COMPOUND POISSON APPROXIMATION FOR DISSOCIATED RANDOM VARIABLES VIA STEIN'S METHOD

PETER EICHELSBACHER AND MAŁGORZATA ROOS

Abstract. In the present paper we consider compound Poisson approximation by Stein's method for dissociated random variables. We present some applications to problems in system reliability such as k-runs, counting isolated vertices in the rectangle lattice on the torus as well as consecutive 2-systems, coloured graphs, birthday problems, connected s-systems and two dimensional consecutive-k-out-of-n systems. In particular our examples have the structure of an incomplete U-statistics. For nonnegative integer valued complete U-statistics improvements of the Poisson approximation results can not be expected in general. We mainly apply techniques from Barbour and Utev, who gave new bounds for the solutions of the Stein equation in compound Poisson approximation in two recent papers.

1. Introduction

Let Γ denote an arbitrary finite collection of indices, usually denoted by α, β and so on. Let X_α be 0-1-valued, possibly dependent, random variables and let $W = \sum_{\alpha\in\Gamma} X_\alpha$. If the X_α are weakly dependent and the value 1 occurs with small probability, Poisson approximation in total variation distance between the law of W, denoted by $\mathcal{L}(W)$, and the Poisson distribution with parameter $\mathbb{E}W$, denoted by $Po(\mathbb{E}W)$, can be successfully established via the Stein-Chen method, which at the same time gives estimates of the approximation error (see the book of Barbour, Holst and Janson [7]). If the random variables X_α can take other positive integer values or if the dependence is stronger (clumps of 1's tend to occur), the compound Poisson approximation should provide better approximation. The importance of developing a compound Poisson approach has been discussed in Aldous [1].

In [3] Barbour, Chen and Loh introduced a Stein equation for compound Poisson approximation. Roos developed the local version of the basic method in [16] and in [15] the coupling approach. The theoretical results were successfully applied to many examples in reliability theory (see [4] and [5]) and to the problem of uniform m-spacings on the unit circle (see [13] and [14]). The solutions of the corresponding Stein equation may in general grow exponentially with the mean of clumps. For a special class of distributions in [3, Theorem 5] a better bound was given which

1991 *Mathematics Subject Classification.* 05C80, 60C05, 60F05, 62E17, 90B25.

Key words and phrases. Stein's method, compound Poisson distribution, dissociated random variables, locally dependent random variables.

is comparable in sharpness to the corresponding bound in the Poisson Stein-Chen method (apart from a logarithmic term in the numerator of the bound). Recently, Barbour and Utev [9] proved bounds for the solutions of the Stein equation with respect to the Kolmogorov distance which enabled them to carry out compound Poisson approximation in this distance with the same efficiency as in the Poisson case. They assumed that the compound Poisson limit distribution has a fourth moment and that it satisfies an aperiodicity condition. In [8] the same authors proved the counterparts in total variation distance under the assumption that the compound Poisson limit distribution is aperiodic of finite exponential moment, which is hardly restrictive.

The contribution of [9] and [8] is fundamental. Barbour and Utev applied the local approach in the way developed by Roos [16] in a more general setting. Consequently, the results derived for indicator random variables in [15] and [16] became a special case of the more general results in [8].

The aim of the present paper is to improve some results stated in [4, 5, 16] by using the new bounds from [9] and [8]. We observe that the random variables in the examples considered up to now are actually dissociated ones. Moreover, we present Stein's method for compound Poisson approximation both in Kolmogorov distance and in total variation distance for this special class of random variables. In view of the theoretical work by Barbour and Utev we provide the following examples and applications: k-runs, isolated vertices in the rectangular lattice on the torus, consecutive 2-systems, coloured graphs, connected-s systems, and the two dimensional consecutive-k-out-of-n system. In [9, Section 5] the k-run example was already discussed briefly. In particular all the examples have the structure of an incomplete U-statistics. After having calculated the bound for compound Poisson approximation for U-statistics, we see that improvements of the Poisson approximation results on U-statistics (see [6]) cannot be expected in general. In the examples we are able to obtain estimates with the right behaviour when the mean of the clump tends to infinity. As a consequence of the good bounds in [9] and [8] no unwanted logarithmic factor and no unwanted e^{λ}, where λ is the mean of the clumps, come into play.

Most of our examples deal with models from reliability theory. In this theory evaluation is an important and integral feature of planing, designing and operating engineering systems. In these models the independent items constitute the system. In the definition of the system there are special critical combinations consisting of underlying independent items. A critical combination is said to fail if all items in it are failed. A failure of a critical combination is dangerous for the system. Let W be the random variable which counts the number of failed critical combinations in the system. The system is said to fail if there are at least m occurrences of failed critical combinations, where $m \geq 1$ is a fixed integer. The reliability of the system is $\mathbb{P}(W \leq m-1)$. Thus estimates of $\mathbb{P}(W \leq m-1)$ are estimates of the reliability

of the system. Since it is not always possible to compute the reliability of a system exactly, it is reasonable to look instead for good approximations. Because the critical combinations may overlap, the indicators of their failure can be dependent, in spite of the independence of individual underlying items. Frequently, the indicators of the failure of critical combinations are dissociated or locally dependent random variables. It is a setting where a compound Poisson approximation is promising.
In our paper we denote 0-1-valued random variables X_α by I_α. In Section 2, we state known theoretical results about Stein's method and give the theorem for dissociated random variables as well as locally dependent ones. In Section 3 examples are presented.

2. Elaboration

2.1. Stein's method, compound Poisson distribution and approximation.

We denote by $CP(\lambda,\mu)$ the compound Poisson distribution of $\sum_{j=1}^N X_j$, where $(X_j, j\geq 1)$ are nonnegative valued, independent and identically distributed variables with distribution μ (a probability measure on $(0,\infty)$) and are independent of N, distributed according to the usual Poisson distribution with mean λ. We briefly summarize Stein's method for this class of distributions. If W is a nonnegative integer valued random variable satisfying

$$\left|\mathbb{E}\biggl(\sum_{i\geq 1} i\lambda_i g(W+i) - Wg(W)\biggr)\right| \leq \varepsilon_0\, M_0(g) + \varepsilon_1\, M_1(g) \tag{2.1}$$

for all bounded $g:\ \mathbb{N}\to\mathbb{R}$ and for some small ε_0 and ε_1, where

$$M_0(g) = \sup_{j\geq 1}|g(j)|, \qquad M_1(g) = \sup_{j\geq 1}|g(j+1)-g(j)|,$$

then it follows that

$$d_{TV}(\mathcal{L}(W), CP(\lambda,\mu)) \leq \varepsilon_0\, H_0(\lambda,\mu) + \varepsilon_1\, H_1(\lambda,\mu), \tag{2.2}$$

where $\lambda = \sum_{i\geq 1}\lambda_i$, $\mu_i = \lambda_i/\lambda$ (that means that the measure μ has the form $\mu = \sum_{i\geq 1}\mu_i\,\delta_i$, where δ_i denotes the Dirac measure concentrated on $i\in\mathbb{N}$), $d_{TV}(P,Q)$ denotes the total variation distance between two probability measures P and Q on $\mathbb{Z}_+$ ($d_{TV}(P,Q) = \sup_{A\subseteq\mathbb{Z}_+}\{|P(A)-Q(A)|\}$) and $H_l(\lambda,\mu) = \sup_{A\subset\mathbb{Z}_+} M_l(g_A)$, $l=0,1$, with g_A being the solution of the Stein equation

$$\sum_{i\geq 1} i\lambda_i g(w+i) - wg(w) = 1_A(w) - CP(\lambda,\mu)\{A\}, \quad w\geq 0, A\subseteq\mathbb{Z}_+. \tag{2.3}$$

The method works as follows: if we can bound the left hand side of (2.1) with suitably small ε_0 and ε_1 and the constants $H_l(\lambda,\mu)$ are not too large, the distribution of W is close to an appropriate compound Poisson distribution. The bounds on $H_l(\lambda,\mu)$ will depend on λ and μ and not on the given random variable W itself.

There may not be any good bounds for the $H_l(\lambda,\mu)$ (see discussion in the introduction of [8]), but for the terms

$$H_l^{(a)}(\lambda,\mu) = \sup_{A\subset\mathbb{Z}_+} M_l(g_A(\cdot + a)),\ l = 0,1,$$

and by inequality (4.4) in [9] (2.2) can be replaced by

$$\begin{aligned} d_{TV}(\mathcal{L}(W), CP(\lambda,\mu)) \le \varepsilon_0\, H_0^{(a)}(\lambda,\mu) + \varepsilon_1 \left(H_1^{(a)}(\lambda,\mu) + \frac{2H_0^{(a)}(\lambda,\mu)}{\lambda m_1(1-c_1)}\right) \\ + \mathbb{P}\big(W \le \frac{1}{2}(1+c_1)\lambda m_1\big)\left(1 + \frac{2m_2 H_0^{(a)}(\lambda,\mu)}{m_1(1-c_1)}\right), \end{aligned} \tag{2.4}$$

where $c_1 \in (0,1)$, $a = c_1\lambda m_1$ and m_l denotes $\sum_{i\ge1} i^l\,\mu_i$ for $l \ge 1$. If $\mu(z) = \sum_{i\ge1}\mu_i\, z^i$ has radius of convergence $R(\mu) > 1$ (meaning that we assume that μ has a finite exponential moment), $\lambda \ge 2$ and

Condition 2.5 (Aperiodicity)**.** *for any* $0 < \zeta \le \pi: \quad \varrho^*(\zeta) > 0,$

where $\varrho^*(\zeta) = \min(\varrho_1^*(\zeta), \frac{1}{2}\varrho_2^*(\zeta), 1)$ with $\varrho_k^*(\zeta) = \inf_{\zeta\le\theta\le\pi}\varrho_k(\theta)$, $k = 1,2$, and $\varrho_1(\theta) = 1 - \sum_{i\ge1}\mu_i\cos i\theta$ and $\varrho_2(\theta) = 1 - \frac{1}{m_1}\sum_{i\ge1} i\mu_i\cos i\theta$, then there exist constants $C_l(\mu)$, $l = 0,1,2$, given explicitly in terms of μ (see [8, (1.20)-(1.28)]) such that for any $a \ge C_2(\mu)\lambda m_1 + 1$

$$H_0^{(a)}(\lambda,\mu) \le \frac{1}{\sqrt{\lambda}}C_0(\mu), \qquad H_1^{(a)}(\lambda,\mu) \le \frac{1}{\lambda}C_1(\mu). \tag{2.6}$$

Note that $C_2(\mu) < 1$, so that there are feasible choices for c_1 in (2.4). Note that $C_2(\mu) = 1 - \varrho^*(\zeta_0)/4$ with $\zeta_0 = \sqrt{m_2/m_4}$.
In [9], where the Kolmogorov distance is considered, the $H_l(\lambda,\mu)$ are replaced by

$$J_l(\lambda,\mu) = \sup_{m\ge0} M_l(g_{[0,m]}),\ l = 0,1,$$

which play the same role as the $H_l(\lambda,\mu)$ if one considers the smaller family of sets $A = [0,m]$, $m \ge 1$, on the right hand side of the Stein equation (2.3). Then the estimate for approximation in terms of Kolmogorov distance is the analogue of (2.2):

$$d_K\big(\mathcal{L}(W), CP(\lambda,\mu)\big) \le \varepsilon_0\, J_0(\lambda,\mu) + \varepsilon_1\, J_1(\lambda,\mu),$$

where $d_K(\mathcal{L}(W), CP(\lambda,\mu)) = \sup_{t\in\mathbb{R}}|\mathbb{P}(W\le t) - CP(\lambda,\mu)\{(-\infty,t)\}|$. The quantities $J_l^{(a)}(\lambda,\mu)$, $l = 0,1$, are obtained in the same way as the $H_l^{(a)}(\lambda,\mu)$ but with $g_{[0,m]}$ instead of g_A, $A \subseteq \mathbb{Z}_+$. Under weaker assumptions, namely Condition 2.5 and the finiteness of the fourth moment, Theorem 4.3 in [9] gives a bound in the Kolmogorov distance:

$$\begin{aligned} d_K(\mathcal{L}(W), CP(\lambda,\mu)) \le \varepsilon_0\, J_0^{(a)}(\lambda,\mu) + \varepsilon_1\left(J_1^{(a)}(\lambda,\mu) + \frac{4J_0^{(a)}(\lambda,\mu)}{\lambda m_1\varrho^*(\zeta_0)}\right) \\ + \mathbb{P}\big(W \le \lambda m_1(1-\varrho^*(\zeta_0)/4)\big)\left(1 + \frac{4m_2 J_0^{(a)}(\lambda,\mu)}{m_1\varrho^*(\zeta_0)}\right), \end{aligned} \tag{2.7}$$

where $\zeta_0 = \sqrt{m_2/m_4}$, ϱ^* is as defined in Condition 2.5 and

$$J_0^{(a)}(\lambda,\mu) \le \frac{1}{\sqrt{\lambda}} D_0(\mu), \quad J_1^{(a)}(\lambda,\mu) \le \frac{1}{\lambda} D_1(\mu),$$

where the constants $D_l(\mu)$, $l = 0, 1$, are given in (4.8) and (4.9) in [9].

2.2. **Dependence structure.** We recall the definition of dissociation introduced by McGinley and Sibson in [12] which can be seen as the most natural and useful generalisation of the assumption of independence. Suppose that Γ is a collection of k-subsets $\alpha = \{\alpha_1, \dots, \alpha_k\}$ of $\{1, 2, \dots, n\}$. Then a family $(X_\alpha, \alpha \in \Gamma)$ of nonnegative integer valued random variables is said to be *dissociated* if the subsets of random variables $(X_\alpha, \alpha \in A)$ and $(X_\beta, \beta \in B)$ are independent whenever $(\cup_{\alpha \in A} \alpha) \cap (\cup_{\beta \in B} \beta) = \emptyset$. If $k = 1$, the notions of dissociated and independent coincide. The most important setting is the one in which $X_\alpha = \phi_\alpha(Y_{\alpha_1}, \dots, Y_{\alpha_k})$ for some nonnegative integer valued symmetric functions ϕ_α, where $Y_1, \dots, Y_n$ are independent random elements of some space $\mathcal{X}$. Here $W = \sum_{\alpha \in \Gamma} X_\alpha$ is called an incomplete U-statistics with kernel class $(\phi_\alpha)_{\alpha \in \Gamma}$. If Γ is the set of all k-subsets of $\{1, \dots, n\}$, W is called a U-statistics. Many statistics commonly used are in fact members of this class (see for example [10] and [11]).

Now we state a result on bounds for ε_0 and ε_1 in the case of nonnegative valued dissociated random variables $(X_\alpha, \alpha \in \Gamma)$. We require some further notation. For Γ fixed, we define the *dissociated partition* for each $\alpha \in \Gamma$:

$$\Gamma = \{\alpha\} \cup \Gamma_\alpha^{vs} \cup \Gamma_\alpha^0 \cup \Gamma_\alpha^w$$

and set

$$Z_\alpha = \sum_{\beta \in \Gamma_\alpha^{vs}} X_\beta, \quad U_\alpha = \sum_{\beta \in \Gamma_\alpha^0} X_\beta, \quad W_\alpha = \sum_{\beta \in \Gamma_\alpha^w} X_\beta.$$

Here Γ_α^{vs} contains indices of those X_β which strongly influence X_α. Depending on the choice of this region we define for each α

$$\Gamma_\alpha^0 = \{\beta \in \Gamma : \ \beta \notin \Gamma_\alpha^{vs} \cup \{\alpha\}, \beta \cap \delta \ne \emptyset \quad \text{for some} \quad \delta \in \Gamma_\alpha^{vs} \cup \{\alpha\}\}$$

and

$$\Gamma_\alpha^w = \{\beta \in \Gamma : \ \beta \cap \delta = \emptyset \quad \text{for all} \quad \delta \in \Gamma_\alpha^{vs} \cup \{\alpha\}\}.$$

Note that, by definition of dissociation, the random variables with indices in the region Γ_α^w do not influence X_α and Z_α. Moreover, the definition of Γ_α^{vs} provide us with enough freedom of choice in defining the clumps. In particular the random variables with indices in Γ_α^0 may be dependent on X_α. Frequently, a "maximal" Γ_α^{vs} is chosen, as for example, in the local dependence below.

Theorem 2.8. *If $(X_\alpha, \alpha \in \Gamma)$ are nonnegative valued dissociated random variables, Γ_α^{vs}, Γ_α^0 and Γ_α^w are as defined above, then we can take*

$$\varepsilon_0 = 0$$

and

$$\varepsilon_1 = \sum_{\alpha \in \Gamma} \Big(\mathbb{E}(X_\alpha)\, \mathbb{E}(X_\alpha + Z_\alpha + U_\alpha) + \mathbb{E}(X_\alpha\, U_\alpha) \Big). \tag{2.9}$$

Proof. We only have to choose the special dissociated partition $\Gamma_\alpha^b = \Gamma_\alpha^0$ in Lemma 1.8 in [8]. Since the X_α are dissociated, W_α is independent of Z_α and X_α and therefore, $\delta_1 = \delta_2 = \delta_3 = 0$ and $\delta_4 = \varepsilon_1$. The theorem follows immediately. □

Remark 2.10. We are able to prove the same result in case of indicator random variables by applying the dissociated partition to Theorem 2 in [16]. We have chosen, however, to use the more general formulation for nonnegative random variables of Lemma 1.8 in [8].

Remark 2.11. With the definition of Γ_α^s and Γ_α^i in [7, Chapter 2.3], taking $\Gamma_\alpha^{vs} = \emptyset$, $\Gamma_\alpha^0 = \Gamma_\alpha^s$ and $\Gamma_\alpha^w = \Gamma_\alpha^i$ we obtain the same result as Theorem 2.N in [7] (where the (X_α) are indicator random variables). For Poisson approximation to be good, the α-clump must be negligible, and $\Gamma_\alpha^{vs} = \emptyset$ is an adequate choice.

Remark 2.12. There are cases when one is able to find some probabilistic structure in Γ_α^0 in the case of dissociated random variables. In such situation it might be advantageous to use the coupling approach instead of the local one, see Remark 2.1.5 in [7].

Let Γ be an arbitrary index set (for applications it is normally sufficient to consider finite sets; in the infinite case one has to ensure that sums $\sum_{\alpha \in \Gamma}$ are convergent). A family of nonnegative integer valued random variables $(X_\alpha, \alpha \in \Gamma)$ is said to be *locally dependent* if for each $\alpha \in \Gamma$, there exist $A_\alpha \subseteq B_\alpha \subseteq \Gamma$ with $\alpha \in A_\alpha$ such that X_α is independent of $(X_\beta, \beta \in A_\alpha^c)$ and $(X_\beta, \beta \in A_\alpha)$ is independent of $(X_\beta, \beta \in B_\alpha^c)$. It is well known (see [3, Section 4]) that m-dependence is a special case of local dependence. Moreover, finite dependence (see Definition in [3]) is a special case of local dependence. If Γ is a collection of k-subsets than a locally dependent family $(X_\alpha, \alpha \in \Gamma)$ is also a dissociated family by definition: for a fixed α we choose $A_\alpha = \{\beta \in \Gamma : \alpha \cap \beta \neq \emptyset\}$ and $B_\alpha = A_\alpha \cup \{\beta \in \Gamma : \beta \notin A_\alpha \;\text{ and }\; \beta \cap \delta \neq \emptyset \text{ for some } \delta \in A_\alpha\}$. Note that, by definition, $\alpha \in A_\alpha$. Actually a dissociated family is also a locally dependent family. For a locally dependent family we obtain

Corollary 2.13. *If $(X_\alpha, \alpha \in \Gamma)$ are nonnegative integer valued locally dependent random variables, then we can take*

$$\varepsilon_0 = 0$$

and

$$\varepsilon_1 = \sum_{\alpha \in \Gamma} \sum_{\beta \in B_\alpha} \mathbb{E}X_\alpha\, \mathbb{E}X_\beta + \sum_{\alpha \in \Gamma} \sum_{\beta \in B_\alpha \setminus A_\alpha} \mathbb{E}X_\alpha\, \mathbb{E}X_\beta. \tag{2.14}$$

Proof. For any $\alpha \in \Gamma$ we choose the partition $\{\alpha\} \cup \Gamma_\alpha^{vs} = A_\alpha$, $\Gamma_\alpha^b = B_\alpha \setminus A_\alpha$ and $\Gamma_\alpha^w = B_\alpha^c$ in Theorem 2.8. By the definition of local dependence X_α and X_β with $\beta \in B_\alpha \setminus A_\alpha$ are independent. □

Remark 2.15. Note that ε_1 in (2.8) and in (2.14) has the form

$$\sum_{\alpha\in\Gamma}\Bigg(\sum_{\beta\in\{\alpha\}\cup\Gamma_\alpha^{vs}\cup\Gamma_\alpha^0} \mathbb{E}X_\alpha \mathbb{E}X_\beta + \sum_{\beta\in\Gamma_\alpha^0} \mathbb{E}(X_\alpha X_\beta)\Bigg).$$

It is easily seen that (2.8) takes the form (2.14) whenever the random variables with indices in Γ_α^0 are independent of X_α (X_α is independent of U_α).

Remark 2.16. Corollary 2.13 is an improvement of Theorem 8 in [3]: we are dealing with the case of nonnegative integer valued random variables. In [3] a slightly weaker bound $\varepsilon_1 = 2\sum_{\alpha\in\Gamma}\sum_{\beta\in B_\alpha} \mathbb{E}X_\alpha\, \mathbb{E}X_\beta$ is proven.

Remark 2.17. Note that the approximating distribution $CP(\lambda,\mu)$ can be calculated explicitly: $\mu_i = \frac{1}{i\,\lambda}\sum_{\alpha\in\Gamma}\mathbb{E}(X_\alpha I[X_\alpha + Z_\alpha = i])$, $i \geq 1$ and

$$\lambda = \sum_{\alpha\in\Gamma}\mathbb{E}\bigg(\frac{X_\alpha}{X_\alpha + Z_\alpha}\bigg).$$

Remark 2.18. One can consider an approximation by a $CP(\lambda,\mu')$ distribution instead of the naturally emerging distribution $CP(\lambda,\mu)$. Theorem 1.10 in [8] states that for any choice of μ':

$$\begin{aligned} d_{TV}(\mathcal{L}(W), CP(\lambda,\mu')) &\leq \varepsilon_0'\, H_0^{(a)}(\lambda,\mu') + \varepsilon_1'\bigg(H_1^{(a)}(\lambda,\mu') + \frac{2H_0^{(a)}(\lambda,\mu')}{\lambda m_1'(1-c_1)}\bigg) \\ &\quad + \mathbb{P}\big(W \leq \frac{1}{2}(1+c_1)\lambda m_1'\big)\bigg(1 + \frac{2m_2' H_0^{(a)}(\lambda,\mu')}{m_1'(1-c_1)}\bigg) \end{aligned} \tag{2.19}$$

with $\varepsilon_0' = \lambda|m_1 - m_1'| + \varepsilon_0$ and $\varepsilon_1' = \lambda\, m_1\, d_W(\mu^*,\mu'^*) + \varepsilon_1$; here * is used to denote a size-biased distribution $\mu_i^* = i\mu_i/m_1$ and d_W denotes the Wasserstein L_1 metric on probability measures over $\mathbb{R}$. One can choose the distribution μ' to be of the following form:

$$\begin{aligned} \mu_1' &= \mu_1 + \sum_{i=l+1}^{\infty} i\,\mu_i, \\ \mu_j' &= \mu_j \text{ for } j = 2,\dots,l \\ \mu_j' &= 0 \text{ for } j \geq l+1. \end{aligned} \tag{2.20}$$

In such a case $m_1' = m_1$. Consequently $\lambda|m_1 - m_1'| = 0$ and $\lambda\, m_1\, d_W(\mu^*,\mu'^*) = \lambda\sum_{i=l+1}^{\infty} i(i-1)\mu_i = \sum_{i=l+1}^{\infty} i(i-1)\lambda_i$. This choice of μ' was implicitly suggested in Theorem 3 of [16]. In her notation $\lambda_i^* = \lambda\,\mu_i'$ for $i \geq 1$.

Remark 2.21. There are different approaches to bound the truncation term $\mathbb{P}\big(W \leq (1-\varepsilon)\mathbb{E}W\big)$; e.g., one might use the Chebyshev-inequality. For a further refinement, Bernstein-type inequalities may be applied. In our examples we mostly apply Janson's inequality which works only for indicator random variables built up from independent ones. For the sake of completeness we state the inequality (see [7, Theorem 2.S]). Consider a collection of independent random indicator variables $(J_i, i \in Q)$ and a finite family of subsets $(Q(\alpha), \alpha \in \Gamma_\alpha)$ of the index set Q and define $I_\alpha = \prod_{i \in Q(\alpha)} J_i$ and $W = \sum_{\alpha \in \Gamma_\alpha} I_\alpha$. Partition Γ_α into $\Gamma_\alpha^+ \cup \Gamma_\alpha^i$, where $\Gamma_\alpha^+ = \{\beta \neq \alpha : Q(\alpha) \cap Q(\beta) \neq \emptyset\}$ and define $\delta = \frac{1}{\mathbb{E}W} \sum_{\alpha \in \Gamma_\alpha} \sum_{\beta \in \Gamma_\alpha^+} \mathbb{E} I_\alpha I_\beta$. Then for any $0 \leq \varepsilon \leq 1$ Janson obtained

$$\mathbb{P}\big(W \leq (1-\varepsilon)\mathbb{E}W\big) \leq \exp\left(-\frac{1}{2(1+\delta)}\varepsilon^2 \mathbb{E}W\right). \tag{2.22}$$

3. Applications

3.1. **k-runs.** Consider the problem of k-runs of 1's in a series of independent identically distributed Bernoulli random variables. When approximating the distribution of the number of overlapping k-runs of heads, in [2, Section 4.1.2] Arratia, Goldstein and Gordon give a bound on the total variation distance from a compound Poisson approximation of order $nkp^{2k}(1-p)$ and Roos improves this in [14] to order $kp^k \log(np^k)$ when $p < 1/2$. Theorem 4.3 in [9] leads to a bound on the Kolmogorov distance from the same compound Poisson distribution of order $kp^k + \exp(-cnp^k)$ for some $c > 0$, as was already mentioned in [9]. We want to show the calculations for the bound on the Kolmogorov distance as well as on the total variation distance. Suppose n satisfies $n > 4k-3$. Consider a collection of independent Bernoulli random variables $(J_j, j \in \{1, \ldots, n\})$ with $\mathbb{P}(J_j = 1) = p$. To avoid edge effects assume that indices of the form $i + nl$ are identified with i whenever $1 \leq i \leq n$ and $l \in \mathbb{Z}$. Using this circular convention define the family $(Q(\alpha), \alpha \in \Gamma)$ of subsets of index set $\{1, \ldots, n\}$ consisting of k consecutive indices. These are the critical combinations in the reliability theory. We identify each index set α consisting of k consecutive indices with the first index on its left hand side. Thus $|\Gamma| = n$. We define $I_\alpha = \prod_{j=\alpha}^{\alpha+k-1} J_j$ and $W = \sum_{\alpha \in \Gamma} I_\alpha$. Then $\mathbb{E}I_\alpha = p^k$ and $\mathbb{E}W = np^k$. The random variables $(I_\alpha, \alpha \in \Gamma)$ are dissociated. Choose $\Gamma_\alpha^{vs} = \{\alpha - (k-1), \ldots, \alpha - 1, \alpha + 1, \ldots, \alpha + k - 1\}$ and Γ_α^0 and Γ_α^w as in the dissociated partition. Then applying Corollary 2.13 we obtain $\varepsilon_1 = (6k-5)n(p^k)^2$. Furthermore we have to calculate the limiting measure $CP(\lambda, \mu)$: Lemma 3.3.4 in [14] states

$$\lambda_i = \begin{cases} np^k p^{i-1}(1-p)^2 & \text{for } i = 1, \ldots, k-1, \\ \frac{np^k}{i}\left(2p^{i-1}(1-p) + (2k-i-2)p^{i-1}(1-p)^2\right) & \text{for } i = k, \ldots, 2k-2, \\ \frac{np^k}{2k-1} p^{2k-2} & \text{for } i = 2k-1. \end{cases} \tag{3.1}$$

Thus we can calculate $\lambda = \sum_{i=1}^{2k-1} \lambda_i$, the μ_i's, and the moments of μ; for example, $m_1 = np^k/\lambda$. It is obvious that m_4 is finite and Condition 2.5 is fulfilled, since

$\mu_1 = \lambda_1/\lambda = np^k(1-p)^2/\lambda > 0$ (see [9, Remark before Lemma 4.1]). So we can apply (2.7): therefore, we have to determine the order of $J_l^{(a)}(\lambda,\mu)$, $l = 0,1$, and to calculate the remaining contribution $\mathbb{P}\big(W \le \lambda m_1(1-\varrho^*(\zeta_0)/4)\big)$ using an appropriate inequality.

Since $\lambda = O(n\,p^k)$, the moments m_1 and m_2 are of order $O(1)$ and the same holds for ζ_0 and $\varrho^*(\zeta_0)$. Therefore, the order of $J_0^{(a)}(\lambda,\mu)$ is $O(1/(n\,p^k)^{1/2})$ and $J_1^{(a)}(\lambda,\mu)$ is of order $O(1/(n\,p^k))$. Since $\varepsilon_1 = O(k\,n(p^k)^2)$, the second summand in (2.7) has order $O(k\,p^k)$. To calculate the order of the remaining contribution (the third summand in (2.7)), we apply Janson's inequality (2.22): With $\Gamma_\alpha^+ = \Gamma_\alpha^{vs}$ we obtain in the k-run example $\delta = (4k-4)p^k$ and get

$$\mathbb{P}\big(W \le \lambda m_1(1-\varrho^*(\zeta_0)/4)\big) \le \exp\left(-\lambda m_1(\varrho^*(\zeta_0)/4)^2\frac{1}{2(1+(4k-4)p^k)}\right).$$

Summarizing we find that in Kolmogorov distance the right hand side of (2.7) has the order

$$O\big(kp^k + \exp(-\text{const.}\,np^k)\big) \tag{3.2}$$

when the mean number of runs $n\,p^k$ satisfies the condition $n\,p^k \ge 1$.

The bound in total variation is only available in the case $\lambda \ge 2$, which roughly means that $n\,p^k \ge 2$. Of course, the assumption $R(\mu) > 1$ is fulfilled. We have to check the influence on the order of the constants $C_l(\mu)$ for $l = 0,1$. We see that this order is dependent only on the order of the moments and therefore in our example $H_0^{(a)}(\lambda,\mu)$ is of order $O(1/(n\,p^k)^{1/2})$ and $H_1^{(a)}(\lambda,\mu)$ of order $O(1/(n\,p^k))$. The main significance is that the constants $C_l(\mu)$ exist. They are not optimal. We use the fact that $C_2(\mu) < 1$ to choose c_1 suitably and therefore, we can apply (2.4) and bound $\mathbb{P}\big(W \le \frac{1}{2}(1+c_1)\lambda m_1\big)$ again by Janson's inequality and obtain the same order (3.2) of the bound in total variation distance, possibly with less favorable constants, and under the condition that $n\,p^k \ge 2$. Note that apart from this one, there is no other restriction on the value p. We see that this bound is quite applicable when $\mathbb{E}W = np^k \to \infty$ which is not the case for the bounds in [2] and [14]. The importance of k-runs for the reliability theory called there consecutive-k-out-of-n systems has been discussed in [4].

3.2. **Isolated vertices in the rectangular lattice on the torus.** Consider a rectangular lattice on the torus with n vertices and $N = 2n$ edges (see [16, Section 3]). Note that it is a 4-regular graph without triangles. Assume that the edges can be deleted independently of each other with a constant probability $1-p=q$. Define a family $(Q(\alpha), \alpha \in \Gamma)$ of subsets of the vertex set $\{1,2,\dots,n\}$ consisting of a vertex and its four neighbours. These are the critical combinations of the system. We identify each index set consisting of 5 vertices with the center. Thus $|\Gamma| = n$ and we define $I_\alpha = I[v_\alpha \text{ is isolated}]$ and $W = \sum_{\alpha\in\Gamma} I_\alpha$. With the above definitions $\mathbb{P}(v_\alpha \text{ is isolated}) = q^4$ and $\mathbb{E}W = nq^4$. The random variable W counts the number of isolated vertices in the graph. The idea is to give a bound for reliability in the system. By construction and the choice of Γ_α^{vs} the random variables $(I_\alpha, \alpha \in \Gamma)$ are dissociated. Choose $\Gamma_\alpha^{vs} = \{\beta \neq \alpha : \, v_\alpha \,\text{and}\, v_\beta \,\text{are neighbours}\}$ and Γ_α^0 and Γ_α^w as in the dissociated partition. Applying Corollary 2.13 we obtain $\varepsilon_1 = 21nq^8$ and by [16, Section 3]

$$\lambda_i = \frac{n}{i}\binom{4}{i-1} q^{3i+1}(1-q^3)^{5-i}, \text{ for } i = 1, \dots, 5.$$

Since $\mu_1 = nq^4(1-q^3)^4/\lambda > 0$ and m_4 is finite, $\lambda = O(nq^4)$, and the moments m_i are of order $O(1)$, the second summand in (2.7) has order $O(q^4)$. We can apply Janson's inequality (2.22) with $\delta = 12q^4$ and obtain in Kolmogorov distance that the upper bound is of order

$$O\big(q^4 + \exp(-\text{const. } nq^4)\big) \tag{3.3}$$

provided $nq^4 \geq 1$. Since $R(\mu) > 1$, the same order holds in total variation distance when $nq^4 \geq 2$.

It is easily seen that our approximating distribution has the form $CP(\lambda, \mu) = \sum_{i=1}^5 i\, Z_i$, where the Z_i are independent random variables distributed according to $Po(\lambda_i)$. It follows from [16, Section 3] that in such a case an order $O\big((\log^+ 2nq^4)q^4\big)$ of the upper bound for $q^3 \leq 1/5$ could be expected when the local approach of [16] would be applied directly. The restrictive condition $q^3 \leq 1/5$ is needed there to assure that $i\lambda_i \searrow 0$. For any q another compound Poisson approximation was suggested in [16]. Her $CP(\lambda^*)$ can be expressed in our notation as $CP(\lambda, \mu')$ for a particular choice of μ' suggested in (2.20) in Remark 2.18 with $l = 2$. In such a case

$$\sum_{i=3}^{5} i(i-1)\, \lambda_i = 12\, n\, q^{10}\big(1 + O(q^3)\big).$$

By (2.19) in Remark 2.18 we obtain for $d_{TV}\big(\mathcal{L}(W), CP(\lambda, \mu')\big)$ the same order as in (3.3). Note that our result improves the one of [16] as she was only able to prove that the order of the upper bound on $d_{TV}\big(\mathcal{L}(W), CP(\lambda, \mu')\big)$ is $O\big((\log^+ 2\, n\, q^4)\, q^4\big)$. Our condition $n\, q^4 \geq 2$ is hardly restrictive. Moreover, our bound (3.3) becomes rapidly smaller than the earlier estimate when $n\, q^4$ increases. Note that for Poisson approximation, when $n\, q^4 \to$ const., the approximation is only accurate to order $O(n^{-3/4})$. Similarly to [16, Section 3] our result allows us to approximate the reliability of the system by the distributions $CP(\lambda, \mu)$ or $CP(\lambda, \mu')$ if q^4 is small.

3.3. **Consecutive 2-systems, coloured graphs, birthday problems.** Given a graph $G(V, E)$ with n vertices and N edges, where the vertices are subject to fail. Take $\Gamma = E$ and associate with each vertex v_j a Bernoulli random variable Y_j where

$$\mathbb{P}(Y_j = 0) = P\big(\text{the item in } v_j \text{ fails (is down)}\big) = q_j$$

for $j = 1, \dots, n$. We consider the case when $Y_1, \dots, Y_n$ are independent. For $\alpha \in \Gamma$ set $I_\alpha = I(Y_i = Y_j = 0) = I$(the two items of the edge α fail) with $\alpha = \{i, j\}$. Edges are critical combinations here. Let $W = \sum_{\alpha\in\Gamma} I_\alpha$ be the number of all edges in the graph, in which both vertices are failed. Of course the family $(I_\alpha, \alpha \in \Gamma)$ is dissociated. Define the strong neighbourhood of dependence by $\Gamma_\alpha^{vs} = \{\beta \in \Gamma \setminus \{\alpha\} : |\beta \cap \alpha| = 1\}$ consisting of edges with one vertex of α and one other vertex. We write $\pi_\alpha = \prod_{i\in\alpha} q_i$ for the probability that both vertices of the edge α fail. Then we obtain from Corollary 2.13

$$\varepsilon_1 = \sum_{\alpha\in\Gamma} \pi_\alpha \left(\pi_\alpha + \sum_{\beta\in\Gamma_\alpha^{vs}} \prod_{j\in\beta} q_j + 2 \sum_{\beta\in\Gamma_\alpha^0} \prod_{j\in\beta} q_j \right).$$

If $D(l)$ denotes the degree of the vertex v_l in the graph and $D = \max_{1\le l\le n} D(l)$, we obtain $|\Gamma_\alpha^{vs}| \le 2(D-1)$ and $|\Gamma_\alpha^0| \le 2(D-1)^2$. With $q_{\max} = \max_{1\le i\le n} q_i$ we get the bound $\varepsilon_1 \le (4D^2 - 6D + 3) q_{\max}^2 \sum_{\alpha\in\Gamma} \pi_\alpha$ (see [5]). If all the failure probabilities are equal to q, it follows that $\varepsilon_1 \le (4D^2 - 6D + 3) q^4 N$. In the notation of [5] we denote for each $\alpha = \{i, j\}$ by $S_2(\alpha)$ the set of vertices s such that $\{s, i\}$ and $\{s, j\}$ are in E and by $S_1(\alpha)$ the set of vertices s such that there is only one edge connecting s with α. So Z_α can be written as $\sum_{\beta\in\Gamma_\alpha^{vs}\,:\,|\beta\cap S_1(\alpha)|=1} I_\beta + \sum_{\beta\in\Gamma_\alpha^{vs}\,:\,|\beta\cap S_2(\alpha)|=1} I_\beta =: Z_\alpha^1 + Z_\alpha^2$ and we have

$$\lambda_i = \frac{1}{i} \sum_{\alpha\in\Gamma} \pi_\alpha \mathbb{P}\big(Z_\alpha^1 + 2Z_\alpha^2 = i - 1 | I_\alpha = 1\big). \tag{3.4}$$

Since Z_α^1 is distributed according to a Binomial distribution with parameters $|S_1(\alpha)|$ and q and Z_α^2 to a Binomial distribution with parameters $|S_2(\alpha)|$ and q and Z_α^1 and Z_α^2 are independent, λ has order $O(Nq^2)$ and therefore the second summand in (2.7) has order $O(\mathrm{poly}(D)\, q^2)$, where $\mathrm{poly}(D)$ denotes a positive polynomial of degree 2 in D. We can apply Janson's inequality (2.22) with $\delta = 2(D-1)^2 q^2$ and thus the order of the bound in total variation is $O\big(\mathrm{poly}(D) q^2 + \exp(-\,\mathrm{const.}\; N\,\mathrm{poly}(D) q^2)\big)$ which holds for $Nq^2 \ge 2$. This again is an improvement of existing results (see [5, Section 2] and [14, Theorem 3.F]). In [5] the underlying graph to be a complete d-regular tree was considered. If d denotes the degree of any non-leaf vertex, the upper bound is $O\big((1 + \log^+(Nq^2)) q^2 (5d^2 - 5d - 1)\big)$, so for constant $\mathbb{E}W = Nq^2$ the order of approximation is $O(q^2)$. Again we get this order even in the case when $\mathbb{E}W \to \infty$. Provided $Nq^2 \ge 1$ we obtain the same order of the bound in Kolmogorov's distance.

The consecutive 2-system has another interpretation: colour the vertices of the fixed graph $G(V, E)$ at random, independently of each other and with the same probability distribution over the colours. Let X_i denote the colour of vertex v_i and set $p_r = \mathbb{P}(X_i = r)$ and take $I_\alpha = I(X_i = X_j)$ with $\alpha = \{i, j\}$. Then $W = \sum_{\alpha\in\Gamma} I_\alpha$ is the number of edges connecting two vertices of the same colour. Else, the vertices of the graph denote people, the colours their birthdays, and the edges pairs of people who are acquainted. Now W counts the number of pairs of people who are acquainted and have the same birthday. Here $\pi = \pi_\alpha = \mathbb{P}(I_\alpha = 1) = \sum_{r=1}^m p_r^2$. So we are in the situation of the consecutive 2-system with equal failure probability π. The bound in total variation distance is of order

$$O\Big(\mathrm{poly}(D) \sum_{r=1}^m p_r^2 + \exp(-\mathrm{const.}\ \mathrm{poly}(D) \sum_{r=1}^m p_r^2)\Big),$$

which is an improvement of [7, Theorem 5.G and 5.H]. The parameters λ_i can be calculated explicitly by (3.4). The Kolmogorov distance has the same order if $N \sum_{r=1}^m p_r^2 \geq 1$.

3.4. **Connected s-systems.** Observe that we can get improved bounds also for the so called connected s-systems which are studied in [5] and [14]. Here one assumes that critical combinations are a subset of all the connected subgraphs with s vertices in the underlying graph. Moreover, one assumes that the system fails if there are at least m ($m \geq 1$) subgraphs on s vertices with all items failed. Here Γ denotes the set of all critical combinations, an element can be denoted as $(k_1, \ldots, k_s)$, the indices of s vertices which span a particular connected subgraph. To get a family of dissociated random variables, we have to consider the family $(Q(\alpha), \alpha \in \Gamma)$ of subsets of the index set $\{1, 2, \ldots, n\}$ consisting of s indices of s vertices which span a particular connected subgraph. Define $I_\alpha = I(\text{all the items of the } s \text{ items of the } \alpha \text{ subgraph fail})$. Defining Γ_α^{vs} we have the choice for the extend of overlap R: for $1 \leq R \leq s-1$ we can define

$$\Gamma_\alpha^{vs}(R) = \{\beta \in \Gamma \setminus \{\alpha\} : |\beta \cap \alpha| = r, \quad r = R, \ldots, s-1\}. \tag{3.5}$$

If the failure probabilities are equal q, one obtains for $|\Gamma| q^{s-R+1} \geq 2$ a bound of the order

$$O\big(\mathrm{const.}\, q^{s-R+1} + \exp(-\mathrm{const.} |\Gamma| q^{s-R+1})\big),$$

where the constants depend on the graph, especially on $|\{\beta : |\beta \cap \alpha| = r\}|$, $r = 1, \ldots, s-1$ and $|\{\gamma \in \Gamma_\alpha^0(R) : |\gamma \cap \alpha| = 0, |\gamma \cap \beta| \neq \emptyset \text{ for some } \beta \in \Gamma_\alpha^{vs}(R)\}|$. For details see [14]. The Kolmogorov distance has the same order if $|\Gamma| q^{s-R+1} \geq 1$.

3.5. **Two dimensional consecutive-k-out-of-n system.** This example is a particular connected-k^2 example. We will discuss it briefly. The system consists of n^2 components placed on a square grid of size n and it fails if and only if there exist at least m possibly overlapping square sub-grids of size k ($1 < k < n$) with all k^2 components failed. The square subgrids of size k are the critical combinations. $\Gamma = \{(r,s) \colon 1 \leq r, s \leq n-k+1\}$ and for $\alpha \in \Gamma$ denote by $A_\alpha = A_{rs}$ the $k \times k$ sub-grid with left lower-most component (r,s). We do not place it on a torus. Consequently, we have to take boundaries into account. Define a family of subsets $(Q(\alpha), \alpha \in \Gamma)$ of the set $\{1, 2, \ldots, n^2\}$ consisting of the points of a $k \times k$ sub-grid and identify each index with the left lower-most component. Define $I_\alpha = I(\text{all items in } A_\alpha \text{ are failed})$ for each $\alpha \in \Gamma$ and $W = \sum_{\alpha \in \Gamma} I_\alpha$. To define the neighbourhood of very strong dependence, one has to decide how large R has to be (see (3.5)).. First of all we take $R=k^2-k$. There are $(n-k+1)^2$ possible positions of the $k\times k$ sub-grid, 4 positions where $|\Gamma_\alpha^{vs}(k^2-k)|=2$ (corners), $4(n-k-1)$ positions where $|\Gamma_\alpha^{vs}(k^2-k)|=3$ (border), and $(n-k-1)^2$ more where $|\Gamma_\alpha^{vs}(k^2-k)|=4$. If the failure probability is q for all items in the grid, we obtain $\mathbb{E}W=(n-k+1)^2q^{k^2}$ and $|\Gamma_\alpha^{vs}(k^2-k)|\le 4$ and $|\Gamma_\alpha^{vs}(k^2-k)|+|\Gamma_\alpha^{0}(k^2-k)|\le(2k+1)^2-1$. The calculations similar to those in [5, Section 4] and in [14, Section 3.7] lead to the following bound for (2.9):

$$\varepsilon_1\le(n-k+1)^2q^{k^2}\left((4k^2+12k-3)q^{k^2}+4\left(\sum_{u=1}^{k-1}\sum_{v=1}^{k-1}q^{k^2-uv}+\sum_{v=1}^{k-2}q^{k^2-kv}\right)\right),$$

which is of order $O\big(q^{2k-1}(n-k+1)^2q^{k^2}\big)$. Moreover, the λ_i's can be computed as follows:

$$\begin{aligned}\lambda_i=\frac{q^{k^2}}{i}\Bigg(&4\binom{2}{i-1}q^{k(i-1)}(1-q^k)^{3-i}+4(n-k-1)\binom{3}{i-1}q^{k(i-1)}(1-q^k)^{4-i}\\&+(n-k-1)^2\binom{4}{i-1}q^{k(i-1)}(1-q^k)^{5-i}\Bigg)\end{aligned}$$

for $i=1,\dots,5$. Again $\mu_1>0$ and $R(\mu)>1$. We obtain $\lambda=O\big((n-k+1)^2\,q^{k^2}\big)$. δ in Janson's inequality can be bounded by $(2k+1)^2-2$ and we obtain in total variation distance that the upper bound is of order

$$O\big(q^{2k-1}+\exp(-\text{const.}\ (n-k+1)^2q^{k^2})\big),$$

whenever $(n-k+1)^2\,q^{k^2}\ge 2$. In Kolmogorov distance we obtain the same order if $(n-k+1)^2\,q^{k^2}\ge 1$. The argument presented here provide an improvement on the order $O\big(q^{2k-1}\log^+((n-k+1)^2q^{k^2})\big)$ obtained in [14, Theorem 3.G]. The Poisson approximation is given in [5, Section 4]. Using the Poisson local approach we obtain an improvement on the order $O(q^k)$ by a factor q^{k-1}. Moreover, the bound is applicable when $\mathbb{E}W\to\infty$ which is not the case for the bounds in [5] and in [14]. The extension of the above result to the case of unequal probabilities of failure of items can be carried out according to Remark 4.2 in [5]. The reliability of the system and bounds on the reliability of the system for $m=1$ can be calculated as in [5]. One might be interested in taking $R=1$ (maximal clump) instead of $R=k^2-k$. It would imply ε_1 to be of order $O\big((n-k+1)^2\,q^{k^2}\,q^{k^2}\big)$. Thus the order of the upper bound would be $O(q^{k^2})$. However, in such a case the computation of the parameters λ_i would be very laborious.

3.6. **U-statistics.** In our last example we study incomplete and complete U-statistics. We observe that for general nonnegative integer valued complete U-statistics the bound in total variation between the law of the statistic and an appropriate compound Poisson distribution introduced in Section 1 has the same order as the one in the Poisson case, provided we choose the neighbourhood of dependence similar to (3.5). This is obvious, since the neighbourhoods are indeed not local. Suppose that Γ is a collection of k-subsets $\alpha = \{\alpha_1, \dots, \alpha_k\}$ of $\{1, 2, \dots, n\}$ and define $X_\alpha = \phi_\alpha(Y_{\alpha_1}, \dots, Y_{\alpha_k})$ for some nonnegative integer valued symmetric (in its arguments) functions ϕ_α, where $Y_1, \dots, Y_n$ are independent random elements of some space $\mathcal{X}$. Let $W = \sum_{\alpha \in \Gamma} X_\alpha$. Given $R \geq 1$, similarly to (3.5), the neighbourhoods of dependence can be defined by

$$\Gamma_\alpha^{vs} = \{\beta : \; R \leq |\beta \cap \alpha| \leq k - 1\}.$$

Thus,

$$\Gamma_\alpha^0 = \{\gamma : \; |\gamma \cap \beta| \geq 1 \text{ for some } \beta \in \{\alpha\} \cup \Gamma_\alpha^{vs}(R)\} \setminus \{\{\alpha\} \cup \Gamma_\alpha^{vs}(R)\}.$$

Here R denotes the minimum of the number of items in common for β and α when $\beta \in \Gamma_\alpha^{vs}(R)$. There are two extreme cases of the above formulae. If one takes $R = k$, then $\Gamma_\alpha^{vs} = \emptyset$ and one considers the Poisson case. The other extreme, $R = 1$, indicates that Γ_α^{vs} is the set of all $\beta \in \Gamma \setminus \{\alpha\}$ which overlap with α. The choice of R affects the order of the term $\mathbb{E}(X_\alpha U_\alpha)$ which is frequently that of the largest order. Now let us calculate ε_1 for W depending on R. Denote $\mathbb{E}(X_\alpha) = \pi_\alpha$. For $1 < R < k$ the terms of (2.9) are of the following form:

$$\sum_{\beta \in \Gamma_\alpha^0(R)} \mathbb{E}(X_\alpha X_\beta) = \sum_{r=1}^{R-1} \sum_{\substack{\beta \in \Gamma_\alpha^0(R) \\ |\beta \cap \alpha| = r}} \mathbb{E}(X_\alpha X_\beta) + \sum_{\substack{\beta \in \Gamma_\alpha^0(R), |\beta \cap \alpha| = 0 \\ |\gamma \cap \beta| \neq 0, \gamma \in \Gamma_\alpha^{vs}(R)}} \pi_\alpha \pi_\beta$$

and

$$\sum_{\beta \in \Gamma_\alpha^{vs}(R) \cup \Gamma_\alpha^0(R)} \pi_\alpha \pi_\beta = \sum_{r=1}^{k-1} \Bigg(\sum_{\substack{\beta \in \Gamma_\alpha^{vs}(R) \\ |\beta \cap \alpha| = r}} \pi_\alpha \pi_\beta \Bigg) + \sum_{\substack{\beta \in \Gamma_\alpha^0(R), |\beta \cap \alpha| = 0 \\ |\gamma \cap \beta| \neq 0, \gamma \in \Gamma_\alpha^{vs}(R)}} \pi_\alpha \pi_\beta.$$

Consider the case $\pi_\alpha = \pi$ for all $\alpha \in \Gamma$. Let $n_r(\alpha) = |\{\beta : \; |\beta \cap \alpha| = r\}|$, $r = 1, \dots, k-1$, $m(\alpha, R) = |\{\gamma \in \Gamma_\alpha^0(R) : \; |\gamma \cap \alpha| = 0, |\gamma \cap \beta| \neq \emptyset \text{ for some } \beta \in \Gamma_\alpha^{vs}(R)\}|$ and $\sigma_r = \mathbb{E}(X_\alpha X_\beta)$ for all pairs α, β satisfying $|\alpha \cap \beta| = r$. Then we obtain

$$\varepsilon_1 = \sum_{\alpha \in \Gamma} \Bigg(\pi^2 + \sum_{r=1}^{R-1} n_r(\alpha) \sigma_r + m(\alpha, R) \pi^2 + \sum_{r=1}^{k-1} \Big(n_r(\alpha) + m(\alpha, R) \Big) \pi^2 \Bigg).$$

In the language of reliability theory a similar formula was obtained in [14, Section 3.5]. In fact, the class considered there was a special class of incomplete U-statistics. Let us consider the case $|\Gamma| = \binom{n}{k}$. We obtain $|\{\beta : \; |\beta \cap \alpha| = r\}| = \binom{k}{r}\binom{n-k}{k-r}$, which is of order $O(n^{k-r})$. The size of the strong dependence region is of order $O(n^{k-R})$. Moreover, it is obvious that $|\Gamma_\alpha^0| = \binom{n}{k} - \sum_{r=R}^{k-1} \binom{k}{r}\binom{n-k}{k-r}$, so $\Gamma_\alpha^w = \emptyset$. For simplicity we assume that the constants $C_l(\lambda, \mu)$ in the total variation distance bound do not influence the order of the bound (as we have observed in all our other examples) and that μ satisfies the assumption $R(\mu) > 1$, $\lambda \geq 2$ and fulfills Condition 2.5. Moreover, apply Chebychev's inequality to establish that

$$\mathbb{P}\big(W \leq 1/2(1+c_1)\lambda m_1\big) \leq \lambda\, m_2 \big(1/2(1-c_1)\lambda\, m_1\big)^{-1} = O(1/\lambda)$$

(use Bernstein-type or Hoeffding-type inequalities for U-statistics for more refinement, see for example [10, Section 2.1, 8.1]). Therefore, for complete U-statistics we find that

$$d_{TV}\big(\mathcal{L}(W), CP(\lambda,\mu)\big) \leq \text{const.}(\mu)\,\pi \binom{n}{k} + \sum_{r=1}^{R-1} \frac{\sigma_r}{\pi} n_r(\alpha) = O\big(\pi n^k + \sum_{r=1}^{R-1} \frac{\sigma_r}{\pi} n^{k-r}\big).$$

The order for Poisson approximation is $O\big(\pi n^{k-1} + \sum_{r=1}^{k-1} \frac{\sigma_r}{\pi} n^{k-r}\big)$ (see [7, Section 2.3, (3.2)]). Since in the compound Poisson case the order is influenced by $\pi\, n^k$, the result is not an improvement in general. Consider the following example: $k = 2$ and $\phi(x,y) = \phi_\alpha(x,y) = x\,y$ and Y_1 is an indicator with $\mathbb{P}(Y_1 = 1) = p$. Then the order of Poisson approximation is $p^2\, n + p\, n$, whereas the order of compound Poisson approximation is $p\, n$.

In the incomplete case our result is applicable, provided the quantities $n_r(\alpha)$ and $m(\alpha, R)$ are uniformly bounded as Γ grows. The smaller R is the better is the order of approximation. Examples 3.1 to 3.5 above are incomplete U-statistics. The choice of Γ, the collection of k-subsets, represents the local character shared by all our examples. In the k-run example the k-subsets are k consecutive indices, thus $|\Gamma| = n$ and the local clump has size k. In the connected s-system we have a fixed set of all critical combinations. These observations lead to a natural question: "how incomplete" should a U-statistics be to get better results when approximating by a compound Poisson distribution rather than a Poisson one.

References

1. D. J. Aldous, *Probability approximations via the Poisson clumping heuristic*, Springer-Verlag, New York, 1989.
2. R. Arratia, L. Goldstein, and L. Gordon, *Poisson approximation and the Chen-Stein method*, Stat. Science **5** (1990), 403–434.
3. A. D. Barbour, L. H. Chen, and W.-L. Loh, *Compound Poisson approximation for nonnegative random variables via Stein's method*, Ann. Probab. **20** (1992), 1843–1866.
4. A. D. Barbour, O. Chryssaphinou, and M. Roos, *Compound Poisson approximation in reliability theory*, IEEE Transactions on Reliability **44** (1995), 398–402.
5. ———, *Compound Poisson approximation in systems reliability*, Naval Research Logistics **43** (1996), 251–264.
6. A. D. Barbour and G. K. Eagleson, *Poisson convergence for dissociated statistics*, Journal of the Royal statistical society **B 46** (1984), 397–402.
7. A. D. Barbour, L. Holst, and S. Janson, *Poisson Approximation*, Oxford University Press, Oxford, 1992.
8. A. D. Barbour and S. Utev, *Compound Poisson approximation in total variation*, preprint, 1997.

9. ———, *Solving the Stein equation in compound Poisson approximation*, Advances in Applied Probability (1998), (to appear).
10. Yu. V. Borovskikh, *U-Statistics in Banach Spaces*, VSP, Utrecht, 1996.
11. A. J. Lee, *U-Statistics: Theory and Practice*, Marcel Dekker, New York, 1990.
12. W. G. McGinley and R. Sibson, *Dissociated random variables*, Mathematical Proceedings of the Cambridge philosophical society **77** (1975), 185–188.
13. M. Roos, *Compound Poisson approximations for the numbers of extreme spacings*, Adv. Appl. Probab. **25** (1993), 847–874.
14. ———, *Stein-Chen method for compound Poisson approximation*, Ph.D. thesis, University of Zürich, Switzerland, 1993.
15. ———, *Stein-Chen method for compound Poisson approximation: the coupling approach*, Probab. Theory and Math. Stat. **4** (1994), 645–660.
16. ———, *Stein's method for compound Poisson approximation: the local approach*, Ann. Appl. Probab. **4** (1994), no. 4, 1177–1187.

(Peter Eichelsbacher) Fakultät für Mathematik, Universität Bielefeld, Postfach, 10-0131, D-33501 Bielefeld, Germany

E-mail address, Peter Eichelsbacher: `peter@mathematik.uni-bielefeld.de`

(Małgorzata Roos) Biostatistik, ISPM, Universität Zürich, Sumatrastrasse 30, CH-8006 Zürich, Schweiz

E-mail address, Małgorzata Roos: `mroos@ifspm.unizh.ch`