

## ON NEGATIVE BINOMIAL APPROXIMATION TO $k$ -RUNS

XIAOXIN WANG\* AND  
AIHUA XIA,\*\* *The University of Melbourne*

### Abstract

The distributions of the run occurrences for a sequence of independent and identically distributed (i.i.d.) experiments are usually obtained by combinatorial methods (see Balakrishnan and Koutras (2002, Chapter 5)) and the resulting formulae are often very tedious, while the distributions for non i.i.d. experiments are generally intractable. It is therefore of practical interest to find a suitable approximate model with reasonable approximation accuracy. In this paper we demonstrate that the negative binomial distribution is the most suitable approximate model for the number of  $k$ -runs: it outperforms the Poisson approximation, the general compound Poisson approximation as observed in Eichelsbacher and Roos (1999), and the translated Poisson approximation in Rollin (2005). In particular, its accuracy of approximation in terms of the total variation distance improves when the number of experiments increases, in the same way as the normal approximation improves in the Berry–Esseen theorem.

*Keywords:* Poisson approximation; translated Poisson approximation; compound Poisson approximation; negative binomial approximation; Stein's method; total variation distance

2000 Mathematics Subject Classification: Primary 60F05  
Secondary 60C05; 60E15

### 1. Introduction

Let  $\{\eta_i, 1 \leq i \leq n\}$  be a sequence of independent Bernoulli random variables with  $E(\eta_i) = p_i, 1 \leq i \leq n$ . Our interest is to study the distribution  $\mathcal{L}(W)$  of  $W = \sum_{i=1}^{n-k+1} \prod_{j=i}^{i+k-1} \eta_j$ . The random variable  $W$  is one of the run statistics (Balakrishnan and Koutras (2002, Section 4.1)) and is often termed as  $k$ -runs.

The run statistics  $W$  were introduced in Mood (1940). The problem is well known in reliability theory as the  $m$ -consecutive- $k$ -out-of- $n$  failure system (Chiang and Niu (1981)), where the major interest is on the reliability of the system. The system has  $n$  items on a line and each item fails with probability  $p \in (0, 1)$ , independently of the others. The system fails whenever there are at least  $k$  consecutive failing items in the system at the same time, and the reliability of the system is determined by the probability that the system works. Using our notation, the reliability of the system is  $P(W = 0)$ .

Run statistics play a critical role in many areas such as nonparametric statistical testing, quality control, and reliability theory (see Balakrishnan and Koutras (2002, Chapters 3, 8, and 12)). The exact distributions of the run occurrences for a sequence of independent and identically distributed (i.i.d.) experiments are usually obtained by combinatorial methods and

Received 1 May 2007; revision received 1 February 2008.

\* Postal address: Department of Mathematics and Statistics, The University of Melbourne, VIC 3010, Australia.  
Email address: danawang\_79@hotmail.com

\*\* Email address: xia@ms.unimelb.edu.au

generating functions (Balakrishnan and Koutras (2002, Chapter 5)), and the resulting formulae are often very tedious. For example, Koutras and Alexandrou (1997) (see also Balakrishnan and Koutras (2002, p. 156)) showed that, if  $p_i = p$  for all  $1 \leq i \leq n$  and  $q = 1 - p$  then

$$\begin{aligned}
 P(W = x) = & \sum_{y=0}^n q^y p^{n-y} \left( \binom{n - k(y + 1)}{y} \right) 1_{\{x+(y+1)k-1=n\}} \\
 & + \sum_{j=0}^{\min(x-1,y)} \sum_{i=0}^{\infty} (-1)^i \binom{x-1}{j} \binom{y+1}{y-j} \binom{y-j}{i} \\
 & \times \binom{n - k(i + j + 1) - x}{y - j - 1},
 \end{aligned}$$

where  $1_{\{\cdot\}}$  is the indicator function. The formula is useful when  $n$  and  $k$  are not too large. However, when  $n$  and  $k$  are relatively large, the computation of the formula is practically very hard and the distributions for non i.i.d. sequences are generally intractable. It is therefore of practical interest to find a suitable approximate distribution for  $\mathcal{L}(W)$  with reasonable approximation accuracy in terms of the total variation distance  $d_{TV}(Q_1, Q_2) = \sup_{A \subset \mathbb{Z}_+} |Q_1(A) - Q_2(A)|$ , where  $Q_1$  and  $Q_2$  are two probability measures on  $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$ .

It seems that von Mises (1921) was the first person to work on this front with the Poisson distribution as the approximate distribution. In the context of reliability theory, Poisson approximation was derived for the consecutive- $k$ -out-of- $n$  failure system in Papastavridis (1987), Chryssaphinou and Papastavridis (1990), and Koutras and Papastavridis (1993). Barbour *et al.* (1992, p. 163) also investigated the same problem with Poisson approximation and obtained the error estimate  $O(p)$ . That is, the approximation accuracy does not improve when  $n$  increases. This is not surprising since, when  $k = 1$ , the lower bound of order  $p$  can also be obtained (Barbour *et al.* (1992, p. 61)).

By viewing the  $k$ -runs as clusters of at least  $k$  1s and considering the clusters as ‘rare events’, Arratia *et al.* (1990) suggested approximating  $\mathcal{L}(W)$  by a compound Poisson and achieved an error bound of  $O(nkp^{2k}(1 - p))$ . Roos (1994) investigated this approach again, with a careful selection of parameters in the compound Poisson approximation, and managed to improve the approximation error to  $O(kp^k \log(np^k))$  when  $p < \frac{1}{2}$ . Eichelsbacher and Roos (1999) obtained the bound  $O(kp^k)$  with respect to the Kolmogorov distance and concluded that a bound in terms of  $d_{TV}$  for large  $p$  could be of order  $O(p^k)$ .

Barbour and Xia (1999) used compound Poisson signed measures with two parameters to approximate the distribution of 2-runs with an approximation error of no more than  $O(n^{-1/2})$ . The problem was also investigated in Rollin (2005) using a translated Poisson approximation with a bound  $K/\sqrt{n}$ , where  $K = K(k, p)$  is a complicated constant independent of  $n$ .

Brown and Xia (2001) showed that, as a special case of the compound Poisson distribution, the negative binomial approximation is sufficiently good for the distribution of 2-runs. The purpose of this note is to show that a suitably chosen negative binomial distribution is a sufficiently good approximate for  $\mathcal{L}(W)$ . The study shows that, amongst these approximating models, the negative binomial approximation works the best since it gives an explicit error estimate without unspecified constants, and numerical calculation also confirms its superiority.

The negative binomial distribution with parameters  $r > 0$  and  $0 < q < 1$  is defined as

$$\text{NB}(r, q)\{j\} = \frac{\Gamma(r + j)}{\Gamma(r)j!} q^r (1 - q)^j, \quad j \in \mathbb{Z}_+.$$

It can be considered as a Poisson distribution mixed with a logarithmic distribution (see Johnson *et al.* (2005, pp. 212–213 and p. 346).

Let  $\vartheta_m$  be the  $m$ th largest number of  $(1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1}$ ,  $1 \leq i \leq n$ , and set

$$\phi := 2 \wedge \frac{4.6}{\sqrt{\sum_{m=4k-1}^n \vartheta_m}}. \tag{4.6}$$

**Theorem 1.1.** *For independent Bernoulli random variables  $\{\eta_i, 1 \leq i \leq n\}$  with  $P(\eta_i = 1) = p_i, 1 \leq i \leq n$ , define  $I_i = \prod_{s=i}^{i+k-1} \eta_s$  and  $W^* = \sum_{i=1}^n I_i$ , where  $\eta_{i+mn}$  is treated as  $\eta_i$  for  $1 \leq i \leq n$  and  $m = \pm 1, \pm 2, \dots$ . Let  $\lambda = E(W^*) = \sum_{i=1}^n \prod_{s=i}^{i+k-1} p_s, \sigma^2 = \text{var}(W^*), r = \lambda^2/(\sigma^2 - \lambda)$ , and  $q = \lambda/\sigma^2$ . If  $\sigma^2 > \lambda$  and  $n > 4k$  then*

$$\begin{aligned} & d_{TV}(\mathcal{L}(W^*), \text{NB}(r, q)) \\ & \leq \frac{\phi}{\lambda} \left( \sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} E I_i \left( \left( \sum_{l=i-2k+2}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) E(I_j I_l) + \sum_{l=i-2k+2}^{i+2k-2} E(I_j) E(I_l) \right) \right. \\ & \quad \left. + \sum_{i=1}^n \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} \left( \left( \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) E(I_i I_j I_l) + \sum_{l=i-2k+2}^{i+2k-2} E(I_i I_j) E(I_l) \right) \right. \\ & \quad \left. + \frac{1}{\lambda} \left( \sum_{i=1}^n \sum_{0 < |l-i| \leq k-1} E(I_i I_l) \right)^2 \right) \tag{1.1} \end{aligned}$$

$$\leq \frac{4.5(4k - 3)(2k - 1)\phi}{\lambda} \sum_{i=1}^n m_i^2 E(I_i), \tag{1.2}$$

where  $m_i = \max\{p_s : i - 2k + 2 \leq s \leq i + 2k - 2\}$ .

**Remark 1.1.** The approximating negative binomial distribution is chosen so that both the mean and variance match with those of  $W^*$ . Bearing in mind that, for a negative binomial, its variance is always greater than its mean, the condition  $\sigma^2 > \lambda$  is also necessary to ensure a valid negative binomial approximation.

If our interest is in  $\mathcal{L}(W)$ , we can use the following proposition and the triangle inequality to bound the approximation error.

**Proposition 1.1.** *Using the notation given in Theorem 1.1, define  $W = \sum_{i=1}^{n-k+1} I_i$ . Then*

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(W^*)) \leq \left( 1 \wedge \frac{2.3}{\sqrt{\sum_{m=k+2}^n \vartheta_m}} \right) \sum_{i=n-k+2}^n E(I_i).$$

**Corollary 1.1.** *Suppose that  $p_i = p$  for all  $1 \leq i \leq n$ . Let*

$$r = \frac{n(1 - p)^k}{2(p - p^k) - (2k - 1)p^k(1 - p)} \quad \text{and} \quad q = \frac{1 - p}{1 - p + 2(p - p^k) - (2k - 1)p^k(1 - p)}.$$

If  $n > 4k$  and if

$$(2k - 1)p^k - (2k + 1)p^{k-1} + 2 > 0, \tag{1.3}$$

TABLE 1.

$k$	$s(k)$
2	0.6667
3	0.8633
4	0.9267
5	0.9545
6	0.9690
7	0.9776
8	0.9830
9	0.9867
10	0.9893
11	0.9912
12	0.9927
13	0.9938
14	0.9947
15	0.9954
20	0.9974
25	0.9984
30	0.9989
35	0.9992

then

$$d_{TV}(\mathcal{L}(W^*), NB(r, q)) \leq 4.5(4k - 3)(2k - 1)p^2 \left( 2 \wedge \frac{4.6}{\sqrt{(n - 4k + 2)p^k(1 - p)^3}} \right). \tag{1.4}$$

*Proof.* Noting that (1.3) is equivalent to the condition that  $\sigma^2 > \lambda$  in Theorem 1.1, (1.4) is an immediate consequence of (1.2).

**Remark 1.2.** Condition (1.3) is equivalent to the condition that  $\sigma^2 > \lambda$  in Theorem 1.1. It defines the range of  $p$  so that it is possible to determine the negative binomial with the same mean and variance as  $W^*$ . For the  $p$  outside the range, it seems better to use the binomial approximation. However, as we will show below,  $p$  outside the range tends to be close to 1 and there is little interest to study the run statistics for big values of  $p$ .

To see how restrictive condition (1.3) is, let  $s(k) \in (0, 1)$  be the solution of the equation  $(2k - 1)p^k - (2k + 1)p^{k-1} + 2 = 0$ . Then (1.3) holds if  $p < s(k)$ . For example, if  $p < \frac{2}{3}$ , (1.3) holds for all  $k \geq 2$  and if  $p < 0.8633$ , (1.3) is valid for all  $k \geq 3$ . In addition, for large  $k$ , the solution  $s(k)$  is close to 1. The values of  $s(k)$  for some  $ks$  are provided in Table 1.

**Remark 1.3.** When  $k = 2$ , the upper bound in (1.4) becomes  $310.5p^2/\sqrt{(n - 6)p^2(1 - p)^3}$ , which is about ten times the bound in Theorem 4.2 of Brown and Xia (2001). The sharper bound in Brown and Xia (2001) is possible because when we apply Stein’s method, it is necessary to consider how the outcome of a typical pair  $(\eta_i, \eta_{i+1})$  affects the distribution of the number of 2-runs in the whole sequence and Brown and Xia (2001) were able to carefully seek more cancellations of errors among the four possible cases based on the outcomes of the pair. Such an ad hoc approach is not suitable for general  $k$ -runs since, instead of considering a typical pair  $(\eta_i, \eta_{i+1})$ , we have to work on a typical sequence  $(\eta_i, \dots, \eta_{i+k-1})$  with  $2^k$  possible outcomes, and the non i.i.d. assumption of the sequence exacerbates the complexity. Our approach in this

paper is relatively crude in the sense that we use the error bound for the worst situation in the  $2^k$  cases to bound all the other cases, resulting in a larger upper bound.

**Remark 1.4.** When  $p$  is small and  $n$  is large, the upper bound of the error is approximately of the order  $k^2 p^2 (1 \wedge 1/\sqrt{np^k})$ , which goes to 0 as  $n$  increases.

**Remark 1.5.** Eichelsbacher and Roos (1999) improved the upper bound of the compound Poisson approximation with respect to the Kolmogorov distance with order  $O(p^k)$ . This bound cannot get better if  $n$  increases. Moreover, there are  $2k - 1$  parameters used in the model fitting. This is in contrast to the two parameters used in the negative binomial approximation, yet, a better bound of order  $O(n^{-1/2})$  is achieved. The bound is of the same order as the Berry–Esseen bound in the normal approximation with respect to the Kolmogorov distance (Petrov (1995, p. 150)).

**Remark 1.6.** The bound of the total variation distance in Corollary 1.1 is explicit and concise. The relationship among  $k$ ,  $p$ , and  $n$  is much more obvious than that for the compound Poisson approximation in Eichelsbacher and Roos (1999) and the translated Poisson approximation in Rollin (2005). Although there is a restriction for  $p$ , it could be negligible in practice when  $k$  is large (see Remark 1.2). Besides, when  $p$  is close to 1 and  $n$  is not large, the distribution of  $W$  is not unimodal (see Balakrishnan and Koutras (2002, pp. 157–160)), hence, we should not expect a good negative binomial approximation in such circumstances.

## 2. Proofs

Brown and Xia (2001) suggested using the following Stein equation for  $NB(r, q)$ :

$$\mathcal{B}g_A(i) := (a + bi)g_A(i + 1) - ig_A(i) = 1_{\{i \in A\}} - NB(r, q)(A), \quad A \subset \mathbb{Z}_+, \quad (2.1)$$

where

$$a = r(1 - q) \quad \text{and} \quad b = 1 - q. \quad (2.2)$$

Theorem 2.10 of Brown and Xia (2001) states that, for  $\Delta g_A(i) := g_A(i + 1) - g_A(i)$ ,

$$\sup_{A \subset \mathbb{Z}_+} |\Delta g_A(i)| \leq \frac{1}{a + bi} \wedge \frac{1}{i} \leq \frac{1}{a}, \quad i \in \mathbb{Z}_+. \quad (2.3)$$

The proof of Theorem 1.1 is similar to that of Theorem 4.2 of Brown and Xia (2001). To achieve this, it is necessary for us to generalize Lemma 5.1 of Barbour and Xia (1999), as summarized below.

**Lemma 2.1.** Let  $\{\eta_i : 1 \leq i \leq n + k - 1\}$  be independent indicator random variables with  $P(\eta_i = 1) = p_i$ , and set  $p_0 = 0$  and  $W_n = \sum_{i=1}^n \eta_i \eta_{i+1} \cdots \eta_{i+k-1}$ . Then, for each  $n \geq k$ ,

$$\begin{aligned} b_n(p_1, p_2, \dots, p_{n+k-1}) &:= d_{TV}(\mathcal{L}(W_n), \mathcal{L}(W_n + 1)) \\ &\leq 1 \wedge \frac{2.3}{\sqrt{\sum_{i=1}^n (1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1}}} \\ &:= C(p_1, p_2, \dots, p_{n+k-1}). \end{aligned} \quad (2.4)$$

*Proof.* The proof is basically a line-by-line repetition of the proof of Lemma 5.1 of Barbour and Xia (1999) with some minor modifications when necessary.

Let  $W_m = \sum_{i=1}^m \eta_i \eta_{i+1} \cdots \eta_{i+k-1}$ , and let  $\{\zeta_i : 1 \leq i \leq n+k-1\}$  be an independent copy of  $\{\eta_i : 1 \leq i \leq n+k-1\}$ . Set  $\eta'_0 = 0$  and, for  $1 \leq i \leq n+k-1$ , define

$$\eta'_i = \begin{cases} \zeta_i & \text{if } \eta_{i-1} = \eta'_{i-1} = 0, \\ \eta_i & \text{otherwise.} \end{cases}$$

Let  $W'_m = 1 + \sum_{i=1}^m \eta'_i \eta'_{i+1} \cdots \eta'_{i+k-1}$ ,  $D_0 = -1$ ,  $D_m = W_m - W'_m$ , and  $\delta_m = D_m - D_{m-1}$ , for  $m \geq 1$ . Then  $\delta_m$  takes values 0 and  $\pm 1$ , and  $\{D_m, 1 \leq m \leq n\}$  is a symmetric random walk. We observe that, for all  $m$ ,  $\delta_m = 1$  if and only if  $\eta_{m-1} = \eta'_{m-1} = 0$ ,  $\eta_m = 1$ ,  $\eta'_m = \zeta_m = 0$ , and  $\eta_{m+1} = 1, \dots, \eta_{m+k-1} = 1$ ; and  $\delta_m = -1$  if and only if  $\eta_{m-1} = \eta'_{m-1} = 0$ ,  $\eta_m = 0$ ,  $\eta'_m = \zeta_m = 1$ , and  $\eta_{m+1} = 1, \dots, \eta_{m+k-1} = 1$ . That is,

$$\{\delta_m \neq 0\} = \{\eta_{m-1} = \eta'_{m-1} = 0, |\eta_m - \eta'_m| = 1, \eta_{m+1} = 1, \dots, \eta_{m+k-1} = 1\}.$$

Now define  $\tau_0 = 0$  and, for  $j \geq 1$ ,

$$\tau_j = \min\{m > \tau_{j-1}, \delta_m \neq 0\}.$$

Set  $R_i = 1_{\{\delta_i \neq 0\}}$  and  $R = \sum_{i=1}^n R_i$ , so  $R = r$  is equivalent to  $\tau_r \leq n < \tau_{r+1}$ , i.e. there are  $r$  jumps in  $\{D_i, 1 \leq i \leq n\}$ . Since  $P(\delta_i = 1) = P(\delta_i = -1)$ , for each  $i \geq 1$ ,

$$\begin{aligned} E(R_i) &= 2 P(\delta_i = 1) \\ &= 2 P(\eta_{i-1} = \eta'_{i-1} = 0) P(\eta_i = 1) P(\zeta_i = 0) P(\eta_{i+1} = 1) \cdots P(\eta_{i+k-1} = 1) \\ &\geq 2(1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1}, \end{aligned} \tag{2.5}$$

where the inequality is because

$$\begin{aligned} P(\eta_{i-1} = \eta'_{i-1} = 0) &= P(\eta_{i-1} = \eta'_{i-1} = 0 \mid \eta_{i-2} = \eta'_{i-2} = 0) P(\eta_{i-2} = \eta'_{i-2} = 0) \\ &\quad + P(\eta_{i-1} = \eta'_{i-1} = 0 \mid \{\eta_{i-2} = \eta'_{i-2} = 0\}^c) P(\{\eta_{i-2} = \eta'_{i-2} = 0\}^c) \\ &= P(\eta_{i-1} = 0) P(\zeta_{i-1} = 0) P(\eta_{i-2} = \eta'_{i-2} = 0) + P(\eta_{i-1} = 0) P(\{\eta_{i-2} = \eta'_{i-2} = 0\}^c) \\ &\geq P(\eta_{i-1} = 0) P(\zeta_{i-1} = 0) \\ &= (1 - p_{i-1})^2, \end{aligned}$$

where the superscript ‘c’ denotes the complement of the set. The conditional distribution  $\mathcal{L}(\delta_{\tau_j} \mid R = r)$ ,  $1 \leq j \leq r$ , is uniformly distributed on  $\{-1, 1\}$ . Let  $Z_j^r$  denote  $D_{\tau_j}$  given  $R = r$ ,  $1 \leq j \leq r$ . Then  $Z_j^r$  is a simple random walk with initial state  $-1$ , i.e.

$$P(Z_j^r - Z_{j-1}^r = 1) = P(Z_j^r - Z_{j-1}^r = -1) = \frac{1}{2}.$$

Let  $J = \min\{j \geq 1 : D_{\tau_j} = 0\}$ , and define

$$W''_m = \begin{cases} W'_m, & m < \tau_J, \\ W_m, & m \geq \tau_J. \end{cases}$$

Therefore, by the above coupling,  $\mathcal{L}(W_n + 1) = \mathcal{L}(W''_n)$  and

$$d_{TV}(\mathcal{L}(W_n), \mathcal{L}(W_n + 1)) = d_{TV}(\mathcal{L}(W_n), \mathcal{L}(W''_n)) \leq P(W_n \neq W''_n).$$

By the reflection principle, we can show that

$$\begin{aligned} P(W_n \neq W_n'' \mid R = r) &= P\left(\max_{1 \leq j \leq r} D_{\tau_j} \leq -1 \mid R = r\right) \\ &= P\left(\max_{1 \leq j \leq r} Z_j^r \leq -1\right) \\ &= P(Z_r^r \in \{-2, -1\}) \\ &= \max_k P(Z_r^r = k) \\ &\leq \sqrt{\frac{2}{\pi r}}. \end{aligned}$$

Then, for all  $r \geq 1$  and a constant  $0 < \kappa < 1$ ,

$$\begin{aligned} P(W_n \neq W_n'') &= \sum_{r=1}^{\infty} P(W_n \neq W_n'' \mid R = r) P(R = r) \\ &\leq \sum_{1 \leq r \leq \kappa E(R)} P(R = r) + \sum_{r \geq \kappa E(R)+1} \frac{0.8}{\sqrt{\kappa E(R)}} P(R = r) \\ &\leq P(|R - E(R)| \geq (1 - \kappa) E(R)) + \frac{0.8}{\sqrt{\kappa E(R)}} \\ &\leq \frac{\text{var } R}{(1 - \kappa)^2 (E(R))^2} + \frac{0.8}{\sqrt{\kappa E(R)}}, \end{aligned} \tag{2.6}$$

by Chebyshev’s inequality. From Lemma 2.2, we have

$$\text{var } R \leq 3 E(R). \tag{2.7}$$

Then, if  $\sum_{i=1}^n (1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1} \leq 5.29$ , (2.4) is obvious. So, from now on, we assume that

$$\sum_{i=1}^n (1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1} > 5.29,$$

which implies from (2.5) that

$$E(R) \geq 2 \sum_{i=1}^n (1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1} > 10.58.$$

Hence, combining (2.6) and (2.7), it follows that

$$\begin{aligned} P(W_n \neq W_n'') &\leq \frac{3}{(1 - \kappa)^2 E(R)} + \frac{0.8}{\sqrt{\kappa E(R)}} \\ &= \frac{1}{\sqrt{E(R)}} \left( \frac{3}{(1 - \kappa)^2 \sqrt{10.58}} + \frac{0.8}{\sqrt{\kappa}} \right) \\ &\leq \frac{3.222}{\sqrt{E(R)}} \\ &\leq \frac{2.3}{\sqrt{\sum_{i=1}^n (1 - p_{i-1})^2 p_i (1 - p_i) p_{i+1} \cdots p_{i+k-1}}}, \end{aligned}$$

where the second inequality is obtained by letting  $\kappa = 0.2197$ .

**Lemma 2.2.** Using the same notation as in Lemma 2.1, and letting  $R = \sum_{i=1}^n R_i = \sum_{i=1}^n 1_{\{\delta_i \neq 0\}}$ , we have

$$\text{var } R \leq 3 \text{E}(R).$$

*Proof.* We have

$$\begin{aligned} \text{var } R &= \sum_{i=1}^n \text{E } R_i (1 - \text{E}(R_i)) + 2 \sum_{i < j} (\text{E}(R_i R_j) - \text{E}(R_i) \text{E}(R_j)) \\ &\leq \text{E}(R) + 2 \sum_{i=1}^n \sum_{j=i+k+1}^n (\text{E}(R_i R_j) - \text{E}(R_i) \text{E}(R_j)), \end{aligned} \tag{2.8}$$

since  $R_i R_j$  is always 0 unless  $j - i \geq k + 1$ . Also,

$$\begin{aligned} \text{E}(R_i R_j) &= \text{P}(R_i = 1, R_j = 1) \\ &= \text{P}(R_j = 1 \mid \eta_{i-1} = \eta'_{i-1} = 0, |\eta_i - \eta'_i| = 1, \\ &\quad \eta_{i+1} = 1, \dots, \eta_{i+k-1} = 1) \text{P}(R_i = 1) \\ &= \text{P}(R_j = 1 \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) \text{P}(R_i = 1) \\ &= \text{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) \text{E}(R_i). \end{aligned}$$

Noting that

$$\begin{aligned} \text{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) &= \text{E}(R_j \mid \eta_{i+k-1} = 0, \eta'_{i+k-1} = 1) \\ &= \text{E}(R_j \mid \eta_{i+k-1} = 1, \eta'_{i+k-1} = 0), \end{aligned}$$

we have, for all  $j - i \geq k + 1$ ,

$$\begin{aligned} |\text{E}(R_i R_j) - \text{E}(R_i) \text{E}(R_j)| &= |\text{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) - \text{E}(R_j)| \text{E}(R_i) \\ &\leq \text{E}(R_i) |\text{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) \\ &\quad - \text{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 0)|. \end{aligned} \tag{2.9}$$

Let  $U = \min\{r \geq i + k : \eta_r = \zeta_r\}$ . Then  $U$  is independent of  $\eta_i$  and  $\eta'_i$  with

$$\begin{aligned} \text{P}(U \geq j) &= \text{P}(\zeta_r \neq \eta_r : i + k \leq r \leq j - 1) \\ &= \prod_{r=i+k}^{j-1} \text{P}(\eta_r \neq \zeta_r) \\ &= \prod_{r=i+k}^{j-1} (2p_r(1 - p_r)) \\ &\leq 2^{-(j-i-k)} \quad \text{for } j \geq i + k + 1. \end{aligned}$$

Note that, when  $U = l$ ,  $i + k \leq l \leq j - 1$ , it is always true that  $\eta_l = \zeta_l$ , which implies that  $\eta_l = \eta'_l$ . However, given  $\eta_l = \eta'_l$ ,  $\{(\eta_r, \eta'_r) : r \geq l\}$  is specified by  $\{(\eta_r, \zeta_r) : r \geq l\}$ , which is

independent of  $(\eta_{i+k-1}, \eta'_{i+k-1})$ . So we have, for  $i + k \leq l \leq j - 1$ ,

$$\begin{aligned} &\mathcal{L}((\eta_r, \eta'_r)_{r \geq l} \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 1) \\ &= \mathcal{L}((\eta_r, \eta'_r)_{r \geq l} \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 0). \end{aligned} \tag{2.10}$$

On the other hand,

$$\begin{aligned} &\mathbb{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) \\ &= \sum_{l=i+k}^{\infty} \mathbb{E}(R_j \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 1) \mathbb{P}(U = l \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) \\ &= \sum_{l=i+k}^{\infty} \mathbb{E}(R_j \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 1) \mathbb{P}(U = l). \end{aligned}$$

Similarly,

$$\mathbb{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 0) = \sum_{l=i+k}^{\infty} \mathbb{E}(R_j \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 0) \mathbb{P}(U = l).$$

Thus, by (2.10),

$$\begin{aligned} &|\mathbb{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 1) - \mathbb{E}(R_j \mid \eta_{i+k-1} = \eta'_{i+k-1} = 0)| \\ &\leq \sum_{l=i+k}^{\infty} |\mathbb{E}(R_j \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 1) \\ &\quad - \mathbb{E}(R_j \mid U = l, \eta_{i+k-1} = \eta'_{i+k-1} = 0)| \mathbb{P}(U = l) \\ &\leq \sum_{l=j}^{\infty} \mathbb{P}(U = l) \\ &= \mathbb{P}(U \geq j). \end{aligned} \tag{2.11}$$

Therefore, it follows from (2.8), (2.9), and (2.11) that

$$\begin{aligned} \text{var } R &\leq \mathbb{E}(R) + 2 \sum_{i=1}^n \sum_{j=i+k+1}^n \mathbb{E}(R_i) \mathbb{P}(U \geq j) \\ &\leq \mathbb{E}(R) + 2 \sum_{i=1}^n \mathbb{E}(R_i) \sum_{j=i+k+1}^n \left(\frac{1}{2}\right)^{j-i-k} \\ &\leq \mathbb{E}(R) + 2 \mathbb{E}(R) \\ &= 3 \mathbb{E}(R). \end{aligned}$$

As we mentioned in Remark 1.3, to implement Stein’s method, we need to consider how the outcome of a typical sequence  $(\eta_i, \dots, \eta_{i+k-1})$  affects the distribution of  $W^*$ . By taking away the elements in  $W^*$  which are dependent on  $(\eta_i, \dots, \eta_{i+k-1})$  in various levels, we define, from

now on, for  $i = 1, 2, \dots, n$  and  $i - k < j < i + k$ ,

$$\begin{aligned} W_i &= W^* - I_i, & U_i &= W^* - \sum_{|s-i| \leq k-1} I_s, \\ V_i &= W^* - \sum_{|i-s| \leq 2k-2} I_s, & X_i &= W^* - \sum_{|i-s| \leq 3k-3} I_s, \\ Y_{i,j} &= \sum_{s=i-k+1}^j I_s, & Y'_{i,j} &= \sum_{\substack{s=i-k+1 \\ s \neq i}}^j I_s. \end{aligned}$$

Then  $U_i$  is independent of  $I_i$ ,  $V_i$  is independent of  $\{I_s : |s - i| \leq k - 1\}$ , and  $X_i$  is independent of  $\{I_s : |s - i| \leq 2k - 2\}$ .

**2.1. Proof of Theorem 1.1**

By (2.1), it suffices to show that, for every  $A \subset \mathbb{Z}_+$ ,  $|\mathbb{E} \mathcal{B}g_A(W^*)|$  is bounded by the right-hand side of (1.1). To this end, we fix  $A \subset \mathbb{Z}_+$  and write  $g$  for  $g_A$ . Then

$$\begin{aligned} \mathbb{E}(\mathcal{B}g(W^*)) &= a \mathbb{E}(g(W^* + 1)) + b \mathbb{E}(W^*g(W^* + 1)) - \mathbb{E}(W^*g(W^*)) \\ &= a \mathbb{E}(g(W^* + 1)) + b \sum_{i=1}^n \mathbb{E}(I_i g(W_i + 2)) - \sum_{i=1}^n \mathbb{E}(I_i g(W_i + 1)). \end{aligned} \tag{2.12}$$

Set  $a = (1 - b)\lambda$ . Then (2.12) becomes

$$\begin{aligned} \mathbb{E}(\mathcal{B}g(W^*)) &= (1 - b) \sum_{i=1}^n \mathbb{E}(I_i) \mathbb{E}(g(W^* + 1) - g(U_i + 1)) \\ &\quad - (1 - b) \sum_{i=1}^n \mathbb{E}(I_i g(W_i + 1) - I_i g(U_i + 1)) + b \sum_{i=1}^n \mathbb{E}(I_i \Delta g(W_i + 1)) \\ &= (1 - b) \sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} \mathbb{E}(I_i) \mathbb{E}(I_j \Delta g(U_i + Y_{i,j-1} + 1)) \\ &\quad - (1 - b) \sum_{i=1}^n \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} \mathbb{E}(I_i I_j \Delta g(U_i + Y'_{i,j-1} + 1)) \\ &\quad + b \sum_{i=1}^n \mathbb{E}(I_i \Delta g(W_i + 1)). \end{aligned} \tag{2.13}$$

Now, let

$$(1 - b) \sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} \mathbb{E}(I_i) \mathbb{E}(I_j) + b \sum_{i=1}^n \mathbb{E}(I_i) - (1 - b) \sum_{i=1}^n \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} \mathbb{E}(I_i I_j) = 0.$$

As  $\mathbb{E}(I_i) = \mathbb{E}(I_i^2)$ , the above equation is equivalent to

$$b \sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} (\mathbb{E}(I_i I_j) - \mathbb{E}(I_i) \mathbb{E}(I_j)) - \sum_{i=1}^n \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} (\mathbb{E}(I_i I_j) - \mathbb{E}(I_i) \mathbb{E}(I_j)) + \sum_{i=1}^n (\mathbb{E}(I_i))^2 = 0,$$

namely,

$$b = \frac{\sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} \sum_{j \neq i} \text{cov}(I_i, I_j) - \sum_{i=1}^n (\mathbb{E}(I_i))^2}{\sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} \text{cov}(I_i, I_j)} = \frac{\sigma^2 - \lambda}{\sigma^2}.$$

Hence, with  $a = (1 - b)\lambda$ , the above yields  $a = \lambda^2/\sigma^2$ , and the values of  $r$  and  $q$  can be obtained by solving the equations in (2.2). With the chosen values of  $a$  and  $b$  above, it follows from (2.13) that

$$\begin{aligned} & \mathbb{E}(\mathcal{B}g(W^*)) \\ &= (1 - b) \sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} (\mathbb{E} I_i \mathbb{E}(I_j \Delta g(U_i + Y_{i,j-1} + 1)) - \mathbb{E}(I_i) \mathbb{E}(I_j) \mathbb{E}(\Delta g(W^* + 1))) \\ & \quad - (1 - b) \sum_{i=1}^n \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} (\mathbb{E}(I_i I_j \Delta g(U_i + Y'_{i,j-1} + 1)) - \mathbb{E}(I_i I_j) \mathbb{E}(\Delta g(W^* + 1))) \\ & \quad + b \sum_{i=1}^n (\mathbb{E}(I_i \Delta g(W_i + 1)) - \mathbb{E}(I_i) \mathbb{E}(\Delta g(W^* + 1))). \end{aligned} \tag{2.14}$$

To further our proof, we need to work on the three terms on the right-hand side of (2.14) and assess the extent to which the impact of the  $k$ -runs at certain locations is on the distribution of  $W^*$ . For ease of reading, we collect the estimates in the next lemma and will continue the proof of Theorem 1.1 after the lemma.

**Lemma 2.3.** *For  $n \geq 4k$ ,  $1 \leq i \leq n$ , and  $i - k + 1 \leq j \leq i + k - 1$ , for any bounded function  $h$ ,*

$$\begin{aligned} & |\mathbb{E}(I_j h(U_i + Y_{i,j-1})) - \mathbb{E}(I_j) \mathbb{E}(h(W^*))| \\ & \leq \phi \|h\| \left( \left( \sum_{l=i-2k+2}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) \mathbb{E}(I_j I_l) + \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_j) \mathbb{E}(I_l) \right), \end{aligned} \tag{2.15}$$

$$|\mathbb{E}(I_i h(W_i)) - \mathbb{E}(I_i) \mathbb{E}(h(W^*))| \leq \phi \|h\| \left( \sum_{0 < |l-i| \leq k-1} \mathbb{E}(I_i I_l) + \sum_{|l-i| \leq k-1} \mathbb{E}(I_i) \mathbb{E}(I_l) \right), \tag{2.16}$$

$$\begin{aligned} & |\mathbb{E}(I_i I_j h(U_i + Y'_{i,j-1})) - \mathbb{E}(I_i I_j) \mathbb{E}(h(W^*))| \\ & \leq \phi \|h\| \left( \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} \mathbb{E}(I_i I_j I_l) + \sum_{l=i+k}^{i+2k-2} \mathbb{E}(I_i I_j I_l) + \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_i I_j) \mathbb{E}(I_l) \right), \end{aligned} \tag{2.17}$$

where (2.17) is valid for  $j \neq i$ .

*Proof.* Since  $V_i$  is independent of  $I_j$ , we have

$$\begin{aligned} & \mathbb{E}(I_j h(U_i + Y_{i,j-1})) - \mathbb{E}(I_j) \mathbb{E}(h(W^*)) \\ & = \mathbb{E}(I_j h(U_i + Y_{i,j-1}) - I_j h(V_i)) + \mathbb{E}(I_j) \mathbb{E}(h(V_i) - h(W^*)). \end{aligned} \tag{2.18}$$

For the first term of (2.18), we obtain

$$\begin{aligned}
 & \mathbb{E}(I_j h(U_i + Y_{i,j-1}) - I_j h(V_i)) \\
 &= \sum_{l=i-2k+2}^{j-1} \mathbb{E} \left( I_j I_l \Delta h \left( V_i + \sum_{s=i-2k+2}^{l-1} I_s \right) \right) \\
 & \quad + \sum_{l=i+k}^{i+2k-2} \mathbb{E} \left( I_j I_l \Delta h \left( V_i + \sum_{s=i-2k+2}^{j-1} I_s + \sum_{s=i+k}^{l-1} I_s \right) \right) \\
 &= \sum_{l=i-2k+2}^{j-1} \mathbb{E}(I_j I_l) \mathbb{E} \left( \Delta h \left( X_i + \sum_{s=i-3k+3}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_j = I_l = 1 \right) \\
 & \quad + \sum_{l=i+k}^{i+2k-2} \mathbb{E}(I_j I_l) \\
 & \quad \times \mathbb{E} \left( \Delta h \left( X_i + \sum_{s=i-3k+3}^{j-1} I_s + \sum_{s=i+k}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_j = I_l = 1 \right).
 \end{aligned} \tag{2.19}$$

For a sequence of independent Bernoulli random variables  $\xi_1, \dots, \xi_m$ , we write  $r(\xi_1, \dots, \xi_m) = \sum_{i=1}^{m-k+1} \prod_{l=i}^{i+k-1} \xi_l$ , and define

$$\mathbf{1}_s := \underbrace{1, \dots, 1}_s.$$

Then direct verification gives

$$\begin{aligned}
 & \mathcal{L} \left( X_i + \sum_{s=i-3k+3}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \middle| I_j = I_l = 1 \right) \\
 &= \mathcal{L}(r(\eta_{i+2k-1}, \dots, \eta_n, \eta_1, \dots, \eta_{l-1}, \mathbf{1}_{k-1})), \\
 & \mathcal{L} \left( X_i + \sum_{s=i-3k+3}^{j-1} I_s + \sum_{s=i+k}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \middle| I_j = I_l = 1 \right) \\
 &= \mathcal{L}(r(\mathbf{1}_{j-i}, \eta_{j \vee i+k}, \dots, \eta_{l-1}, \mathbf{1}_{k-1}, 0, \mathbf{1}_{l-i-k+1}, \eta_{l+k}, \dots, \eta_n, \eta_1, \dots, \eta_{j-1}, \mathbf{1}_{k-1})),
 \end{aligned}$$

where here and in the sequel,  $\mathbf{1}_s$  is interpreted as an empty set if  $s \leq 0$ . Applying Lemma 2.1 yields

$$\begin{aligned}
 & \left| \mathbb{E} \left( \Delta h \left( X_i + \sum_{s=i-3k+3}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_j = I_l = 1 \right) \right| \\
 &= \left| \mathbb{E}(\Delta h(r(\eta_{i+2k-1}, \dots, \eta_n, \eta_1, \dots, \eta_{l-1}, \mathbf{1}_{k-1}))) \right| \\
 &\leq 2 \|h\| d_{TV}(\mathcal{L}(r(\eta_{i+2k-1}, \dots, \eta_n, \eta_1, \dots, \eta_{l-1}, \mathbf{1}_{k-1})), \\
 & \quad \mathcal{L}(r(\eta_{i+2k-1}, \dots, \eta_n, \eta_1, \dots, \eta_{l-1}, \mathbf{1}_{k-1}) + 1)) \\
 &\leq 2 \|h\| C(p_{i+2k-1}, \dots, p_n, p_1, \dots, p_{l-1}, \mathbf{1}_{k-1}) \\
 &\leq \phi \|h\|,
 \end{aligned} \tag{2.20}$$

and, likewise, for  $i - k + 1 \leq j \leq i + k - 1$ ,

$$\begin{aligned} & \left| \mathbb{E} \left( \Delta h \left( X_i + \sum_{s=i-3k+3}^{j-1} I_s + \sum_{s=i+k}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_j = I_l = 1 \right) \right| \\ & \leq 2 \|h\| \\ & \quad \times C(\mathbf{1}_{j-i}, p_{j \vee i+k}, \dots, p_{l-1}, \mathbf{1}_{k-1}, 0, \mathbf{1}_{l-i-k+1}, p_{l+k}, \dots, p_n, p_1, \dots, p_{j-1}, \mathbf{1}_{k-1}) \\ & \leq \phi \|h\|. \end{aligned} \tag{2.21}$$

Combining (2.19), (2.20), and (2.21) gives

$$|\mathbb{E}(I_j h(U_i + Y_{i,j-1}) - I_j h(V_i))| \leq \phi \|h\| \left( \sum_{l=i-2k+2}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) \mathbb{E}(I_j I_l). \tag{2.22}$$

On the other hand,

$$\begin{aligned} & |\mathbb{E}(h(V_i) - h(W^*))| \\ & = \left| \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E} \left( I_l \Delta h \left( V_i + \sum_{s=i-2k+2}^{l-1} I_s \right) \right) \right| \\ & = \left| \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E} I_l \mathbb{E} \left( \Delta h \left( V_i + \sum_{s=i-2k+2}^{l-1} I_s \right) \middle| I_l = 1 \right) \right| \\ & = \left| \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_l) \mathbb{E}(\Delta h(r(\mathbf{1}_{l-i-k+1}, \eta_{l \vee (i+k-1)+k}, \dots, \eta_n, \eta_1, \dots, \eta_{l-1}, \mathbf{1}_{k-1}))) \right| \\ & \leq \phi \|h\| \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_l). \end{aligned} \tag{2.23}$$

Thus, (2.15) is obtained from (2.18) and the estimates in (2.22) and (2.23).

Apropos of (2.16), because of the independence between  $I_i$  and  $U_i$ , using arguments similar to (2.20), we have

$$\begin{aligned} & |\mathbb{E}(I_i h(W_i)) - \mathbb{E}(I_i) \mathbb{E}(h(W^*))| \\ & = |\mathbb{E}(I_i h(W_i) - I_i h(U_i)) + \mathbb{E}(I_i) \mathbb{E}(h(U_i) - h(W^*))| \\ & \leq \sum_{\substack{l=i-k+1 \\ l \neq i}}^{i+k-1} |\mathbb{E}(I_i I_l \Delta h(U_i + Y'_{i,l-1}))| + \sum_{l=i-k+1}^{i+k-1} |\mathbb{E}(I_i) \mathbb{E}(I_l \Delta h(U_i + Y_{i,l-1}))| \\ & = \sum_{\substack{l=i-k+1 \\ l \neq i}}^{i+k-1} \mathbb{E}(I_i I_l) \left| \mathbb{E} \left( \Delta h \left( V_i + \sum_{\substack{s=i-2k+2 \\ s \neq i}}^{l-1} I_s + \sum_{s=i+k}^{i+2k-2} I_s \right) \middle| I_i = I_l = 1 \right) \right| \\ & \quad + \sum_{l=i-k+1}^{i+k-1} \mathbb{E} I_i \mathbb{E} I_l \left| \mathbb{E} \left( \Delta h \left( V_i + \sum_{s=i-2k+2}^{l-1} I_s + \sum_{s=i+k}^{i+2k-2} I_s \right) \middle| I_l = 1 \right) \right| \\ & \leq \phi \|h\| \left( \sum_{l=i-k+1}^{i-1} \mathbb{E}(I_i I_l) + \sum_{l=i+1}^{i+k-1} \mathbb{E}(I_i I_l) + \sum_{l=i-k+1}^{i+k-1} \mathbb{E}(I_i) \mathbb{E}(I_l) \right). \end{aligned}$$

Finally, for (2.17), because  $(I_i, I_j)$  is independent of  $V_i$ , it follows that

$$\begin{aligned}
 & \mathbb{E}(I_i I_j h(U_i + Y'_{i,j-1})) - \mathbb{E}(I_i I_j) \mathbb{E} h(W^*) \\
 &= \mathbb{E}(I_i I_j (h(U_i + Y'_{i,j-1}) - h(V_i))) + \mathbb{E}(I_i I_j) \mathbb{E}(h(V_i) - h(W^*)) \\
 &= \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} \mathbb{E} \left( I_i I_j I_l \Delta h \left( V_i + \sum_{\substack{s=i-2k+2 \\ s \neq i}}^{l-1} I_s \right) \right) \\
 &+ \sum_{l=i+k}^{i+2k-2} \mathbb{E} \left( I_i I_j I_l \Delta h \left( V_i + \sum_{\substack{s=i-2k+2 \\ s \neq i}}^{j-1} I_s + \sum_{s=i+k}^{l-1} I_s \right) \right) \\
 &- \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_i I_j) \mathbb{E} \left( I_l \Delta h \left( V_i + \sum_{s=i-2k+2}^{l-1} I_s \right) \right) \\
 &= \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} \mathbb{E}(I_i I_j I_l) \mathbb{E} \left( \Delta h \left( X_i + \sum_{\substack{s=i-3k+3 \\ s \neq i}}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_i = I_j = I_l = 1 \right) \\
 &+ \sum_{l=i+k}^{i+2k-2} \mathbb{E}(I_i I_j I_l) \mathbb{E} \left( \Delta h \left( X_i + \sum_{\substack{s=i-3k+3 \\ s \neq i}}^{j-1} I_s + \sum_{s=i+k}^{l-1} I_s \right. \right. \\
 &\qquad \qquad \qquad \left. \left. + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_i = I_j = I_l = 1 \right) \\
 &- \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_i I_j) \mathbb{E}(I_l) \mathbb{E} \left( \Delta h \left( X_i + \sum_{s=i-3k+3}^{l-1} I_s + \sum_{s=i+2k-1}^{i+3k-3} I_s \right) \middle| I_l = 1 \right).
 \end{aligned} \tag{2.24}$$

Using Lemma 2.1 and arguing as for (2.20), we obtain (2.17) from (2.24).

We now continue the proof of Theorem 1.1. Applying Lemma 2.3 to (2.14) with  $h(\cdot) = \Delta g(\cdot + 1)$ , we obtain, from (2.3),

$$\begin{aligned}
 & |\mathbb{E}(\mathcal{B}g(W^*))| \\
 &\leq \frac{\phi}{\lambda} \sum_{i=1}^n \sum_{j=i-k+1}^{i+k-1} \mathbb{E}(I_i) \left( \left( \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) \mathbb{E}(I_j I_l) + \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_j) \mathbb{E}(I_l) \right) \\
 &+ \frac{\phi}{\lambda} \sum_{i=1}^n \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} \left( \left( \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) \mathbb{E}(I_i I_j I_l) + \sum_{l=i-2k+2}^{i+2k-2} \mathbb{E}(I_i I_j) \mathbb{E}(I_l) \right) \\
 &+ \frac{\phi}{\lambda} \frac{b}{1-b} \sum_{i=1}^n \left( \sum_{0 < |l-i| \leq k-1} \mathbb{E}(I_i I_l) + \sum_{|l-i| \leq k-1} \mathbb{E}(I_i) \mathbb{E}(I_l) \right),
 \end{aligned}$$

which implies (1.1) since

$$\begin{aligned} & \frac{b}{1-b} \sum_{i=1}^n \left( \sum_{0 < |l-i| \leq k-1} E(I_i I_l) + \sum_{|l-i| \leq k-1} E(I_i) E(I_l) \right) \\ &= \frac{1}{\lambda} \left( \left( \sum_{i=1}^n \sum_{0 < |l-i| \leq k-1} E(I_i I_l) \right)^2 - \left( \sum_{i=1}^n \sum_{|l-i| \leq k-1} E(I_i) E(I_l) \right)^2 \right). \end{aligned}$$

Concerning (1.2), we combine (1.1) with the following estimates:

$$\begin{aligned} \sum_{j=i-k+1}^{i+k-1} \left( \sum_{l=i-2k+2}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) E(I_j I_l) &\leq (4k-3) \sum_{j=i-k+1}^{i+k-1} E(I_j) \leq (4k-3)(2k-1)m_i^2, \\ \sum_{j=i-k+1}^{i+k-1} \sum_{l=i-2k+2}^{i+2k-2} E(I_j) E(I_l) &\leq (4k-3)(2k-1)m_i^2, \\ \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} \left( \sum_{\substack{l=i-2k+2 \\ l \neq i}}^{j-1} + \sum_{l=i+k}^{i+2k-2} \right) E(I_i I_j I_l) &\leq (4k-3)(2k-1)m_i^2 E(I_i), \\ \sum_{\substack{j=i-k+1 \\ j \neq i}}^{i+k-1} \sum_{l=i-2k+2}^{i+2k-2} E(I_i I_j) E(I_l) &\leq (4k-3)(2k-1)m_i^2 E(I_i), \end{aligned}$$

and, by Cauchy–Bunyakovskii’s inequality,

$$\begin{aligned} \left( \sum_{i=1}^n \sum_{0 < |l-i| \leq k-1} E(I_i I_l) \right)^2 &= \left( \sum_{i=1}^n \sum_{0 < |l-i| \leq k-1} E(I_l \mid I_i = 1) E(I_i) \right)^2 \\ &\leq \left( \sum_{i=1}^n \sum_{0 < |l-i| \leq k-1} E(I_l \mid I_i = 1)^2 E(I_i) \right) \sum_{i=1}^n \sum_{0 < |j-i| \leq k-1} E(I_i) \\ &\leq (2k-2)^2 \lambda \sum_{i=1}^n m_i^2 E(I_i). \end{aligned}$$

This completes the proof.

**2.2. Proof of Proposition 1.1**

For each  $f : \mathbb{Z}_+ \rightarrow [0, 1]$ ,

$$\begin{aligned} |E f(W) - E f(W^*)| &= \left| E \sum_{i=n-k+2}^n I_i \Delta f \left( W + \sum_{j=n-k+2}^{i-1} I_j \right) \right| \\ &\leq \sum_{i=n-k+2}^n E(I_i) \left| E \left( \Delta f \left( W + \sum_{j=n-k+2}^{i-1} I_j \right) \mid I_i = 1 \right) \right| \end{aligned}$$

$$\begin{aligned}
 &= \sum_{i=n-k+2}^n E(I_i) |E(\Delta f(r(\mathbf{1}_{i+k-1-n}, \eta_{i+k-n}, \dots, \eta_{i-1}, \mathbf{1}_{k-1})))| \\
 &\leq \sum_{i=n-k+2}^n E(I_i) C(\mathbf{1}_{i+k-1-n}, p_{i+k-n}, \dots, p_{i-1}, \mathbf{1}_{k-1}) \quad (\text{see (2.4)}) \\
 &\leq \sum_{i=n-k+2}^n E(I_i) \left(1 \wedge \frac{2.3}{\sqrt{\sum_{m=k+2}^n \vartheta_m}}\right),
 \end{aligned}$$

which implies Proposition 1.1.

### Acknowledgements

This research was supported by the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems (AX) and a Melbourne University scholarship (XW).

### References

- ARRATIA, R., GOLDSTEIN, L. AND GORDON, L. (1990). Poisson approximation and the Chen–Stein method. *Statist. Sci.* **5**, 403–434.
- BALAKRISHNAN, N. AND KOUTRAS, M. V. (2002). *Runs and Scans with Applications*. John Wiley, New York.
- BARBOUR, A. D. AND XIA, A. (1999). Poisson Perturbations. *ESAIM Prob. Statist.* **3**, 131–150.
- BARBOUR, A. D., HOLST, L. AND JANSON, S. (1992). *Poisson Approximation*. Oxford University Press.
- BROWN, T. C. AND XIA, A. (2001). Stein’s method and birth–death processes. *Ann. Prob.* **29**, 1373–1403.
- CHIANG, D. T. AND NIU, S. C. (1981). Reliability of consecutive- $k$ -out-of- $n$ :  $F$  system. *IEEE Trans. Reliab.* **30**, 87–89.
- CHRYSSAPHINOY, O. AND PAPANASTAVRIDIS, S. (1990). Limit distributions for a consecutive- $k$ -out-of- $n$ :  $F$  system. *Adv. Appl. Prob.* **22**, 491–493.
- EICHELSBACHER, P. AND ROOS, M. (1999). Compound Poisson approximation for dissociated random variables via Stein’s method. *Combin. Prob. Comput.* **8**, 335–346.
- JOHNSON, N. L., KEMP, A. W. AND KOTZ, S. (2005). *Univariate Discrete Distributions*. John Wiley, New York.
- KOUTRAS, M. V. AND ALEXANDROU, V. A. (1997). Non-parametric randomness tests based on success runs of fixed length. *Statist. Prob. Lett.* **32**, 393–404.
- KOUTRAS, M. V. AND PAPANASTAVRIDIS, S. G. (1993). Application of the Stein–Chen method for bounds and limit theorems in the reliability of coherent structures. *Naval Res. Logistics* **40**, 617–631.
- MOOD, A. M. (1940). The distribution theory of runs. *Ann. Math. Statist.* **11**, 367–392.
- PAPANASTAVRIDIS, S. (1987). A limit theorem for the reliability of a consecutive- $k$ -out-of- $n$  system. *Adv. Appl. Prob.* **19**, 746–748.
- PETROV, V. V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*. Clarendon Press, Oxford.
- ROLLIN, A. (2005). Approximation of sum of conditionally independent variables by the translated Poisson distribution. *Bernoulli* **11**, 1115–1128.
- ROOS, M. (1994). Stein–Chen method for compound Poisson approximation. Doctoral Thesis, University of Zürich.
- VON MISES, R. (1921). Das Problem der Iterationen. *Z. Angewandte Math. Mech.* **1**, 298–307.