Refined total variation bounds in the multivariate and compound Poisson approximation

Bero Roos

FB IV – Department of Mathematics,
University of Trier,
54286 Trier, Germany.
E-mail address: bero.roos@uni-trier.de
URL: http://www.math.uni-trier.de/~roos/

Abstract. We consider the approximation of a convolution of possibly different probability measures by (compound) Poisson distributions and also by related signed measures of higher order. We present new total variation bounds having a better structure than those from the literature. A numerical example illustrates the usefulness of the bounds, and an application in the Poisson process approximation is given. The proofs use arguments from Kerstan (1964) and Roos (1999b) in combination with new smoothness inequalities, which could be of independent interest.

1. Introduction

1.1. *Aim of the paper.* Nowadays, there are numerous results in the (compound) Poisson approximation of convolutions of probability distributions (cf. Arak and Zaĭtsev, 1986; Barbour et al., 1992). However, it turned out that the investigation in the multidimensional case is somewhat difficult. Even in the simple case of Poisson approximation of the generalized multinomial distribution, the correct order of approximation is not exactly known. Indeed, to the best of our knowledge, the literature does not contain any lower and upper total variation bounds differing only by an absolute constant factor. The problem here is that not only the number of convolution factors but also the dimension can be arbitrarily large. But there are useful approximation results, see, e.g., Franken (1963), U. Herrmann (1965b), Deheuvels and Pfeifer (1988), Roos (1999b), Barbour (2005). In this paper, we show how some bounds from Roos (1999b) can be further substantially improved.

We also indicate how these improved bounds in combination with ideas in Roos (2007) can be useful in the compound Poisson approximation.

The paper is organized as follows. In the next three subsections, we explain the notation, comment on the method used, give a review of some results from the literature and discuss the benefits of some result of the present paper. Sections 2 and 3 are devoted to the main results and an application in the Poisson process approximation. In Section 4, we present some auxiliary norm estimates including smoothness inequalities as well as the proofs of the results.

1.2. *Notation.* In what follows, let $(\mathfrak{X}, +, \mathcal{A})$ be a measurable Abelian semigroup with zero element, that is, $(\mathfrak{X}, +)$ is a commutative semigroup with identity element 0 and $\mathcal{A}$ is a σ-algebra of subsets of $\mathfrak{X}$ such that the mapping $\mathfrak{X} \times \mathfrak{X} \ni (x, y) \mapsto x + y \in \mathfrak{X}$ from $(\mathfrak{X} \times \mathfrak{X}, \mathcal{A} \otimes \mathcal{A})$ to $(\mathfrak{X}, \mathcal{A})$ is measurable. In particular this implies that, for arbitrary $y \in \mathfrak{X}$, the mapping $\mathfrak{X} \ni x \mapsto x + y \in \mathfrak{X}$ is measurable as well. The approach used in this paper requires a measure theoretic setting. Random variables are rarely needed or used. Let $\mathcal{F}$ (resp. $\mathcal{M}$) be the set of all probability distributions (resp. finite signed measures) on $(\mathfrak{X}, \mathcal{A})$. Products and powers of finite signed measures in $\mathcal{M}$ are defined in the convolution sense, that is, for $V, W \in \mathcal{M}$ and $A \in \mathcal{A}$, we write

$$VW(A) = \int_{\mathfrak{X}} V(\{y \in \mathfrak{X} \mid x + y \in A\}) \, \mathrm{d}W(x).$$

Empty products and powers of signed measures in $\mathcal{M}$ are defined to be δ_0, where δ_x is the Dirac measure at point $x \in \mathfrak{X}$. Let $V = V^+ - V^-$ denote the Hahn-Jordan decomposition of $V \in \mathcal{M}$ and let $|V| = V^+ + V^-$ be its total variation measure. The total variation norm of V is defined by $\|V\| = |V|(\mathfrak{X})$. We note that the total variation distance between two finite signed measures $V, W \in \mathcal{M}$ is usually defined by $d_{\mathrm{TV}}(V, W) = \sup_{A \in \mathcal{A}} |V(A) - W(A)|$. However, this distance is rarely needed or used here, since $d_{\mathrm{TV}}(V, W) = \frac{1}{2}\|V - W\|$ provided that $V(\mathfrak{X}) = W(\mathfrak{X})$, which in concrete situations is often the case. With the usual operations of real scalar multiplication, addition, together with convolution and the total variation norm, $\mathcal{M}$ is a real commutative Banach algebra with unity δ_0, see for example Section 2 in Liese (1987). For $V \in \mathcal{M}$ and a power series $g(z) = \sum_{m=0}^{\infty} a_m z^m$ with $a_m \in \mathbb{R}$ converging absolutely for each complex $z \in \mathbb{C}$ with $|z| \leqslant \|V\|$, we set $g(V) = \sum_{m=0}^{\infty} a_m V^m \in \mathcal{M}$. The exponential of $V \in \mathcal{M}$ is defined by the finite signed measure

$$\mathrm{e}^V = \exp(V) = \sum_{m=0}^{\infty} \frac{V^m}{m!} \in \mathcal{M}.$$

In particular, $\mathrm{CPo}(t, F) := \exp(t(F - \delta_0))$ is the compound Poisson distribution with parameters $t \in [0, \infty)$, $F \in \mathcal{F}$. In other words, this is the distribution of the random sum $\sum_{j=1}^{N} X_j$, where $N, X_j, (j \in \mathbb{N})$ are independent random variables, N has values in $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$ and has Poisson distribution $\mathrm{Po}(t) := \exp(t(\delta_1 - \delta_0)) = \mathrm{CPo}(t, \delta_1)$ with mean t, whereas the $\mathfrak{X}$-valued X_j are identically distributed with distribution F. We denote the counting density of $\mathrm{Po}(t)$ by $\mathrm{po}(\cdot, t) : \mathbb{Z} \longrightarrow [0, 1]$, where $\mathbb{Z}$ is the set of integers, $\mathrm{po}(m, t) = \mathrm{e}^{-t} \frac{t^m}{m!}$ for $m \in \mathbb{Z}_+$ and $\mathrm{po}(m, t) = 0$ otherwise. If F and G are non-negative measures on $(\mathfrak{X}, \mathcal{A})$ and F is absolutely continuous with respect to G, we write $F \ll G$. For a set A, let $\mathbb{1}_A(x) = 1$ if $x \in A$ and $\mathbb{1}_A(x) = 0$ otherwise. Set $\underline{0} = \emptyset$ and $\underline{n} = \{1, \dots, n\}$ for $n \in \mathbb{N} = \{1, 2, \dots\}$;

further, for $n \in \mathbb{Z}_+$, set $\underline{n}_0 = \{0, \ldots, n\}$. For a finite set J, let $|J|$ be the number of its elements. Always, let $0^0 = 1$, $\frac{1}{0} = \infty$ and, for $k \in \mathbb{Z}$, $\sum_{m=k}^{k-1} = 0$ be the empty sum. For $k \in \mathbb{N}$, let $\underline{k}_{\neq}^k = \{(\ell_1, \ldots, \ell_k) \in \underline{k}^k \,|\, \ell_i \neq \ell_j \text{ for all } i, j \in \underline{k} \text{ with } i \neq j\}$ be the set of all permutations on the set $\underline{k}$. We use the standard multi-index notation: For $d \in \mathbb{N}$, $z = (z_1, \ldots, z_d) \in \mathbb{C}^d$ and $m = (m_1, \ldots, m_d) \in \mathbb{Z}_+^d$, set $z^m = \prod_{r=1}^d z_r^{m_r}$, $|m| = \sum_{r=1}^d m_r$ and $m! = \prod_{r=1}^d m_r!$.

1.3. *On the method used.* In our proofs, we make use of the fact that $\mathcal{M}$ is a real commutative Banach algebra with unity. In the (compound) Poisson approximation such an approach has already been used by various authors. For example, Le Cam (1960), Chen (1975a), Chen and Roos (1995), Borisov (2003) considered measures on a general measurable Abelian group. Other authors also used Banach algebra properties under other assumptions, see, for instance, Kerstan (1964) using a complex variable approach, Deheuvels and Pfeifer (1986) using an operator semigroup framework, and Witte (1990) for a unification of these methods.

Our proofs are based on ideas of Kerstan (1964) in combination with arguments given in Roos (1999b) as well as new auxiliary norm estimates. For further papers using Kerstan's method, see H. Herrmann (1965a), U. Herrmann (1965b), Kruopis and Čekanavičius (2014), Upadhye and Vellaisamy (2014), and also some of the references cited therein.

The main idea of Kerstan (1964) was to expand the difference of two univariate distributions in a certain way and to estimate the norm terms involved using the Cauchy integral formula. In Roos (1999b), the corresponding multidimensional generalization was studied, which made it necessary to slightly modify the expansion. The norm terms have been estimated using the Cauchy-Schwarz inequality without using integrals. In the present paper, we use a different expansion (see formulas (4.6) and (4.7) below) and use new norm term estimates using Charlier polynomials and the Cauchy-Schwarz inequality (see Subsection 4.1).

We note that in this paper it suffices to consider measures on a measurable Abelian semigroup with zero element rather than a measurable Abelian group. This makes it possible to use our results in the Poisson point process approximation, see Section 3.

1.4. *Review of some known results.* Let us consider some important results for discrete distributions on $(\mathfrak{X}, +, \mathcal{A}) = (\mathbb{R}^d, +, \mathcal{B}^d)$ for $d \in \mathbb{N}$, where $\mathcal{B}^d$ is the Borel σ-Algebra over $\mathbb{R}^d$. During this subsection, let $n \in \mathbb{N}$ and, for $j \in \underline{n}$ and $r \in \underline{d}$,

$$p_j, q_{j,r} \in [0,1] \text{ with } \sum_{r=1}^d q_{j,r} = 1 \text{ and } \lambda_r = \sum_{j=1}^n p_j q_{j,r} > 0, \quad \lambda = \sum_{j=1}^n p_j = \sum_{r=1}^d \lambda_r,$$

$$U_r = \delta_{e_r}, \quad Q_j = \sum_{r=1}^d q_{j,r} U_r, \quad F_j = \delta_0 + p_j(Q_j - \delta_0),$$

$$Q = \frac{1}{\lambda} \sum_{j=1}^n p_j Q_j = \frac{1}{\lambda} \sum_{r=1}^d \lambda_r U_r, \quad F = \prod_{j=1}^n F_j, \quad G = \mathrm{CPo}(\lambda, Q) = \exp(\lambda(Q - \delta_0)).$$

Here, $e_r \in \mathbb{R}^d$ is the unit vector with 1 at position r and 0 otherwise. In what follows, we discuss some bounds in the approximation of the distribution F by G.

1.4.1. *The one-dimensional case $d = 1$.* Here, we have $Q_1 = \cdots = Q_n = Q = \delta_1$, such that $F = \prod_{j=1}^n (\delta_0 + p_j(\delta_1 - \delta_0))$ is a so-called Bernoulli convolution and $G = \mathrm{Po}(\lambda)$ is the Poisson distribution with mean λ. In this situation, one of the most remarkable results is the following:

$$\frac{1}{7} \min\Big\{\frac{1}{\lambda}, 1\Big\} \sum_{j=1}^n p_j^2 \leqslant \|F - G\| \leqslant 2 \frac{1 - \mathrm{e}^{-\lambda}}{\lambda} \sum_{j=1}^n p_j^2 \leqslant 2 \min\Big\{\frac{1}{\lambda}, 1\Big\} \sum_{j=1}^n p_j^2. \quad (1.1)$$

The upper bounds of $\|F - G\|$ are due to Barbour and Hall (1984, Theorem 1), who used Stein's method to improve results of Le Cam (1960, Theorem 2), Kerstan (1964, formula (1) on page 174) and Chen (1975b, formula (4.23)). In their Theorem 2, they also showed a comparable lower bound with constant $\frac{1}{16}$ instead of $\frac{1}{7}$. The lower bound with the better constant was mentioned in Remark 3.2.2 of Barbour et al. (1992). The estimates in (1.1) depend on the behavior of the so-called magic factor $\frac{1}{\lambda}$ (cf. Introduction in Barbour et al., 1992) and on the smallness of all p_j, $j \in \underline{n}$, which is reflected by $\sum_{j=1}^n p_j^2$. It is easily seen that the leading constant 2 in front of $\theta := \frac{1}{\lambda} \sum_{j=1}^n p_j^2$, resp. in front of $\lambda\theta$, in the upper bound (1.1) is optimal. However, formula (32) in Roos (1999a) implies that

$$\Big| \|F - G\| - \sqrt{\frac{2}{\pi \mathrm{e}}} \, \theta \Big| \leqslant C \theta \, \min\Big\{1, \frac{1}{\sqrt{\lambda}} + \theta\Big\}, \quad (1.2)$$

where C denotes an absolute constant. In particular, this implies that $\|F - G\| \sim \sqrt{\frac{2}{\pi \mathrm{e}}} \, \theta$ as $\theta \to 0$ and $\lambda \to \infty$. Here, $\sim$ means that the quotient of both sides tends to one. We note that the bound (1.2) is a generalization, resp. refinement, of results of Prohorov (1953, Theorem 2) and Deheuvels and Pfeifer (1986, Theorem 1.2); see also Barbour et al. (1992, page 2). In Čekanavičius and Roos (2006, formula (30)), it was shown that, in the case $\theta < 1$,

$$\|F - G\| \leqslant \frac{3\theta}{2\mathrm{e}(1 - \sqrt{\theta})^{3/2}}, \quad (1.3)$$

which is an improvement of formula (10) in Roos (2001). The more general Theorem 1, resp. Corollary 1, of the latter paper implies the sharpness of the constant $\frac{3}{2\mathrm{e}}$. In fact, we have

$$\lim_{t \downarrow 0} \Big(\sup \frac{1}{\theta} \|F - G\| \Big) = \frac{3}{2\mathrm{e}}, \quad (1.4)$$

where the sup is taken over all $n \in \mathbb{N}$, $p_1, \ldots, p_n \in [0, 1]$ such that $\lambda = \sum_{j=1}^n p_j > 0$ and $\theta = \frac{1}{\lambda} \sum_{j=1}^n p_j^2 \leqslant t$ (or, alternatively, such that $\max_{j \in \underline{n}} p_j \leqslant t$).

1.4.2. *The multi-dimensional case $d \in \mathbb{N}$.* Here

$$F = \prod_{j=1}^n \Big(\delta_0 + p_j \sum_{r=1}^d q_{j,r}(\delta_{e_r} - \delta_0) \Big)$$

is a generalized multinomial distribution, which we wish to approximate by a product of Poisson distributions

$$G = \mathrm{CPo}(\lambda, Q) = \exp\Big(\sum_{r=1}^d \lambda_r(\delta_{e_r} - \delta_0) \Big) = \bigotimes_{r=1}^d \exp(\lambda_r(\delta_1 - \delta_0)) = \bigotimes_{r=1}^d \mathrm{Po}(\lambda_r),$$

i.e. G is a multivariate Poisson distribution with mean vector $(\lambda_1, \ldots, \lambda_d)$. In this context, there are two papers by Franken (1963) and U. Herrmann (1965b), which unfortunately have been largely overlooked in subsequent publications. Both papers considered more general convolution factors. Under our assumptions, some of the results are as follows. Franken (1963, formula (1) on page 102) used direct calculations to show a multivariate version of Proposition 1 of Le Cam (1960). His inequality reads as

$$\|F - G\| \leqslant 2 \sum_{j=1}^{n} p_j^2 \tag{1.5}$$

and was later rediscovered by McDonald (1980, Theorem 1) using coupling arguments. We note that Franken (1963, formulas (2) and (3) on page 102) also proved two bounds for the point metric; one of these however can, under the present assumptions, be replaced by a bound of a better order, cf. Roos (1998, Theorem 2). U. Herrmann (1965b, formula (0) on page 18) proved a bound containing a magic factor by using the method of Kerstan (1964): If $\max_{j \in \underline{n}} p_j \leqslant \frac{1}{4}$, then

$$\|F - G\| \leqslant 9 \sum_{j=1}^{n} p_j^2 \left(\sum_{r=1}^{d} \frac{q_{j,r}}{\sqrt{\lambda_r}} \right)^2. \tag{1.6}$$

Consequently, in view of (1.1), we see that, in order to obtain a new bound, which is of the right order in the case $d = 1$, one could simply take the minimum of the right-hand sides of (1.5) and (1.6). But, as is shown below, it is possible to get bounds having a better structure concerning the minimum term. Indeed, the following interesting bound containing a magic factor was shown by Barbour (1988, Theorem 1) using Stein's method:

$$\|F - G\| \leqslant 2 \sum_{j=1}^{n} p_j^2 \min \left\{ c_\lambda \sum_{r=1}^{d} \frac{q_{j,r}^2}{\lambda_r}, 1 \right\}, \tag{1.7}$$

where $c_\lambda = \frac{1}{2} + \max\{\log(2\lambda), 0\}$. Unfortunately, the term c_λ is logarithmically increasing in λ and therefore the upper bound in (1.7) does not have the correct order in the case $d = 1$, see (1.1).

An improvement of (1.7) without logarithmic factor was shown in Roos (1999b, Theorem 1) using some modifications in the method of Kerstan (1964). Let

$$g(z) = \frac{2e^z}{z^2}(e^{-z} - 1 + z) = 2 \sum_{m=2}^{\infty} \frac{m-1}{m!} z^{m-2}, \quad (z \in \mathbb{C}), \tag{1.8}$$

$$\alpha_0 = \sum_{j=1}^{n} g(2p_j) p_j^2 \min \left\{ \frac{1}{2^{3/2}} \sum_{r=1}^{d} \frac{q_{j,r}^2}{\lambda_r}, 1 \right\}, \quad \beta_0 = \sum_{j=1}^{n} p_j^2 \min \left\{ \sum_{r=1}^{d} \frac{q_{j,r}^2}{\lambda_r}, 1 \right\}. \tag{1.9}$$

We note that

$$1 \leqslant g(x) \leqslant e^x \quad (x \in [0, \infty)), \qquad \max_{j \in \underline{n}} g(2p_j) \leqslant g(2) \leqslant 4.195. \tag{1.10}$$

If $\alpha_0 < \frac{1}{2e}$, then

$$\|F - G\| \leqslant \frac{2\alpha_0}{1 - 2\alpha_0 e}. \tag{1.11}$$

The following estimate is valid without any restrictions:

$$\|F - G\| \leqslant 17.6\,\beta_0. \tag{1.12}$$

It is clear that (1.11) or (1.12) should be preferred over (1.6), because of the term $\sum_{r=1}^{d} \frac{q_{j,r}^2}{\lambda_r}$ in the representations of α_0 and β_0. Indeed, if $q_{j,r} = \frac{1}{d}$ for all $j \in \underline{n}$ and $r \in \underline{d}$, then $\lambda_1 = \cdots = \lambda_d$ and hence $(\sum_{r=1}^{d} \frac{q_{j,r}}{\sqrt{\lambda_r}})^2 = \frac{1}{\lambda_1} = d\sum_{r=1}^{d} \frac{q_{j,r}^2}{\lambda_r}$, so that the difference in the order is the factor d, if we consider the first entry in the minimum terms in (1.9). On the other hand, for a precise comparison of (1.11) with (1.6), let us assume that $\max_{j\in\underline{n}} p_j \leqslant \frac{1}{4}$ and that $\gamma := \sum_{j=1}^{n} p_j^2(\sum_{r=1}^{d} \frac{q_{j,r}}{\sqrt{\lambda_r}})^2 < \frac{2}{9}$, such that the right-hand side of (1.6) is smaller than the trivial bound 2. If we now use the crude estimate $2\alpha_0 \leqslant g(\frac{1}{2})\frac{\gamma}{\sqrt{2}}$, then, since $\gamma \leqslant \frac{2}{9}$, (1.11) implies the bound $\frac{5}{2}\gamma$, which is better than the one in (1.6).

In the present paper, among other results, we show the following further improvement of (1.11) and (1.12).

Theorem 1.1. *Let the function g be defined as in* (1.8). *Write*

$$\alpha_1 = \sum_{j=1}^{n} g(2p_j)p_j^2 \sum_{r=1}^{d} q_{j,r}\min\left\{\frac{q_{j,r}}{2^{3/2}\lambda_r}, 2\right\}, \quad \beta_1 = \sum_{j=1}^{n} p_j^2 \sum_{r=1}^{d} q_{j,r}\min\left\{\frac{q_{j,r}}{\lambda_r}, 1\right\}. \tag{1.13}$$

If $\alpha_1 < \frac{1}{2^{3/2}}$, then

$$\|F - G\| \leqslant \frac{2\alpha_1}{1 - 2^{3/2}\alpha_1}. \tag{1.14}$$

Without any restrictions, we have

$$\|F - G\| \leqslant 15.6\,\beta_1. \tag{1.15}$$

Remark 1.2.

(a) Let us explain the bounds in Theorem 1.1 with the help of random variables. We assume the notation as given above. Furthermore, for $j \in \underline{n}$, let $X_j = (X_{j,1}, \ldots, X_{j,d})$ be d-dimensional independent Bernoulli random vectors with $P(X_j = (0, \ldots, 0)) = 1 - p_j$ and $P(X_j = e_r) = p_j q_{j,r}$ for $r \in \underline{d}$. Let $T = (T_1, \ldots, T_d)$, where T_r, $(r \in \underline{d})$ are independent one-dimensional Poisson $\mathrm{Po}(\lambda_r)$ distributed random variables. Let P^{S_n} and P^T denote the distribution of $S_n = (S_{n,1}, \ldots, S_{n,d}) = \sum_{j=1}^{n} X_j$ and T, respectively. Then $F = P^{S_n}$, $G = P^T$ and

$$d_{\mathrm{TV}}(P^{S_n}, P^T) \leqslant \frac{\alpha_1}{1 - 2^{3/2}\alpha_1}, \text{ if } \alpha_1 < \frac{1}{2^{3/2}}; \quad d_{\mathrm{TV}}(P^{S_n}, P^T) \leqslant 7.8\,\beta_1.$$

(b) The structure of the term β_1 is better than that of β_0, since we always have $\beta_1 \leqslant \beta_0$ and there are examples in which β_1 is significantly smaller than β_0 (see Example 2.5). In particular, (1.15) is always better than (1.12). It should be mentioned that, if $q_{1,r} = \cdots = q_{n,r}$ for all $r \in \underline{d}$, then $\lambda_r = q_{1,r}\lambda$ for all $r \in \underline{d}$ and $\beta_1 = \min\{\frac{1}{\lambda}, 1\}\sum_{j=1}^{n} p_j^2 = \beta_0$. Similarly, the structure of α_1 is better than that of α_0. However, α_1 is not always smaller than or equal to α_0, since α_1 contains an additional factor 2 in the second entry of the minimum term.

(c) In practical applications, (1.14) often leads to smaller values than (1.15).

(d) Generally, an inequality $\|F - G\| \leqslant C\, d^c \beta_1'$ with

$$\beta_1' = \sum_{j=1}^{n} p_j^2 \sum_{r=1}^{d} q_{j,r}^2 \min\left\{\frac{1}{\lambda_r}, 1\right\}$$

and absolute constants $C \in (0, \infty)$ and $c \in [0, 1)$ cannot hold, see the remark after Corollary 1 in Roos (1998). Consequently, there is no hope of a bound of order β_1'.

(e) In view of (1.8), we see that, if $j \in \underline{n}$ and p_j is small, then $g(2p_j) \approx 1$. Hence, if α_1 and $\max_{j \in \underline{n}} p_j$ are small, then

$$\|F - G\| \leqslant c \sum_{j=1}^{n} p_j^2 \sum_{r=1}^{d} q_{j,r} \min\left\{\frac{q_{j,r}}{2^{3/2}\lambda_r}, 2\right\}$$

with $c \approx 2$. In (1.15), the factor 15.6 cannot be replaced by a constant smaller than 2, which follows from the remark after (1.1). Relation (1.4) implies that (1.14) cannot generally hold when the factor $\frac{1}{2^{3/2}}$ in the representation of α_1 (see (1.13)) is replaced by a constant smaller than $\frac{3}{4e}$.

(f) All upper bounds in (1.1), (1.3), (1.5), (1.6), (1.7), (1.11), (1.12), (1.14) and (1.15) remain valid, if, in the definition of F and G, we generalize U_r to $U_r \in \mathcal{F}$ for $r \in \underline{d}$, which follows from the definition of the total variation norm, see, e.g., Le Cam (1965, page 187) or Michel (1987, page 167).

The next proposition provides lower bounds in the multi-dimensional case.

Proposition 1.3. *Let $J \subseteq \underline{d}$, $y_j = \sum_{r \in J} q_{j,r}$, $\widetilde{p}_j = p_j y_j$ for all $j \in \underline{n}$ and $\widetilde{\lambda} = \sum_{j=1}^{n} \widetilde{p}_j$. Then*

$$\|F - G\| \geqslant \left\| \prod_{j=1}^{n}(\delta_0 + \widetilde{p}_j(\delta_1 - \delta_0)) - \mathrm{Po}(\widetilde{\lambda}) \right\| \geqslant \frac{1}{7} \min\left\{\frac{1}{\widetilde{\lambda}}, 1\right\} \sum_{j=1}^{n} \widetilde{p}_j^2. \tag{1.16}$$

In particular,

$$\|F - G\| \geqslant \frac{1}{7} \min\left\{\frac{1}{\lambda}, 1\right\} \sum_{j=1}^{n} p_j^2 \tag{1.17}$$

and

$$\|F - G\| \geqslant \frac{1}{7} \max_{r \in \underline{d}} \left(\min\left\{\frac{1}{\lambda_r}, 1\right\} \sum_{j=1}^{n} p_j^2 q_{j,r}^2\right). \tag{1.18}$$

The second inequality in (1.16) is taken from (1.1). In the case $J = \underline{d}$, the first lower bound in (1.16) is the same as the one in Deheuvels and Pfeifer (1988, Remark 2.5), who used a maximal coupling for a proof. The generalization with arbitrary $J \subseteq \underline{d}$ is shown analogously. However, in order to keep the paper self-contained, we give a further simple proof, which avoids the coupling method, see Section 4.2. The bounds in (1.17) and (1.18) follow from (1.16) with $J = \underline{d}$ and $J = \{r\}$ for all $r \in \underline{d}$, respectively. Consequently, the lower bound in (1.1) still holds in the multi-dimensional case. The bound in (1.18) is a slight improvement of the first inequality in Corollary 1 of Roos (1998).

Let us compare the bounds in (1.15), (1.17) and (1.18).

Remark 1.4. (a) Suppose that, for all $r \in \underline{d}$, $a_r, b_r \in (0,1]$ exist, such that $a_r \leqslant q_{j,r} \leqslant b_r$ for all $j \in \underline{n}$. Then $\min\{\frac{1}{\lambda}, 1\} \sum_{j=1}^{n} p_j^2 \geqslant \frac{\beta_1}{\eta}$ with $\eta = \max_{r \in \underline{d}} \frac{b_r}{a_r}$. Here, the bounds in (1.15) and (1.17) differ at most by a constant multiple of $\frac{1}{\eta}$. If $q_{1,r} = \cdots = q_{n,r} = a_r = b_r$ for all $r \in \underline{d}$, then $\eta = 1$. We note that, in this case, (1.18) is worse than (1.17).

(b) Assume now that $c \in (0,1)$, $\kappa \in [0,\infty)$, $d = n$, $p_j = \frac{c}{j^\kappa}$, $q_{j,r} = \mathbb{1}_{\{j\}}(r)$ for all $j, r \in \underline{n}$.

Let us first assume that $\kappa = 1$. Then (1.18) implies that $\|F - G\| \geqslant \frac{1}{7} \max_{j \in \underline{n}} p_j^2 = \frac{c^2}{7}$, whereas (1.15) gives $\|F - G\| \leqslant 15.6 \sum_{j=1}^{n} p_j^2 \leqslant 15.6 \frac{\pi^2}{6} c^2$. Hence, in this case, (1.15) and (1.18) have the same order as $c \to 0$. The bound (1.17) gives $\|F - G\| \geqslant \frac{c^2}{7} \min\{\frac{1}{c \sum_{j=1}^{n} 1/j}, 1\} \sum_{j=1}^{n} \frac{1}{j^2}$, which is worse than (1.18) as $n \to \infty$ if c is fixed. This together with (a) shows that the bounds in (1.17) and (1.18) are not comparable in general.

Let us now consider the case $\kappa = 0$. Then (1.17) and (1.18) imply that $\|F - G\| \geqslant \frac{1}{7} \min\{\frac{1}{nc}, 1\} nc^2$ and $\|F - G\| \geqslant \frac{c^2}{7}$, respectively, whereas (1.15) gives $\|F - G\| \leqslant 15.6 nc^2$, having a different order as $n \to \infty$ if c is fixed.

2. Main results

Theorem 2.1. *Let $d, n \in \mathbb{N}$ and $\ell \in \underline{n}_0$. For $j \in \underline{n}$ and $r \in \underline{d}$, let*

$$p_j, q_{j,r} \in [0,1] \text{ with } \sum_{r=1}^{d} q_{j,r} = 1 \text{ and } \lambda_r = \sum_{j=1}^{n} p_j q_{j,r} > 0, \quad \lambda = \sum_{j=1}^{n} p_j, \quad U_r \in \mathcal{F},$$

$$Q_j = \sum_{r=1}^{d} q_{j,r} U_r, \quad Q = \frac{1}{\lambda} \sum_{j=1}^{n} p_j Q_j, \quad R_j = p_j(Q_j - \delta_0), \quad F_j = \delta_0 + R_j,$$

$$V_j = F_j e^{-R_j} - \delta_0, \quad F = \prod_{j=1}^{n} F_j.$$

For $k \in \underline{n}_0$, let

$$M_k = \sum_{J \subseteq \underline{n}: |J| = k} \prod_{j \in J} V_j, \qquad H_k = M_k \exp(\lambda(Q - \delta_0))$$

and set $G_\ell = \sum_{k=0}^{\ell} H_k$. Let the function g be defined as in (1.8). Write

$$\alpha_1 = \sum_{j=1}^{n} g(2p_j) p_j^2 \sum_{r=1}^{d} q_{j,r} \min\left\{\frac{q_{j,r}}{2^{3/2} \lambda_r}, 2\right\}.$$

If $\alpha_1 < \frac{1}{2^{3/2}}$, then

$$\|F - G_\ell\| \leqslant \frac{\sqrt{(2(\ell+1))!}}{(\ell+1)!} 2^{(\ell+1)/2} \frac{\alpha_1^{\ell+1}}{1 - 2^{3/2} \alpha_1}. \tag{2.1}$$

Remark 2.2. Consider the assumptions of Theorem 2.1. In order to give an alternative formula for G_ℓ for the first few $\ell \in \underline{n}_0$, let $\Gamma_k = \sum_{j=1}^{n} V_j^k$ for $k \in \mathbb{N}$. Then $M_0 = \delta_0$ and, for $k \in \underline{n}$, Newton's identity (see Bourbaki (1990, A.IV.70,

Lemma 4)) gives

$$M_k = \frac{1}{k}\sum_{j=1}^{k}(-1)^{j-1}M_{k-j}\Gamma_j.$$

In particular, if $n \geqslant 3$,

$$M_1 = \Gamma_1, \quad M_2 = \frac{1}{2}(\Gamma_1^2 - \Gamma_2), \quad M_3 = \frac{1}{6}(\Gamma_1^3 - 3\Gamma_1\Gamma_2 + 2\Gamma_3)$$

and consequently

$$G_0 = \exp(\lambda(Q - \delta_0)), \quad G_1 = (\delta_0 + \Gamma_1)\exp(\lambda(Q - \delta_0)),$$

$$G_2 = \left(\delta_0 + \Gamma_1 + \frac{1}{2}(\Gamma_1^2 - \Gamma_2)\right)\exp(\lambda(Q - \delta_0)),$$

$$G_3 = \left(\delta_0 + \Gamma_1 + \frac{1}{2}(\Gamma_1^2 - \Gamma_2) + \frac{1}{6}(\Gamma_1^3 - 3\Gamma_1\Gamma_2 + 2\Gamma_3)\right)\exp(\lambda(Q - \delta_0)).$$

We note that, in Roos (1999b, formulas (10), (28)), the signed measure

$$\left(\delta_0 - \frac{1}{2}\sum_{j=1}^{n}R_j^2\right)\exp(\lambda(Q - \delta_0)) \tag{2.2}$$

as approximation of F was used. The corresponding total variation bound has a somewhat complicated form and is of worse order than β_0^2, the definition of which can be found in (1.9). In comparison, our signed measure

$$G_1 = \left(\delta_0 + \sum_{j=1}^{n}(F_j e^{-R_j} - \delta_0)\right)\exp(\lambda(Q - \delta_0))$$

is slightly more complicated than (2.2), but gives a total variation bound of order α_1^2.

In the following result, we present approximation bounds without a singularity as in (2.1).

Theorem 2.3. *Let the notation of Theorem 2.1 be valid. Let $D_1' = 3.11$ and $D_k' = D_k(\frac{g(2)}{2})^k$ for $k \in \mathbb{N} \setminus \{1\}$, where D_k is defined as in Corollary 4.5 below (see Table 2). Let $h_1(x) = h_{1,\ell}(x) = \sum_{k=\ell+1}^{\infty} D_k' x^k$, $h_2(x) = h_{2,\ell}(x) = 2 + \sum_{k=1}^{\ell} D_k' x^k$ for $x \in [0, \infty)$. Write*

$$\beta_1 = \sum_{j=1}^{n} p_j^2 \sum_{r=1}^{d} q_{j,r}\min\left\{\frac{q_{j,r}}{\lambda_r}, 1\right\}.$$

Without any restrictions, we have

$$\|F - G_\ell\| \leqslant c_\ell \beta_1^{\ell+1}, \tag{2.3}$$

where $c_\ell = \frac{h_2(x_\ell)}{x_\ell^{\ell+1}}$ and $x_\ell \in (0, \infty)$ is the unique positive solution of the equation $h_1(x_\ell) = h_2(x_\ell)$. In particular, we have $c_0 \leqslant 15.6$, $c_1 \leqslant 113.0$, $c_2 \leqslant 633.8$, $c_3 \leqslant 3204.8$, $c_4 \leqslant 15945.6$.

Remark 2.4. Theorem 1.1 is a direct consequence of Theorems 2.1 and 2.3 for $\ell = 0$.

Example 2.5. In order to compare the bounds in a numerical example, let us consider the assumptions of Theorem 2.1, 2.3 with $d = n = 1000$ and $p_{j,r} = \frac{10^{-4}}{|j-r|^{1/2}+0.1}$, $p_j = \sum_{r=1}^{d} p_{j,r}$ and $q_{j,r} = \frac{p_{j,r}}{p_j}$ for $j, r \in \underline{n}$. This implies that $\beta_0 = 0.081578\ldots$,

$\beta_1 = 0.022183\ldots$, $\alpha_0 = 0.044626\ldots$, $\alpha_1 = 0.023037\ldots$, $\lambda = 9.01\ldots$, $\max_{j\in\underline{n}} p_j = 0.009521\ldots$. Here α_0 and β_0 are as in (1.9). Table 1 below shows that, in the approximation of F by $G_0 = \exp(\lambda(Q - \delta_0))$, the bound in (1.14) is the smallest one. In the approximation of F by the signed measure G_ℓ for $\ell \in \underline{4}$, we expect that the accuracy increases as ℓ increases. Indeed, this is reflected in the bounds as well. Furthermore, we see that here (2.1) is better than (2.3).

Table 1: Numerical comparison of the bounds in Example 2.5

Approximation by $\exp(\lambda(Q - \delta_0))$		Approximation by signed meas. G_ℓ, ($\ell \in \underline{4}$)		
number of formula	upper bound	number of formula	ℓ	upper bound
(1.5)	0.163157	(2.1)	1	0.002782
(1.6)	81.3	(2.3)	1	0.055608
(1.7)	0.163157	(2.1)	2	0.000166
(1.11)	0.117843	(2.3)	2	0.006919
(1.12)	1.435779	(2.1)	3	0.000011
(1.14)	0.049286	(2.3)	3	0.000777
(1.15)	0.346060	(2.1)	4	6.24×10^{-7}
		(2.3)	4	0.000086

We note that the value of the bound in (1.6) exceeds by far the trivial bound 2, which however depends on the kind of example. The lower bounds in (1.17) and (1.18) give 0.001292 and 1.60×10^{-7}, respectively.

Remark 2.6. Let the notation of Theorems 2.1, 2.3 be valid and assume that there exist pairwise disjoint sets $A_1, \ldots, A_d \in \mathcal{A}$ with $U_r(\mathfrak{X} \setminus A_r) = 0$ for all $r \in \underline{d}$. Let $\mathbb{1}_{A_r} : \mathfrak{X} \longrightarrow \mathfrak{X}$ be the indicator function of A_r. Then, for $j \in \underline{n}$, $f_j := \sum_{r=1}^{d} \frac{\lambda}{\lambda_r} q_{j,r} \mathbb{1}_{A_r}$ is a Q-density of Q_j, since, for $B \in \mathcal{A}$, we have

$$\int_B f_j \, dQ = \sum_{r=1}^{d} \int_{B \cap A_r} \frac{\lambda}{\lambda_r} q_{j,r} \, d\Big(\frac{\lambda_r}{\lambda} U_r\Big) = \sum_{r=1}^{d} q_{j,r} U_r(B) = Q_j(B).$$

Furthermore, for $c \in (0, \infty)$,

$$\sum_{r=1}^{d} q_{j,r} \min\Big\{c\frac{q_{j,r}}{\lambda_r}, 1\Big\} = \sum_{r=1}^{d} \int_{A_r} \frac{\lambda}{\lambda_r} q_{j,r} \min\Big\{c\frac{q_{j,r}}{\lambda_r}, 1\Big\} \, dQ = \int f_j \min\Big\{c\frac{f_j}{\lambda}, 1\Big\} \, dQ,$$

such that

$$\alpha_1 = \sum_{j=1}^{n} g(2p_j) p_j^2 \int f_j \min\Big\{\frac{f_j}{2^{3/2}\lambda}, 2\Big\} \, dQ, \quad \beta_1 = \sum_{j=1}^{n} p_j^2 \int f_j \min\Big\{\frac{f_j}{\lambda}, 1\Big\} \, dQ.$$

The next result shows that Theorems 2.1 and 2.3 can be generalized using the ideas of Remark 2.6. In fact, the Q_j, ($j \in \underline{n}$) are now general probability measures and the U_r for $r \in \underline{d}$ are no longer needed.

Corollary 2.7. *Let $n \in \mathbb{N}$, $\ell \in \underline{n}_0$. For $j \in \underline{n}$, let $p_j \in (0,1]$, $Q_j \in \mathcal{F}$, $R_j = p_j(Q_j - \delta_0)$, $F_j = \delta_0 + R_j$ and $V_j = F_j \mathrm{e}^{-R_j} - \delta_0$. Set $\lambda = \sum_{j=1}^{n} p_j$, $F = \prod_{j=1}^{n} F_j$, $Q = \frac{1}{\lambda} \sum_{j=1}^{n} p_j Q_j$. For $j \in \underline{n}$, let f_j be a Radon-Nikodým density of Q_j with respect to Q, which exists since $Q_j \ll Q$. For $k \in \underline{n}_0$, let*

$$M_k = \sum_{J \subseteq \underline{n}:\, |J|=k} \prod_{j \in J} V_j, \qquad H_k = M_k \exp(\lambda(Q - \delta_0))$$

and set $G_\ell = \sum_{k=0}^{\ell} H_k$. Let the function g be defined as in (1.8). *Set*

$$\widetilde{\alpha}_1 = \sum_{j=1}^{n} g(2p_j)p_j^2 \int f_j \min\Big\{\frac{f_j}{2^{3/2}\lambda}, 2\Big\} \, \mathrm{d}Q, \quad \widetilde{\beta}_1 = \sum_{j=1}^{n} p_j^2 \int f_j \min\Big\{\frac{f_j}{\lambda}, 1\Big\} \, \mathrm{d}Q.$$

If $\widetilde{\alpha}_1 < \frac{1}{2^{3/2}}$, then

$$\|F - G_\ell\| \leqslant \frac{\sqrt{(2(\ell+1))!}}{(\ell+1)!} 2^{(\ell+1)/2} \frac{\widetilde{\alpha}_1^{\ell+1}}{1 - 2^{3/2}\widetilde{\alpha}_1}. \tag{2.4}$$

The following bound is generally valid:

$$\|F - G_\ell\| \leqslant c_\ell \, \widetilde{\beta}_1^{\ell+1}, \tag{2.5}$$

where the constant c_ℓ is the same as in Theorem 2.3.

Remark 2.8.

(a) Let us explain the bounds of Corollary 2.7 in the case $\ell = 0$ with the help of random variables. We assume the notation as in that corollary. Let $\{0\} \in \mathcal{A}$ and $S_n = \sum_{j=1}^{n} X_j$ be the sum of independent $\mathfrak{X}$-valued random variables $X_1, \ldots, X_n$ with $P(X_j \neq 0) = p_j > 0$ and $Q_j = P(X_j \in \cdot \mid X_j \neq 0)$. Let $T = \sum_{m=1}^{N} Y_m$, where $N, Y_m, (m \in \mathbb{N})$ are independent random variables, N is $\mathbb{Z}_+$-valued and has Poisson distribution $\mathrm{Po}(\lambda)$, whereas the $\mathfrak{X}$-valued Y_m are identically distributed with distribution Q. Then we have

$$d_{\mathrm{TV}}(P^{S_n}, P^T) \leqslant \frac{\widetilde{\alpha}_1}{1 - 2^{3/2}\widetilde{\alpha}_1}, \text{ if } \widetilde{\alpha}_1 < \frac{1}{2^{3/2}}; \quad d_{\mathrm{TV}}(P^{S_n}, P^T) \leqslant 7.8\,\widetilde{\beta}_1.$$

(b) For $\ell = 0$, (2.4) and (2.5) are refinements of (10) and (11) in Roos (2007).

(c) Let the assumptions of Corollary 2.7 hold. Further suppose that $Q_j \ll \mu$ for all $j \in \underline{n}$, where μ is a σ-finite measure on $(\mathfrak{X}, \mathcal{A})$. Let $\widetilde{f}_j$ be a Radon-Nikodým density of Q_j with respect to μ. Then $\widetilde{f} := \frac{1}{\lambda} \sum_{j=1}^{n} p_j \widetilde{f}_j$ is a μ-density of Q. For $j \in \underline{n}$, we get a Q-density of Q_j by defining $f_j(x) = \frac{\widetilde{f}_j(x)}{\widetilde{f}(x)}$ for $x \in \{\widetilde{f} > 0\}$ and $f_j(x) = 0$ otherwise. This gives the possibility to evaluate $\widetilde{\alpha}_1$ and $\widetilde{\beta}_1$ by using $\widetilde{f}_j$ for $j \in \underline{n}$, $\widetilde{f}$ and μ. In fact,

$$\widetilde{\alpha}_1 = \sum_{j=1}^{n} g(2p_j)p_j^2 \int_{\{\widetilde{f}>0\}} \widetilde{f}_j \min\Big\{\frac{\widetilde{f}_j}{2^{3/2}\lambda\widetilde{f}}, 2\Big\} \, \mathrm{d}\mu,$$

$$\widetilde{\beta}_1 = \sum_{j=1}^{n} p_j^2 \int_{\{\widetilde{f}>0\}} \widetilde{f}_j \min\Big\{\frac{\widetilde{f}_j}{\lambda\widetilde{f}}, 1\Big\} \, \mathrm{d}\mu.$$

If, for example, $(\mathfrak{X}, +, \mathcal{A}) = (\mathbb{R}^1, +, \mathcal{B}^1)$, $\mu = \lambda^1$ is the Lebesgue measure on $(\mathbb{R}^1, \mathcal{B}^1)$ and Q_j is the exponential distribution with λ^1-density $\widetilde{f}_j(x) = t_j \mathrm{e}^{-t_j x} \mathbb{1}_{(0,\infty)}(x)$ for $x \in \mathbb{R}$, $j \in \underline{n}$, $t_j \in (0,\infty)$, then we obtain

$$\widetilde{\alpha}_1 = \sum_{j=1}^{n} g(2p_j)p_j^2 \int_{(0,\infty)} t_j \mathrm{e}^{-t_j x} \min\Big\{\frac{t_j \mathrm{e}^{-t_j x}}{2^{3/2} \sum_{i=1}^{n} p_i t_i \mathrm{e}^{-t_i x}}, 2\Big\} \, \mathrm{d}\lambda^1(x)$$

and a similar formula for $\widetilde{\beta}_1$.

3. Application in the Poisson point process approximation

Let $(S, \mathcal{S})$ be a measurable space and $\mathfrak{X} = \mathfrak{X}(S, \mathcal{S})$ be the set of all point measures of the form $\mu = \sum_{i \in I} \delta_{x_i}$, where $I \subseteq \mathbb{N}$ and $x_i \in S$ for all $i \in I$. Further, let $\mathcal{A} = \sigma((\pi_B \,|\, B \in \mathcal{S}))$ be the smallest σ-algebra over $\mathfrak{X}$ such that all the evaluation maps $\pi_B : \mathfrak{X} \longrightarrow \mathbb{Z}_+ \cup \{\infty\} =: \overline{\mathbb{Z}}_+, \mu \mapsto \mu(B)$ for $B \in \mathcal{S}$ are measurable with respect to the power set $2^{\overline{\mathbb{Z}}_+}$ of $\overline{\mathbb{Z}}_+$, see e.g. Reiss (1993). The mapping $\mathfrak{s} : \mathfrak{X} \times \mathfrak{X} \longrightarrow \mathfrak{X}$, $(\mu, \nu) \mapsto \mu + \nu$ is measurable with respect to $\mathcal{A} \otimes \mathcal{A}$ and $\mathcal{A}$. Indeed, for $B \in \mathcal{S}$ and $k \in \mathbb{Z}_+$, we have $\mathfrak{s}^{-1}(\pi_B^{-1}(\{k\})) = \bigcup_{j \in \underline{k}_0} (\pi_B^{-1}(\{j\}) \times \pi_B^{-1}(\{k - j\})) \in \mathcal{A} \otimes \mathcal{A}$ and $\mathfrak{s}^{-1}(\pi_B^{-1}(\{\infty\})) = (\pi_B^{-1}(\{\infty\}) \times \mathfrak{X}) \cup (\mathfrak{X} \times \pi_B^{-1}(\{\infty\})) \in \mathcal{A} \otimes \mathcal{A}$. Therefore, $\mathfrak{s}^{-1}(\pi_B^{-1}(C)) \in \mathcal{A} \otimes \mathcal{A}$ for all $C \subseteq \overline{\mathbb{Z}}_+$. Consequently, $(\mathfrak{X}, +, \mathcal{A})$ is a measurable Abelian semigroup, where the zero element is the zero measure 0.

Let $n \in \mathbb{N}$ be fixed and $N_j, X_j, X_{j,k}, Z_j, (j \in \underline{n}, k \in \mathbb{N})$ be independent random variables, where the $X_j, X_{j,k}$ are S-valued with distributions $P^{X_j} = P^{X_{j,k}}$, the Z_j are Bernoulli random variables with $P(Z_j = 1) = 1 - P(Z_j = 0) = p_j \in (0, 1]$ and the N_j are Poisson $\mathrm{Po}(p_j)$ distributed. Suppose that, for all $j \in \underline{n}$, P^{X_j} has a density $\widetilde{h}_j$ with respect to a σ-finite measure ν on $(S, \mathcal{S})$. Set $Q_j = P^{\delta_{X_j}}$ for $j \in \underline{n}$ and let $\lambda = \sum_{j=1}^n p_j, Q = \frac{1}{\lambda} \sum_{j=1}^n p_j Q_j, \eta = \frac{1}{\lambda} \sum_{j=1}^n p_j P^{X_j}$ and $\widetilde{h} = \frac{1}{\lambda} \sum_{j=1}^n p_j \widetilde{h}_j$. Then the point process $\xi = \sum_{j=1}^n Z_j \delta_{X_j}$ has distribution $F = \prod_{j=1}^n (\delta_0 + p_j(Q_j - \delta_0))$. The approximating $G = \exp(\lambda(Q - \delta_0))$ is the distribution of the Poisson point process $\zeta = \sum_{j=1}^n \sum_{k=1}^{N_j} \delta_{X_{j,k}}$ with intensity measure $E\zeta = E\xi = \lambda\eta$ with ν-density $\lambda\widetilde{h}$.

Proposition 3.1. *Under the assumptions above, we have*

$$d_{\mathrm{TV}}(P^\xi, P^\zeta) \leqslant \frac{\widetilde{\alpha}_1}{1 - 2^{3/2}\widetilde{\alpha}_1}, \; \text{if } \widetilde{\alpha}_1 < \frac{1}{2^{3/2}}; \quad d_{\mathrm{TV}}(P^\xi, P^\zeta) \leqslant 7.8\,\widetilde{\beta}_1, \qquad (3.1)$$

where

$$\widetilde{\alpha}_1 = \sum_{j=1}^n g(2p_j)p_j^2 \int_{\{\widetilde{h}>0\}} \widetilde{h}_j \min\left\{\frac{\widetilde{h}_j}{2^{3/2}\lambda\widetilde{h}}, 2\right\} \mathrm{d}\nu,$$

$$\widetilde{\beta}_1 = \sum_{j=1}^n p_j^2 \int_{\{\widetilde{h}>0\}} \widetilde{h}_j \min\left\{\frac{\widetilde{h}_j}{\lambda\widetilde{h}}, 1\right\} \mathrm{d}\nu,$$

and g is defined as in (1.8).

Remark 3.2. In the literature, there are two inequalities, which are comparable with those of Proposition 3.1. The simple one is the Le Cam type bound

$$d_{\mathrm{TV}}(P^\xi, P^\zeta) \leqslant \sum_{j=1}^n p_j^2 \qquad (3.2)$$

and is comparable to (1.5). A proof can, for example, be found in Matthes et al. (1978, 1.11.2 on p. 81).

A more interesting bound is given in Theorem 2 of Barbour (1988), which reads in our notation as

$$d_{\mathrm{TV}}(P^\xi, P^\zeta) \leqslant \frac{c_\lambda}{\lambda} \sum_{j=1}^n p_j^2 \varphi_j^2 \qquad (3.3)$$

with $c_\lambda = \frac{1}{2} + \max\{\log(2\lambda), 0\}$ and

$$\varphi_j = \sup_{C \in \mathcal{S}:\, \eta(C) > 0} \frac{P(X_j \in C)}{\eta(C)}, \qquad (j \in \underline{n}).$$

We note that the change of the notation is justified by Reiss (1993, Theorem 1.4.1, p. 29). In fact, for $j \in \underline{n}$, Barbour's term $\delta_{Y_j}(\cdot \cap B)$ can be replaced with $Z_j \delta_{X_j}$ where $P^{Z_j} = \delta_0 + P(Y_j \in B)(\delta_1 - \delta_0)$ and $P^{X_j} = P(Y_j \in \cdot \,|\, Y_j \in B)$.

For a comparison of the bounds in (3.2) and (3.3) with those of Proposition 3.1, we note that

$$\widetilde{h}_j \leqslant \varphi_j \widetilde{h} \quad \nu\text{-almost everywhere for all } j \in \underline{n}. \tag{3.4}$$

Indeed, if $C \in \mathcal{S}$ with $\eta(C) > 0$ then $\int_C \widetilde{h}_j \, d\nu = P(X_j \in C) = \int_C \frac{P(X_j \in C)}{\eta(C)} \widetilde{h} \, d\nu \leqslant \int_C \varphi_j \widetilde{h} \, d\nu$; on the other hand, if $\eta(C) = 0$, then $\int_C \widetilde{h}_j \, d\nu = P(X_j \in C) = 0 = \int_C \varphi_j \widetilde{h} \, d\nu$. Now, (3.4) follows from 3.17 in Hoffmann-Jørgensen (1994). Therefore, $\int_{\{\widetilde{h}>0\}} \frac{\widetilde{h}_j^2}{\widetilde{h}} \, d\nu \leqslant \varphi_j \leqslant \varphi_j^2$. Consequently, if λ is large and the $\widetilde{h}_j$ for $j \in \underline{n}$ are not too different, then the bounds in (3.1) are preferable to the ones in (3.2) and (3.3).

Further results in the Poisson process approximation can, for example, be found in Barbour et al. (1992, Chapter 10) and Reiss (1993) and in the works cited there.

4. Proofs

4.1. *Auxiliary norm estimates.* The proofs of the theorems require some upper bounds of certain norm terms, which measure the smoothness of compound Poisson distributions. In fact, in the simplest case terms like $\|(U-\delta_0)^k \exp(\lambda(U-\delta_0))\|$ have to be considered for $U \in \mathcal{F}$ and $k \in \mathbb{N}$. For some properties of such norm terms, see, e.g., Čekanavičius (1995), Roos (1999a, Proposition 4), Roos (2001, Lemma 3), Čekanavičius and Roos (2006, Lemmata 3.4, 3.12) and the references cited therein. In the following lemma, we present preliminary norm estimates, which will be used in the proof of Lemma 4.4. A related bound can be found in Roos (2003, Lemma 2).

Lemma 4.1. *Let $d, k \in \mathbb{N}$, $p_{j,r} \in \mathbb{R}$ for $j \in \underline{k}$ and $r \in \underline{d}$, $\Lambda = (\lambda_1, \ldots, \lambda_d) \in (0, \infty)^d$. For $r \in \underline{d}$, let $U_r \in \mathcal{F}$, $W_r = U_r - \delta_0$. Set $R_j = \sum_{r=1}^d p_{j,r} W_r$ for $j \in \underline{k}$ and $G = \exp(\sum_{r=1}^d \lambda_r W_r)$. Then, we have*

$$\Big\| \Big(\prod_{j=1}^k R_j \Big) G \Big\| \leqslant \Big(\frac{1}{k!} \sum_{r \in \underline{d}^k} \Big(\sum_{\ell \in \underline{k}_{\neq}^k} \prod_{j=1}^k \frac{p_{j,r_{\ell(j)}}}{\sqrt{\lambda_{r_{\ell(j)}}}} \Big)^2 \Big)^{1/2} \leqslant \sqrt{k!} \prod_{j=1}^k \Big(\sum_{r=1}^d \frac{p_{j,r}^2}{\lambda_r} \Big)^{1/2}. \tag{4.1}$$

Proof: We need some preparations. For $j \in \mathbb{N}$, $m \in \mathbb{Z}$ and $t \in [0, \infty)$, let $\Delta^j \mathrm{po}(m, t) = \Delta^{j-1} \mathrm{po}(m-1, t) - \Delta^{j-1} \mathrm{po}(m, t)$, $\Delta^0 \mathrm{po}(m, t) = \mathrm{po}(m, t)$. It is well-known that $\Delta^j \mathrm{po}(m, t) = \frac{1}{t^j} \mathrm{po}(m, t) \, \mathrm{Ch}(j, m, t)$, $(j, m \in \mathbb{Z}_+,\ t \in (0, \infty))$ (cf. Roos (1999a)), where

$$\mathrm{Ch}(j, x, t) = \sum_{i=0}^j \binom{j}{i} \binom{x}{i} i! \, (-t)^{j-i}, \qquad (j \in \mathbb{Z}_+,\ t, x \in \mathbb{R})$$

denotes the Charlier polynomial of degree j and $\binom{x}{i} = \prod_{j=1}^i \frac{x-j+1}{j}$ for $i \in \mathbb{Z}_+$ and $x \in \mathbb{R}$. Further, the Charlier polynomials are orthogonal with respect to the

Poisson distribution (see, e.g., Chihara, 1978, formula (1.14), page 4), that is

$$\sum_{m=0}^{\infty} \mathrm{po}(m,t)\,\mathrm{Ch}(i,m,t)\,\mathrm{Ch}(j,m,t) = \mathbb{1}_{\{i\}}(j)\,i!\,t^i, \quad (i,j \in \mathbb{Z}_+,\, t \in (0,\infty)). \quad (4.2)$$

It is easily shown that, for $j \in \mathbb{Z}_+$ and $r \in \underline{d}$, we have

$$W_r^j \exp(\lambda_r W_r) = \sum_{m=0}^{\infty} \Delta^j \mathrm{po}(m,\lambda_r)\,U_r^m.$$

For $r \in \underline{d}^k$ and $s \in \underline{d}$, let $v_s(r) = \sum_{j=1}^k \mathbb{1}_{\{s\}}(r_j)$ and set $v(r) = (v_1(r), \ldots, v_d(r)) \in \mathbb{Z}_+^d$. Clearly, $|v(r)| = k$. For $r \in \underline{d}^k$, we obtain

$$\prod_{j=1}^k W_{r_j} = \prod_{j=1}^k \Big(\prod_{s=1}^d W_s^{\mathbb{1}_{\{s\}}(r_j)} \Big) = \prod_{s=1}^d W_s^{v_s(r)}$$

and similarly $\prod_{j=1}^k \lambda_{r_j} = \Lambda^{v(r)}$. Therefore, letting $\mathrm{po}(m,\Lambda) = \prod_{r=1}^d \mathrm{po}(m_r,\lambda_r)$ for $m \in \mathbb{Z}_+^d$, we get

$$\Big\| \Big(\prod_{j=1}^k R_j \Big) G \Big\| = \Big\| \sum_{r \in \underline{d}^k} \Big(\prod_{j=1}^k p_{j,r_j} \Big) \prod_{s=1}^d (W_s^{v_s(r)} \exp(\lambda_s W_s)) \Big\|$$

$$= \Big\| \sum_{r \in \underline{d}^k} \Big(\prod_{j=1}^k p_{j,r_j} \Big) \sum_{m \in \mathbb{Z}_+^d} \Big(\prod_{s=1}^d (\Delta^{v_s(r)} \mathrm{po}(m_s,\lambda_s) U_s^{m_s}) \Big) \Big\|$$

$$= \Big\| \sum_{m \in \mathbb{Z}_+^d} \mathrm{po}(m,\Lambda) \sum_{r \in \underline{d}^k} \frac{1}{\Lambda^{v(r)}} \Big(\prod_{j=1}^k p_{j,r_j} \Big) \prod_{s=1}^d (\mathrm{Ch}(v_s(r),m_s,\lambda_s) U_s^{m_s}) \Big\|$$

$$\leqslant \sum_{m \in \mathbb{Z}_+^d} \mathrm{po}(m,\Lambda) \Big| \sum_{r \in \underline{d}^k} \frac{1}{\Lambda^{v(r)}} \Big(\prod_{j=1}^k p_{j,r_j} \Big) \prod_{s=1}^d \mathrm{Ch}(v_s(r),m_s,\lambda_s) \Big|.$$

For $j \in \underline{k}$ and $r \in \underline{d}$, set $a_{j,r} = \frac{p_{j,r}}{\sqrt{\lambda_r}}$. Hence, using the Cauchy-Schwarz inequality, we obtain

$$\Big\| \Big(\prod_{j=1}^k R_j \Big) G \Big\|^2 \leqslant \sum_{m \in \mathbb{Z}_+^d} \mathrm{po}(m,\Lambda) \Big(\sum_{r \in \underline{d}^k} \frac{1}{\Lambda^{v(r)}} \Big(\prod_{j=1}^k p_{j,r_j} \Big) \prod_{s=1}^d \mathrm{Ch}(v_s(r),m_s,\lambda_s) \Big)^2$$

$$= \sum_{r \in \underline{d}^k} \sum_{\widetilde{r} \in \underline{d}^k} \frac{1}{\Lambda^{v(r)+v(\widetilde{r})}} \prod_{j=1}^k (p_{j,r_j} p_{j,\widetilde{r}_j})$$

$$\times \prod_{s=1}^d \Big(\sum_{m_s=0}^{\infty} \mathrm{po}(m_s,\lambda_s) \mathrm{Ch}(v_s(r),m_s,\lambda_s) \mathrm{Ch}(v_s(\widetilde{r}),m_s,\lambda_s) \Big).$$

The application of (4.2) now gives

$$\left\|\Big(\prod_{j=1}^{k} R_j\Big) G\right\|^2 \leqslant \sum_{r\in\underline{d}^k} \sum_{\widetilde{r}\in\underline{d}^k:\, v(r)=v(\widetilde{r})} \frac{1}{\Lambda^{v(r)+v(\widetilde{r})}} \Big(\prod_{j=1}^{k} (p_{j,r_j} p_{j,\widetilde{r}_j})\Big) \prod_{s=1}^{d} \Big(v_s(r)!\, \lambda_s^{v_s(r)}\Big)$$

$$= \sum_{m\in\mathbb{Z}_+^d:\, |m|=k} m! \sum_{r\in\underline{d}^k:\, v(r)=m} \sum_{\widetilde{r}\in\underline{d}^k:\, v(\widetilde{r})=m} \prod_{j=1}^{k} (a_{j,r_j} a_{j,\widetilde{r}_j})$$

$$= \sum_{m\in\mathbb{Z}_+^d:\, |m|=k} m! \Big(\sum_{r\in\underline{d}^k:\, v(r)=m} \prod_{j=1}^{k} a_{j,r_j} \Big)^2.$$

For $z\in\mathbb{C}^d$, we have

$$\sum_{m\in\mathbb{Z}_+^d:\, |m|=k} \Big(\sum_{s\in\underline{d}^k:\, v(s)=m} 1 \Big) z^m = \sum_{s\in\underline{d}^k} \sum_{m\in\mathbb{Z}_+^d:\, m=v(s)} z^{v(s)}$$

$$= \sum_{s\in\underline{d}^k} \prod_{j=1}^{k} z_{s_j} = \Big(\sum_{s=1}^{d} z_s\Big)^k = \sum_{m\in\mathbb{Z}_+^d:\, |m|=k} \frac{k!}{m!} z^m,$$

which implies that, for $m\in\mathbb{Z}_+^d$ with $|m|=k$,

$$\sum_{s\in\underline{d}^k:\, v(s)=m} 1 = \frac{k!}{m!}. \tag{4.3}$$

Consequently

$$\left\|\Big(\prod_{j=1}^{k} R_j\Big) G\right\|^2 \leqslant \sum_{m\in\mathbb{Z}_+^d:\, |m|=k} m! \Big(\sum_{r\in\underline{d}^k:\, v(r)=m} \prod_{j=1}^{k} a_{j,r_j} \Big)^2$$

$$= \frac{1}{k!} \sum_{m\in\mathbb{Z}_+^d:\, |m|=k} \sum_{s\in\underline{d}^k:\, v(s)=m} (v(s)!)^2 \Big(\sum_{r\in\underline{d}^k:\, v(r)=v(s)} \prod_{j=1}^{k} a_{j,r_j} \Big)^2$$

$$= \frac{1}{k!} \sum_{s\in\underline{d}^k} \Big(v(s)! \sum_{r\in\underline{d}^k:\, v(r)=v(s)} \prod_{j=1}^{k} a_{j,r_j} \Big)^2.$$

For $m\in\mathbb{Z}_+^d$ and $r,s\in\underline{d}^k$ with $v(r)=v(s)=m$, it easily follows from the definition of $v(r)$ that $\sum_{\ell\in\underline{k}_{\neq}^k} \mathbb{1}_{\{r\}}(s_{\ell(1)},\ldots,s_{\ell(k)}) = m!$. However, a more explicit proof is as follows: Since the left-hand side clearly only depends on m, we obtain by using (4.3) that

$$\sum_{\ell\in\underline{k}_{\neq}^k} \mathbb{1}_{\{r\}}(s_{\ell(1)},\ldots,s_{\ell(k)}) = \frac{m!}{k!} \sum_{\widetilde{r}\in\underline{d}^k:\, v(\widetilde{r})=m} \sum_{\ell\in\underline{k}_{\neq}^k} \mathbb{1}_{\{r\}}(s_{\ell(1)},\ldots,s_{\ell(k)})$$

$$= \frac{m!}{k!} \sum_{\ell\in\underline{k}_{\neq}^k} \sum_{\widetilde{r}\in\underline{d}^k:\, v(\widetilde{r})=m} \mathbb{1}_{\{\widetilde{r}\}}(s_{\ell(1)},\ldots,s_{\ell(k)}) = \frac{m!}{k!} \sum_{\ell\in\underline{k}_{\neq}^k} 1 = m!.$$

Hence, for $s \in \underline{d}^k$,

$$v(s)! \sum_{r \in \underline{d}^k : v(r)=v(s)} \prod_{j=1}^k a_{j,r_j} = \sum_{r \in \underline{d}^k : v(r)=v(s)} \sum_{\ell \in \underline{k}^k_{\neq}} \mathbb{1}_{\{r\}}(s_{\ell(1)}, \ldots, s_{\ell(k)}) \prod_{j=1}^k a_{j,r_j}$$

$$= \sum_{\ell \in \underline{k}^k_{\neq}} \Big(\sum_{r \in \underline{d}^k : v(r)=v(s)} \mathbb{1}_{\{r\}}(s_{\ell(1)}, \ldots, s_{\ell(k)}) \Big) \prod_{j=1}^k a_{j,s_{\ell(j)}} = \sum_{\ell \in \underline{k}^k_{\neq}} \prod_{j=1}^k a_{j,s_{\ell(j)}}.$$

Using the Cauchy-Schwarz inequality again,

$$\Big\| \Big(\prod_{j=1}^k R_j \Big) G \Big\|^2 \leqslant \frac{1}{k!} \sum_{r \in \underline{d}^k} \Big(\sum_{\ell \in \underline{k}^k_{\neq}} \prod_{j=1}^k a_{j,r_{\ell(j)}} \Big)^2 = \frac{1}{k!} \sum_{\ell \in \underline{k}^k_{\neq}} \sum_{\widetilde{\ell} \in \underline{k}^k_{\neq}} \sum_{r \in \underline{d}^k} \prod_{j=1}^k (a_{j,r_{\ell(j)}} a_{j,r_{\widetilde{\ell}(j)}})$$

$$\leqslant \frac{1}{k!} \sum_{\ell \in \underline{k}^k_{\neq}} \sum_{\widetilde{\ell} \in \underline{k}^k_{\neq}} \Big(\sum_{r \in \underline{d}^k} \prod_{j=1}^k a_{j,r_{\ell(j)}}^2 \Big)^{1/2} \Big(\sum_{r \in \underline{d}^k} \prod_{j=1}^k a_{j,r_{\widetilde{\ell}(j)}}^2 \Big)^{1/2}$$

$$= \frac{1}{k!} \sum_{\ell \in \underline{k}^k_{\neq}} \sum_{\widetilde{\ell} \in \underline{k}^k_{\neq}} \prod_{j=1}^k \Big(\Big(\sum_{r_{\ell(j)}=1}^d a_{j,r_{\ell(j)}}^2 \Big) \Big(\sum_{r_{\widetilde{\ell}(j)}=1}^d a_{j,r_{\widetilde{\ell}(j)}}^2 \Big) \Big)^{1/2} = k! \prod_{j=1}^k \Big(\sum_{r=1}^d a_{j,r}^2 \Big),$$

which proves (4.1). □

Corollary 4.2. *Under the assumptions of Lemma 4.1, we obtain, for $k = 1$, resp. $k = 2$, that*

$$\|R_1 G\| \leqslant \Big(\sum_{r=1}^d \frac{p_{1,r}^2}{\lambda_r} \Big)^{1/2}, \tag{4.4}$$

$$\|R_1 R_2 G\| \leqslant \Big(\frac{1}{2} \sum_{(r,s) \in \underline{d}^2} \frac{(p_{1,r} p_{2,s} + p_{1,s} p_{2,r})^2}{\lambda_r \lambda_s} \Big)^{1/2} \leqslant \sqrt{2} \prod_{j=1}^2 \Big(\sum_{r=1}^d \frac{p_{j,r}^2}{\lambda_r} \Big)^{1/2}. \tag{4.5}$$

We note that (4.4) was shown in Roos (1999b, formula (18)), whereas (4.5) is a generalization of one part of (19) of that paper. The next lemma is needed in the proof of Lemma 4.4 below.

Lemma 4.3. *Let $k \in \mathbb{N}$, $m \in \mathbb{Z}_+^2$ with $|m| \leqslant k$. Then*

$$(2m_1 + m_2)! \leqslant ((2k)!)^{m_1/k} ((2k-1)!)^{m_2/(2k)},$$

where equality holds in the case $k = 1$.

Proof: For $\ell \in \mathbb{N}$, we have $\frac{(\ell!)^{\ell+1}}{((\ell+1)!)^\ell} = \frac{\ell!}{(\ell+1)^\ell} \leqslant 1$ and $\frac{((2\ell-1)!)^{\ell+1}}{((2\ell+1)!)^\ell} = \frac{(2\ell-1)!}{(2\ell(2\ell+1))^\ell} \leqslant 1$. Therefore $(\ell!)^{1/\ell}$ and $((2\ell-1)!)^{1/\ell}$ are both increasing in $\ell \in \mathbb{N}$. Hence we may assume that $m_2 \geqslant 1$ and $|m| = k$. Using that $\ell! \leqslant \ell^{\ell-1}$ for $\ell \in \mathbb{N}$, we get

$$\frac{((2m_1 + m_2)!)^{2k}}{((2k)!)^{2m_1}((2k-1)!)^{m_2}} = ((2k-m_2)!)^{m_2} \Big(\frac{(2k-m_2)!}{(2k)!} \Big)^{2k-2m_2} \Big(\frac{(2k-m_2)!}{(2k-1)!} \Big)^{m_2}$$

$$\leqslant \frac{(2k-m_2)^{m_2(2k-m_2-1)}}{(2k-m_2)^{m_2(2k-2m_2)}(2k-m_2)^{m_2(m_2-1)}} = 1,$$

which implies the assertion. □

Lemma 4.4. *Let the assumptions of Lemma 4.1 hold and let $p_j = \sum_{r=1}^{d} |p_{j,r}|$ for $j \in \underline{k}$. Set $c_k = ((2k)!)^{1/(2k)}$ and $c'_k = ((2k-1)!)^{1/(4k)}$. Then, for all $u \in [0, \frac{1}{2}]^k$, $v, w \in (0, \infty)^k$, we have*

$$\left\| \left(\prod_{j=1}^{k} R_j^2 \right) G \right\| \leqslant \prod_{j=1}^{k} \left(C_j \sum_{r=1}^{d} |p_{j,r}| \min\left\{ \frac{|p_{j,r}|}{\lambda_r}, \frac{4}{w_j} p_j \right\} \right),$$

where $C_j = \max\{c_k + c'_k \frac{u_j}{v_j}, (2(1-u_j) + c'_k u_j v_j) w_j\}$ for $j \in \underline{k}$. In particular, for $u = 0$ and $w_j = \frac{c_k}{2}$ for $j \in \underline{k}$, we obtain

$$\left\| \left(\prod_{j=1}^{k} R_j^2 \right) G \right\| \leqslant \sqrt{(2k)!} \prod_{j=1}^{k} \left(\sum_{r=1}^{d} |p_{j,r}| \min\left\{ \frac{|p_{j,r}|}{\lambda_r}, \frac{8}{c_k} p_j \right\} \right).$$

We note that $c_k \geqslant 8$, if $k \geqslant 10$.

Proof: We may assume that $p_j > 0$ for all $j \in \underline{k}$; further, set

$$I_j = \left\{ r \in \underline{d} \,\middle|\, \frac{|p_{j,r}|}{\lambda_r} \leqslant \frac{4}{w_j} p_j \right\}, \qquad I_j^c = \underline{d} \setminus I_j,$$

$$a_j = \sum_{r \in I_j} \frac{p_{j,r}^2}{\lambda_r} = \sum_{r \in I_j} |p_{j,r}| \min\left\{ \frac{|p_{j,r}|}{\lambda_r}, \frac{4}{w_j} p_j \right\},$$

$$b_j = 2 \sum_{r \in I_j^c} |p_{j,r}| = \frac{w_j}{2p_j} \sum_{r \in I_j^c} |p_{j,r}| \min\left\{ \frac{|p_{j,r}|}{\lambda_r}, \frac{4}{w_j} p_j \right\},$$

$$R'_j = \sum_{r \in I_j} p_{j,r} W_r = \sum_{r=1}^{d} \mathbb{1}_{I_j}(r) p_{j,r} W_r, \quad R''_j = \sum_{r \in I_j^c} p_{j,r} W_r,$$

$$Y_j = 2(1 - u_j) R'_j R''_j + (R''_j)^2.$$

In particular, we have

$$b_j \leqslant 2p_j, \quad a_j + \frac{2p_j}{w_j} b_j = \sum_{r=1}^{d} |p_{j,r}| \min\left\{ \frac{|p_{j,r}|}{\lambda_r}, \frac{4}{w_j} p_j \right\},$$

$$\|R'_j\| \leqslant 2p_j - b_j, \ \|R''_j\| \leqslant b_j, \ \|Y_j\| \leqslant 2(1 - u_j)(2p_j - b_j) b_j + b_j^2 \leqslant 4(1 - u_j) p_j b_j.$$

Further, for $J_1, J_2 \subseteq \underline{k}$ with $J_1 \cap J_2 = \emptyset$, $|J_1| = m_1$, $|J_2| = m_2$, we have $m_1 + m_2 \leqslant k$ and Lemmata 4.1 and 4.3 imply that

$$\left\| \left(\prod_{j \in J_1} (R'_j)^2 \right) \left(\prod_{j \in J_2} R'_j \right) G \right\| \leqslant \sqrt{(2m_1 + m_2)!} \left(\prod_{j \in J_1} a_j \right) \prod_{j \in J_2} \sqrt{a_j}$$

$$\leqslant \left(\prod_{j \in J_1} (c_k a_j) \right) \prod_{j \in J_2} (c'_k \sqrt{a_j}).$$

Therefore

$$\Big\|\Big(\prod_{j=1}^{k} R_j^2\Big)G\Big\| = \Big\|\Big(\prod_{j=1}^{k}\big((R_j')^2 + 2u_j R_j' R_j'' + Y_j\big)\Big)G\Big\|$$

$$= \Big\| \sum_{m\in\mathbb{Z}_+^2:\,|m|\leqslant k} \sum_{J_1\subseteq\underline{k}:\,|J_1|=m_1} \sum_{J_2\subseteq\underline{k}\setminus J_1:\,|J_2|=m_2} \Big(\prod_{j\in J_1}(R_j')^2\Big)$$

$$\times \Big(\prod_{j\in J_2}(2u_j R_j' R_j'')\Big)\Big(\prod_{j\in\underline{k}\setminus(J_1\cup J_2)} Y_j\Big)G\Big\|$$

$$\leqslant \sum_{m\in\mathbb{Z}_+^2:\,|m|\leqslant k} \sum_{J_1\subseteq\underline{k}:\,|J_1|=m_1} \sum_{J_2\subseteq\underline{k}\setminus J_1:\,|J_2|=m_2} \Big\|\Big(\prod_{j\in J_1}(R_j')^2\Big)\Big(\prod_{j\in J_2} R_j'\Big)G\Big\|$$

$$\times \Big(\prod_{j\in J_2}(2u_j\|R_j''\|)\Big)\prod_{j\in\underline{k}\setminus(J_1\cup J_2)}\|Y_j\|$$

$$\leqslant \sum_{m\in\mathbb{Z}_+^2:\,|m|\leqslant k} \sum_{J_1\subseteq\underline{k}:\,|J_1|=m_1} \sum_{J_2\subseteq\underline{k}\setminus J_1:\,|J_2|=m_2} \Big(\prod_{j\in J_1}(c_k a_j)\Big)$$

$$\times \Big(\prod_{j\in J_2}(2c_k' u_j\sqrt{a_j}\,b_j)\Big)\prod_{j\in\underline{k}\setminus(J_1\cup J_2)}(4(1-u_j)p_j b_j),$$

giving

$$\Big\|\Big(\prod_{j=1}^{k} R_j^2\Big)G\Big\| \leqslant \prod_{j=1}^{d}\Big(c_k a_j + 2c_k' u_j\sqrt{\frac{a_j}{v_j}\,b_j^2 v_j} + 4(1-u_j)p_j b_j\Big).$$

Using that $2\sqrt{xy}\leqslant x+y$ for $x,y\in[0,\infty)$, we obtain, for $j\in\underline{d}$,

$$c_k a_j + 2c_k' u_j\sqrt{\frac{a_j}{v_j}\,b_j^2 v_j} + 4(1-u_j)p_j b_j$$

$$\leqslant \Big(c_k + c_k'\frac{u_j}{v_j}\Big)a_j + (2(1-u_j) + c_k' u_j v_j)2p_j b_j$$

$$\leqslant \max\Big\{c_k + c_k'\frac{u_j}{v_j},\ (2(1-u_j) + c_k' u_j v_j)w_j\Big\}\Big(a_j + \frac{2p_j}{w_j}b_j\Big)$$

$$= C_j \sum_{r=1}^{d}|p_{j,r}|\,\min\Big\{\frac{|p_{j,r}|}{\lambda_r},\ \frac{4}{w_j}p_j\Big\},$$

which implies the assertion. $\square$

Corollary 4.5. *Under the assumptions of Lemma 4.4, we have*

$$\Big\|\Big(\prod_{j=1}^{k} R_j^2\Big)G\Big\| \leqslant D_k\,k! \prod_{j=1}^{k}\Big(\sum_{r=1}^{d}|p_{j,r}|\,\min\Big\{\frac{|p_{j,r}|}{\lambda_r},\ p_j\Big\}\Big),$$

where $w_j = 4$, $(j\in\underline{k})$, if $k\in\underline{9}$, and the values of u_j, v_j, D_k are given in Table 2 below.

Table 2: Explicit values of the constants D_k in Corollary 4.5

k	1	2	3	4	5
u_j	0.5000	0.5000	0.5000	0.5000	0.4500
v_j	0.1708	0.2574	0.3589	0.4666	0.5192
D_k	4.342	10.784	21.721	40.687	74.672

k	6	7	8	9	$\geqslant 10$
u_j	0.3000	0.1996	0.1500	0.0500	0
v_j	0.4414	0.4099	0.5002	0.4560	1
D_k	125.448	186.872	253.020	305.314	$\frac{\sqrt{(2k)!}}{k!}$

Lemma 4.6. *Let $d \in \mathbb{N}$. For $r \in \underline{d}$, let $p_r \in [0,1]$, $\lambda_r \in (0,\infty)$ with $\lambda_r \geqslant p_r$, $U_r \in \mathcal{F}$. We assume that $p := \sum_{r=1}^d p_r \leqslant 1$. Let $R = \sum_{r=1}^d p_r(U_r - \delta_0)$, $G = \exp(\sum_{r=1}^d \lambda_r(U_r - \delta_0))$. Let $u \in [0, \frac{1}{2}]$, $v, w \in (0,\infty)$ and $w_0 \in (1,\infty)$ be the unique solution of $f(w_0) = \frac{2}{w}$, where $f(x) = x\log(1 + \frac{1}{x-1}) - 1 = \int_0^1 \frac{t}{x-t}\,dt$ for $x \in (1,\infty)$. Then, letting $C = \max\{(\sqrt{2} + \frac{u}{v})\frac{2}{w}, 4(1-u) + 2uv\}$,*

$$\|((\delta_0 + R)\mathrm{e}^{-R} - \delta_0)G\| \leqslant C \sum_{r=1}^d p_r \min\Big\{w_0 \frac{p_r}{\lambda_r}, p\Big\}.$$

In particular, if $u = \frac{1}{2}$, $v = 0.47248$ and $w = 2$, then $C \leqslant 2.473$ and $w_0 \leqslant 1.256$, giving

$$\|((\delta_0 + R)\mathrm{e}^{-R} - \delta_0)G\| \leqslant 3.11 \sum_{r=1}^d p_r \min\Big\{\frac{p_r}{\lambda_r}, p\Big\}.$$

Proof: We may assume that $p_r > 0$ for all $r \in \underline{d}$. It is easily shown that

$$((\delta_0 + R)\mathrm{e}^{-R} - \delta_0)G = -\int_0^1 tR^2 \exp(-tR)G\,dt,$$

where the equality holds setwise. From Lemma 4.4, we obtain for $t \in (0,1)$ that

$$\|R^2 \exp(-tR)G\| \leqslant C\frac{w}{2}\sum_{r=1}^d p_r \min\Big\{\frac{p_r}{\lambda_r - tp_r}, \frac{4}{w}p\Big\}.$$

Consequently

$$\|((\delta_0 + R)\mathrm{e}^{-R} - \delta_0)G\| \leqslant C\frac{w}{2}\int_0^1 t\sum_{r=1}^d p_r \min\Big\{\frac{p_r}{\lambda_r - tp_r}, \frac{4}{w}p\Big\}\,dt$$

$$\leqslant C\sum_{r=1}^d p_r \min\Big\{\frac{w}{2}f\Big(\frac{\lambda_r}{p_r}\Big), p\Big\}.$$

Let $r \in \underline{d}$. If $\frac{w}{2}f(\frac{\lambda_r}{p_r}) \leqslant p$, then $f(\frac{\lambda_r}{p_r}) \leqslant \frac{2}{w} = f(w_0)$, giving $\frac{\lambda_r}{p_r} \geqslant w_0$, since f is decreasing. Further, $xf(x) = \int_0^1 \frac{t}{1-t/x}\,dt$ is decreasing in $x \in (1,\infty)$, giving $\frac{w}{2}f(\frac{\lambda_r}{p_r}) \leqslant \frac{w}{2}\frac{p_r}{\lambda_r}w_0 f(w_0) = w_0\frac{p_r}{\lambda_r}$. On the other hand, if $p \leqslant \frac{w}{2}f(\frac{\lambda_r}{p_r})$, then $f(w_0) = \frac{2}{w} \leqslant \frac{1}{p}f(\frac{\lambda_r}{p_r}) \leqslant f(p\frac{\lambda_r}{p_r})$ and so $p\frac{\lambda_r}{p_r} \leqslant w_0$, which implies that $p \leqslant w_0\frac{p_r}{\lambda_r}$. Therefore, in any case $\min\{\frac{w}{2}f(\frac{\lambda_r}{p_r}), p\} \leqslant \min\{w_0\frac{p_r}{\lambda_r}, p\}$. Together with the above, we obtain the assertion. $\square$

4.2. *Remaining proofs.*

Proof of the first inequality in Proposition 1.3: Let the notation of Remark 1.2(a) be valid. Then $P(\sum_{r\in J} X_{j,r} = 1) = 1 - P(\sum_{r\in J} X_{j,r} = 0) = \widetilde{p}_j$ for $j \in \underline{n}$ and $\sum_{r\in J} \lambda_r = \widetilde{\lambda}$. Consequently $\sum_{r\in J} S_{n,r}$ and $\sum_{r\in J} T_r$ have the distributions $\prod_{j=1}^n (\delta_0 + \widetilde{p}_j(\delta_1 - \delta_0))$ and $\mathrm{Po}(\widetilde{\lambda})$, respectively, and hence

$$\|F - G\| = 2 \sup_{A\subseteq\mathbb{Z}_+^d} |P(S_n \in A) - P(T \in A)|$$

$$\geqslant 2 \sup_{B\subseteq\mathbb{Z}_+} \left| P\Big(\sum_{r\in J} S_{n,r} \in B\Big) - P\Big(\sum_{r\in J} T_r \in B\Big) \right|$$

$$= \|P^{\sum_{r\in J} S_{n,r}} - P^{\sum_{r\in J} T_r}\| = \left\| \prod_{j=1}^n (\delta_0 + \widetilde{p}_j(\delta_1 - \delta_0)) - \mathrm{Po}(\widetilde{\lambda}) \right\|,$$

which implies the first inequality in Proposition 1.3. $\square$

Proof of Theorem 2.1: We first note that

$$F = \prod_{j=1}^n F_j = \prod_{j=1}^n ((V_j + \delta_0)\mathrm{e}^{R_j}) = \sum_{k=0}^n H_k = G_n, \qquad (4.6)$$

which implies that $F - G_\ell = \sum_{k=\ell+1}^n H_k$. For $j \in \underline{n}$, we have $V_j = F_j \mathrm{e}^{-R_j} - \delta_0 = -\frac{g(-R_j)}{2} R_j^2$. Hence, for $k \in \underline{n}_0$,

$$H_k = (-1)^k \sum_{J\subseteq\underline{n}:\, |J|=k} \Big(\prod_{j\in J} \frac{g(-R_j)}{2} \Big) \Big(\prod_{j\in J} R_j^2 \Big) \exp(\lambda(Q - \delta_0)). \qquad (4.7)$$

If $k \in \underline{n}$, $J \subseteq \underline{n}$ with $|J| = k$, then Lemma 4.4 with $u = 0$, $w_j = \frac{1}{\sqrt{2}}$ for $j \in J$ implies that

$$\left\| \Big(\prod_{j\in J} R_j^2 \Big) \exp(\lambda(Q - \delta_0)) \right\| \leqslant \sqrt{(2k)!} \prod_{j\in J} \Big(\sum_{r=1}^d p_j q_{j,r} \min\Big\{ \frac{p_j q_{j,r}}{\lambda_r}, 2^{5/2} p_j \Big\} \Big), \quad (4.8)$$

since $((2k)!)^{1/(2k)} \geqslant \sqrt{2}$. On the other hand, for $j \in \underline{n}$, $\|R_j\| \leqslant 2p_j$ and therefore

$$\|g(-R_j)\| = \left\| 2\sum_{m=2}^\infty \frac{m-1}{m!} (-R_j)^{m-2} \right\| \leqslant 2\sum_{m=2}^\infty \frac{m-1}{m!} \|R_j\|^{m-2} \leqslant g(2p_j). \quad (4.9)$$

By (4.7), (4.8), (4.9) and the polynomial theorem, we derive for $k \in \underline{n}$,

$$\|H_k\| \leqslant \sum_{J\subseteq\underline{n}:\, |J|=k} \Big(\prod_{j\in J} \frac{\|g(-R_j)\|}{2} \Big) \left\| \Big(\prod_{j\in J} R_j^2 \Big) \exp(\lambda(Q - \delta_0)) \right\|$$

$$\leqslant \frac{\sqrt{(2k)!}}{2^k} \sum_{J\subseteq\underline{n}:\, |J|=k} \prod_{j\in J} \Big(g(2p_j) p_j^2 \sum_{r=1}^d q_{j,r} \min\Big\{ \frac{q_{j,r}}{\lambda_r}, 2^{5/2} \Big\} \Big)$$

$$\leqslant \frac{\sqrt{(2k)!}}{k!\, 2^k} \Big(\sum_{j=1}^n g(2p_j) p_j^2 \sum_{r=1}^d q_{j,r} \min\Big\{ \frac{q_{j,r}}{\lambda_r}, 2^{5/2} \Big\} \Big)^k = \frac{\sqrt{(2k)!}}{k!\, 2^k} (2^{3/2} \alpha_1)^k.$$

It is easily shown that $\frac{\sqrt{(2k)!}}{k!\,2^k}$ is decreasing in $k \in \mathbb{Z}_+$. Consequently, if $\alpha_1 < \frac{1}{2^{3/2}}$, then

$$\|F - G_\ell\| \leqslant \sum_{k=\ell+1}^{n} \|H_k\| \leqslant \sum_{k=\ell+1}^{n} \frac{\sqrt{(2k)!}}{k!\,2^k} (2^{3/2}\,\alpha_1)^k$$

$$\leqslant \frac{\sqrt{(2(\ell+1))!}}{(\ell+1)!\,2^{\ell+1}} 2^{3(\ell+1)/2} \frac{\alpha_1^{\ell+1}}{1 - 2^{3/2}\alpha_1},$$

which proves (2.1). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Theorem 2.3: We need a further bound for $\|H_k\|$, $(k \in \underline{n})$ in terms of β_1. Lemma 4.6 gives

$$\|H_1\| = \left\|\sum_{j=1}^{n}((\delta_0 + R_j)\mathrm{e}^{-R_j} - \delta_0)\exp(\lambda(Q - \delta_0)))\right\| \leqslant D_1'\beta_1.$$

For $k \in \underline{n} \setminus \{1\}$, Corollary 4.5 and the polynomial theorem imply that

$$\|H_k\| \leqslant \sum_{J \subseteq \underline{n}:\,|J|=k} \left(\prod_{j \in J} \frac{g(2p_j)}{2}\right)\left\|\left(\prod_{j \in J} R_j^2\right)\exp(\lambda(Q - \delta_0))\right\|$$

$$\leqslant D_k\left(\frac{g(2)}{2}\right)^k k! \sum_{J \subseteq \underline{n}:\,|J|=k} \prod_{j \in J} \left(p_j^2 \sum_{r=1}^{d} q_{j,r}\min\left\{\frac{q_{j,r}}{\lambda_r}, 1\right\}\right) \leqslant D_k'\beta_1^k.$$

Hence

$$\|F - G_\ell\| \leqslant \sum_{k=\ell+1}^{\infty} D_k'\beta_1^k = h_1(\beta_1)$$

and, alternatively,

$$\|F - G_\ell\| \leqslant \|F\| + \|G_\ell\| \leqslant 2 + \sum_{k=1}^{\ell}\|H_k\| \leqslant 2 + \sum_{k=1}^{\ell} D_k'\beta_1^k = h_2(\beta_1).$$

By the definition of D_k' for $k \geqslant 10$, we know that $h_1(x) < \infty$ for $x \in [0, \frac{1}{g(2)})$. Further, it is easily seen that $\frac{h_1(x)}{h_2(x)}$ is increasing in $x \in [0, \frac{1}{g(2)})$ with $\lim_{x\uparrow 1/g(2)} \frac{h_1(x)}{h_2(x)} = \infty$. Therefore, for all $\ell \in \underline{n}_0$, there exists a unique $x_\ell \in (0, \infty)$ with $h_1(x_\ell) = h_2(x_\ell)$. If $\beta_1 \leqslant x_\ell$ then $\|F - G_\ell\| \leqslant \frac{h_1(\beta_1)}{\beta_1^{\ell+1}}\beta_1^{\ell+1} \leqslant \frac{h_1(x_\ell)}{x_\ell^{\ell+1}}\beta_1^{\ell+1} = c_\ell\beta_1^{\ell+1}$. If $\beta_1 > x_\ell$, then $\|F - G_\ell\| \leqslant \frac{h_2(\beta_1)}{\beta_1^{\ell+1}}\beta_1^{\ell+1} \leqslant c_\ell\beta_1^{\ell+1}$. Hence, generally we have $\|F - G_\ell\| \leqslant c_\ell\beta_1^{\ell+1}$. In particular, $x_0 \in (0.128316, 0.128317)$, $x_1 \in (0.147522, 0.147523)$, $x_2 \in (0.189075, 0.189076)$, $x_3 \in (0.215065, 0.215066)$, $x_4 \in (0.226773, 0.226774)$, which implies the remaining part of the assertion. $\qquad\qquad\qquad\qquad\square$

Proof of Corollary 2.7: The proof follows arguments very similar to those used in the proofs of Theorems 1 and 2 in Roos (2007), where a comparable result was shown, generalizing (1.11) and (1.12). The idea here is a standard approximation procedure: In the first step, construct a new set of distributions $\widetilde{Q}_1, \ldots, \widetilde{Q}_n$ of the form used in Theorems 2.1 and 2.3, such that all the norms $\|Q_j - \widetilde{Q}_j\|$, $(j \in \underline{n})$ are small. This also leads to corresponding new (signed) measures $\widetilde{F}$ and $\widetilde{G}_\ell$. In the

second step, use the properties of the total variation distance to show that $\|F - \widetilde{F}\|$ and $\|\widetilde{G}_\ell - G_\ell\|$ are both small. Finally, use Theorems 2.1 and 2.3 to estimate $\|\widetilde{F} - \widetilde{G}_\ell\|$ and prove that the resulting bounds are close to the bounds in (2.4) and (2.5). We omit the details. □

Proof of Proposition 3.1: Under the assumptions of Section 3, let $\tau : S \longrightarrow \mathfrak{X}$, $x \mapsto \delta_x$. For arbitrary $B \in \mathcal{S}$, we then have $\pi_B \circ \tau = \mathbb{1}_B$ and $B = \tau^{-1}(\pi_B^{-1}(\{1\}))$, and hence $\{\tau^{-1}(A) \,|\, A \in \mathcal{A}\} = \mathcal{S}$. In particular, τ is $\mathcal{S}$-$\mathcal{A}$-measurable. Let $\mu = \nu^\tau$ be the image measure of ν under τ defined on $(\mathfrak{X}, \mathcal{A})$. For $B \in \mathcal{S}$, we have $\{\delta_x \,|\, x \in B\} = \pi_B^{-1}(\{1\}) \cap \pi_{S \backslash B}^{-1}(\{0\}) \in \mathcal{A}$ and $\mu(\{\delta_x \,|\, x \in B\}) = \nu(B)$. This shows that, since ν is σ-finite, this holds for μ as well. If $A \in \mathcal{A}$ with $\mu(A) = 0$, then $\nu(\tau^{-1}(A)) = 0$, and in turn $0 = P^{X_j}(\tau^{-1}(A)) = P((\tau \circ X_j)^{-1}(A)) = Q_j(A)$ and hence $Q_j \ll \mu$ for all $j \in \underline{n}$. Let $\widetilde{f}_j$ be a Radon-Nikodým density of Q_j with respect to μ and set $\widetilde{f} = \frac{1}{\lambda} \sum_{j=1}^n p_j \widetilde{f}_j$. As has been observed in Remark 2.8(c), for $j \in \underline{n}$, $f_j = \frac{\widetilde{f}_j}{\widetilde{f}} \mathbb{1}_{\{\widetilde{f}>0\}}$ is a Radon-Nikodým density of Q_j with respect to Q. From the above, we get that, for each $B \in \mathcal{S}$, a set $A \in \mathcal{A}$ exists such that $B = \{\tau \in A\}$ and hence

$$\int_B \widetilde{f}_j \circ \tau \, \mathrm{d}\nu = \int_A \widetilde{f}_j \, \mathrm{d}\mu = Q_j(A) = P^{X_j}(B) = \int_B \widetilde{h}_j \, \mathrm{d}\nu.$$

Therefore $\widetilde{f}_j \circ \tau = \widetilde{h}_j$, $\widetilde{f} \circ \tau = \widetilde{h}$ and $f_j \circ \tau = \frac{\widetilde{f}_j \circ \tau}{\widetilde{f} \circ \tau} \mathbb{1}_{\{\widetilde{f} \circ \tau > 0\}} = \frac{\widetilde{h}_j}{\widetilde{h}} \mathbb{1}_{\{\widetilde{h}>0\}}$ ν-almost everywhere. The assertion now follows from Corollary 2.7 and Remark 2.8(c) using that

$$\widetilde{\beta}_1 = \sum_{j=1}^n p_j^2 \int_{\{\widetilde{f}>0\}} \widetilde{f}_j \min\Big\{\frac{\widetilde{f}_j}{\lambda \widetilde{f}}, 1\Big\} \, \mathrm{d}\mu = \sum_{j=1}^n p_j^2 \int_{\{\widetilde{h}>0\}} \widetilde{h}_j \min\Big\{\frac{\widetilde{h}_j}{\lambda \widetilde{h}}, 1\Big\} \, \mathrm{d}\nu$$

and a similar calculation for $\widetilde{\alpha}_1$. □

Acknowledgment

References

T. V. Arak and A. Yu. Zaĭtsev. Uniform limit theorems for sums of independent random variables. *Trudy Mat. Inst. Steklov.* **174**, 214 (1986). ISSN 0371-9685. MR871856. English translation in *Proc. Steklov Inst. Math.* **1(174)**, viii+222 pp. (1988).

A. D. Barbour. Stein's method and Poisson process convergence. *J. Appl. Probab.* (Special Vol. 25A), 175–184 (1988). MR974580.

A. D. Barbour. Multivariate Poisson-binomial approximation using Stein's method. In *Stein's method and applications*, volume 5 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 131–142. Singapore Univ. Press, Singapore (2005). MR2205332.

A. D. Barbour and P. Hall. On the rate of Poisson convergence. *Math. Proc. Cambridge Philos. Soc.* **95** (3), 473–480 (1984). MR755837.

A. D. Barbour, L. Holst and S. Janson. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York (1992). ISBN 0-19-852235-5. MR1163825.

I. S. Borisov. A remark on a theorem of R. L. Dobrushin, and couplings in the Poisson approximation in Abelian groups. *Teor. Veroyatnost. i Primenen.* **48** (3), 576–583 (2003). MR2141351. English translation in *Theory Probab. Appl.* **48** (3), 521–528 (2004).

N. Bourbaki. *Algebra. II. Chapters 4–7.* Elements of Mathematics. Springer-Verlag, Berlin (1990). ISBN 3-540-19375-8. MR1080964.

V. Čekanavičius. On the smoothing properties of generalized Poisson distributions. *Liet. Mat. Rink.* **35** (2), 152–170 (1995). MR1368759. English translation in *Lithuanian Math. J.* **35** (2), 121–135 (1995).

V. Čekanavičius and B. Roos. An expansion in the exponent for compound binomial approximations. *Liet. Mat. Rink.* **46** (1), 67–110 (2006). MR2251442. Reprinted in *Lithuanian Math. J.* **46** (1), 54–91 (2006).

L. H. Y. Chen. An approximation theorem for convolutions of probability measures. *Ann. Probability* **3** (6), 992–999 (1975a). MR0383483.

L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability* **3** (3), 534–545 (1975b). MR0428387.

L. H. Y. Chen and M. Roos. Compound Poisson approximation for unbounded functions on a group, with application to large deviations. *Probab. Theory Related Fields* **103** (4), 515–528 (1995). MR1360203.

T. S. Chihara. *An introduction to orthogonal polynomials.* Gordon and Breach Science Publishers, New York-London-Paris (1978). ISBN 0-677-04150-0. MR0481884.

P. Deheuvels and D. Pfeifer. A semigroup approach to Poisson approximation. *Ann. Probability* **14** (2), 663–676 (1986). MR832029.

P. Deheuvels and D. Pfeifer. Poisson approximations of multinomial distributions and point processes. *J. Multivariate Anal.* **25** (1), 65–89 (1988). MR935295.

P. Franken. Approximation durch Poissonsche Prozesse. *Math. Nachr* **26**, 101–114 (1963). MR0160279.

H. Herrmann. Variationsabstand zwischen der Verteilung einer Summe unabhängiger nichtnegativer ganzzahliger Zufallsgrößen und Poissonschen Verteilungen. *Math. Nachr.* **29**, 265–289 (1965a). MR0190979.

U. Herrmann. Verallgemeinerung eines Satzes von Prochorow und Le Cam auf zufällige Vektoren. *Math. Nachr.* **29**, 17–24 (1965b). ISSN 0025-584X. MR0179820.

J. Hoffmann-Jørgensen. *Probability with a view toward statistics. Vol. I.* Chapman & Hall Probability Series. Chapman & Hall, New York (1994). ISBN 0-412-05221-0. MR1278485.

J. Kerstan. Verallgemeinerung eines Satzes von Prochorow und Le Cam. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **2**, 173–179 (1964). MR0165555.

J. Kruopis and V. Čekanavičius. Compound Poisson approximations for symmetric vectors. *J. Multivariate Anal.* **123**, 30–42 (2014). MR3130419.

L. Le Cam. An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.* **10**, 1181–1197 (1960). MR0142174.

L. Le Cam. On the distribution of sums of independent random variables. In *Proc. Internat. Res. Sem., Statist. Lab., Univ. California, Berkeley, Calif*, pages

179–202. Springer-Verlag, New York (1965). MR0199871.

F. Liese. Estimates of Hellinger integrals of infinitely divisible distributions. *Kybernetika (Prague)* **23** (3), 227–238 (1987). MR900332.

K. Matthes, J. Kerstan and J. Mecke. *Infinitely divisible point processes.* John Wiley & Sons, Chichester-New York-Brisbane (1978). ISBN 0-471-99460-X. MR0517931.

D. R. McDonald. On the Poisson approximation to the multinomial distribution. *Canad. J. Statist.* **8** (1), 115–118 (1980). MR595721.

R. Michel. An improved error bound for the compound Poisson approximation of a nearly homogeneous portfolio. *ASTIN Bulletin* **17** (2), 165–169 (1987). DOI: 10.2143/AST.17.2.2014971.

Yu. V. Prohorov. Asymptotic behavior of the binomial distribution. (Russian). *Uspehi Matem. Nauk (N.S.)* **8** (3), 135–142 (1953). MR0056861. English translation in *Select. Transl. Math. Statist. and Probability*, **1**, 87–95 (1961). Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I.

R.-D. Reiss. *A course on point processes.* Springer Series in Statistics. Springer-Verlag, New York (1993). ISBN 0-387-97924-7. MR1199815.

B. Roos. Metric multivariate Poisson approximation of the generalized multinomial distribution. *Teor. Veroyatnost. i Primenen.* **43** (2), 404–413 (1998). MR1679016. Reprinted in *Theory Probab. Appl.* **43** (2), 306–316 (1999).

B. Roos. Asymptotic and sharp bounds in the Poisson approximation to the Poisson-binomial distribution. *Bernoulli* **5** (6), 1021–1034 (1999a). MR1735783.

B. Roos. On the rate of multivariate Poisson convergence. *J. Multivariate Anal.* **69** (1), 120–134 (1999b). MR1701409.

B. Roos. Sharp constants in the Poisson approximation. *Statist. Probab. Lett.* **52** (2), 155–168 (2001). MR1841404.

B. Roos. Poisson approximation of multivariate Poisson mixtures. *J. Appl. Probab.* **40** (2), 376–390 (2003). MR1978097.

B. Roos. On variational bounds in the compound Poisson approximation of the individual risk model. *Insurance Math. Econom.* **40** (3), 403–414 (2007). MR2310979.

N. S. Upadhye and P. Vellaisamy. Compound Poisson approximation to convolutions of compound negative binomial variables. *Methodol. Comput. Appl. Probab.* **16** (4), 951–968 (2014). MR3270603.

H.-J. Witte. A unification of some approaches to Poisson approximation. *J. Appl. Probab.* **27** (3), 611–621 (1990). MR1067026.