Z}_+$-valued r.v. with $\mathbb{E}(|Y|) < \infty$, $\mathcal{F} := \{f | f : \mathbb{Z}_+ \to \mathbb{R} \text{ and is bounded}\}$ and $\mathcal{G}_Y = \{g \in \mathcal{F} | g(0) = 0 \text{ and } g(x) = 0 \text{ for } x \notin \operatorname{supp}(Y)\}$, where $\operatorname{supp}(Y)$ denotes the support of r.v. Y. We want to bound $\mathbb{E}f(Z) - \mathbb{E}f(Y)$ for some r.v. Z concentrated on $\mathbb{Z}_+$ and $f \in \mathcal{F}$. Stein's method is then realized in three consecutive steps. First, for any $g \in \mathcal{G}_Y$, a linear operator $\mathcal{A}$ satisfying $\mathbb{E}(\mathcal{A}g)(Y) = 0$ is established and is called a Stein operator. For a general framework of Stein operators, the reader is referred to [17,20,23,27,28,37,38] and [26].

In the next step, the so-called Stein equation

$$(\mathcal{A}g)(j) = f(j) - \mathbb{E}f(Y), \qquad j \in \mathbb{Z}_+, f \in \mathcal{F} \tag{1}$$

is solved with respect to $g(j)$ in terms of f and is referred to as a solution to the Stein equation (1). As a rule, solutions to the Stein equations have useful properties, such as $\|\Delta g\| := \sup_{j\in\mathbb{Z}_+} |\Delta g(j)|$ is small, where $\Delta g(j) := g(j+1) - g(j)$ denotes the first forward difference. Note that the properties of Δg depend on the form of $\mathcal{A}$ and some properties of Y. Finally, taking expectations on both sides of (1), we get

$$\mathbb{E}f(Z) - \mathbb{E}f(Y) = \mathbb{E}(\mathcal{A}g)(Z) \tag{2}$$

and bounds for $\mathbb{E}(\mathcal{A}g)(Z)$ are established through the bounds for Δg and $\Delta^{k+1}g(j) := \Delta^k(g(j+1) - g(j))$, $k = 1, 2, \dots$. For more details on Stein's method under a general setup, we refer the readers to [5,22,23,26] and the references therein.

For some standard distributions, a Stein operator can be established easily. Indeed, let $\mu_j := P(Y = j) > 0$, $j \in \mathbb{Z}_+$. Then $\sum_{j=0}^\infty \mu_j(\frac{(j+1)\mu_{j+1}}{\mu_j} g(j+1) - jg(j)) = 0$. Therefore,

$$(\mathcal{A}g)(j) = \frac{(j+1)\mu_{j+1}}{\mu_j} g(j+1) - jg(j), \qquad j \in \mathbb{Z}_+, \tag{3}$$

and it can be easily verified that $\mathbb{E}(\mathcal{A}g)(Y) = 0$. Some well-known examples are listed below.

(1) For $\alpha > 0$, let Y_1 be a Poisson $P(\alpha)$ r.v. with $\mu_j = P(Y_1 = j) = \alpha^j \mathrm{e}^{-\alpha}/j!$. Then

$$(\mathcal{A}g)(j) = \alpha g(j+1) - jg(j), \qquad j \in \mathbb{Z}_+. \tag{4}$$

(2) Let $0 < p < 1$, $q = 1 - p$, $\widetilde{M} > 1$, and Y_2 have the pseudo-binomial distribution (see [11], page 370) so that

$$\mu_j = P(Y_2 = j) = \frac{1}{\widetilde{C}} \binom{\widetilde{M}}{j} p^j q^{\widetilde{M}-j}, \qquad j \in \{0, 1, \dots, \lfloor \widetilde{M} \rfloor\},$$

where $\widetilde{C} = \sum_{j=0}^{\lfloor \widetilde{M} \rfloor} \binom{\widetilde{M}}{j} p^j q^{\widetilde{M}-j}$, $\lfloor \widetilde{M} \rfloor$ denotes integer part of $\widetilde{M}$ and $\binom{\widetilde{M}}{j} = \frac{\widetilde{M}(\widetilde{M}-1)\cdots(\widetilde{M}-j+1)}{j!}$. If $\widetilde{M}$ is an integer, then Y_2 is a binomial r.v. Suppose now $g(0) = 0$ and $g(\lfloor \widetilde{M} \rfloor + 1) = g(\lfloor \widetilde{M} \rfloor + 2) = \cdots = 0$. Then, from (3)

$$(\mathcal{A}g)(j) = \frac{(\widetilde{M} - j)p}{q} g(j+1) - jg(j), \qquad j = 0, 1, \dots, \lfloor \widetilde{M} \rfloor.$$

Multiplying the above expression by q, we can get the following Stein operator:

$$(\mathcal{A}g)(j) = (\widetilde{M} - j)pg(j+1) - jqg(j), \qquad j = 0, 1, \dots, \lfloor \widetilde{M} \rfloor. \tag{5}$$

(3) Let $Y_3 \sim \mathrm{NB}(r, \bar{p})$, $0 < \bar{p} < 1$, be negative binomial distribution with $\mu_j = P(Y_3 = j) = \Gamma(r+j)/(\Gamma(r)j!)\bar{p}^r \bar{q}^j$, for $j \in \mathbb{Z}_+$, $r > 0$ and $\bar{q} = 1 - \bar{p}$. Then (3) reduces to

$$(\mathcal{A}g)(j) := \bar{q}(r+j)g(j+1) - jg(j), \qquad j \in \mathbb{Z}_+. \tag{6}$$

Observe that equation (3) is not that useful if we do not have simple expressions for μ_j and especially for μ_j/μ_{j+1}. One such class is the Ord family of cumulative distributions (see [1]). For example, if we consider compound distribution or convolution of two or more distributions, then μ_j's are usually expressed through sums or converging series of probabilities. Therefore, some other refined approaches for obtaining Stein operator(s) are needed.

The paper is organized as follows. In Section 2, we use the *PGF* approach to obtain general expressions for Stein operators arising out of convolution of r.v.s and random sums that satisfy Panjer's recursive relation. These operators are then seen as perturbations of known operators for standard distributions which motivate the discussion about perturbation approach and its applications. In Section 3, some facts about the perturbation approach to a solution of Stein equation are discussed and applied to the operators derived in Section 2. In Section 4, as an application, an approximation problem for the distribution of the sum of possibly dependent indicator variables by the convolution of Poisson and binomial distribution is considered. We show that such approximations can be treated either as Poisson perturbation or as binomial perturbation, leading to two different bounds. Finally, we mention that though the approach is restricted to distributional approximations, its ideas can be extended for approximations to signed measures as well.

2. Stein operators via *PGF*

In this section, the *PGF* approach is used to derive the operators satisfying $\mathbb{E}(\mathcal{A}g)(Y) = 0$ for $g \in \mathcal{G}_Y$. The construction of $\mathcal{A}$ is well known if probabilities of approximating distribution satisfy some recursive relation and it can be easily verified by using this approach. Indeed, the *PGF* has been used as a tool for establishing Panjer's recurrence relations; see, for example, [39] and [24]. Note also that, strictly speaking, $\mathcal{A}$ can be called a Stein operator only if it is used in (2) with g satisfying (1). Moreover, one expects g to have some useful properties. In Section 3, we show that, the majority of operators considered below have solutions to (1) with properties typical for the Stein method.

Next, we use the *PGF* approach to derive the Stein operators for compound Poisson distribution, certain convolution of distributions and a compound distribution where the summand satisfy the Panjer's recurrence relation.

2.1. The general idea

Let N be a $\mathbb{Z}_+$-valued r.v. with $\mu_k = P(N = k)$ and finite mean. Then its *PGF*

$$G_N(z) = \sum_{k=0}^{\infty} \mu_k z^k \tag{7}$$

satisfies

$$G'_N(z) = \frac{d}{dz} G_N(z) = \sum_{k=1}^{\infty} k\mu_k z^{k-1} = \sum_{k=0}^{\infty} (k+1)\mu_{k+1} z^k, \tag{8}$$

where prime denotes the derivative with respect to z. If we can express $G'_N(z)$ through $G_N(z)$ then, by collecting factors corresponding to z^k, the recursion follows. One can easily verify the Stein operators derived for standard distributions in the previous section, using this approach.

Next, we start with the derivation of a Stein operator for a compound Poisson distribution.

Let $\{X_j\}$ be an i.i.d. sequence of random variables with $P(X_j = k) = p_k$, $k = 0, 1, 2, \ldots$. Also, let $N \sim P(\lambda)$ and be independent of the $\{X_j\}$. Then the distribution of $Y_4 := \sum_{j=1}^{N} X_j$ is known as the compound Poisson distribution with the *PGF*

$$G_{\mathrm{cp}}(z) = \exp\left\{ \sum_{j=1}^{\infty} \lambda_j \left(z^j - 1\right) \right\}, \tag{9}$$

where $\lambda_j = \lambda p_j$ and $\sum_{j=1}^{\infty} j|\lambda_j| < \infty$. Then

$$G'_{\mathrm{cp}}(z) = G_{\mathrm{cp}}(z) \sum_{j=1}^{\infty} j\lambda_j z^{j-1} = \sum_{k=0}^{\infty} \mu_k z^k \sum_{j=1}^{\infty} j\lambda_j z^{j-1} = \sum_{k=0}^{\infty} z^k \sum_{m=0}^{k} \mu_m (k-m+1)\lambda_{k-m+1}.$$

Comparing the last expression to the right-hand side of (8), we obtain the recursive relation, for all $k \in \mathbb{Z}_+$, as

$$\sum_{m=0}^{k} \mu_m (k-m+1)\lambda_{k-m+1} - (k+1)\mu_{k+1} = 0.$$

Then, for $g \in \mathcal{G}_{Y_4}$, we have

$$\begin{aligned}
0 &= \sum_{k=0}^{\infty} g(k+1) \left[\sum_{m=0}^{k} \mu_m (k-m+1)\lambda_{k-m+1} - (k+1)\mu_{k+1} \right] \\
&= \sum_{m=0}^{\infty} \mu_m \left[\sum_{k=m}^{\infty} g(k+1)(k-m+1)\lambda_{k-m+1} - mg(m) \right] \\
&= \sum_{m=0}^{\infty} \mu_m \left[\sum_{j=1}^{\infty} j\lambda_j g(j+m) - mg(m) \right].
\end{aligned}$$

Therefore, a Stein operator for the compound Poisson distribution, defined in (9), is

$$\begin{aligned}(\mathcal{A}g)(j) &= \sum_{l=1}^{\infty} l\lambda_l g(j+l) - jg(j) \\ &= \sum_{l=1}^{\infty} l\lambda_l g(j+1) - jg(j) + \sum_{m=2}^{\infty} m\lambda_m \sum_{l=1}^{m-1} \Delta g(j+l), \qquad j \in \mathbb{Z}_+,\end{aligned} \tag{10}$$

since $\mathbb{E}(\mathcal{A}g)(Y_4) = 0$. This operator coincides with the one from [6].

Next, we derive multiple Stein operators for convolution of standard distributions discussed above.

2.2. Convolutions of distributions

Recall that $Y_1 \sim P(\alpha)$ ($\alpha > 0$), $Y_2 \sim \mathrm{Bi}(M, p)$ ($M \in \mathbb{N}$, $0 < p < 1$), $Y_3 \sim \mathrm{NB}(r, \bar{p})$ ($0 < \bar{p} < 1$, $r > 0$) and Y_4 follows the compound Poisson distribution defined in (9). We assume that Y_1, Y_2, Y_3 and Y_4 are independent. Then the *PGF*'s of $Y_1 + Y_2$, Y_2 and Y_3 are given by

$$\begin{aligned} &G_{12}(z) = (q + pz)^M \exp\{\alpha(z-1)\}, \qquad G_2(z) = (q+pz)^M, \\ &G_3(z) = \left(\frac{\bar{p}}{1 - \bar{q}z}\right)^r, \end{aligned} \tag{11}$$

respectively. Here $\bar{q} = 1 - \bar{p}$ and $q = 1 - p$. We now derive the Stein operators for the convolutions of various combinations of Y_1, Y_2, Y_3 and Y_4.

Proposition 2.1. *Let $G_{\mathrm{cp}}(z)$ be the PGF of Y_4 and $\lambda = \sum_{j=1}^{\infty} j\lambda_j$. Then we have the following results*:

(i) *The r.v. $Y_{24} = Y_2 + Y_4$ has the PGF $G_2(z)G_{\mathrm{cp}}(z)$ and its Stein operator, for $g \in \mathcal{G}_{Y_{24}}$, is*

$$\begin{aligned}(\mathcal{A}g)(j) &= \left(M + \frac{\lambda}{p} - j\right) pg(j+1) - qjg(j) \\ &\quad + \sum_{m=2}^{\infty} \big(qm\lambda_m + p(m-1)\lambda_{m-1}\big) \sum_{l=1}^{m-1} \Delta g(j+l).\end{aligned} \tag{12}$$

(ii) *The r.v. $Y_{34} = Y_3 + Y_4$ has the PGF $G_3(z)G_{\mathrm{cp}}(z)$ and has a Stein operator, for $g \in \mathcal{G}_{Y_{34}}$,*

$$\begin{aligned}(\mathcal{A}g)(j) &= \left(\frac{\lambda\bar{p}}{\bar{q}} + r + j\right) \bar{q}g(j+1) - jg(j) \\ &\quad + \sum_{m=2}^{\infty} \big(m\lambda_m - \bar{q}(m-1)\lambda_{m-1}\big) \sum_{l=1}^{m-1} \Delta g(j+l).\end{aligned} \tag{13}$$

Proof. Write $G_2(z)G_{\text{cp}}(z) = \sum_{k=0}^{\infty} \mu_k z^k$. Differentiating with respect to z, we get the identity

$$\sum_{k=0}^{\infty} \mu_k z^k \left(\frac{Mp}{q+pz} + \sum_{j=1}^{\infty} \lambda_j j z^{j-1} \right) = \sum_{k=0}^{\infty} k\mu_k z^{k-1}.$$

Multiplying both sides by $(q + pz)$ and collecting the terms corresponding to z^k, we obtain the recursive relation

$$\sum_{m=0}^{k} \mu_m \big(q\lambda_{k-m+1}(k-m+1) + p(k-m)\lambda_{k-m} \big) - (k+1)\mu_{k+1} q + (Mp - pk)\mu_k = 0.$$

Multiplying the last equation by $g(k+1)$ and summing over all nonnegative integer k leads to (12).

To prove (13), let $G_3(z)G_{\text{cp}}(z) = \sum_{k=0}^{\infty} \mu_k z^k$. Differentiating with respect to z gives the identity

$$\sum_{k=0}^{\infty} \mu_k z^k \left(\frac{r\bar{q}}{1-\bar{q}z} + \sum_{j=1}^{\infty} \lambda_j j z^{j-1} \right) = \sum_{k=0}^{\infty} k\mu_k z^{k-1}.$$

Multiplying both sides by $(1 - \bar{q}z)$ and collecting the terms corresponding to z^k, we obtain

$$\sum_{m=0}^{k} \mu_m \big(\lambda_{k-m+1}(k-m+1) - \bar{q}(k-m)\lambda_{k-m} \big) - (k+1)\mu_{k+1} + \bar{q}(k+r)\mu_k = 0.$$

Multiply the above equation by $g(k+1)$ and then sum over $k \in \mathbb{Z}_+$ to obtain the result. $\square$

Proposition 2.2. *Let $Y_{12} = Y_1 + Y_2$ have PGF $G_{12}(z)$ as defined in (11). Then, for $j \in \mathbb{Z}_+$ and $g \in \mathcal{G}_{Y_{12}}$, a Stein operator for Y_{12} is*

$$(\mathcal{A}g)(j) = (Mp + \alpha - jp)g(j+1) - jqg(j) + p\alpha\Delta g(j+1). \tag{14}$$

If in addition $p < q$, then

$$(\mathcal{A}g)(j) = (\alpha + Mp)g(j+1) - jg(j) + M\sum_{l=2}^{\infty}(-1)^{l+1}\left(\frac{p}{q}\right)^l \sum_{k=1}^{l-1} \Delta g(j+k). \tag{15}$$

Proof. Observe that (15) follows from (10) and the expansion

$$(q+pz)^M = \exp\left\{ M\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} \left(\frac{p}{q}\right)^i (z^i - 1) \right\}. \tag{16}$$

Note that (14) is a special case of (12). $\square$

Remark 2.3. (i) As is known in the literature (see [22]), we have two significantly different Stein operators (see (14) and (15)) for the approximation problem.

(ii) Observe that, the operator given in (14) is similar to the operator given in (5), where $\widetilde{M}$ is replaced by $M + \alpha/p$, except for the last term, and hence is known as a binomial perturbation.

(iii) Similarly, the operator given in (15) is similar to the operator given in (4), where α is replaced by $Mp + \alpha$, except for the last sum, leading to a Poisson perturbation.

Next, we demonstrate that the number of such operators might be even larger. We consider the convolution of negative binomial and binomial distributions. It is logical to use the binomial approximation for sums of r.v.'s with variances smaller than their means and the negative binomial approximation if variances are larger than means. Therefore, one can expect that the convolution of a binomial with a negative binomial r.v. to be a more versatile discrete approximation, as it gives more flexibility in the choice of parameters to match the second moment, for example.

Proposition 2.4. *Let $Y_{23} = Y_2 + Y_3$ have PGF $G_{23}(z) = G_2(z)G_3(z)$ and $p < q$. Then, for $j \in \mathbb{Z}_+$ and $g \in \mathcal{G}_{Y_{23}}$, the r.v. Y_{23} has the following Stein operators*:

$$\begin{aligned}(\mathcal{A}_1 g)(j) &= (Mp + rq\bar{q} - pj + q\bar{q}j)g(j+1)\\ &\quad + (r\bar{q}p - Mp\bar{q} + p\bar{q}j)g(j+2) - qjg(j),\end{aligned} \tag{17}$$

$$\begin{aligned}(\mathcal{A}_2 g)(j) &= p\left(\frac{r\bar{q}}{p\bar{p}} + M - j\right)g(j+1) - qjg(j)\\ &\quad + r(q\bar{q} + p)\sum_{m=2}^{\infty}\bar{q}^{m-1}\sum_{l=1}^{m-1}\Delta g(j+l),\end{aligned} \tag{18}$$

$$\begin{aligned}(\mathcal{A}_3 g)(j) &= \bar{q}\left(\frac{Mp\bar{p}}{\bar{q}} + r + j\right)g(j+1) - jg(j)\\ &\quad + M\left(\frac{p}{q} + \bar{q}\right)\sum_{m=2}^{\infty}(-1)^{m+1}\left(\frac{p}{q}\right)^{m-1}\sum_{l=1}^{m-1}\Delta g(j+l),\end{aligned} \tag{19}$$

$$\begin{aligned}(\mathcal{A}_4 g)(j) &= \left(Mp + \frac{r\bar{q}}{\bar{p}}\right)g(j+1) - jg(j)\\ &\quad + \sum_{m=2}^{\infty}\left(M(-1)^{m+1}\left(\frac{p}{q}\right)^{m} + r\bar{q}^{m}\right)\sum_{l=1}^{m-1}\Delta g(j+l).\end{aligned} \tag{20}$$

Proof. Differentiating $G_{23}(z) = G_2(z)G_3(z)$ with respect to z, we obtain

$$\sum_{k=0}^{\infty}\mu_k z^k\left(\frac{Mp}{q+pz} + \frac{r\bar{q}}{1-\bar{q}z}\right) = \sum_{k=0}^{\infty}k\mu_k z^{k-1}.$$

Multiplying both sides by $(q + pz)(1 - \bar{q}z)$ and collecting the terms corresponding to z^k, we obtain the recursive relation

$$\mu_k(Mp + rq\bar{q} - pk + q\bar{q}k) + \mu_{k-1}\big(rp\bar{q} - Mp\bar{q} + p\bar{q}(k-1)\big) - q\mu_{k+1}(k+1) = 0.$$

Multiplying the last equation by $g(k+1)$ and summing over all nonnegative k, we obtain (17). Observe next that

$$\left(\frac{\bar{p}}{1-\bar{q}z}\right)^r = \exp\left\{r\sum_{i=1}^{\infty}\frac{\bar{q}^i}{i}\big(z^i - 1\big)\right\}.$$

Therefore, (18) follow from (12). Similarly, (19) follows from (13) and (16), and (20) follows from (10) and (16). □

Remark 2.5. As discussed earlier, the operators $\mathcal{A}_2$, $\mathcal{A}_3$, and $\mathcal{A}_4$ are binomial, negative binomial and Poisson perturbations, respectively. Note, however, $\mathcal{A}_1$ cannot be seen as a perturbation operator.

2.3. Compound distributions

Next, we extend the *PGF* technique for finding Stein operators for a general class of compound distributions. Let $S_N = \sum_{j=1}^{N} X_j$, where N is a $\mathbb{Z}_+$-valued r.v. with $\mu_k = P(N = k)$ and the X_j are i.i.d. r.v.s, independent of N, with $P(X_j = k) = p_k$ for $k \in \mathbb{Z}_+$. Here and henceforth, S_0 is treated as a degenerate r.v. concentrated at zero. Then the *PGF* of S_N is given by

$$G_{S_N}(z) = G_N\big(G_{X_1}(z)\big) = \sum_{j=0}^{\infty}\pi_j z^j,$$

where

$$\pi_j = P(S_N = j) = \sum_{k=0}^{\infty}P(N = k)P(S_k = j) = \sum_{k=0}^{\infty}\mu_k p_{k,j}, \tag{21}$$

and $p_{k,j} = P(S_k = j)$ denotes the k-fold convolution of $\{p_j\}_{j\ge 0}$. Thus,

$$G_N\big(G_{X_1}(z)\big) = \sum_{j=0}^{\infty}\left(\sum_{k=0}^{\infty}\mu_k p_{k,j}\right)z^j.$$

Further on, we assume that $\mathbb{E}(S_N) < \infty$. Then

$$G'_{S_N}(z) = \sum_{j=1}^{\infty} j\pi_j z^{j-1} = \sum_{j=0}^{\infty}(j+1)\pi_{j+1}z^j = \sum_{j=0}^{\infty}(j+1)\left(\sum_{k=0}^{\infty}\mu_k p_{k,j+1}\right)z^j. \tag{22}$$

Similarly,

$$\begin{aligned} G'_{S_N}(z) &= \frac{d}{dG_{X_1}(z)} \sum_{k=0}^{\infty} \mu_k \big(G_{X_1}(z)\big)^k \left(\frac{d}{dz} \sum_{m=0}^{\infty} p_{1,m} z^m \right) \\ &= \sum_{k=0}^{\infty} (k+1)\mu_{k+1} \big(G_X(z)\big)^k \sum_{m=0}^{\infty} (m+1) p_{m+1} z^m. \end{aligned} \tag{23}$$

Noting that $(G_{X_1}(z))^k = \sum_{s=0}^{\infty} p_{k,s} z^s$, we get

$$\begin{aligned} G'_{S_N}(z) &= \sum_{k=0}^{\infty} (k+1)\mu_{k+1} \sum_{s=0}^{\infty} p_{k,s} z^s \sum_{m=0}^{\infty} (m+1) p_{m+1} z^m \\ &= \sum_{s=0}^{\infty} \left\{ \sum_{k=0}^{\infty} (k+1)\mu_{k+1} \sum_{m=0}^{s} p_{k,m} (s-m+1) p_{s-m+1} \right\} z^s. \end{aligned} \tag{24}$$

Comparing (24) with (22), we obtain the required recursion relation, for $s \in \mathbb{Z}_+$, as

$$(s+1) \sum_{k=0}^{\infty} \mu_k p_{k,s+1} = \sum_{k=0}^{\infty} (k+1)\mu_{k+1} \sum_{m=0}^{s} p_{k,m} (s-m+1) p_{s-m+1}. \tag{25}$$

Next, we derive a Stein operator. So far, some μ_j's were allowed to be equal to zero. Now we restrict ourselves to the case $\mu_j > 0$, $j = 0, 1, 2, \dots, K$ ($K = \infty$ is also allowed) and assume that $\mu_{K+1} = \mu_{K+2} = \cdots = 0$, when $K < \infty$. Multiplying (25) by $g(s+1)$ and summing over $s \in \mathbb{Z}_+$, we obtain

$$\sum_{s=0}^{\infty} s g(s) \sum_{k=0}^{K} \mu_k p_{k,s} = \sum_{s=0}^{\infty} g(s+1) \sum_{k=0}^{K} (k+1)\mu_{k+1} \sum_{m=0}^{s} p_{k,m} (s-m+1) p_{s-m+1},$$

or equivalently

$$\sum_{k=0}^{K} \mu_k \sum_{m=0}^{\infty} p_{k,m} \left(a_k \sum_{s=m}^{\infty} g(s+1)(s-m+1) p_{s-m+1} - m g(m) \right) = 0,$$

where $a_k = (k+1)\mu_{k+1}/\mu_k$. Changing the order of summation in the above equation and setting $l = s - m + 1$, we obtain

$$\sum_{m=0}^{\infty} \sum_{k=0}^{K} \mu_k p_{k,m} \left(a_k \sum_{l=1}^{\infty} g(l+m) l p_l - m g(m) \right) = 0. \tag{26}$$

Next, let us assume that a_k's satisfy Panjer's recursion: $a_k = a + bk$ (see [30]). From (21) and (26),

$$\sum_{m=0}^{\infty} \pi_m \left(a \sum_{l=1}^{\infty} g(l+m) l p_l - m g(m) \right) + b \sum_{m=0}^{\infty} \sum_{k=0}^{K} k \mu_k p_{k,m} \sum_{l=1}^{\infty} g(l+m) l p_l = 0. \tag{27}$$

Let X be an independent copy of X_1. Then $\mathbb{E} g(S_k + X) X = \mathbb{E} g(S_k + X) X_i$, $(i = 1, 2, \dots, k)$. Therefore,

$$\sum_{m=0}^{\infty} k p_{k,m} \sum_{l=1}^{\infty} g(l+m) l p_l = k \mathbb{E} g(S_k + X) X = \sum_{i=1}^{k} \mathbb{E} g(S_k + X) X_i = \mathbb{E} S_k g(S_k + X)$$

and

$$\begin{aligned}
\sum_{m=0}^{\infty} \sum_{k=0}^{K} k \mu_k p_{k,m} \sum_{l=1}^{\infty} g(l+m) l p_l &= \sum_{k=0}^{K} \mu_k \mathbb{E} S_k g(S_k + X) \\
&= \sum_{k=0}^{K} \mu_k \sum_{m=0}^{\infty} m p_{k,m} \sum_{l=0}^{\infty} g(l+m) p_l \\
&= \sum_{m=0}^{\infty} \pi_m m \sum_{l=0}^{\infty} g(l+m) p_l.
\end{aligned}$$

Substituting the last expression into (27), we obtain a Stein operator as

$$(\mathcal{A} g)(j) = \sum_{l=1}^{\infty} (al + bj) g(l+j) p_l - (1 - b p_0) j g(j), \qquad j \in \mathbb{Z}_+. \tag{28}$$

Thus, we have proved the following result.

Theorem 2.6. *Let N be r.v. concentrated on $\{0, 1, 2, \dots, K\}$ (K may be infinite) with distribution $\mu_k = P(N = k)$ satisfying Panjer's recursion, for some $a, b \in \mathbb{R}$,*

$$\frac{(k+1)\mu_{k+1}}{\mu_k} = a + bk, \qquad k = 0, 1, \dots, K,$$

with $\mu_{K+1} = 0$. Let $S_N = \sum_{j=1}^{N} X_j$, where the X_j are i.i.d. r.v.s independent of N and concentrated on $\mathbb{Z}_+$ with probabilities $P(X_1 = k) = p_k$. If $\mathbb{E}(S_N) < \infty$ and $g \in \mathcal{G}_{S_N}$, then a Stein operator for S_N is given by (28).

2.4. Some examples

(a) Let $N \sim P(\lambda)$, $\lambda > 0$. Applying Theorem 2.6 with $K = \infty$, $a = \lambda$ and $b = 0$, we obtain

$$(\mathcal{A}g)(j) = \lambda \sum_{l=1}^{\infty} lg(l+j)p_j - jg(j),$$

which coincides with the one given in (10) with $\lambda_j = \lambda p_j$.

(b) Let $N \sim NB(r, \bar{p})$, the negative binomial distribution, $r > 0$ and $0 < \bar{p} < 1$. Then $K = \infty$, $a = r\bar{q}$, $b = \bar{q}$ and a Stein operator for the compound negative binomial distribution is

$$\begin{aligned}(\mathcal{A}g)(j) &= \bar{q} \sum_{m=1}^{\infty} (rm + j)g(j+m)p_m - (1 - \bar{q}p_0)jg(j) \\ &= \sum_{m=1}^{\infty} p_m\big\{\bar{q}(rm + j)g(j+m) - jg(j)\big\} - \bar{p}p_0 jg(j) \\ &= \bar{q}(r\mathbb{E}X_1 + j) - jg(j) - p_0\bar{q}j\Delta g(j) \\ &\quad + \bar{q}\sum_{m=2}^{\infty}(rm + j)p_m \sum_{k=1}^{m-1} \Delta g(j+k).\end{aligned} \tag{29}$$

Note that the *PGF* of S_N is

$$G_{S_N}(z) = \left(\frac{\bar{p}}{1 - \bar{q}G_X(z)}\right)^r = \left(\frac{\bar{p}}{1 - \bar{q}\sum_{j=0}^{\infty} p_j z^j}\right)^r.$$

(c) Let $N \sim \mathrm{Bi}(n, p)$, the binomial distribution, where $n \in \mathbb{N}$ (the set of natural numbers) and $0 < p < 1$. Then $K = n$, $a = np/q$, $b = -p/q$ and a Stein operator for the compound binomial distribution is given by

$$(\mathcal{A}g)(j) = (p/q)\sum_{m=1}^{\infty}(nm - j)g(j+m)p_m - \big(1 + (p/q)p_0\big)jg(j)$$

which can be written, in a form similar to (5), as

$$\begin{aligned}(\mathcal{A}g)(j) &= p\sum_{m=1}^{\infty}(nm - j)g(j+m)p_m - (q + pp_0)jg(j) \\ &= p(n\mathbb{E}X_1 - j)g(j+1) - qjg(j) \\ &\quad + pp_0 j\Delta g(j) + \sum_{m=2}^{\infty}(nm - j)p_m \sum_{k=1}^{m-1}\Delta g(j+k).\end{aligned} \tag{30}$$

Also, in this case

$$G_{S_N}(z) = \big(1 + p\big(G_X(z) - 1\big)\big)^n = \left(1 + p\sum_{j=0}^{\infty} p_j\big(z^j - 1\big)\right)^n.$$

Remark 2.7. (i) If we take $p_1 = 1$ in the examples above, we obtain the standard Stein operators for Poisson, binomial and negative binomial distributions, as given by (4), (5) and (6), respectively.

(ii) Sometimes the form of *PGF* allows to establish recursive relations without differentiation. For example, the *PGF* for the compound geometric distribution is of the form

$$\frac{p}{1 - q\sum_{m=1}^{\infty} p_m z^m} = \sum_{k=0}^{\infty} \mu_k z^k.$$

Multiplying both sides by $1 - q\sum_{m=1}^{\infty} p_m z^m$ and collecting factors corresponding to z^k, we obtain

$$(\mathcal{A}g)(j) = q\sum_{m=1}^{\infty} p_m g(j+m) - g(j).$$

This operator coincides with the one from [14]. Note in this example $p_0 = 0$.

3. Perturbed solutions to the Stein equation

In this section, we discuss some known facts and explore properties of exact and approximate solutions to the Stein equation. Assume that Y and Z are r.v.s concentrated on $\mathbb{Z}_+$, $f \in \mathcal{F}$ and $g \in \mathcal{G}_Y$. Henceforth, $\|f\| = \sup_k |f(k)|$. As mentioned in Section 1, the second step in Stein's method is solving the equation (1). Suppose a Stein operator for Y is given by

$$(\mathcal{A}g)(j) = \alpha_j g(j+1) - \beta_j g(j), \tag{31}$$

where $\beta_0 = 0$ and $\alpha_k - \alpha_{k-1} \le \beta_k - \beta_{k-1}$ $(k = 1, 2, \ldots)$. Then a solution g to (1) satisfies

$$\big|\Delta g(j)\big| \le 2\|f\| \min\left\{\frac{1}{\alpha_j}, \frac{1}{\beta_j}\right\}, \qquad j \in \mathbb{Z}_+, f \in \mathcal{F}. \tag{32}$$

Define g_i as a solution to (1) for the choice $f(j) = I(j = i)$, where $I(A)$ denotes the indicator function of A. Then, from (2.18) and Theorem 2.10 of [10], we have

$$\big|\Delta g(i)\big| = \left|\sum_{j=0}^{\infty} f(j)\Delta g_j(i)\right| \le \sup_{j\ge 0} f(j)\big|\Delta g_i(i)\big| \le \sup_{j\ge 0} f(j)\min\big\{\alpha_i^{-1}, \beta_i^{-1}\big\}, \tag{33}$$

for nonnegative functions f. The proof of (32) can now be completed by following steps similar to that of Lemma 2.2 from [2], by noting the fact Stein equations with $f^+(j)(:=$

$f(j) - \inf_k f(k) \geq 0$) and $f(j)$ on the right-hand side of (1) have the same solution. If f is nonnegative, then $f^+(j)$ is not needed and $2\|f\|$ in (32) can be replaced by $\|f\|$. Therefore, if $f : \mathbb{Z}_+ \to [0, 1]$, then $2\|f\|$ in (32) should be replaced by 1.

Note that different choices of f lead to different probabilistic metrics. In this paper, we consider total variation norm which is twice the total variation metric. That is,

$$\begin{aligned}\big\|\mathcal{L}(Y) - \mathcal{L}(Z)\big\|_{\mathrm{TV}} &= \sum_{j=0}^{\infty}\big|P(Y = j) - P(Z = j)\big| = \sup_{\|f\|\leq 1}\big|\mathbb{E}f(Y) - \mathbb{E}f(Z)\big| \\ &= 2\sup_{f\in\mathcal{F}_1}\big|\mathbb{E}f(Y) - \mathbb{E}f(Z)\big| = 2\sup_A\big|P(Y \in A) - P(Z \in A)\big|,\end{aligned}$$

where $\mathcal{F}_1 = \{f | f : \mathbb{Z}_+ \to [0, 1]\}$, and the supremum is taken over all Borel sets in the last equality.

Let g be the solution to (1) for Poisson or negative binomial or pseudo-binomial r.v. with Stein operator given by (4) or (6) or (5), respectively. Then the corresponding bounds are given respectively, as

$$\|\Delta g\| \leq \frac{2\|f\|}{\max(1, \lambda)}, \qquad \|\Delta g\| \leq \frac{2\|f\|}{r\bar{q}}, \qquad \|\Delta g\| \leq \frac{2\|f\|}{\lfloor\widetilde{N}\rfloor pq}. \tag{34}$$

The first two bounds follow directly from (32). Observe that for pseudo-binomial distribution, the assumptions of (32) are not always satisfied. The last bound of (34) follows from Lemma 9.2.1 in [7], and using similar arguments as above.

If a Stein operator has a form different from (31), then solving (1) and checking properties similar to (32) becomes rather tedious. Apart from the solution for compound geometric distribution by [14], some partial success has been achieved for compound Poisson distribution by [8]. In such situations, one can try the perturbation technique introduced in [9] and further developed in [3] and [4]. Roughly, the basic idea of perturbation can be summarized in the following way: good properties of the solution of (1) can be carried over to solutions of Stein operators in similar forms.

Next, we formulate a partial case of Lemma 2.3 and Theorem 2.4 from [4] under following setup.

Let $\mathcal{A}_0$ be a Stein operator for r.v. Y with support $\{0, 1, 2, \dots, K\}$ ($K = \infty$ is allowed) and g_0 be the solution of the Stein equation

$$(\mathcal{A}_0 g_0)(j) = f(j) - \mathbb{E}f(Y), \qquad f \in \mathcal{F}, g_0 \in \mathcal{G}_Y.$$

Also, let there exist $\omega_1, \gamma > 0$ such that $\|\Delta g_0\| \leq \omega_1\|f\|\min(1, \gamma^{-1})$. Let $\mathcal{A}$ denote a Stein operator for r.v. Z and $U := \mathcal{A} - \mathcal{A}_0$ be the perturbed part of $\mathcal{A}$ with respect to $\mathcal{A}_0$.

The following lemma establishes, under certain conditions, an approximation result between any two r.v.s W and Z, using the observation that a Stein operator for the r.v. Z can be seen as perturbation of a Stein operator for r.v. Y.

Lemma 3.1. *Let Z be a r.v. with a Stein operator $\mathcal{A} = \mathcal{A}_0 + U$ and W be another r.v., both concentrated on $\mathbb{Z}_+$. Also, assume that, for $g \in \mathcal{G}_Y \cap \mathcal{G}_Z$, there exist $\omega_2, \varepsilon > 0$ such that*

$$\|Ug\| \le \omega_2 \|\Delta g\|, \qquad \left|\mathbb{E}(\mathcal{A}g)(W)\right| \le \varepsilon \|\Delta g\|,$$

and $\omega_1 \omega_2 < \gamma$. Then

$$\left\|\mathcal{L}(W) - \mathcal{L}(Z)\right\|_{\mathrm{TV}} \le \frac{\gamma}{\gamma - \omega_1 \omega_2}\left(\varepsilon \omega_1 \min\left(1, \gamma^{-1}\right) + 2P(Z > K) + 2P(W > K)\right).$$

Next, using the assumptions of Lemma 3.1 and (34), we evaluate the values of ω_1, ω_2 and γ to the various Stein operators derived in Section 2. Our observations are as follows:

(O1) If a Stein operator is given by (10), then we have the Poisson perturbation with $\omega_1 = 2$, $\gamma = \sum_{m=1}^{\infty} m\lambda_m$,

$$\|Ug\| \le \|\Delta g\| \sum_{m=2}^{\infty} m(m-1)|\lambda_m| = \|\Delta g\| \omega_2$$

and $\omega_1 \omega_2 < \gamma$, provided $\{\lambda_m\}_{m \ge 2}$ is sufficiently small. For a general description of the problem, see [6].

(O2) For the Stein operator given by (14), we have the pseudo-binomial perturbation with $\omega_1 = 2/pq$, $\gamma = \lfloor M + \alpha/p \rfloor$, $\omega_2 = p\alpha$ and $\omega_1 \omega_2 < \gamma$, if p is sufficiently small (see Theorem 4.4).

(O3) Consider the Stein operator given by (15). Then we have the Poisson perturbation with $\omega_1 = 2$, $\gamma = Mp + \alpha$, $\omega_2 = Mp^2/(q-p)^2$ and $\omega_1 \omega_2 < \gamma$, whenever p is sufficiently small (see Theorem 4.1).

(O4) For the Stein operator given by (18), we have the pseudo-binomial perturbation with $\omega_1 = 2$, $\gamma = \lfloor M + r\bar{q}/(p\bar{p}) \rfloor pq$ and $\omega_2 = \frac{r\bar{q}(q\bar{q}+p)}{\bar{p}^2}$. The condition $\omega_1 \omega_2 < \gamma$ is satisfied if p and $\bar{q}$ are sufficiently small.

(O5) If the Stein operator is given by (19), then we have the negative binomial perturbation with $\omega_1 = 2$, $\gamma = Mp\bar{p} + r\bar{q}$, $\omega_2 = Mpq(p/q + \bar{q})(q-p)^{-2}$ and $\omega_1 \omega_2 < \gamma$, provided p and $\bar{q}$ are sufficiently small.

(O6) Finally, consider the Stein operator given by (20). Then we have the Poisson perturbation, $\omega_1 = 2$, $\gamma = Mp + r\bar{q}/\bar{p}$, $\omega_2 = Mp^2/(q-p)^2 + r\bar{q}^2/\bar{p}^2$ and $\omega_1 \omega_2 < \gamma$, whenever p and $\bar{q}$ are sufficiently small.

Remark 3.2. (i) Note that, for the Stein operator in (17), the perturbation approach is not applicable. Also, for compound negative binomial or compound binomial distributions, the perturbation part of the operator contains j, which makes the perturbation technique inapplicable, as the upper bound for $\|Ug\|$ cannot be established. Consequently, either a new version of perturbation technique with nonuniform bounds should be developed or a different approach should be devised.

(ii) We also remark here that once a Stein operator is derived (as discussed in Section 2), the properties of the associated exact solution to the Stein equation must be derived and this can be quite difficult. The perturbation approach, as discussed in some examples above (see (O1)–(O6)), can be useful to get an upper bound on approximate solution to the Stein equation.

4. Application to sums of indicator variables

In this section, we exploit the different forms of Stein operator to obtain better bounds for the approximation problems to sums of possibly dependent indicator r.v.s. In particular, we consider Stein operators derived in (14) and (15) along with the corresponding observations (O2) and (O3) and establish the approximation results to the sums of independent and dependent indicators.

Consider the sum $W = \sum_{i=1}^{n} \mathbb{I}_i$ of possibly dependent indicator variables and let $W^{(i)} = W - \mathbb{I}_i$, $P(\mathbb{I}_i = 1) = p_i = 1 - P(\mathbb{I}_i = 0) = 1 - q_i$ $(i = 1, 2, \dots, n)$. Assume also $\widetilde{W}^{(i)}$ satisfy $P(\widetilde{W}^{(i)} = k) = P(W^{(i)} = k|\mathbb{I}_i = 1)$, for all k. We choose $Y_{12} = Y_1 + Y_2$ as the approximating variable, where $Y_1 \sim P(\alpha)$, $Y_2 \sim \mathrm{Bi}(M, p)$ and are independent. Denote its distribution by BCP whose *PGF* is given in (11). Poisson, signed compound Poisson and translated Poisson, binomial and negative binomial approximations have been applied to the sums of independent and dependent Bernoulli variables in numerous papers; see, for example, [7,9,16,31,33,34,36] and [41]. Unlike asymptotic expansions or a signed compound Poisson measure, BCP is a distribution. This might be an added advantage in practical applications.

4.1. The choice of parameters

Note that the BCP is a three-parametric distribution. We choose the parameters p, M and α to ensure the almost matching of the first three moments of W. Denoting as before the integral part by $\lfloor\cdot\rfloor$, we define

$$M := \left\lfloor \left(\sum_{i=1}^{n} p_i^2\right)^3 \left(\sum_{i=1}^{n} p_i^3\right)^{-2} \right\rfloor, \tag{35}$$

$$\delta := \left(\sum_{i=1}^{n} p_i^2\right)^3 \left(\sum_{i=1}^{n} p_i^3\right)^{-2} - M, \qquad 0 \le \delta < 1, \tag{36}$$

$$p := \left(\sum_{i=1}^{n} p_i^3\right)\left(\sum_{i=1}^{n} p_i^2\right)^{-1}; \qquad \alpha := \sum_{i=1}^{n} p_i - Mp. \tag{37}$$

Then the following relations hold:

$$Mp^2 = \sum_{i=1}^{n} p_i^2 - \delta p^2, \qquad Mp^3 = \sum_{i=1}^{n} p_i^3 - \delta p^3. \tag{38}$$

Observe also that

$$\left(\sum_{i=1}^{n} p_i^2\right)^2 \le \sum_{i=1}^{n} p_i \sum_{i=1}^{n} p_i^3.$$

Therefore, for $\alpha > 0$, the BCP is not a signed measure, but a distribution. Similar to [36], we choose parameters to match the three moments for the sum of independent Bernoulli variables.

Thus, only weak dependence of r.v.s is assumed. Note that the additional information about the dependence of r.v.s can significantly alter the choice of parameters, see, for example, [16] and Corollary 4.8. Observe also that α and Mp can be of the same order. Indeed, let n be even and $p_1 = p_2 = \cdots = p_{n/2} = 1/6$, $p_{n/2+1} = \cdots = p_n = 1/12$. Then $Mp = O(n) = \alpha$.

4.2. Poisson perturbation

We start with the Stein operator given in (15). Some additional notations are needed. Henceforth, let I_1 and I denote the degenerate distributions concentrated at 1 and 0, respectively. The convolution operator is denoted by $*$. Also, let

$$d := \big\|\mathcal{L}(W) * (I_1 - I)^{*2}\big\|_{\mathrm{TV}} = \sum_{k=0}^{n} \big|\Delta^2 P(W = k)\big|, \tag{39}$$

$$d_1 := \max_i \big\|\mathcal{L}\big(W^{(i)}\big) * (I_1 - I)^{*2}\big\|_{\mathrm{TV}} = \max_i \sum_{k=0}^{n} \big|\Delta^2 P\big(W^{(i)} = k\big)\big|, \tag{40}$$

$$\widehat{\lambda} = \sum_{i=1}^{n} p_i, \qquad \sigma^2 = \sum_{i=1}^{n} p_i q_i, \qquad \tau = \max_i p_i q_i,$$

$$\eta_1 := \sum_{i=1}^{n} p_i\big(1 + 2p_i + 4p_i^2\big)\mathbb{E}\big|\widetilde{W}^{(i)} - W^{(i)}\big|,$$

$$\theta_1 := \frac{Mp^2}{(1-2p)^2(Mp+\alpha)} = \frac{\sum_{i=1}^{n} p_i^2 - \delta p^2}{(1-2p)^2 \sum_{i=1}^{n} p_i}. \tag{41}$$

Now, we have the following BCP approximation result for the sum of weakly dependent indicator r.v.s.

Theorem 4.1. *Let* $\max(p, \theta_1) < 1/2$. *Then*

$$\big\|\mathcal{L}(W) - \mathrm{BCP}\big\|_{\mathrm{TV}} \le \frac{2}{(1-2\theta_1)\widehat{\lambda}} \left\{ d_1 \sum_{i=1}^{n} p_i^4 + \frac{dMp^4}{(1-2p)^2} + (1+2p)\delta p^2 + \eta_1 \right\}.$$

If the indicator variables are dependent, then obtaining the bounds for d and d_1 is difficult; see Lemma 4.7 and [15] for some partial cases and the history of the problem. On the other hand, if the r.v.s are independent, then by the unimodality of W (see [42]), we obtain

$$P(W = k) \le \frac{1}{2\sigma}, \qquad \big\|\mathcal{L}(W) * (I_1 - I)\big\|_{\mathrm{TV}} \le \frac{1}{\sigma}.$$

Now let S_1 and S_2 be the sets of indices such that

$$S_1 \cup S_2 = \{1, 2, \dots, n\}, \qquad \sum_{i \in S_1} p_i q_i \geq \frac{\sigma^2}{2}, \qquad \sum_{i \in S_2} p_i q_i \geq \frac{\sigma^2 - \tau}{2}.$$

Then, by the properties of total variation,

$$d \leq \left\| \mathcal{L}\left(\sum_{i \in S_1} \mathbb{I}_i \right) * (I_1 - I) \right\|_{\mathrm{TV}} \left\| \mathcal{L}\left(\sum_{i \in S_2} \mathbb{I}_i \right) * (I_1 - I) \right\|_{\mathrm{TV}} \leq \frac{2}{\sigma \sqrt{\sigma^2 - \tau}}. \tag{42}$$

Similarly,

$$d_1 \leq \frac{2}{\sqrt{(\sigma^2 - \tau)(\sigma^2 - 3\tau)}}. \tag{43}$$

Thus, we have the following corollary for independent r.v.s.

Corollary 4.2. *Let W be the sum of n independent Bernoulli r.v.s with success probabilities p_i,* $\max(p, \theta_1) < 1/2$ *and $\sigma^2 > 3\tau$. Then*

$$\begin{aligned} &\left\| \mathcal{L}(W) - \mathrm{BCP} \right\|_{\mathrm{TV}} \\ &\quad \leq \frac{2}{(1 - 2\theta_1)\widehat{\lambda}} \left\{ \frac{2 \sum_{i=1}^n p_i^4}{\sqrt{(\sigma^2 - \tau)(\sigma^2 - 3\tau)}} + \frac{2Mp^4}{(1 - 2p)^2 \sigma \sqrt{\sigma^2 - \tau}} + (1 + 2p)\delta p^2 \right\}. \end{aligned} \tag{44}$$

Remark 4.3. (i) Observe that $\theta_1 < p(1 - 2p)^{-2} \leq \max_i p_i (1 - 2\max_i p_i)^{-2}$. Therefore, a sufficient condition for $\max(p, \theta_1) < 1/2$ is $\max_i p_i < (3 - \sqrt{5})/4 = 0.19098\dots$.

(ii) If all $p_i \asymp C$, then the order of accuracy of the bound in (44) is $O(n^{-1})$. In comparison to the Edgeworth expansion, the BCP is more advantageous since the approximation holds for the total variation norm and no additional measures compensating for the difference in supports are needed.

(iii) Also, one can compare (44) with the classical Poisson approximation result (see [13], equations (1.1)–(1.2)), where for $p_i \asymp C$ and the order of accuracy is $O(1)$.

Proof of Theorem 4.1. Applying Newton's expansion, similar to [3], page 518, we get

$$\begin{aligned} &\left| \mathbb{E}\Delta g(W + k) - \mathbb{E}\Delta g(W + 1) - (k - 1)\mathbb{E}\Delta^2 g(W + 1) \right| \\ &\quad \leq \sum_{s=1}^{k-2} (k - 1 - s) \left| \mathbb{E}\Delta^3 g(W + s) \right| \\ &\quad \leq \sum_{s=1}^{k-2} (k - 1 - s) \left| \sum_{j=0}^{\infty} \Delta g(j + s) \Delta^2 P(W = j - 2) \right| \\ &\quad \leq \frac{(k - 1)(k - 2)}{2} \|\Delta g\| d. \end{aligned} \tag{45}$$

By the definition of M and p, defined respectively, in (35) and (37),

$$\begin{aligned}
-M\sum_{l=2}^{\infty}\left(\frac{-p}{q}\right)^{l}(l-1) &= -\sum_{k=1}^{n}p_k^2+\delta p^2,\\
-M\sum_{l=2}^{\infty}\left(\frac{-p}{q}\right)^{l}\sum_{k=1}^{l-1}(k-1) &= \sum_{k=1}^{n}p_k^3-\delta p^3,\\
M\sum_{k=2}^{\infty}\left(\frac{p}{q}\right)^{l}\sum_{k=1}^{l-1}(k-1)(k-2) &= \frac{2Mp^4}{(1-2p)^4}.
\end{aligned}\tag{46}$$

Therefore, from (45) and (46), we get

$$\begin{aligned}
&\left|-M\sum_{l=2}^{\infty}\left(\frac{-p}{q}\right)^{l}\sum_{k=1}^{l-1}\mathbb{E}\Delta g(W+k)+\sum_{i=1}^{n}p_i^2\mathbb{E}\Delta g(W+1)-\sum_{i=1}^{n}p_i^3\mathbb{E}\Delta^2 g(W+1)\right|\\
&\quad\le \frac{Mp^4}{(1-2p)^4}\|\Delta g\|d+\left|\delta p^2\mathbb{E}\Delta g(W+1)\right|+\left|\delta p^3\mathbb{E}\Delta^2 g(W+1)\right|\\
&\quad\le \frac{Mp^4}{(1-2p)^4}\|\Delta g\|d+\delta p^2(1+2p)\|\Delta g\|.
\end{aligned}\tag{47}$$

Taking into account (15) and (47), we obtain

$$\begin{aligned}
\left|\mathbb{E}(\mathcal{A}g)(W)\right| &\le \left|\mathbb{E}\left\{\sum_{i=1}^{n}p_i g(W+1)-Wg(W)\right\}-\sum_{i=1}^{n}p_i^2\mathbb{E}\Delta g(W+1)\right.\\
&\qquad\left.+\sum_{i=1}^{n}p_i^3\mathbb{E}\Delta^2 g(W+1)\right|+\|\Delta g\|\left(\frac{Mp^4 d}{(1-2p)^2}+\delta p^2(1+2p)\right)\\
&\le J_1+J_2+J_3+\|\Delta g\|\left(\frac{Mp^4 d}{(1-2p)^2}+\delta p^2(1+2p)\right)\qquad\text{(say)}.
\end{aligned}\tag{48}$$

Here,

$$\begin{aligned}
J_1 &= \left|\mathbb{E}\left\{\sum_{i=1}^{n}p_i g(W+1)-Wg(W)\right\}-\sum_{i=1}^{n}p_i^2\mathbb{E}\left\{\Delta g\left(W^{(i)}+1\right)\middle|\mathbb{I}_i=1\right\}\right|\\
&\le \left|\sum_{i=1}^{n}p_i q_i\left(\mathbb{E}\left\{g\left(W^{(i)}+1\right)\middle|\mathbb{I}_i=0\right\}-\mathbb{E}\left\{g\left(W^{(i)}+1\right)\middle|\mathbb{I}_i=1\right\}\right)\right|\\
&= \left|\sum_{i=1}^{n}p_i\mathbb{E}\left(g\left(W^{(i)}+1\right)-g\left(\widetilde{W}^{(i)}+1\right)\right)\right|
\end{aligned}\tag{49}$$

$$\leq \|\Delta g\| \sum_{i=1}^{n} p_i \mathbb{E}\big|W^{(i)} - \widetilde{W}^{(i)}\big|.$$

Similarly,

$$\begin{aligned} J_2 &= \Bigg| \sum_{i=1}^{n} p_i^2 \mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\} - \sum_{i=1}^{n} p_i^2 \mathbb{E}\Delta g(W+1) \\ &\quad + \sum_{i=1}^{n} p_i^3 \mathbb{E}\big\{\Delta^2 g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\}\Bigg| \\ &= \Bigg| \sum_{i=1}^{n} p_i^2 q_i \big(\mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 0\big\} - \mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\}\big)\Bigg| \\ &\leq 2\|\Delta g\| \sum_{i=1}^{n} p_i^2 \mathbb{E}\big|W^{(i)} - \widetilde{W}^{(i)}\big| \end{aligned} \tag{50}$$

and

$$\begin{aligned} J_3 &= \Bigg| \sum_{i=1}^{n} p_i^3 \mathbb{E}\Delta^2 g(W+1) - \sum_{i=1}^{n} p_i^3 \mathbb{E}\big\{\Delta^2 g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\}\Bigg| \\ &\leq \sum_{i=1}^{n} p_i^3 q_i \big|\mathbb{E}\big\{\Delta^2 g\big(W^{(i)}+1\big)|\mathbb{I}_i = 0\big\} - \mathbb{E}\big\{\Delta^2 g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\}\big| \\ &\quad + \sum_{i=1}^{n} p_i^4 \big|\mathbb{E}\Delta^3 g\big(\widetilde{W}^{(i)}+1\big)\big| \\ &\leq \big\|\Delta^3 g\big\| \sum_{i=1}^{n} p_i^3 \mathbb{E}\big|W^{(i)} - \widetilde{W}^{(i)}\big| + \sum_{i=1}^{n} p_i^4 \|\Delta g\| d_1. \end{aligned} \tag{51}$$

Collecting the bounds in (47)–(51), applying Lemma 3.1 and (O3) with $T = \infty$, the proof is completed. □

4.3. Binomial perturbation

Here, we approximate W using Stein operator in (14). In addition to the notations used above, let

$$d_2 := \max_{i,j} \big\|\mathcal{L}\big(W^{(ij)}\big) * (I_1 - I)\big\|_{\mathrm{TV}} = \max_{i,j} \sum_{k} \big|\Delta P\big(W^{(ij)} = k\big)\big|,$$

$$\widehat{T} := \lfloor M + \alpha/p \rfloor, \qquad \theta_2 := \frac{\alpha}{q\widehat{T}}, \qquad W^{(ij)} = W - \mathbb{I}_i - \mathbb{I}_j.$$

Also, let the distribution of $\widetilde{W}_i^{(ij)}$ satisfy $P(\widetilde{W}_i^{(ij)} = k) = P(W^{(ij)} = k|\mathbb{I}_i = 1)$, for all k.

Theorem 4.4. *Let $\theta_2 < 1/2$. Then*

$$\begin{aligned}
&\|\mathcal{L}(W) - \mathrm{BCP}\|_{\mathrm{TV}} \\
&\quad \le \frac{2}{pq\widehat{T}(1-2\theta_2)}\Bigg\{d_2\Bigg(\sum_{i=1}^n p_i^4 - p\sum_{i=1}^n p_i^3\Bigg) + \delta p^2 \\
&\qquad + \sum_{i=1}^n p_i\big(2 + 2|p_i - p|\big)\mathbb{E}\big|\widetilde{W}^{(i)} - W^{(i)}\big| \\
&\qquad + \Bigg(2\sum_{k=1}^n p_k^2\Bigg)^{-1} \sum_{i,j=1}^n p_i p_j |p_i - p_j| \big[d_2|p_i - p_j|\big|\mathrm{Cov}(\mathbb{I}_i, \mathbb{I}_j)\big| \\
&\qquad + 4p_i p_j \mathbb{E}\big|\widetilde{W}_i^{(ij)} - \widetilde{W}^{(ij)}\big|\big]\Bigg\} + \frac{2}{1-2\theta_2}\big(P(Y_1 + Y_2 > \widehat{T}) + P(W > \widehat{T})\big).
\end{aligned} \tag{52}$$

When the indicator r.v.s are independent, a bound for the term d_2, similar to the one in (43) for d_1, can be obtained. This leads to the following corollary.

Corollary 4.5. *Let W be the sum of n independent indicator r.v.s, $\theta_2 < 1/2$ and $\sigma^2 > 3\tau$. Then*

$$\begin{aligned}
\|\mathcal{L}(W) - \mathrm{BCP}\|_{\mathrm{TV}} \le \frac{2}{1-2\theta_2}\Bigg\{&\frac{4}{pq\widehat{T}(\sigma^2 - 3\tau)}\Bigg(\sum_{i=1}^n p_i^4 - p\sum_{i=1}^n p_i^3\Bigg) \\
&+ \delta p^2 + P(W > \widehat{T}) + P(Y_1 + Y_2 > \widehat{T})\Bigg\}.
\end{aligned} \tag{53}$$

Remark 4.6. (i) If the r.v.s are independent, then

$$P(W > \widehat{T}) + P(Y_1 + Y_2 > \widehat{T}) \le \exp\big\{-\widehat{\lambda}\psi(p)\big\},$$

where $\psi(p) = (p)^{-1}(-\ln p - 1) + 1$. Indeed,

$$\begin{aligned}
P(Y_1 + Y_2 \ge \widehat{T} + 1) &\le \mathrm{e}^{-x(\widehat{T}+1)}\mathbb{E}\mathrm{e}^{xY_1}\mathbb{E}\mathrm{e}^{xY_2} \le \mathrm{e}^{-x(\widehat{T}+1)}\exp\big\{\alpha\big(\mathrm{e}^x - 1\big)\big\}\big(q + p\mathrm{e}^x\big)^M \\
&\le \exp\big\{-x\widehat{\lambda}/p + (Mp + \alpha)\big(\mathrm{e}^x - 1\big)\big\} \le \exp\big\{-\widehat{\lambda}\big(x/p + 1 - \mathrm{e}^x\big)\big\}.
\end{aligned}$$

Now it suffices to take $x = -\ln p$. Similarly, one can obtain a bound for $P(W > \widehat{T})$. Observe, that $\psi(p) > 0$ for any $p < 1$.

(ii) If $p_i = C$, then the bound in (53) is at least of the order $O(n^{-1})$. The corresponding bounds for the binomial approximation as given in Corollary 1.3 of [36] are of order $O(n)$ and the ones in Remarks 2 of [34] are of order $O(n^{-1/2})$. Also, see Theorem 1 of [18] where the bound is of order $O(1)$.

(iii) If all the p_i are equal, then both sides of (53) are equal to zero, as is the case for the binomial approximation (see [36]).

(iv) Comparing Theorem 4.4 with Theorem 4.1, we observe that both have similar accuracy with respect to $\widehat{\lambda}$. On the other hand, Theorem 4.4 reflects the closeness of p_i and, in this sense, is more accurate than Theorem 4.1.

(v) The BCP approximation (matching the first three moments) provides bounds with better accuracy (see Theorems 4.1 and 4.4) than the bounds obtained (matching the two moments) for the binomial approximation (see [36] and [34]).

Proof of Theorem 4.4. Using (14) and (35)–(37), we get

$$(\mathcal{A}g)(j) = \sum_{i=1}^{n} p_i g(j+1) - jg(j) - pj\Delta g(j) + p\alpha\Delta g(j+1). \tag{54}$$

Therefore,

$$\begin{aligned}
&\big|\mathbb{E}(\mathcal{A}g)(W)\big| \\
&\quad = \Bigg|\sum_{i=1}^{n} p_i \mathbb{E}g(W+1) - \sum_{i=1}^{n} \mathbb{E}\mathbb{I}_i g(W) - p\sum_{i=1}^{n} p_i \mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\} \\
&\qquad + \Bigg(\sum_{i=1}^{n} p_i(p-p_i) + \delta p^2\Bigg)\mathbb{E}\Delta g(W+1)\Bigg| \\
&\quad \le \delta p^2 \big|\mathbb{E}\Delta g(W+1)\big| \qquad (55)\\
&\qquad + \Bigg|\sum_{i=1}^{n} p_i(p_i-p)\mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\} + \sum_{i=1}^{n} p_i(p-p_i)\mathbb{E}\Delta g(W+1)\Bigg| \\
&\qquad + \sum_{i=1}^{n} p_i q_i \big|\mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 0\big\} - \mathbb{E}\big\{\Delta g\big(W^{(i)}+1\big)|\mathbb{I}_i = 1\big\}\big| \\
&\quad = R_1 + R_2 + R_3 \qquad \text{(say)}.
\end{aligned}$$

It is easy to check that

$$|R_1| \le \|\Delta g\|\delta p^2, \tag{56}$$

$$\begin{aligned}
|R_3| &= \sum_{i=1}^{n} p_i \big|\mathbb{E}\Delta g\big(W^{(i)}+1\big) - \mathbb{E}g\big(\Delta\widetilde{W}^{(i)}+1\big)\big| \\
&\le 2\|\Delta g\| \sum_{i=1}^{n} p_i \mathbb{E}\big|\widetilde{W}^{(i)} - W^{(i)}\big|,
\end{aligned} \tag{57}$$

$$
\begin{aligned}
|R_2| \le & \left| \sum_{i=1}^{n} p_i^2 (p - p_i) \mathbb{E}\{\Delta^2 g(W^{(i)} + 1) | \mathbb{I}_i = 1\} \right| \\
& + \sum_{i=1}^{n} p_i q_i |p_i - p| \left| \mathbb{E}\{\Delta g(W^{(i)} + 1) | \mathbb{I}_i = 0\} - \mathbb{E}\{\Delta g(W^{(i)} + 1) | \mathbb{I}_i = 1\} \right|.
\end{aligned} \tag{58}
$$

The second summand in (58) is less than or equal to

$$
\|\Delta^2 g\| \sum_{i=1}^{n} p_i |p_i - p| \mathbb{E}\left|\widetilde{W}^{(i)} - W^{(i)}\right| \le 2\|\Delta g\| \sum_{i=1}^{n} p_i |p_i - p| \mathbb{E}\left|\widetilde{W}^{(i)} - W^{(i)}\right|. \tag{59}
$$

Also, the first term in (58) is

$$
\begin{aligned}
& \sum_{i=1}^{n} p_i^2 (p_i - p) \mathbb{E}\{\Delta^2 g(W^{(i)} + 1) | \mathbb{I}_i = 1\} \\
& \quad = \sum_{k=1}^{n} \Delta^2 g(k) \sum_{i=1}^{n} p_i (p_i - p) P(\mathbb{I}_i = 1, W = k).
\end{aligned} \tag{60}
$$

Moreover,

$$
\begin{aligned}
& \sum_{i=1}^{n} p_i (p_i - p) P(\mathbb{I}_i = 1, W = k) \\
& \quad = \left(\sum_{k=1}^{n} p_k^2 \right)^{-1} \sum_{i,j} p_i p_j^2 (p_i - p_j) P(\mathbb{I}_i = 1, W = k) \\
& \quad = \left(2 \sum_{k=1}^{n} p_k^2 \right)^{-1} \left\{ \sum_{i,j} p_i p_j^2 (p_i - p_j) P(\mathbb{I}_i = 1, W = k) \right. \\
& \qquad \left. + \sum_{i,j} p_j p_i^2 (p_j - p_i) P(\mathbb{I}_j = 1, W = k) \right\}.
\end{aligned} \tag{61}
$$

Set

$$
P_{ij}(k) = P(W = k + 1 | \mathbb{I}_i = 1, \mathbb{I}_j = 1) - P(W = k | \mathbb{I}_i = 1, \mathbb{I}_j = 1). \tag{62}
$$

Then it can be seen (see [36], page 709) that

$$
\begin{aligned}
& P(\mathbb{I}_i = 1, W = k) \\
& \quad = P(\mathbb{I}_j = 1, W = k) + (p_i - p_j) P\left(W^{(ij)} = k - 1\right) + \operatorname{Cov}\left(\mathbb{I}_i - \mathbb{I}_j, \mathbb{I}\{W^{(ij)} = k - 1\}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
& p_i P\big(W^{(ij)} = k-1\big) - P(\mathbb{I}_i = 1, W = k) \\
&\quad = \big(p_i p_j + \operatorname{Cov}(\mathbb{I}_i, \mathbb{I}_j)\big) P_{ij}(k) - \operatorname{Cov}\big(\mathbb{I}_i, \mathbb{I}\big\{W^{(ij)} = k-1\big\}\big).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \sum_{i=1}^{n} p_i (p_i - p) P(\mathbb{I}_i = 1, W = k) \\
&\quad = \frac{1}{2\sum_k p_k^2} \bigg\{ \sum_{i,j} p_i^2 p_j (p_i - p_j)^2 P\big(W^{(ij)} = k-1\big) - \sum_{i,j} p_i p_j (p_j - p_i)^2 P(\mathbb{I}_i = 1, W = k) \\
&\qquad + \sum_{i,j} p_i^2 p_j (p_i - p_j) \operatorname{Cov}\big(\mathbb{I}_i - \mathbb{I}_j, \mathbb{I}\big\{W^{(ij)} = k-1\big\}\big) \bigg\} \\
&\quad = \frac{1}{2\sum_k p_k^2} \bigg\{ \sum_{i,j} p_i^2 p_j^2 (p_i - p_j)^2 P_{ij}(k) + \sum_{i,j} p_i p_j (p_j - p_i)^2 \operatorname{Cov}(\mathbb{I}_i, \mathbb{I}_j) P_{ij}(k) \\
&\qquad - \sum_{i,j} p_i p_j (p_i - p_j)^2 \operatorname{Cov}\big(\mathbb{I}_i, \mathbb{I}\big\{W^{(ij)} = k-1\big\}\big) \\
&\qquad + \sum_{i,j} p_i^2 p_j (p_i - p_j) \operatorname{Cov}\big(\mathbb{I}_i - \mathbb{I}_j, \mathbb{I}\big\{W^{(ij)} = k-1\big\}\big) \bigg\} \\
&\quad = \frac{1}{2\sum_k p_k^2} \bigg\{ \sum_{i,j} p_i^2 p_j^2 (p_i - p_j)^2 P_{ij}(k) + \sum_{i,j} p_i p_j (p_i - p_j)^2 \operatorname{Cov}(\mathbb{I}_i, \mathbb{I}_j) P_{ij}(k) \\
&\qquad + \sum_{i,j} p_i p_j^2 (p_i - p_j) \operatorname{Cov}\big(\mathbb{I}_i, \mathbb{I}\big\{W^{(ij)} = k-1\big\}\big) \bigg\}.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
& \left| \sum_{i=1}^{n} p_i^2 (p_i - p) \mathbb{E}\big\{\Delta^2 g\big(W^{(i)} + 1\big) | \mathbb{I}_i = 1\big\} \right| \\
&\quad \le \frac{1}{2\sum_k p_k^2} \Bigg\{ \left| \sum_{k=1}^{n} \Delta^2 g(k) \sum_{i,j} p_i^2 p_j^2 (p_i - p_j)^2 P_{ij}(k) \right| \\
&\qquad + \left| \sum_{k=1}^{n} \Delta^2 g(k) \sum_{i,j} p_i p_j (p_i - p_j)^2 \operatorname{Cov}(\mathbb{I}_i, \mathbb{I}_j) P_{ij}(k) \right|
\end{aligned} \tag{63}
$$

$$+2\left|\sum_{k=1}^{n}\Delta^2 g(k)\sum_{i,j} p_i p_j^2(p_i-p_j)\operatorname{Cov}\big(\mathbb{I}_i,\mathbb{I}\{W^{(ij)}=k-1\}\big)\right|\Bigg\}$$

$$= R_4+R_5+R_6 \qquad \text{(say)}.$$

We next derive upper bounds for R_4, R_5 and R_6 separately. First,

$$\begin{aligned}
R_4 &\le \frac{1}{2\sum_k p_k^2}\sum_{i,j} p_i^2 p_j^2 (p_i-p_j)^2\left|\sum_{k=1}^{n}\Delta^2 g(k)P_{ij}(k)\right| \\
&\le \frac{1}{2\sum_k p_k^2}\sum_{i,j} p_i^2 p_j^2 (p_i-p_j)^2\|\Delta g\|\sum_{k=1}^{n}\big|\Delta P_{ij}(k-1)\big| \qquad (64)\\
&\le d_2\|\Delta g\|\left(\sum_{k=1}^{n} p_k^4 - p\sum_{i=1}^{n} p_i^3\right).
\end{aligned}$$

Secondly,

$$R_5 \le \frac{d_2\|\Delta g\|}{2\sum_k p_k^2}\sum_{i,j} p_i p_j (p_i-p_j)^2\big|\operatorname{Cov}(\mathbb{I}_i,\mathbb{I}_j)\big|. \qquad (65)$$

Finally,

$$\begin{aligned}
\operatorname{Cov}\big(\mathbb{I}_i,\mathbb{I}\{W^{(ij)}=k-1\}\big) &= \mathbb{E}\mathbb{I}_i\mathbb{I}\{W^{(ij)}=k-1\} - p_i P\big(W^{(ij)}=k-1\big)\\
&= p_i P\big(W^{(ij)}=k-1|\mathbb{I}_i=1\big) - p_i P\big(W^{(ij)}=k-1\big).
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\left|\sum_{k=1}^{n}\Delta^2 g(k)\sum_{i,j} p_i p_j^2(p_i-p_j)\operatorname{Cov}\big(\mathbb{I}_i,\mathbb{I}\{W^{(ij)}=k-1\}\big)\right|\\
&\quad = \left|\sum_{i,j} p_i^2 p_j^2(p_i-p_j)\big(\mathbb{E}\Delta^2 g\big(\widetilde{W}^{(ij)}+1\big)-\mathbb{E}\Delta^2 g\big(W^{(ij)}+1\big)\big)\right|\\
&\quad \le 4\|\Delta g\|\sum_{i,j} p_i^2 p_j^2|p_i-p_j|\mathbb{E}\big|\widetilde{W}^{(ij)}-W^{(ij)}\big|
\end{aligned}$$

and

$$R_6 \le 2\frac{\|\Delta g\|}{\sum_k p_k^2}\sum_{i,j} p_i^2 p_j^2|p_i-p_j|\mathbb{E}\big|\widetilde{W}^{(ij)}-W^{(ij)}\big|. \qquad (66)$$

Collecting the bounds in (55)–(66), we get the required bound for the Stein operator defined in (54). Applying Lemma 3.1 and (O2), the proof is completed. $\square$

4.4. Application to (1, 1)-runs

We consider here a dependent setup arising out of independent Bernoulli trials. Let $\{X_j\}$ be a sequence of independent Be(p^*) variables and $a(p^*) = p^*(1-p^*)$. Define, for $j \geq 2$,

$$\mathbb{I}_j = X_j(1 - X_{j-1}) \quad \text{and} \quad W = \sum_{j=2}^{n} \mathbb{I}_j. \tag{67}$$

Then, it can be easily seen that

$$\mathbb{E}(W) = \sum_{j=2}^{n} P(\mathbb{I}_j = 1) = (n-1)\bigl(1-p^*\bigr)p^* = (n-1)a\bigl(p^*\bigr) \qquad \text{(say)}, \tag{68}$$

$$\mathbb{V}(W) = (n-1)a\bigl(p^*\bigr) + (5-3n)\bigl(a\bigl(p^*\bigr)\bigr)^2 \tag{69}$$

and

$$\mathbb{E}(W - \mathbb{E}W)^3 = (n-1)a\bigl(p^*\bigr) + (15-9n)a\bigl(p^*\bigr)^2 + 4(5n-11)a\bigl(p^*\bigr)^3. \tag{70}$$

This leads to the following choice of parameters:

$$M := \left\lfloor \frac{(3n-5)^2}{10n-22} \right\rfloor, \tag{71}$$

$$\delta := \frac{(3n-5)^2}{10n-22} - M, \qquad 0 \leq \delta < 1, \tag{72}$$

$$p := \left(\frac{10n-22}{3n-5}\right)a\bigl(p^*\bigr); \qquad \alpha := (n-1)a\bigl(p^*\bigr) - Mp. \tag{73}$$

Let us define

$$K_1 = \frac{288(1-3a(p^*))}{a(p^*)} \quad \text{and} \quad K_2 = \frac{4}{a(p^*)\sqrt{\min\{1-a(p^*), 1/2\}}}. \tag{74}$$

To apply Theorem 4.1, we need the following lemma.

Lemma 4.7. *Let $\{\mathbb{I}_j\}_{j\geq 2}$ and W be as defined in (67), d and d_1 be respectively defined in (39) and (40). Then, for $(n-2)a(p^*) \geq 8$,*

$$d \leq \frac{K_1}{n-1} + \frac{K_2}{\sqrt{n-1}} := \gamma(n-1), \tag{75}$$

$$d_1 \leq \frac{K_1}{n-2} + \frac{K_2}{\sqrt{n-2}}, \tag{76}$$

where K_1 and K_2 are as defined in (74).

An application of Theorem 4.1 leads to the following corollary.

Corollary 4.8. *Let W be as defined in* (67) *and θ_1 be as defined in* (41). *Assume* $\max(p,\theta_1)\leq 1/2$ *and* $(n-2)p^*(1-p^*)\geq 8$. *Then*

$$\begin{aligned}\|\mathcal{L}(W)-\mathrm{BCP}\|_{\mathrm{TV}}\leq \frac{2}{(1-2\theta_1)\widehat{\lambda}}\Bigg\{&\left(na^4(p^*)+\frac{Mp^4}{(1-2p)^2}\right)\left(\frac{K_1}{n-2}+\frac{K_2}{\sqrt{n-2}}\right)\\&+(1+2p)\delta p^2+(n-1)C_1\Bigg\},\end{aligned}\tag{77}$$

where $C_1=2\max\{1,2(1-a(p^*))\}a(p^*)(1+2a(p^*)+4a^2(p^*))(1-a(p^*)(1-a(p^*)))$.

Remark 4.9. The bound given in (77) is of order $O(1)$ and comparable to the existing bounds for Poisson approximation given in Theorem 2.1 of [40]. Also, it is an improvement over the bound given in Theorem 2.1 of [21] which is of order $O(n)$.

Proof of Lemma 4.7. Let $\rho_0=0$ and define the stopping times

$$\rho_j=\min\{l>\rho_{j-1}|\mathbb{I}_l=1\}.$$

From [25], the $T_j=\rho_j-\rho_{j-1}$ are i.i.d. having the *PGF*

$$\mathbb{E}(z^T)=\frac{a(p^*)z^2}{1-z+a(p^*)z^2}.$$

Hence, $\mathbb{E}(T)=1/a(p^*)$ and $\mathbb{V}(T)=1-3a(p^*)/(a^2(p^*))$. Observe now that $\rho_j=\sum_{i=1}^{j}T_i$ is the waiting time for jth occurrence of $\mathbb{I}_l$. Then it follows that the average number of occurrences in a sequence $\{\mathbb{I}_j\}_{2\leq j\leq n}$ is $(n-1)/\mathbb{E}(T)=(n-1)a(p^*)$. Suppose now $k=\lfloor(n-1)a(p^*)\rfloor+1$. Then $\rho_k=\sum_{j=1}^{k}T_j$ and by Proposition 4.6 of [9], we get

$$\|\mathcal{L}(\rho_k)*(I_1-I)\|_{\mathrm{TV}}\leq\frac{2}{(ka(p^*)\min\{u_1,1/2\})^{\frac{1}{2}}},$$

where $u_1=1-(1/2)\|\mathcal{L}(T)*(I_1-I)\|_{\mathrm{TV}}$. Now, it can be easily seen that $\|\mathcal{L}(T)*(I_1-I)\|_{\mathrm{TV}}=2a(p^*)$ which implies

$$\begin{aligned}\|\mathcal{L}(\rho_k)*(I_1-I)\|_{\mathrm{TV}}&\leq\frac{2}{(ka(p^*)\min\{1-a(p^*),1/2\})^{\frac{1}{2}}}\\&\leq\frac{2}{((n-1)(a(p^*))^2\min\{1-a(p^*),1/2\})^{\frac{1}{2}}}.\end{aligned}$$

Define maximal coupling (see [7], page 254)

$$2P(\rho_k\neq\rho_k')=\|\mathcal{L}(\rho_k)*(I_1-I)\|_{\mathrm{TV}}\leq\frac{2}{((n-1)(a(p^*))^2\min\{1-a(p^*),1/2\})^{\frac{1}{2}}}.\tag{78}$$

Let now $\rho'_k = \sum_{j=1}^{k} T'_j$ such that T_j's are i.i.d. and $\rho'_j = \rho'_{j-1} + T'_j$ with $\rho'_0 = 0$. Define now

$$\mathbb{I}_i = \begin{cases} 0, & \rho'_{j-1} < i < \rho'_j;\ 1 \le j \le k, \\ 1, & \rho'_j = i;\ 1 \le j \le k, \\ \mathbb{I}_i, & \rho'_k < i. \end{cases}$$

Then, for $\rho_k \le (n-1)$ and $\rho_k = \rho'_k + 1$, we have $W = W' + 1$. Hence,

$$P\bigl(W' + 1 \ne W\bigr) \le P(\rho_k > n-1) + P\bigl(\rho_k \ne \rho'_k + 1\bigr). \tag{79}$$

Using Chebyshev's inequality, we get

$$P(\rho_k > n-1) \le \frac{\mathbb{V}(\rho_k)}{(n-1-\mathbb{E}(\rho_k))^2}.$$

As seen earlier,

$$\mathbb{E}(\rho_k) = \frac{k}{a(p^*)}; \qquad \mathbb{V}(\rho_k) = \frac{k(1-3a(p^*))}{a^2(p^*)}.$$

Assume now, without loss of generality, $(n-1)a(p^*) \ge 8$. Then

$$\begin{aligned} P(\rho_k > n-1) &\le \frac{k(1-3a(p^*))}{((n-1)a(p^*)-k)^2} \\ &\le \frac{1.125(1-3a(p^*))}{(n-1)a(p^*)(0.125)^2} \\ &= \frac{72(1-3a(p^*))}{(n-1)a(p^*)} = K_1/(n-1) \qquad \text{(say)}. \end{aligned} \tag{80}$$

Hence, we obtain from (78), (79) and (80)

$$d \le 2\bigl\|\mathcal{L}(W) * (I_1 - I)\bigr\|_{\mathrm{TV}} \le \frac{K_1}{n-1} + \frac{K_2}{\sqrt{n-1}}.$$

This proves (75). □

Using similar arguments and the fact that T_j's are i.i.d., (76) immediately follows.

Proof of Corollary 4.8. The bounds for d and d_1 in Theorem 4.1 are given by Lemma 4.7. Next, to compute $\mathbb{E}|\widetilde{W}^{(i)} - W^{(i)}|$, construct the following two-dimensional stochastic process $\{(Z_l^{i1}, Z_l^{i0})\}_{l \ge i}$ with initial state $(Z_i^{i1}, Z_i^{i0}) = (1, 0)$, where $\mathcal{L}(Z_l^{ij}) = \mathcal{L}(\mathbb{I}_l | \mathbb{I}_i = j)$, for $j = 0, 1$, having following marginal distributions.

(i) For $l \ge i+2$,

$$P\bigl((Z_l^{i1}, Z_l^{i0}) = (0, 0)\bigr) = 1 - a\bigl(p^*\bigr),$$

$$
\begin{aligned}
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (0,1)\big) = 0,\\
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (1,0)\big) = 0,\\
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (1,1)\big) = a\big(p^*\big).
\end{aligned}
$$

(ii) For $i < l \le i+1$,

$$
\begin{aligned}
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = 1 - \frac{a(p^*)}{1-a(p^*)},\\
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (0,1)\big) = \frac{a(p^*)}{1-a(p^*)},\\
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (1,0)\big) = 0,\\
&P\big(\big(Z_l^{i1}, Z_l^{i0}\big) = (1,1)\big) = 0.
\end{aligned}
$$

Also, the joint distributions satisfy:

(i) For $l = i$

$$
\begin{aligned}
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,0), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = 1 - 2\frac{a(p^*)}{1-a(p^*)},\\
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,1), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = \frac{a(p^*)}{1-a(p^*)},\\
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,0), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,1)\big) = \frac{a(p^*)}{1-a(p^*)},
\end{aligned}
$$

and zero otherwise.

(ii) For $l = i+1$,

$$
\begin{aligned}
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,0), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = 1 - \big(2 - a\big(p^*\big)\big)\frac{a(p^*)}{1-a(p^*)},\\
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (1,1), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = a\big(p^*\big),\\
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,0), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,1)\big) = \frac{a(p^*)}{1-a(p^*)},
\end{aligned}
$$

and zero otherwise.

(iii) For $l \ge i+2$,

$$
\begin{aligned}
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,0), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = 1 - 2a\big(p^*\big),\\
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (1,1), \big(Z_l^{i1}, Z_l^{i0}\big) = (0,0)\big) = a\big(p^*\big),\\
&P\big(\big(Z_{l+1}^{i1}, Z_{l+1}^{i0}\big) = (0,0), \big(Z_l^{i1}, Z_l^{i0}\big) = (1,1)\big) = a\big(p^*\big),
\end{aligned}
$$

and zero otherwise.

Let us now define the random variables

$$\zeta = \min\{k - i | Z_k^{i1} = Z_k^{i0} = 1\}, \qquad \text{for } k \geq i$$

and

$$\widetilde{\zeta} = \min\{i - k | Z_k^{i1} = Z_k^{i0} = 1\}, \qquad \text{for } i \leq k.$$

Due to symmetry of the stochastic process about i, we have suppressed the index i. The distribution of ζ is given by

$$P(\zeta = k) = \begin{cases} a(p^*), & \text{for } 2 \leq k \leq 3, \\ a(p^*)\left(\dfrac{1-2a(p^*)}{1-a(p^*)}\right)^{k-3}, & \text{for } k \geq 4. \end{cases}$$

Therefore,

$$\mathbb{E}(\zeta) = a(p^*) + \frac{1}{a(p^*)}.$$

Also, due to symmetry, we have $\zeta \stackrel{\mathcal{L}}{=} \widetilde{\zeta}$.

Define now

$$W_l^{(i)} = \sum_{j=2}^{i-\widetilde{\zeta}} Z_j^{i1} = \sum_{j=2}^{i-\widetilde{\zeta}} Z_j^{i0}, \qquad W_r^{(i)} = \sum_{j=i+\zeta}^{n} Z_j^{i1} = \sum_{j=i+\zeta}^{n} Z_j^{i0}$$

and

$$\xi^{i1} = \sum_{j=(i-\widetilde{\zeta}+1)\vee 2}^{(i+\zeta-1)\wedge n} Z_j^{i1} - Z_i^{i1}.$$

Thus,

$$\widetilde{W}^{(i)} = W_l^{(i)} + W_r^{(i)} + \xi^{i1}.$$

Let now

$$\mathbb{I}'_j = \begin{cases} Z_j^{i1}, & \text{with probability } a(p^*), \\ Z_j^{i0}, & \text{with probability } 1 - a(p^*), \end{cases}$$

and $\mathbb{I}''_j \stackrel{\mathcal{L}}{=} \mathbb{I}_j$, but $\mathbb{I}''_j$ is independent of $\{(Z_j^{i1}, Z_j^{i0}) | j \in [i - \widetilde{\zeta}, i + \zeta]\}$. Then

$$Z_j := \begin{cases} \mathbb{I}''_j, & \text{if } j \in [i - \widetilde{\zeta}, i + \zeta], \\ \mathbb{I}'_j, & \text{if } j > i + \zeta \text{ or } j < i - \widetilde{\zeta}. \end{cases}$$

Define

$$\xi^i = \sum_{j=(i-\tilde{\zeta}+1)\vee m}^{(i+\zeta-1)\wedge n} Z_j - Z_i; \qquad W^{(i)\prime} = W_l^{(i)} + W_r^{(i)} + \xi^i$$

so that $W^{(i)} \overset{\mathcal{L}}{=} W^{(i)\prime}$. Note that

$$\mathbb{E}(\xi^i) \le \mathbb{E}(\zeta + \tilde{\zeta} - 1) = \frac{2}{a(p^*)} + 2a(p^*) - 1,$$
$$\mathbb{E}(\xi^{i1}) \le \mathbb{E}(\zeta + \tilde{\zeta} - 2) = \frac{2}{a(p^*)} + 2a(p^*) - 2.$$

Therefore,

$$\begin{aligned}\mathbb{E}\big|\widetilde{W}^{(i)} - W^{(i)}\big| &\le \mathbb{E}\big|\xi^i - \xi^{i1}\big| \\ &\le a(p^*)\max\{2(1-a(p^*)),1\}\mathbb{E}(\zeta + \tilde{\zeta} - 2) \\ &= 2\max\{2(1-a(p^*)),1\}(1-a(p^*)(1-a(p^*))).\end{aligned}$$

Thus, the bound given in Theorem 4.1 becomes

$$\begin{aligned}\|\mathcal{L}(W) - \mathrm{BCP}\|_{\mathrm{TV}} \le \frac{2}{(1-2\theta_1)\widehat{\lambda}}\bigg\{&\bigg(na(p^*)^4 + \frac{Mp^4}{(1-2p)^2}\bigg)\bigg(\frac{K_1}{n-2} + \frac{K_2}{\sqrt{n-2}}\bigg) \\ &+ (1+2p)\delta p^2 + (n-1)C_1\bigg\}.\end{aligned}$$

This proves the corollary. □

Acknowledgments

The authors are grateful to the reviewers for several suggestions, critical comments and especially for pointing out some relevant references related to this work.

References

[1] Afendras, G., Balakrishnan, N. and Papadatos, N. (2014). Orthogonal polynomials in the cumulative ord family and its application to variance bounds. Preprint. Available at arXiv:1408.1849.
[2] Barbour, A.D. (1987). Asymptotic expansions in the Poisson limit theorem. *Ann. Probab.* **15** 748–766. MR0885141
[3] Barbour, A.D. and Čekanavičius, V. (2002). Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.* **30** 509–545. MR1905850
[4] Barbour, A.D., Čekanavičius, V. and Xia, A. (2007). On Stein's method and perturbations. *ALEA Lat. Am. J. Probab. Math. Stat.* **3** 31–53. MR2324747

[5] Barbour, A.D. and Chen, L.H.Y. (2014). Stein's (magic) method. Preprint. Available at arXiv:1411.1179.
[6] Barbour, A.D., Chen, L.H.Y. and Loh, W.-L. (1992). Compound Poisson approximation for nonnegative random variables via Stein's method. *Ann. Probab.* **20** 1843–1866. MR1188044
[7] Barbour, A.D., Holst, L. and Janson, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. New York: Oxford Univ. Press. MR1163825
[8] Barbour, A.D. and Utev, S. (1998). Solving the Stein equation in compound Poisson approximation. *Adv. in Appl. Probab.* **30** 449–475. MR1642848
[9] Barbour, A.D. and Xia, A. (1999). Poisson perturbations. *ESAIM Probab. Stat.* **3** 131–150. MR1716120
[10] Brown, T.C. and Xia, A. (2001). Stein's method and birth–death processes. *Ann. Probab.* **29** 1373–1403. MR1872746
[11] Čekanavičius, V. and Roos, B. (2004). Two-parametric compound binomial approximations. *Lith. Math. J.* **44** 354–373.
[12] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications* (*New York*). Heidelberg: Springer. MR2732624
[13] Chen, L.H.Y. and Röllin, A. (2013). Approximating dependent rare events. *Bernoulli* **19** 1243–1267. MR3102550
[14] Daly, F. (2010). Stein's method for compound geometric approximation. *J. Appl. Probab.* **47** 146–156. MR2654764
[15] Daly, F. (2011). On Stein's method, smoothing estimates in total variation distance and mixture distributions. *J. Statist. Plann. Inference* **141** 2228–2237. MR2775201
[16] Daly, F., Lefèvre, C. and Utev, S. (2012). Stein's method and stochastic orderings. *Adv. in Appl. Probab.* **44** 343–372. MR2977399
[17] Döbler, C. (2012). Stein's method of exchangeable pairs for absolutely continuous, univariate distributions with applications to the polya urn model. Preprint. Available at arXiv:1207.0533.
[18] Ehm, W. (1991). Binomial approximation to the Poisson binomial distribution. *Statist. Probab. Lett.* **11** 7–16. MR1093412
[19] Eichelsbacher, P. and Reinert, G. (2008). Stein's method for discrete Gibbs measures. *Ann. Appl. Probab.* **18** 1588–1618. MR2434182
[20] Fulman, J. and Goldstein, L. (2015). Stein's method and the rank distribution of random matrices over finite fields. *Ann. Probab.* **43** 1274–1314. MR3342663
[21] Godbole, A.P. (1993). Approximate reliabilities of m-consecutive-k-out-of-n: Failure systems. *Statist. Sinica* **3** 321–327. MR1243390
[22] Goldstein, L. and Reinert, G. (2005). Distributional transformations, orthogonal polynomials, and Stein characterizations. *J. Theoret. Probab.* **18** 237–260. MR2132278
[23] Goldstein, L. and Reinert, G. (2013). Stein's method for the beta distribution and the Pólya-Eggenberger urn. *J. Appl. Probab.* **50** 1187–1205. MR3161381
[24] Hess, K.Th., Liewald, A. and Schmidt, K.D. (2002). An extension of Panjer's recursion. *Astin Bull.* **32** 283–297. MR1942940
[25] Huang, W.T. and Tsai, C.S. (1991). On a modified binomial distribution of order k. *Statist. Probab. Lett.* **11** 125–131. MR1092971
[26] Ley, C., Reinert, G. and Swan, Y. (2014). Approximate computation of expectations: A canonical Stein operator. Preprint. Available at arXiv:1408.2998.
[27] Ley, C. and Swan, Y. (2013). Stein's density approach and information inequalities. *Electron. Commun. Probab.* **18** 1–14.
[28] Ley, C. and Swan, Y. (2013). Local Pinsker inequalities via Stein's discrete density approach. *IEEE Trans. Inform. Theory* **59** 5584–5591. MR3096942

[29] Norudin, I. and Peccati, G. (2012). Normal approximations with Malliavin calculus. From Stein's method to universality. Cambridge Tracts in Mathematics No. 192.
[30] Panjer, H.H. and Wang, Sh. (1995). Computational aspects of Sundt's generalized class. *Astin Bull*. **25** 5–17.
[31] Peköz, A., Röllin, A., Čekanavičius, V. and Shwartz, M. (2009). A three-parameter binomial approximation. *J. Appl. Probab.* **46** 1073–1085.
[32] Reinert, G. (2005). Three general approaches to Stein's method. In *An Introduction to Stein's Method*. *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap*. **4** 183–221. Singapore: Singapore Univ. Press. MR2235451
[33] Röllin, A. (2005). Approximation of sums of conditionally independent variables by the translated Poisson distribution. *Bernoulli* **11** 1115–1128. MR2189083
[34] Roos, B. (2000). Binomial approximation to the Poisson binomial distribution: The Krawtchouk expansion. *Theory Probab. Appl*. **45** 258–272.
[35] Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293. MR2861132
[36] Soon, S.Y.T. (1996). Binomial approximation for dependent indicators. *Statist. Sinica* **6** 703–714. MR1410742
[37] Stein, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **7**. Hayward, CA: IMS. MR0882007
[38] Stein, C., Diaconis, P., Holmes, S. and Reinert, G. (2004). Use of exchangeable pairs in the analysis of simulations. In *Stein's Method: Expository Lectures and Applications*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **46** 1–26. Beachwood, OH: IMS. MR2118600
[39] Sundt, B. (1992). On some extensions of Panjer's class of counting distributions. *Astin Bull*. **22** 61–80.
[40] Vellaisamy, P. (2004). Poisson approximation for (k_1, k_2)-events via the Stein–Chen method. *J. Appl. Probab.* **41** 1081–1092. MR2122802
[41] Vellaisamy, P., Upadhye, N.S. and Čekanavičius, V. (2013). On negative binomial approximation. *Theory Probab. Appl*. **57** 97–109. MR3201640
[42] Xia, A. (1997). On using the first difference in the Stein–Chen method. *Ann. Appl. Probab.* **7** 899–916. MR1484790

Received June 2014 and revised October 2015