Electronic Journal of Probability

Vol. 13 (2008), Paper no. 24, pages 756–776.

Journal URL
http://www.math.washington.edu/~ejpecp/

# Symmetric and Centered Binomial Approximation of Sums of Locally Dependent Random Variables

Adrian Röllin*

**Abstract**

Stein's method is used to approximate sums of discrete and locally dependent random variables by a centered and symmetric binomial distribution, serving as a natural alternative to the normal distribution in discrete settings. The bounds are given with respect to the total variation and a local limit metric. Under appropriate smoothness properties of the summands, the same order of accuracy as in the Berry-Essen Theorem is achieved. The approximation of the total number of points of a point processes is also considered. The results are applied to the exceedances of the $r$-scans process and to the Matérn hardcore point process type I to obtain explicit bounds with respect to the two metrics.

**Key words:** Stein's method; total variation metric; binomial distribution; local dependence.

**AMS 2000 Subject Classification:** Primary 60F05.

Submitted to EJP on February 12, 2007, final version accepted April 1, 2008.

*Departement of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, United Kingdom; roellin(at)stats.ox.ac.uk; supported by Swiss National Science Foundation projects 20-107935/1 and PBZH2-117033

# 1 Introduction

The approximation of sums of dependent random variables by the standard normal distribution has been investigated in a large variety of settings. The accuracy of approximation is most often measured by the Kolmogorov and Kantorovich metrics (see Sunklodas [29] for an overview with many references).

The use of stronger metrics typically requires that some 'smoothness' condition must be satisfied. In this paper, under the assumption of a general local dependence structure, we study the approximation of sums of discrete random variables by a symmetric and centered binomial distribution. In a discrete setting this distribution will serve us as a natural alternative to the normal distribution and allow for an approximation in total variation, which would always equal to 1 if the normal distribution were directly used to approximate integer valued random variables. Under some general smoothness property of the summands, the same order of accuracy as in the Berry-Essen Theorem can be achieved, but now, instead of the Kolmogorov metric, for the stronger total variation metric. We also examine another metric, from which local limit approximations can be obtained.

In the setting of independent summands, approximation by a centered Poisson distribution has been successfully adopted by Čekanavičius and Vaĭtkus [11] and Barbour and Čekanavičius [4]. However, for dependent summands, applications were limited to simple examples; first attempts were made by Barbour and Xia [5] and Čekanavičius and Vaĭtkus [11]. In contrast, the results in this paper are of general nature and allow a wide range of applications.

The proofs are based on Stein's method for distributional approximation. A main idea, introduced in Röllin [27], is to use interpolation functions to represent the Stein operator of a discrete distribution as the Stein operator of a continuous distribution. In the case of the binomial, this then allows the application of standard techniques in Stein's method for normal approximation. A careful analysis of the remainder terms then shows how a suitable smoothness condition can be exploited, to obtain total variation error bounds.

The paper is organized as follows. In the next section, we introduce the main technique in the simple case of independent summands. In Section 3, these results are extended to locally dependent summands and Section 4 shows their application to some simple examples. These applications are instances of $m$-dependent sequences and finite-range random fields for which a large amount of literature is available (see again Sunklodas [29] for references). We would like to point out the local limit CLTs and Edgeworth expansions in Götze and Hipp [18, 19]) and Heinrich [20, 21], which also cover the examples in Section 4. In Section 5 we prove some technical lemmas.

## 1.1 Notation

Denote by $\mathrm{Bi}(n,p)$ the binomial distribution with $n$ trials of probability $p$ each. Denote by $\widehat{\mathrm{Bi}}(n,p)$ the centered binomial distribution, i.e. a binomial distribution shifted by $-np$. Note that this distribution does not necessarily lie on the integers, but on a lattice of $\mathbb{R}$ with span 1.

Throughout the paper, we shall be concerned with two metrics for probability distributions, the total variation metric $d_{\mathrm{TV}}$ and the local limit metric $d_{\mathrm{loc}}$, where, for two probability distributions

$P$ and $Q$,

$$d_{\mathrm{TV}}\big(P,Q\big) := \sup_{A\subset\mathbb{R}}\big|P(A)-Q(A)\big|,$$
$$d_{\mathrm{loc}}\big(P,Q\big) := \sup_{x\in\mathbb{R}}\big|P\big([x,x+1)\big)-Q\big([x,x+1)\big)\big|.$$

For simplicity, we will often use the notation $d_l$, where $l=1$ will stand for $d_{\mathrm{TV}}$ and $l=2$ for $d_{\mathrm{loc}}$.

We denote by $\|\cdot\|$ the supremum norm if applied to functions, and the variation norm if applied to measures, where for any signed measure $M$, $\|M\| = \int |M|(dx)$ and thus $\|P-Q\| = 2\,d_{\mathrm{TV}}(P,Q)$. Let $\delta_x$ denote the unit mass at $x\in\mathbb{R}$, and $*$ the convolution of measures. Define for any measure $\mu$ and any $l\in\mathbb{N}:=\{1,2,\dots\}$

$$D^l(\mu) = \big\|\mu*(\delta_1-\delta_0)^{*l}\big\|.$$

Note that for measures $\mu$ and $\lambda$,

$$D^1\big(\mu\big) = 2\,d_{\mathrm{TV}}\big(\mu,\mu*\delta_1\big), \tag{1.1}$$
$$D^2(\mu*\lambda) \leqslant D^1(\mu)D^1(\lambda), \tag{1.2}$$

where the later inequality can be deduced from the fact that for signed measures $M_1$ and $M_2$, $\|M_1*M_2\|\leqslant\|M_1\|\cdot\|M_2\|$. Furthermore, define $\langle x\rangle := x-\lfloor x\rfloor$ to be the fractional part of $x\in\mathbb{R}$, and $(x)_+ = x\vee 0$.

## 1.2 Basic setup

Consider a sum of the form $W=\sum_{i\in J}\xi_i$, where $W$ takes its values in a lattice of $\mathbb{R}$ with span 1. The expectation of $W$ has no influence on the quality of the approximation, and we therefore assume without loss of generality that $\mathbb{E}W=0$; this can always be accomplished by subtracting the expectation from each individual summand. Each of the summands may now take its values on a different lattice; this, however, will result in no further complications.

To approximate $W$ by a centered binomial distribution, we have to choose $n$ in such a way that the variance of $\widehat{\mathrm{Bi}}(n,1/2)$ is as close to the variance of $W$ as possible. As $n$ has to be integer, this is only possible up to a rounding error. However, the symmetric and centered binomial distribution thus chosen will in general take its values on a different lattice from $W$ and the total variation distance will become 1. To circumvent this problem, we introduce an additional parameter $t$ and approximate $W$ by a centered binomial distribution with success probability $1/2-t$ instead ($t$ being small), to be able to match not only the variance but also the lattice.

Hence, to put the above in a rigorous form, we will make the following assumptions if not otherwise stated:

*Assumptions G:* Let $J$ be a finite set and let $\{\xi_i, i\in J\}$ be a collection of random variables with $\mathbb{E}\xi_i=0$ for all $i\in J$ and assume that there are numbers $\{a_i\in\mathbb{R}; i\in J\}$ such that almost surely $\xi_i\in\mathbb{Z}+a_i$. Let $W=\sum_{i\in J}\xi_i$; then $\mathbb{E}W=0$ and almost surely $W\in\mathbb{Z}+a$ for $a:=\sum_{i\in J}a_i$. Assume that $\sigma^2:=\operatorname{Var}W>1$. Define now $\delta:=\langle -4\sigma^2\rangle$ and $t:=\langle a+2\sigma^2+\delta/2\rangle/(4\sigma^2+\delta)$. Clearly, $4\sigma^2+\delta=\lceil 4\sigma^2\rceil$, and by definition the distribution $\widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil,1/2-t\big)$ has expectation 0; it is also easy to check that it takes values in $\mathbb{Z}+a$.

From the above definition, we see that $t$ is only of order $\mathrm{O}(\sigma^{-2})$ as $\sigma^2 \to \infty$, which is rather small in the setting that we are concerned with; Corollary 2.4 shows how to obtain results without $t$, using Lemma 5.2.

## 2 Sum of Independent Random Variables

First, we examine the case of independent discrete summands. Previous work on total variation approximation has been concerned with the compound Poisson distribution (see Le Cam [23] and Roos [28] and references therein), the signed compound Poisson distribution (see Čekanavičius [9] and references therein), the Poisson distribution (see Barbour et al. [7]), the centered Poisson distribution (see Čekanavičius [10], Čekanavičius and Vaĭtkus [11], Barbour and Xia [5] and Barbour and Čekanavičius [4]) and some more general distributions (see Brown and Xia [8]).

We present the theorem below to demonstrate the main technique in a simple setting, noting that it also follows as a consequence of Theorem 3.1.

**Theorem 2.1.** *Let $\{\xi_i; i \in J\}$ be independent and satisfy Assumptions G. Then, if the $\xi_i$ have finite third moments,*

$$d_l\big(\mathscr{L}(W), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil, 1/2 - t\big)\big) \leqslant \sigma^{-2}\Big(\sum_{i\in J} c_{l,i}\rho_i + 1.75\Big), \qquad l = 1, 2,$$

*where $\rho_i = \sigma_i^3 + \frac{1}{2}\mathbb{E}|\xi_i|^3$, $\sigma_i^2 = \operatorname{Var}\xi_i$ and $c_{l,i} = D^l\big(\mathscr{L}(W - \xi_i)\big)$.*

**Remark 2.2.** It is clear that the above bound is useful only if the $c_{l,i}$ are small. In the case of $n$ identically distributed random variables, we need $c_{1,i} = \mathrm{o}(1)$ as $n \to \infty$ for asymptotic approximation in total variation, and in order to deduce a local limit theorem we must have $c_{2,i} = \mathrm{o}(n^{-1/2})$. This is however always the case if $D^1(\xi_1) < 2$ (which corresponds to the usual condition in the local limit theorem that $\xi_1$ must not be concentrated on a lattice with span greater than 1; see e.g. Gnedenko [17]), as can be seen from (5.9)–(5.10), and we then even have $c_{l,i} = \mathrm{O}(n^{-l/2})$ for $l = 1, 2$.

Before proving the theorem, we start with a short summary of Stein's method for binomial approximation; for details see also Stein [31] and Ehm [16]. Denote by $F(M)$ the set of all real valued measurable functions on some given measure space $M$. A Stein operator $\mathcal{B} : F(\mathbb{Z}) \to F(\mathbb{Z})$ for the binomial distribution $\mathrm{Bi}(n, p)$ is characterized by the fact that, for any integer valued random variable $W$,

$$\mathbb{E}(\mathcal{B}g)(W) = 0 \text{ for all bounded } g \in F(\mathbb{Z}) \iff W \sim \mathrm{Bi}(n, p), \tag{2.1}$$

and a possible choice is

$$(\mathcal{B}g)(z) = qzg(z-1) - p(n-z)g(z), \qquad \text{for all } z \in \mathbb{Z}, \tag{2.2}$$

where, as usual, we put $q = 1 - p$.

Let $h \in F(\mathbb{Z})$ be a bounded function. Then, the solution $g = g_h$ to the Stein equation

$$(\mathcal{B}g)(z) = I[0 \leqslant z \leqslant n]\big\{h(z) - \mathbb{E}h(Y)\big\}, \qquad \text{for all } z \in \mathbb{Z}, \tag{2.3}$$

where $Y \sim \mathrm{Bi}(n,p)$, is also bounded. If the functions $h$ are of the form $h(z) = h_A(z) = I[z \in A]$, $A \subset \mathbb{Z}$, we have the uniform bound

$$\|\Delta g_A\| \leqslant \frac{1 - p^{n+1} - q^{n+1}}{(n+1)pq}, \tag{2.4}$$

where $\Delta g(z) := g(z+1) - g(z)$, and the same bound holds for $\|g_{\{b\}}\|$, $b \in \mathbb{Z}$; see Ehm [16]. Now, for all $z \in \mathbb{Z}$, we can write

$$I[z \in A] - \mathbb{P}[Y \in A] = (\mathcal{B}g_A)(z) + I[z \notin \{0 \dots n\}]\big(I[z \in A] - \mathbb{P}[Y \in A]\big),$$

and thus, for any integer valued random variable $V$,

$$\begin{aligned} d_{\mathrm{TV}}\big(\mathscr{L}(V), \mathrm{Bi}(n,p)\big) &= \sup_{A \subset \mathbb{Z}} \big|\mathbb{P}[V \in A] - \mathbb{P}[Y \in A]\big| \\ &\leqslant \sup_{A \subset \mathbb{Z}} \big|\mathbb{E}(\mathcal{B}g_A)(V)\big| + \mathbb{P}\big[|V - n/2| > n/2\big]. \end{aligned} \tag{2.5}$$

We now construct a Stein operator for the centered binomial distribution $\widehat{\mathrm{Bi}}(n,p)$ on the lattice $\mathbb{Z} - np$. For any function $g \in F(\mathbb{Z})$ define the function $\hat{g} \in F(\mathbb{Z} - np)$ by $\hat{g}(w) := g(w + np)$ for $w \in \mathbb{Z} - np$. Then the Stein operator is defined as

$$\begin{aligned} (\hat{\mathcal{B}}\hat{g})(w) &:= (\mathcal{B}g)(w + np) \\ &= p(w+np)g(w+np) + q(w+np)g(w-1+np) - npg(w+np) \\ &= w\big(p\hat{g}(w) + q\hat{g}(w-1)\big) - npq\Delta\hat{g}(w-1). \end{aligned} \tag{2.6}$$

for all $w \in \mathbb{Z} - np$. Thus, for $W = V - np$, an inequality corresponding to (2.5) holds, namely

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathscr{L}(W), \widehat{\mathrm{Bi}}(n,p)\big) \\ &\quad\leqslant \sup_{B \subset \mathbb{Z} - np} \big|\mathbb{E}(\hat{\mathcal{B}}\hat{g}_B)(W)\big| + \mathbb{P}\big[|W + n(p - 1/2)| > n/2\big]. \end{aligned} \tag{2.7}$$

An equivalent inequality holds for the $d_{\mathrm{loc}}$ metric, but the supremum is taken only over the sets $\{b\}$, $b \in \mathbb{Z} - np$.

Under the assumptions of the theorem, $n = \lceil 4\sigma^2 \rceil = 4\sigma^2 + \delta$ and $p = 1/2 - t$, and (2.6) becomes

$$(\hat{\mathcal{B}}\hat{g})(w) = w\Theta\hat{g}(w-1) - \sigma^2\Delta\hat{g}(w-1) + \big(t^2(4\sigma^2 + \delta) - wt - \delta/4\big)\Delta\hat{g}(w-1), \tag{2.8}$$

where $\Theta\hat{g}(w) := \frac{1}{2}\big(\hat{g}(w+1) + \hat{g}(w)\big)$. Since $\sigma^2 > 1$, the bound (2.4) simplifies to

$$\|\Delta\hat{g}_B\| \leqslant \frac{1}{\sigma^2}. \tag{2.9}$$

To see this, note that $t < 1/\lceil 4\sigma^2 \rceil = 1/n$ and $n = \lceil 4\sigma^2 \rceil \geqslant 5$. Then from (2.4) we have

$$\|\Delta\hat{g}_B\| \leqslant \frac{1}{(n+1)pq} = \frac{1}{(n+1)(1/4 - t^2)} \leqslant \frac{4n^2}{(n+1)(n^2-4)} \leqslant \frac{4}{n} \leqslant \frac{1}{\sigma^2}.$$

**Lemma 2.3.** *Assume the conditions of Theorem 2.1. Define* $\mathcal{A} : F(\mathbb{Z}+a) \to F(\mathbb{Z}+a)$ *by*

$$(\mathcal{A}\hat{g})(w) := w\Theta\hat{g}(w-1) - \sigma^2\Delta\hat{g}(w-1), \qquad w \in \mathbb{Z}+a,\ \hat{g} \in F(\mathbb{Z}+a).$$

*Then,*

$$\big|\mathbb{E}(\mathcal{A}\hat{g})(W)\big| \leqslant \Big(\|\Delta\hat{g}\| \sum_{i\in J} c_{1,i}\rho_i\Big) \wedge \Big(\|\hat{g}\| \sum_{i\in J} c_{2,i}\rho_i\Big). \tag{2.10}$$

*Proof.* For every $w \in \mathbb{Z}+a$ and $x \in [0,1)$ define

$$f(w+x) := \Theta\hat{g}(w-1) + x\Delta\hat{g}(w-1) + \tfrac{1}{2}x^2\Delta^2\hat{g}(w-1). \tag{2.11}$$

One easily checks that $f \in C^1$ and $f(w) = \Theta\hat{g}(w-1)$ and $f'(w) = \Delta\hat{g}(w-1)$, hence

$$(\mathcal{A}\hat{g})(w) = wf(w) - \sigma^2 f'(w), \tag{2.12}$$

for all $w \in \mathbb{Z}+a$. Furthermore, $f'$ is absolutely continuous, hence $f''$ exists almost everywhere. Choose $f''$ to be the function

$$f''(w+x) = \Delta^2\hat{g}(w-1) \tag{2.13}$$

for all $w \in \mathbb{Z}+a$, $0 \leqslant x < 1$.

We can now apply the usual Taylor expansion (cf. Reinert [25], Theorem 2.1), but with a refined estimate of the remainder terms. Write $W_i = W - \xi_i$, $i \in J$; then

$$\xi_i f(W) = \xi_i f(W_i) + \xi_i^2 f'(W_i) + \xi_i^3 \int_0^1 (1-s) f''(W_i + s\xi_i)\, ds,$$
$$\sigma_i^2 f'(W) = \sigma_i^2 f'(W_i) + \xi_i\sigma_i^2 \int_0^1 f''(W_i + s\xi_i)\, ds,$$

and hence, using the independence of $\xi_i$ and $W_i$ and that $\mathbb{E}\xi_i = 0$,

$$\begin{aligned}\big|\mathbb{E}\big\{\xi_i f(W) - \sigma_i^2 f'(W)\big\}\big| \leqslant \mathbb{E}\Big|&\xi_i^3 \int_0^1 (1-s)\mathbb{E}\big[f''(W_i + s\xi_i) \bigm| \xi_i\big]\, ds \\ &-\xi_i\sigma_i^2 \int_0^1 \mathbb{E}\big[f''(W_i + s\xi_i) \bigm| \xi_i\big]\, ds\Big|.\end{aligned} \tag{2.14}$$

Note now that for any real valued random variable $U$ taking values on a lattice with span 1, we obtain together with (2.13)

$$\big|\mathbb{E}\big(f''(U+z)\big)\big| \leqslant \Big(\|\Delta\hat{g}\| D^1\big(\mathscr{L}(U)\big)\Big) \wedge \Big(\|\hat{g}\| D^2\big(\mathscr{L}(U)\big)\Big), \tag{2.15}$$

for all $z \in \mathbb{R}$. Thus, from (2.14) and (2.15),

$$\begin{aligned}&\big|\mathbb{E}\big\{\xi_i f(W) - \sigma_i^2 f'(W)\big\}\big| \\ &\leqslant \Big(\|\Delta\hat{g}\| D^1\big(\mathscr{L}(W_i)\big)\big(\sigma_i^3 + \tfrac{1}{2}\mathbb{E}|\xi_i|^3\big)\Big) \wedge \Big(\|\hat{g}\| D^2\big(\mathscr{L}(W_i)\big)\big(\sigma_i^3 + \tfrac{1}{2}\mathbb{E}|\xi_i|^3\big)\Big).\end{aligned} \tag{2.16}$$

Now, using (2.12) we have

$$\big|\mathbb{E}\big\{\mathcal{A}\hat{g}(W)\big\}\big| \leqslant \sum_{i\in J}\big|\mathbb{E}\big\{\xi_i f(W) - \sigma_i^2 f'(W)\big\}\big|$$

and with (2.16) the lemma is proved. □

*Proof of Theorem 2.1.* Recall that, by Assumptions G, the distributions $\mathscr{L}(W)$ and $\widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil, 1/2-t\big)\big)$ are concentrated on the same lattice. Thus, using (2.7) and the form (2.8) of the Stein operator, and applying the left side of the minimum in (2.10) to the first part of (2.8) with the bound (2.9) gives

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathscr{L}(W), \widehat{\mathrm{Bi}}(4\sigma^2+\delta, 1/2-t)\big) \\ &\quad\leqslant \frac{\sum_{i\in J} c_{1,i}\rho_i}{\sigma^2} + \frac{t^2(4\sigma^2+\delta)+\sigma t+\delta/4}{\sigma^2} + \mathbb{P}\big[|W|\geqslant 2\sigma^2-1\big]. \end{aligned} \tag{2.17}$$

To bound the middle part of (2.17) note that $0\leqslant t<(4\sigma^2+\delta)^{-1}$ and $0\leqslant\delta<1$. Thus, recalling that $\sigma^2>1$, we obtain the simple bounds

$$t^2(4\sigma^2+\delta)<(4\sigma^2+\delta)^{-1}\leqslant 1/4, \qquad \sigma t\leqslant \sigma/(4\sigma^2+\delta)\leqslant 1/4, \qquad \delta/4\leqslant 1/4.$$

Applying Chebyshev's inequality on the last term of (2.17) we obtain

$$\mathbb{P}\big[|W|\geqslant 2\sigma^2-1\big]\leqslant \frac{\sigma^2}{(2\sigma^2-1)^2}\leqslant\frac{1}{\sigma^2}.$$

The $d_{\mathrm{loc}}$ case is analogous, using the right side of the minimum in (2.10) instead and the remark after (2.4). □

Note that in the next corollary we do not assume that the $\xi_i$ have expectation zero.

**Corollary 2.4.** *Let $W$ be the sum of independent and integer valued random variables $\{\xi_i, i\in J\}$ with $\sigma_i^2=\operatorname{Var}\xi_i$ and*

$$v_i=\min\big\{1/2, 1-d_{\mathrm{TV}}\big(\mathscr{L}(\xi_i),\mathscr{L}(\xi_i+1)\big)\big\}.$$

*Then, if $\sigma^2>1$,*

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathscr{L}(W), \mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2)*\delta_s\big) \\ &\qquad\leqslant \frac{2\sum\big(\sigma_i^3+\frac{1}{2}\mathbb{E}|\xi_i|^3\big)}{\sigma^2(V-v^*)^{1/2}} + \frac{1+2.25\sigma^{-1}+0.25\sigma^{-2}}{\sigma}, \\ &d_{\mathrm{loc}}\big(\mathscr{L}(W), \mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2)*\delta_s\big) \\ &\qquad\leqslant \frac{8\sum\big(\sigma_i^3+\frac{1}{2}\mathbb{E}|\xi_i|^3\big)}{\sigma^2(V-4v^*)_+} + \frac{3.25+0.25\sigma^{-1}}{\sigma^2}, \end{aligned}$$

*where $s:=\lceil\mu-\lceil 4\sigma^2\rceil/2\rceil$, $\mu=\mathbb{E}W$, $V=\sum_{i\in J}v_i$ and $v^*=\max_{i\in J}v_i$.*

*Proof.* Define $W_0=W-\mu$, and let $t$ be defined with respect to $W_0$, taking $a=-\mu$. Then, as the metrics $d_l$ are shift invariant,

$$\begin{aligned} d_l\big(\mathscr{L}(W), \mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2)*\delta_s\big) &= d_l\big(\mathscr{L}(W_0), \mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2)*\delta_{s-\mu}\big) \\ &\leqslant d_l\big(\mathscr{L}(W_0), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil,1/2-t\big)\big) + d_l\big(\mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2-t), \mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2)\big) \\ &=: R_1^l + R_2^l, \end{aligned}$$

since $\mathrm{Bi}(\lceil 4\sigma^2\rceil,1/2-t)*\delta_s*\delta_{-\mu}=\widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil,1/2-t\big)$.

Applying Lemma 5.2 to $R_2^l$ with the fact that $0 \leqslant t \leqslant (4\sigma^2+\delta)^{-1}$ gives

$$R_2^1 \leqslant \sigma^{-1}\big(1+(2\sigma)^{-1}+(4\sigma^2)^{-1}\big), \qquad R_2^2 \leqslant \sigma^{-2}\big(1.5+(4\sigma)^{-1}\big). \tag{2.18}$$

Define now $c_l = \max_{i\in J}\big\{D^l\big(\mathscr{L}(W-\xi_i)\big)\big\}$. Application of (5.9)-(5.10) yields

$$c_1 \leqslant \frac{2}{(V-v^*)^{1/2}}, \qquad c_2 \leqslant \frac{8}{(V-4v^*)_+}. \tag{2.19}$$

Thus, application of Theorem 2.1 to $R_1^l$ proves the corollary. □

# 3 Locally dependent random variables

In this section we present the main results of the paper. We exploit a finite local dependence structure as presented in Chen and Shao [12]. In the context of Stein's method for normal approximation, it has been successfully applied to a variety of problems; see for example Stein [31] (Lecture 10), Barbour et al. [6], Rinott and Rotar [26], Dembo and Rinott [15] and Barbour and Xia [3]. Note that Barbour et al. [6] use a slightly more general dependence structure, often yielding crucial improvement when approximating sums of dissociated random variables by the normal distribution. The generalization of Theorem 3.1 is straightforward, yet somewhat tedious, and we therefore use the simpler dependence structure of Chen and Shao [12]; see the Appendix for the more general version, but without proof. The dependence structure used by Rinott and Rotar [26] is more general in the sense that the set of neighbours is allowed be random itself; they obtain multivariate normal approximation in the uniform and more general metrics under the assumption of bounded summands.

Let $\{\xi_i; i\in J\}$ be a collection of random variables satisfying Assumptions G. For convenience, let $\xi_A$ denote $\{\xi_i; i\in A\}$ for every subset $A\subset J$. Assume further the following dependence structure: For every $i\in J$ there are subsets $A_i \subset B_i \subset J$ such that $\xi_i$ is independent of $\xi_{A_i^c}$, and $\xi_{A_i}$ is independent of $\xi_{B_i^c}$. Define $\eta_i = \sum_{j\in A_i}\xi_j$ and $\tau_i = \sum_{j\in B_i}\xi_j$.

**Theorem 3.1.** *With $W$ as above,*

$$d_l\big(\mathscr{L}(W), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil, 1/2-t\big)\big) \leqslant \sigma^{-2}\bigg(\sum_{i\in J}\vartheta_{l,i} + 1.75\bigg), \qquad l=1,2, \tag{3.1}$$

*where*

$$\begin{aligned}\vartheta_{l,i} = {}& \tfrac{1}{2}\mathbb{E}\big\{|\xi_i|\eta_i^2 D^l\big(\mathscr{L}(W|\xi_i,\eta_i)\big)\big\} + \mathbb{E}\big\{|\xi_i\eta_i(\tau_i-\eta_i)|D^l\big(\mathscr{L}(W|\xi_i,\eta_i,\tau_i)\big)\big\} \\ & + |\mathbb{E}\xi_i\eta_i|\mathbb{E}\big\{|\tau_i|D^l\big(\mathscr{L}(W|\tau_i)\big)\big\}\end{aligned} \tag{3.2}$$

*If further there are constants $c_{l,i}$ such that almost surely*

$$D^l\big(\mathscr{L}(W|\xi_{B_i})\big) \leqslant c_{l,i}, \tag{3.3}$$

*then*

$$\vartheta_{l,i} \leqslant c_{l,i}\big(\tfrac{1}{2}\mathbb{E}|\xi_i\eta_i^2| + \mathbb{E}|\xi_i\eta_i(\tau_i-\eta_i)| + |\mathbb{E}\xi_i\eta_i|\mathbb{E}|\tau_i|\big). \tag{3.4}$$

*Proof.* Estimate (3.4) is immediate. Following the proof of Theorem 2.1 and using Lemma 3.2 below, (3.1) is proved. □

Note that Theorem 2.1 follows from Theorem 3.1 with the choices $A_i = B_i = \{i\}$.

**Lemma 3.2.** *Assume the conditions of Theorem 3.1. Define $\mathcal{A} : F(\mathbb{Z}+a) \to F(\mathbb{Z}+a)$ as in Lemma 2.3. Then,*

$$\big|\mathbb{E}(\mathcal{A}\hat{g})(W)\big| \leqslant \Big(\|\Delta\hat{g}\| \sum_{i\in J} \vartheta_{1,i}\Big) \wedge \Big(\|\hat{g}\| \sum_{i\in J} \vartheta_{2,i}\Big). \tag{3.5}$$

*Proof.* We follow the proof of Lemma 2.3 right up to the end of the paragraph of (2.13). Note now that

$$\sigma^2 = \sum_{i\in J} \mathbb{E}\{\xi_i\eta_i\} \tag{3.6}$$

and that, by Taylor expansion, almost surely

$$\begin{aligned}
\xi_i f(W) &= \xi_i f(W-\eta_i) + \xi_i\eta_i f'(W-\eta_i) + \xi_i\eta_i^2 \int_0^1 (1-s) f''(W-\eta_i+s\eta_i)\, ds,\\
\xi_i\eta_i f'(W-\eta_i) &= \xi_i\eta_i f'(W-\tau_i) + \xi_i\eta_i(\tau_i-\eta_i) \int_0^1 f''(W-\tau_i+s(\tau_i-\eta_i))\, ds,\\
\mathbb{E}\{\xi_i\eta_i\} f'(W) &= \mathbb{E}\{\xi_i\eta_i\} f'(W-\tau_i) + \mathbb{E}\{\xi_i\eta_i\}\tau_i \int_0^1 f''(W-\tau_i+s\tau_i)\, ds.
\end{aligned}$$

Now, using the facts that $\mathbb{E}\xi_i = 0$, that $\xi_i$ is independent of $W-\eta_i$ and that $\eta_i$ is independent of $W-\tau_i$, we obtain from (3.6) and (3) that

$$\begin{aligned}
\mathbb{E}\big\{W f(W) - \sigma^2 f'(W)\big\} &= \sum_{i\in J} \mathbb{E}\big\{\xi_i f(W) - \mathbb{E}\{\xi_i\eta_i\} f'(W)\big\}\\
&= \sum_{i\in J} \mathbb{E}\Big\{ \xi_i\eta_i^2 \int_0^1 (1-s)\mathbb{E}\big\{ f''(W-\eta_i+s\eta_i) \;\big|\; \xi_i,\eta_i\big\}\, ds\\
&\qquad + \xi_i\eta_i(\tau_i-\eta_i) \int_0^1 \mathbb{E}\big\{ f''(W-\tau_i+s(\tau_i-\eta_i)) \;\big|\; \xi_i,\eta_i,\tau_i\big\}\, ds\\
&\qquad - \mathbb{E}\big\{\xi_i\eta_i\big\}\tau_i \int_0^1 \mathbb{E}\big\{ f''(W-\tau_i+s\tau_i)) \;\big|\; \tau_i\big\}\, ds\Big\}.
\end{aligned}$$

With (2.12) and (2.15) the lemma follows. □

We now give a point process version of Theorem 3.1, exploiting mainly the same dependence structure as before. Recall that, heuristically, the Palm process $\Phi_\alpha$ of a simple point process $\Phi$ on a set $J$ has the conditional distribution of $\Phi$ given that there is a point at $\alpha \in J$; for a rigorous definition see e.g. Kallenberg [22]. The Palm processes are characterized by the equality

$$\mathbb{E}\int f(\alpha,\Phi)\Phi(d\alpha) = \int \mathbb{E} f(\alpha,\Phi_\alpha)\mu(d\alpha),$$

for measurable functions $f$, where $\mu$ is the mean measure of $\Phi$; this equality will be used throughout the proof of the next theorem.

**Theorem 3.3.** *Let $\Phi$ be a point process on a Polish space $J$ with mean measure $\mu$ such that $\mu(J) < \infty$. For all points $\alpha \in J$, assume that there are measurable subsets $A_\alpha \subset B_\alpha \subset J$, such that for every $\alpha \in J$,*

$$\mathscr{L}\big(\Phi_\alpha(A_\alpha^c)\big) = \mathscr{L}\big(\Phi(A_\alpha^c)\big), \tag{3.7}$$

$$\Phi_\alpha(A_\alpha) \text{ and } \Phi_\alpha(B_\alpha^c) \text{ are independent,} \tag{3.8}$$

$$\Phi(A_\alpha) \text{ and } \Phi(B_\alpha^c) \text{ are independent,} \tag{3.9}$$

*where $\Phi_\alpha$ denotes the Palm process of $\Phi$ at point $\alpha$. Then, for $W = \Phi(J) - \mu(J)$ and if $\sigma^2 = \operatorname{Var} W > 1$,*

$$\begin{aligned} &d_l\big(\mathscr{L}(W), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil, 1/2 - t\big)\big) \\ &\qquad \leqslant \sigma^{-2} \int_{\alpha\in J} \vartheta_l(\alpha)\mu(d\alpha) + 1.75\sigma^{-2}, \qquad l = 1, 2. \end{aligned} \tag{3.10}$$

*where, with $\Phi' := \Phi - \mu$ and $\Phi'_\alpha := \Phi_\alpha - \mu$,*

$$\begin{aligned} \vartheta_l(\alpha) = {} & \big|\mathbb{E}\Phi'_\alpha(A_\alpha)\big|\mathbb{E}\big\{|\Phi'(B_\alpha)|D^l\big(\Phi(B_\alpha^c) \bigm| \Phi(B_\alpha)\big)\big\} \\ & + \tfrac{1}{2}\mathbb{E}\big\{\Phi'_\alpha(A_\alpha)^2 D^l\big(\Phi_\alpha(A_\alpha^c) \bigm| \Phi_\alpha(A_\alpha)\big)\big\} \\ & + \tfrac{1}{2}\mathbb{E}\big\{\Phi'(A_\alpha)^2 D^l\big(\Phi(A_\alpha^c) \bigm| \Phi(A_\alpha)\big)\big\} \\ & + \mathbb{E}\big\{|\Phi'_\alpha(A_\alpha)\Phi'_\alpha(B_\alpha \setminus A_\alpha)|D^l\big(\Phi_\alpha(B_\alpha^c) \bigm| \Phi_\alpha(A_\alpha), \Phi_\alpha(B_\alpha)\big)\big\} \\ & + \mathbb{E}\big\{|\Phi'(A_\alpha)\Phi'(B_\alpha \setminus A_\alpha)|D^l\big(\Phi(B_\alpha^c) \bigm| \Phi(A_\alpha), \Phi(B_\alpha)\big)\big\}. \end{aligned} \tag{3.11}$$

*Proof.* Following the proof of Theorem 2.1 and Lemma 2.3, it is clear that we only have to bound $\mathbb{E}\big\{Wf(W) - \sigma^2 f'(W)\big\}$ for $f$ defined as in (2.11).

In what follows, all integrals are taken over $\{\alpha \in J\}$ if not otherwise stated. Note first that, because of (3.7),

$$\sigma^2 = \mathbb{E}\{\Phi(J)\Phi'(J)\} = \int \mu(d\alpha)\mathbb{E}\{\Phi'_\alpha(A_\alpha) + \Phi'_\alpha(A_\alpha^c)\} = \int \mu(d\alpha)\mathbb{E}\Phi'_\alpha(A_\alpha)$$

and hence, by Taylor's expansion,

$$\begin{aligned} \sigma^2\mathbb{E}f'(W) = {} & \int \mu(d\alpha)\mathbb{E}\Phi'_\alpha(A_\alpha)\mathbb{E}f'(\Phi'(B_\alpha^c)) \\ & + \int \mu(d\alpha)\mathbb{E}\Phi'_\alpha(A_\alpha)\mathbb{E}\bigg\{\Phi'(B_\alpha)\int_0^1 f''(\Phi'(B_\alpha^c) + t\Phi'(B_\alpha))\,dt\bigg\} \\ =: {} & R_1 + R_2. \end{aligned}$$

Now, again by Taylor,

$$\begin{aligned}
\mathbb{E}\{Wf(W)\} &= \int \mu(d\alpha)\big[\mathbb{E}f(\Phi'_\alpha(J)) - \mathbb{E}f(\Phi'(J))\big] \\
&= \int \mu(d\alpha)\big[\mathbb{E}f(\Phi'_\alpha(A^c_\alpha)) - \mathbb{E}f(\Phi'(A^c_\alpha))\big] \\
&\quad + \int \mu(d\alpha)\Big[\mathbb{E}\big\{\Phi'_\alpha(A_\alpha)f'(\Phi'_\alpha(A^c_\alpha))\big\} - \mathbb{E}\big\{\Phi'(A_\alpha)f'(\Phi'(A^c_\alpha))\big\}\Big] \\
&\quad + \int \mu(d\alpha)\Big[\mathbb{E}\Big\{\Phi'_\alpha(A_\alpha)^2\int_0^1 (1-t)f''\big(\Phi'_\alpha(A^c_\alpha) + t\Phi'_\alpha(A_\alpha)\big)\,dt\Big\} \\
&\qquad\qquad - \mathbb{E}\Big\{\Phi'(A_\alpha)^2\int_0^1 (1-t)f''\big(\Phi'(A^c_\alpha) + t\Phi'(A_\alpha)\big)\,dt\Big\}\Big] \\
&=: R_3 + R_4 + R_5
\end{aligned}$$

and

$$\begin{aligned}
R_4 &= \int \mu(d\alpha)\big[\mathbb{E}\big\{\Phi'_\alpha(A_\alpha)f'(\Phi'_\alpha(B^c_\alpha))\big\} - \mathbb{E}\big\{\Phi'(A_\alpha)f'(\Phi'(B^c_\alpha))\big\}\big] \\
&\quad + \int \mu(d\alpha)\Big[\mathbb{E}\Big\{\Phi'_\alpha(A_\alpha)\Phi'_\alpha(B_\alpha\setminus A_\alpha)\int_0^1 f''\big(\Phi'_\alpha(B^c_\alpha) + t\Phi'_\alpha(B_\alpha\setminus A_\alpha)\big)\,dt\Big\} \\
&\qquad\qquad - \mathbb{E}\Big\{\Phi'(A_\alpha)\Phi'(B_\alpha\setminus A_\alpha)\int_0^1 f''\big(\Phi'(B^c_\alpha) + t\Phi'(B_\alpha\setminus A_\alpha)\big)\,dt\Big\}\Big] \\
&=: R_6 + R_7.
\end{aligned}$$

Using (3.7)–(3.9), we see that $R_3 = 0$ and $R_1 = R_6$, hence

$$\big|\mathbb{E}\big\{Wf(W) - \sigma^2 f'(W)\big\}\big| \leqslant |R_2| + |R_5| + |R_7|.$$

With (2.15) we finally obtain

$$\begin{aligned}
|R_2| &\leqslant \|\Delta\hat{g}\| \int \mu(d\alpha)\big|\mathbb{E}\Phi'_\alpha(A_\alpha)\big|\mathbb{E}\big\{|\Phi'(B_\alpha)|D^1\big[\mathscr{L}\big(\Phi(B^c_\alpha)\;\big|\;\Phi(B_\alpha)\big)\big]\big\}, \\
|R_5| &\leqslant \tfrac{1}{2}\|\Delta\hat{g}\| \int \mu(d\alpha)\Big[\mathbb{E}\big\{\Phi'_\alpha(A_\alpha)^2 D^1\big[\mathscr{L}\big(\Phi_\alpha(A^c_\alpha)\;\big|\;\Phi_\alpha(A_\alpha)\big)\big]\big\} \\
&\qquad\qquad + \mathbb{E}\big\{\Phi'(A_\alpha)^2 D^1\big[\mathscr{L}\big(\Phi(A^c_\alpha)\;\big|\;\Phi(A_\alpha)\big)\big]\big\}\Big], \\
|R_7| &\leqslant \|\Delta\hat{g}\| \int \Big[\mathbb{E}\big\{|\Phi'_\alpha(A_\alpha)\Phi'_\alpha(B_\alpha\setminus A_\alpha)|D^1\big[\mathscr{L}\big(\Phi_\alpha(B^c_\alpha)\;\big|\;\Phi_\alpha(A_\alpha),\Phi_\alpha(B_\alpha)\big)\big]\big\} \\
&\qquad + \mathbb{E}\big\{|\Phi'(A_\alpha)\Phi'(B_\alpha\setminus A_\alpha)|D^1\big[\mathscr{L}\big(\Phi(B^c_\alpha)\;\big|\;\Phi(A_\alpha),\Phi(B_\alpha)\big)\big]\big\}\Big]\mu(d\alpha).
\end{aligned}$$

To obtain $\vartheta_2$, just replace $\|\Delta g\|$ by $\|g\|$ and $D^1$ by $D^2$ in the above bounds. □

**Corollary 3.4.** *Let $\Phi$ be a simple point process satisfying (3.7)–(3.9). If there is further a function $c_l(\alpha)$, such that for $\mu$-almost every $\alpha \in J$ almost surely*

$$D^l\big(\mathscr{L}\big(\Phi(J)\;\big|\;\Phi|_{B_\alpha}\big)\big) \vee D^l\big(\mathscr{L}\big(\Phi_\alpha(J)\;\big|\;\Phi_\alpha|_{B_\alpha}\big)\big) \leqslant c_l(\alpha), \qquad l = 1, 2, \tag{3.12}$$

*then the $\vartheta_l$ from* (3.11) *satisfy*

$$\begin{aligned}\vartheta_l(\alpha) &\leqslant c_l(\alpha)\Big[\big|\mathbb{E}\Phi'_\alpha(A_\alpha)\big|\mathbb{E}|\Phi'(B_\alpha)| + \tfrac{1}{2}\mathbb{E}\Phi'_\alpha(A_\alpha)^2 + \tfrac{1}{2}\mathbb{E}\Phi'(A_\alpha)^2 \\ &\qquad + \mathbb{E}|\Phi'_\alpha(A_\alpha)\Phi'_\alpha(B_\alpha \setminus A_\alpha)| + \mathbb{E}|\Phi'(A_\alpha)\Phi'(B_\alpha \setminus A_\alpha)|\Big] \\ &\leqslant c_l(\alpha)\Big[1.5\mathbb{E}\big\{\Phi_\alpha(A_\alpha)\Phi_\alpha(B_\alpha)\big\} + 1.5\mathbb{E}\big\{\Phi(A_\alpha)\Phi(B_\alpha)\big\} \\ &\qquad + 6\mu(A_\alpha)\mu(B_\alpha) + 4\mu(B_\alpha)\mathbb{E}\Phi_\alpha(B_\alpha)\Big].\end{aligned} \tag{3.13}$$

# 4 Applications

In what follows, we calculate only rough bounds, leaving much scope for improvement. In particular, we replace the moments in the estimates by almost sure bounds. To estimate the terms describing the conditional smoothness, such as $D^l(W|\tau_i)$, we essentially use the a blocking argument and Remark 2.2, see also Subsections 5.3 and 5.4. Note that the following examples are cases of $m$-dependent random variables (in particular, finite range functionals of i.i.d. random variables), for which a large literature exists, see Sunklodas [29] for an excellent overview. We point out Kolmogorov distance approximations such as Stein [30] and Tikhomirov [32], and Edgeworth expansions and local limit CLTs such as Götze and Hipp [18, 19] and Heinrich [20, 21].

We provide approximations in both the total variation and the local limit metrics, where all constants can easily be extracted from the proofs and, hence, be made explicit.

## 4.1 Exceedances of the $r$-scans process

We follow the notation of Dembo and Karlin [14]. Let $X_1, X_2, \dots, X_{n+r-1}$ be independent and identically distributed random variables with distribution function $F$. Define the $r$-scan process $R_i = \sum_{k=0}^{r-1} X_{i+k}$, $i = 1, 2, \dots, n$ and further $W_i^- = I[R_i \leqslant a]$ for $a \in \mathbb{R}$. We are interested in the number $N^- = \sum_{i=1}^n W_i^-$, that is the number of $R_i$ not exceeding $a$. With $p = \mathbb{E}W_i^- = \mathbb{P}[R_1 \leqslant a]$, we have $\mathbb{E}N^- = np$ and

$$\sigma^2 = \operatorname{Var} W = np\Big(1 - p + 2\sum_{d=1}^{r-1}(1 - d/n)\psi(d)\Big), \tag{4.1}$$

where $\psi(d) := \mathbb{P}[R_{d+1} \leqslant a | R_1 < a] - p \geqslant 0$.

Poisson approximations for the $r$-scan process have been extensively studied by Dembo and Karlin [14]. Normal approximation has been considered by Dembo and Rinott [15]; in particular they show, that, for fixed $r$ and $a$, $N^-$ converges in the Kolmogorov metric to the normal distribution with rate $\mathrm{O}(n^{-1/2})$. In the next theorem we achieve the same rate in total variation, and also a rate for the corresponding local limit approximation.

**Theorem 4.1.** *Assume that $F$ is continuous, $F(0) = 0$, and $0 \leqslant F(x) < F(y)$ for all $x < y$, and let $a > 0$ be fixed. Then, for all $n$ such that $\sigma^2 > 1$,*

$$d_l\big(\mathscr{L}(N^- - np), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2\rceil, 1/2 - t\big)\big) \leqslant C_l n^{-l/2}, \qquad l = 1, 2,$$

*where the constants $C_1$ and $C_2$ are independent of $n$ and can be extracted from the proof.*

*Proof.* We apply Theorem 3.1 for $W = \sum_{i=1}^n \xi_i = \sum_{i=1}^n (W_i^- - p)$. We can set

$$A_i = \{i - r + 1, \dots, i + r - 1\} \cap \{1, \dots, n\},$$
$$B_i = \{i - 2r + 2, \dots, i + 2r - 2\} \cap \{1, \dots, n\}.$$

Then, as $|A_i| \leqslant 2r - 1$, $|B_i| \leqslant 4r - 3$ and $|B_i \setminus A_i| \leqslant 2r - 2$, the following rough bounds are obvious:

$$\mathbb{E}|\xi_i \eta_i^2| \leqslant (2r-1)^2, \quad \mathbb{E}|\xi_i \eta_i (\tau_i - \eta_i)| \leqslant (2r-1)(2r-2),$$
$$|\mathbb{E}\xi_i \eta_i| \mathbb{E}|\tau_i| \leqslant (2r-1)(4r-3),$$

thus

$$\vartheta_{l,i} \leqslant c_{l,i}\big(16r^2 - 20r + 6\big) \tag{4.2}$$

Consider now the block $B_1 = \sum_{i=1}^{3r-2} W_i^-$, and assume that the values $\partial B_1 = (X_1, \dots, X_{r-1})$ and $\partial B_2 = (X_{3r-1}, \dots, X_{4r-2})$ are given. Define the events

$$\mathcal{A} := \Big\{ \frac{a}{r} < X_r, \dots, X_{2r-2}, X_{2r+1}, \dots, X_{3r-2} \leqslant \frac{a(r+1)}{r^2},\, 0 < X_{2r} \leqslant \frac{a}{2r^2} \Big\}$$
$$\mathcal{A}_0 := \Big\{ \frac{a}{r} < X_{2r-1} \leqslant \frac{a(r+1)}{r^2} \Big\}, \quad \mathcal{A}_1 := \Big\{ 0 < X_{2r-1} \leqslant \frac{a}{2r^2} \Big\}.$$

Due to the conditions on $F$ and independence it is clear that $p_j := \mathbb{P}[\mathcal{A} \cap \mathcal{A}_j] > 0$ for $j = 0, 1$. Note now that

$$R_r = \sum_{i=r}^{2r-1} X_i > a \quad \text{on } \mathcal{A} \cap \mathcal{A}_0, \qquad R_r < a \quad \text{on } \mathcal{A} \cap \mathcal{A}_1.$$

Note further that $R_s < a$ for all $s = r+1, \dots, 2r-1$ on $\mathcal{A} \cap (\mathcal{A}_0 \cup \mathcal{A}_1)$. Hence

$$\sum_{i=r}^{2r-1} W_i^- = r - 1 \quad \text{on } \mathcal{A} \cap \mathcal{A}_0, \qquad \sum_{i=r}^{2r-1} W_i^- = r \quad \text{on } \mathcal{A} \cap \mathcal{A}_1.$$

It easy to see now by a coupling argument that

$$\tfrac{1}{2} D^1\big(\mathscr{L}(B_1)\big) \leqslant 1 - (p_0 \wedge p_1) < 1.$$

Noting that by sequentially stringing together blocks like $B_1$, we can have $m := \lfloor n/(3r-2) \rfloor$ such blocks, which are independent given all the borders $\partial B_i$. Furthermore, for every $i$, the $R_j$ in $B_i$ depend on the $X_k$ of at most two such blocks. Therefore, defining $Z = (\partial B_1, \dots, \partial B_m)$ and using (5.11) and (5.12),

$$D^1\big(\mathscr{L}(W | \partial B_i, i = 1, \dots, m)\big) \leqslant \frac{2}{\big(\min\{1/2, p_0, p_1\}(m-2)\big)^{1/2}} =: c_{1,i},$$
$$D^2\big(\mathscr{L}(W | \partial B_i, i = 1, \dots, m)\big) \leqslant \frac{8}{\min\{1/2, p_0, p_1\}(m-4)_+} =: c_{2,i}.$$

Clearly, $c_{l,i} = \mathrm{O}(n^{-l/2})$. Hence, putting this, (4.1) and (4.2) into (3.1), the theorem follows. □

## 4.2 Matérn hard-core process type I

We approximate the total number of points of the Matérn hard-core process type I introduced by Matérn [24]. We use rectangular instead of the usual circular neighborhoods. Let $\Phi$ be the process on the $d$-dimensional cube $J = [0,1)^d \subset \mathbb{R}^d$ defined as

$$\Phi(B) = \sum_{i=1}^{\tau} I[X_i \in B] I[X_j \notin K_r(X_i) \text{ for all } j = 1, \dots, \tau,\ j \neq i],$$

where $\tau \sim \mathrm{Po}(\lambda)$ and $\{X_i;\ i \in \mathbb{N}\}$ is a sequence of independent and uniformly distributed random variables on $J$ and where, for $x = (x_1, \dots, x_d) \in J$ and $r > 0$, $K_r(x)$ denotes the $d$-dimensional closed cube with center $x$ and side length $r$. To avoid edge effects, we treat $J$ as a $d$-dimensional torus, thus identifying any point outside $J$ by the point in $J$ which results in coordinate-wise shifting by 1. The process $\Phi$ is thus a thinned Poisson point process with rate $\lambda$ having all points deleted which contain another point in their $K_r$ neighborhood. For the mean measure $\mu$ of $\Phi$ we obtain

$$\frac{d\mu(x)}{dx} = \lambda e^{-\lambda r^d}. \tag{4.3}$$

We are now interested in the distribution of $\Phi(B)$ when $r$ is small an $\lambda$ large.

**Theorem 4.2.** *Put $W := \Phi(J) - \mu(J)$ and let $a > 0$ be a fixed real number. Then, for every $\lambda$ and $r$ such that $\lambda r^d = a$ and $\sigma^2 := \operatorname{Var} W > 1$,*

$$d_l\big(\mathscr{L}(\Phi(J) - \mu(J)), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2 \rceil, 1/2 - t\big)\big) \leqslant C_l \lambda^{-l/2}, \qquad l = 1, 2,$$

*for constants $C_1$ and $C_2$ which are independent of $\lambda$ and can be extracted from the proof.*

*Proof.* We apply Corollary 3.4. We can take $A_x = K_{2r}(x)$ and $B_x = K_{4r}(x)$ and check that the conditions (3.7)–(3.9) are fulfilled. Some calculations show that the reduced second factorial moment measure $M$ satisfies

$$\frac{dM(x)}{dx} = \begin{cases} 0 & \text{if } x \in K_r(0), \\ \lambda^2 e^{-\lambda |K_r(0) \cup K_r(x)|} & \text{if } x \in K_{2r}(0) \setminus K_r(0), \\ \lambda^2 e^{-2a} & \text{if } x \notin K_{2r}(0), \end{cases}$$

compare with Daley and Vere-Jones [13], pp. 367, 373. Thus, $M(J) \geqslant \lambda^2 e^{-2a}(1 - r^d)$ and

$$\sigma^2 = \lambda e^{-a} + M(J) - \mu(J)^2 \geqslant \lambda e^{-a}(1 - a e^{-a}). \tag{4.4}$$

Since we can have at most $7^d$ points of $\Phi$ in $B_x$, we obtain from (3.13) the rough estimate

$$\vartheta_l(x) \leqslant 26 \cdot 7^d c_l(x), \tag{4.5}$$

where $c_l(\cdot)$ is as in (3.12). To estimate $c_l(x)$ write $K_r = K_r(0)$. We have

$$\mathbb{P}\big[\Phi(K_r) = 0 \,\big|\, \Phi_{K_{7r} \setminus K_{5r}}\big] \geqslant \mathrm{Po}\big(\lambda |K_r|\big)\{0\} = e^{-a} =: p_0,$$
$$\mathbb{P}\big[\Phi(K_r) = 1 \,\big|\, \Phi_{K_{7r} \setminus K_{5r}}\big] \geqslant \mathrm{Po}\big(\lambda |K_{3r} \setminus K_r|\big)\{0\} \cdot \mathrm{Po}\big(\lambda |K_r|\big)\{1\} = a e^{-3^d a} =: p_1.$$

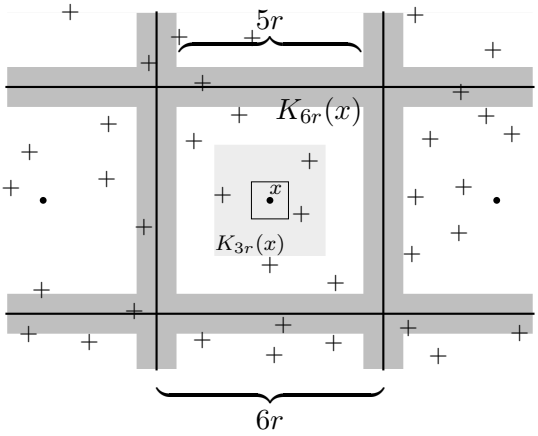


Figure 1: Matérn hard-core process type I: Given that the process $\Phi$ is known on the borders $\cup_{l\in\mathcal{M}} K_{6r}(x_l) \setminus K_{5r}(x_l)$ (grey area), the boxes $\Phi|_{K_{6r}(x_l)}$, $l \in \mathcal{M}$, are independent.

Hence, by a coupling argument,

$$\tfrac{1}{2} D^1\big(\mathscr{L}(\Phi(K_{6r})|\Phi_{K_{7r}\setminus K_{5r}})\big) \leqslant 1 - (p_0 \wedge p_1) < 1. \tag{4.6}$$

Let now $x$ be arbitrary. Divide the space $J$ into boxes of side length $6r$, centered around $x$ (see Figure 1). With $m := \lfloor 1/(6r) \rfloor$, we can have $m^d$ such boxes plus a remainder. Denote this remainder by $J^R$ and denote by $x_l$, $l \in \{1,\dots,m\}^d =: \mathcal{M}$ the centers of the boxes where $x_{1,\dots,1} = x$. Note now that, given $\Phi$ on all the borders $K_{6r}(x_l)\setminus K_{5r}(x_l)$, $l \in \mathcal{M}$ (grey area in Figure 1), the random variables $\Phi(K_{6r}(x_l))$, $l \in \mathcal{M}$, are independent and satisfy Inequality (4.6). Furthermore, $\Phi|_{J\setminus K_{6r}(x)}$ is independent of $\Phi|_{B_x}$, and therefore, defining $Z = \big((\Phi|_{K_{6r}(x_l)\setminus K_{5r}(x_l)})_{l\in\mathcal{M}}, \Phi|_{J^R}\big)$ and using (5.11) and (5.12), we obtain

$$D^1\big(\mathscr{L}\big(\Phi(J) \;\big|\; \Phi|_{B_x}\big)\big) \leqslant \frac{2}{\big(\min\{1/2, p_0, p_1\}(m^d-2)_+\big)^{1/2}} =: c_1(x), \tag{4.7}$$

$$D^2\big(\mathscr{L}\big(\Phi(J) \;\big|\; \Phi|_{B_x}\big)\big) \leqslant \frac{8}{\min\{1/2, p_0, p_1\}(m^d-3)_+} =: c_2(x). \tag{4.8}$$

Noting that almost surely $\mathscr{L}\big[\Phi_x\big(J\setminus K_{6r}(x)\big) \;\big|\; \Phi_x|_{B_x}\big] = \mathscr{L}\big[\Phi\big(J\setminus K_{6r}(x)\big)\big]$, we see that (4.7) and (4.8) hold also for $\Phi_x$, thus $c_l(x)$ satisfies (3.12). Now, recalling that $a = \lambda r^d$ is constant, we have $c_l(x) = \mathrm{O}(\lambda^{-l/2})$. Hence, putting this and (4.3)–(4.5) into (3.10), the theorem follows. □

# 5 Appendix

## 5.1 Properties of the solution to the Stein equation

**Lemma 5.1.** *For any indicator function $h(z) = I[z \in A]$, $z \in \mathbb{Z}$, $A \subset \mathbb{Z}$, the solution $g = g_h$ to the Stein Equation (2.3) satisfies*

$$\|g\| \leqslant 1 \wedge (npq)^{-1/2}. \tag{5.1}$$

*Proof.* We apply the generator method introduced by Barbour [1]. For any function $f : \{0, \dots, n\} \to \mathbb{R}$, define

$$(\mathcal{A}f)(z) = \big(\mathcal{B}(-\Delta f)\big)(z) = qzf(z-1) - \big(qz + p(n-z)\big)f(z) + p(n-z)f(z+1), \tag{5.2}$$

which is the infinitesimal generator of a pure jump Markov process. A solution $g$ to (2.3) is now given through

$$\psi(z) = -\int_0^\infty \mathbb{E}\big\{h\big(Y_z(t)\big) - h(Y)\big\}\, dt, \qquad \text{for } z \in \{0, \dots, n\},$$

and $g(z) = -\Delta\psi(z)$ for $z \in \{0, \dots, n-1\}$ and $g(z) = 0$ else, where $Y_z$ is a Markov process with generator $\mathcal{A}$ starting at point $z$, and $Y$ is a random variable having the stationary distribution $\mathrm{Bi}(n, p)$. Now, we have for $z \in \{0, \dots, n-1\}$,

$$\Delta\psi(z) = \int_0^\infty \mathbb{E}\big\{h\big(Y_z(t)\big) - h\big(Y_{z+1}(t)\big)\big\}\, dt. \tag{5.3}$$

We now fix $z$ and construct a coupling of $Y_z$ and $Y_{z+1}$ to bound (5.3). Let thereto $X_k^{(i)}(t)$, $k \in \{1, \dots, n\}$, $i \in \{0, 1\}$, be independent Markov processes with state space $\{0, 1\}$, starting in point $i$ and having jump rate $p$ if the process is in 0 and $q$ otherwise. It is easy to see by the Kolmogorov differential equations that

$$X_k^{(1)}(t) \sim \mathrm{Be}(p + qe^{-t}), \qquad X_k^{(0)}(t) \sim \mathrm{Be}(p - pe^{-t}) \tag{5.4}$$

where $\mathrm{Be}(p)$ denotes the Bernoulli distribution with success probability $p$. Let $\tau$ be the minimum of the first jump times of the two processes $X_{z+1}^{(0)}$ and $X_{z+1}^{(1)}$, and define a new process

$$X(t) = \begin{cases} X_{z+1}^{(1)} & \text{if } \tau > t, \\ X_{z+1}^{(0)} & \text{if } \tau \leqslant t, \end{cases}$$

describing the well-known Doeblin coupling. Then, let

$$Y_z = \sum_{k=1}^{z} X_k^{(1)} + \sum_{k=z+1}^{n} X_k^{(0)}, \qquad Y_{z+1} = Y_z - X_{z+1}^{(0)} + X(t), \tag{5.5}$$

and one proves that $Y_z$ and $Y_{z+1}$ are Markov processes with generator (5.2). Hence, we can write (5.3) as

$$-\Delta\psi(z) = \int_0^\infty e^{-t}\mathbb{E}\big\{\Delta h(Y_z)\big\}\, dt, \tag{5.6}$$

since $\tau$ is exponentially distributed with rate 1. The bound $\|g\| \leqslant 1$ is now immediate from (5.6), thus we may assume that $npq > 1$. Note that, from (5.4) and (5.5),

$$\mathscr{L}(Y_z) = \mathrm{Bi}(z, p + qe^{-t}) * \mathrm{Bi}(n - z, p - pe^{-t}),$$

and hence, from Barbour and Jensen [2], Lemma 1,

$$\begin{aligned} D^1\big(\mathscr{L}(Y_z)\big) &\leqslant \mathrm{Var}(Y_k)^{-1/2} \\ &\leqslant \big(z(p+qe^{-t})(q-qe^{-t}) + (n-z)(p-pe^{-t})(q+pe^{-t})\big)^{-1/2} \\ &\leqslant \big(npq(1-e^{-t})\big)^{-1/2}. \end{aligned} \tag{5.7}$$

Note also that for $\tilde{h} := h - 1/2$

$$\big|\mathbb{E}\{\Delta h(Y_z)\}\big| = \big|\mathbb{E}\{\Delta\tilde{h}(Y_z)\}\big| \leqslant D^1\big(\mathscr{L}(Y_z)\big)/2. \tag{5.8}$$

Thus, applying (5.8) on (5.6) and using (5.7),

$$\big|\Delta\psi\big| \leqslant \int_0^s e^{-t}\,dt + \frac{1}{2\sqrt{npq}} \int_s^\infty \frac{e^{-t}}{\sqrt{1-e^{-t}}}\,dt.$$

Choosing $s = -\ln\big(1 - (npq)^{-1}\big)$ and computing the integrals proves the lemma. □

## 5.2 Change of the success probabilities

**Lemma 5.2.** *For every $n \in \mathbb{N}$, $0 < p < 1$ and $-(1-p) < t < p$*

$$\begin{aligned} d_{\mathrm{TV}}\big(\mathrm{Bi}(n, p-t), \mathrm{Bi}(n,p)\big) &\leqslant |t|\left(\frac{\sqrt{n}}{\sqrt{pq}} + \frac{p-t}{pq} + \frac{\sqrt{(p-t)(q+t)}}{pq\sqrt{n}}\right) \\ d_{\mathrm{loc}}\big(\mathrm{Bi}(n, p-t), \mathrm{Bi}(n,p)\big) &\leqslant |t|\left(\frac{1+p-t}{pq} + \frac{\sqrt{(p-t)(q+t)}}{pq\sqrt{n}}\right) \end{aligned}$$

*Proof.* We use Stein's method. If $W \sim \mathrm{Bi}(n, p-t)$, we obtain from (2.1) and (2.2)

$$\mathbb{E}\big\{(1-p)Wg(W-1) - p(n-W)g(W)\big\} = \mathbb{E}\big\{tW\Delta g(W-1) - tng(W)\big\}$$

for every bounded function $g \in F(\mathbb{Z})$. The left side is just the Stein operator for $\mathrm{Bi}(n,p)$ hence, taking $g = g_A$ obtained by solving (2.3) for $\mathrm{Bi}(n,p)$, with the bounds (2.4) and (5.1) the $d_{\mathrm{TV}}$-bound follows, noting also that $\mathbb{E}|W| \leqslant |\mathbb{E}W| + \sqrt{\mathrm{Var}\,W}$. With the remark after (2.4), the $d_{\mathrm{loc}}$-bound is proved. □

## 5.3 Smoothing properties of independent random variables

In several parts of this article we need to estimate $D^m(U)$, $m = 1, 2$, for some integer valued random variable $U$, being a sum of some other random variables. If the $U$ is a sum of independent random variables, we can proceed as follows. Assume that $U = \sum_{i=1}^n X_i$, where the $X_i$ are

independent. Defining $v_i = \min\{\frac{1}{2}, 1 - \frac{1}{2}D^1(X_i)\}$ and $V = \sum_i v_i$ we obtain from Barbour and Xia [5], Proposition 4.6 the bound

$$D^1(U) \leqslant \frac{2}{V^{1/2}}. \tag{5.9}$$

Define further $v^* = \max_i v_i$. Now it is always possible to write $U = U^{(1)} + U^{(2)}$ in such a way that the analogously defined numbers $V^{(1)}$ and $V^{(2)}$ satisfy $V^{(k)} \geqslant V/2 - v^*$, $k = 1, 2$. Using (1.2) and (5.9), we obtain

$$D^2(U) \leqslant D^1\big(U^{(1)}\big)D^1\big(U^{(2)}\big) \leqslant \frac{4}{\big(V^{(1)}V^{(2)}\big)^{1/2}} \leqslant \frac{8}{(V - 2v^*)_+}. \tag{5.10}$$

### 5.4 Smoothing properties of conditional independent random variables

In most applications, $U$ is a sum of dependent summands and we can not apply (5.9) and (5.10) directly. However, assuming that there is a random variable $Z$ on the same probability space as $U$ such that $\mathscr{L}(U|Z = z)$ can be represented as a sum of independend summands, say $X_i^{(z)}$, $i = 1, \dots, n_z$, for each $z$ that $Z$ can attain, we can still apply (5.9) and (5.10), and we obtain

$$D^1(U) \leqslant \mathbb{E}\big\{\mathbb{E}[D^1(U)|Z]\big\} \leqslant \mathbb{E}\Big\{\frac{2}{V_Z^{1/2}}\Big\}, \tag{5.11}$$

$$D^2(U) \leqslant \mathbb{E}\big\{\mathbb{E}[D^1(U)|Z]\big\} \leqslant \mathbb{E}\Big\{\frac{8}{(V_Z - 2v_Z^*)_+}\Big\}, \tag{5.12}$$

where, for each $z$, $V_z$ and $v_z^*$ are the corresponding values as defined in Subsection 5.3 with respect to the $X_i^{(z)}$.

### 5.5 Generalization of the local dependence structure

We now give a generalization of Theorem 3.1. The proof is omitted, because it runs analogously to the proof of Theorem 3.1; see also Barbour et al. [6].

Suppose that a random variable $W$ satisfies Assumptions G and assume that there are sets $K_i \subset J$, $i \in I$, and square integrable random variables $Z_i$, $Z_{ik}$ and $V_{ik}$, $k \in K_i$ and $i \in I$, as follows:

$$W = W_i + Z_i,\ i \in I, \text{ where } W_i \text{ is independent of } \xi_i, \tag{5.13}$$

$$Z_i = \sum_{k \in K_i} Z_{ik}, \tag{5.14}$$

$$W_i = W_{ik} + V_{ik},\ i \in I,\ k \in K_i, \tag{5.15}$$

where $W_{ik}$ is independent of the pair $(X_i, Z_{ik})$.

**Theorem 5.3.** *With $W$ as above,*

$$d_l\big(\mathscr{L}(W), \widehat{\mathrm{Bi}}\big(\lceil 4\sigma^2 \rceil, 1/2 - t\big)\big) \leqslant \sigma^{-2}\Big(\sum_{i \in I} \vartheta_{l,i} + 1.75\Big), \qquad l = 1, 2, \tag{5.16}$$

*where*

$$\begin{aligned}\vartheta_{l,i} = \frac{1}{2}\mathbb{E}\big\{|\xi_i| Z_i^2\, D^l\big(\mathscr{L}(W_i|\xi_i, Z_i)\big)\big\} \\ + \sum_{k\in K_i} \mathbb{E}\big\{\big|\xi_i Z_{ik} V_{ik}\big| D^l\big(\mathscr{L}(W_{ik}|\xi_i, Z_{ik}, V_{ik})\big)\big\} \\ + \sum_{k\in K_i} \big|\mathbb{E}\{\xi_i Z_{ik}\}\big|\, \mathbb{E}\big\{|Z_i + V_{ik}| D^l\big(\mathscr{L}(W_{ik}|Z_i, V_{ik})\big)\big\}.\end{aligned} \tag{5.17}$$

## Acknowledgments

I thank Andrew Barbour, Dominic Schuhmacher and Bruno Nietlispach for many helpful discussions. I also thank Peter Eichelsbacher for pointing out some mistakes in the manuscript and the anonymous referee for helpful references.

## References

[1] A. D. Barbour (1988). Stein's method and Poisson process convergence. A celebration of applied probability. *J. Appl. Probab.* **25A**, 175–184. MR0974580

[2] A. D. Barbour and J. L. Jensen (1989). Local and tail approximation near the Poisson limit. *Scand. J. Statist.* **16**, 75–87. MR1003970

[3] A. D. Barbour and A. Xia (2001). The number of two-dimensional maxima. *Adv. in Appl. Probab.* **33**, 727–750. MR1875775

[4] A. D. Barbour and V. Čekanavičius (2002). Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.* **30**, 509–545. MR1905850

[5] A. D. Barbour and A. Xia (1999). Poisson perturbations. *ESAIM Probab. Statist.* **3**, 131–159. MR1716120

[6] A. D. Barbour, M. Karoński, and A. Ruciński (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47**, 125–145. MR1047781

[7] A. D. Barbour, L. Holst, and S. Janson (1992). *Poisson approximation.* Oxford University Press. MR1163825

[8] T. C. Brown and A. Xia (2001). Stein's method and birth-death processes. *Ann. Probab.* **29**, 1373–1403. MR1872746

[9] V. Čekanavičius (1997). Asymptotic expansions in the exponent: a compound Poisson approach. *Adv. in Appl. Probab.* **29**, 374–387. MR1450935

[10] V. Čekanavičius (1998). Poisson approximations for sequences of random variables. *Statist. Probab. Lett.* **39**, 101–107. MR1652516

[11] V. Čekanavičius and P. Vaĭtkus (2001). Centered Poisson approximation by the Stein method. *Lithuanian Math. J.* **41**, 319–329. MR1903485

[12] L. H. Y. Chen and Q.-M. Shao (2004). Normal approximation under local dependence. *Ann. Probab.* **32**, 1985–2028. MR2073183

[13] D. J. Daley and D. Vere-Jones (1988). *An introduction to the theory of point processes.* Springer Series in Statistics. Springer-Verlag, New York. MR0950166

[14] A. Dembo and S. Karlin (1992). Poisson approximations for $r$-scan processes. *Ann. Appl. Probab.* **2**, 329–357. MR1161058

[15] A. Dembo and Y. Rinott (1996). Some examples of normal approximations by Stein's method. In *Random Discrete Structures*, *IMA Vol. Math. Appl.* **76**, 25–44. Springer, New York. MR1395606

[16] W. Ehm (1991). Binomial approximation to the Possion-binomial distribution. *Statist. Probab. Lett.* **11**, 7–16, 1991. MR1093412

[17] B. V. Gnedenko (1948). On a local limit theorem of the theory of probability *Uspehi Matem. Nauk (N. S.)* **3**, 187–194. MR0026275

[18] F. Götze and C. Hipp (1989). Asymptotic expansions for potential functions of i.i.d. random fields. *Probab. Theory Related Fields* **82**, 349–370. MR1001518

[19] F. Götze and C. Hipp (1990). Local limit theorems for sums of finite range potentials of a Gibbsian random field. *Ann. Probab.* **18**, 810–828. MR1055434

[20] L. Heinrich (1987). Asymptotic expansions in the central limit theorem for a special class of $m$-dependent random fields. I. *Math. Nachr.* **134**, 83–106. MR0918670

[21] L. Heinrich (1990). Asymptotic expansions in the central limit theorem for a special class of $m$-dependent random fields. II. Lattice case. *Math. Nachr.* **145**, 309–327. MR1069038

[22] O. Kallenberg (1986). *Random measures.* 4th edition, Akademie Verlag, Berlin. MR0854102

[23] L. Le Cam (1965). On the distribution of sums of independent random variables. In *Proc. Internat. Res. Sem., Statist. Lab., Univ. California, Berkeley, Calif.*, 179–202. Springer-Verlag, New York. MR0199871

[24] B. Matérn (1960). Spatial variation: Stochastic models and their application to some problems in forest surveys and other sampling investigations. *Meddelanden Fran Statens Skogsforskningsinstitut*, Band 49, Nr. 5. Stockholm. MR0169346

[25] G. Reinert (1998). Couplings for normal approximations with Stein's method. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **41**, 193–207. MR1630415

[26] Y. Rinott and V. Rotar (1996). A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56**, 333–350. MR1379533

[27] A. Röllin (2005). Approximation of sums of conditionally independent variables by the translated Poisson distribution. *Bernoulli* **11**, 1115–1128. MR2189083

[28] B. Roos (2003). Kerstan's method for compound Poisson approximation. *Ann. Probab.* **31**, 1754–1771. MR2016599

[29] Ĭ. Sunklodas (1991). Approximation of distributions of sums of weakly dependent random variables by the normal distribution. In *Probability theory, 6 (Russian)*, Itogi Nauki i Tekhniki, pages 140–199. Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow. MR1157208

[30] C. Stein (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602, Berkeley, Calif. Univ. California Press. MR0402873

[31] C. Stein (1986). *Approximate computation of expectation.* IMS Lecture Notes Monogr. Ser. 7. MR0882007

[32] A. N. Tikhomirov (1980). Convergence rate in the central limit theorem for weakly dependent random variables. *Teor. Veroyatnost. i Primenen.* **25**, 800–818. MR0595140