

## Article

# On the Convergence Rate for the Longest at Most $T$ -Contaminated Runs of Heads

István Fazekas <sup>1,\*</sup> , Borbála Fazekas <sup>2</sup> and László Fórián <sup>1</sup>

<sup>1</sup> Faculty of Informatics, University of Debrecen, Kassai Street 26, 4028 Debrecen, Hungary; forian.laszlo@inf.unideb.hu

<sup>2</sup> Institute of Mathematics, University of Debrecen, Egyetem Square 1, 4032 Debrecen, Hungary; borbala.fazekas@science.unideb.hu

\* Correspondence: fazekas.istvan@inf.unideb.hu

**Abstract:** In this paper, we study the usual coin tossing experiment. We call a run at most  $T$ -contaminated, if it contains at most  $T$  tails. We approximate the distribution of the length of the longest at most  $T$ -contaminated runs. We offer a more precise approximation than the previous one.

**Keywords:** coin tossing; longest head run; asymptotic distribution; rate of convergence

**MSC:** 60F05

## 1. Introduction

Consider the usual coin tossing experiment. Let  $p$  be the probability of heads and  $q = 1 - p$  be the probability of tails. Here,  $p$  is a fixed number with  $0 < p < 1$ . We toss a coin  $N$  times independently. We write 1 for heads and 0 for tails. Therefore, we consider independent identically distributed random variables  $X_1, X_2, \dots, X_N$  with distribution  $P(X_i = 1) = p$  and  $P(X_i = 0) = q = 1 - p, i = 1, 2, \dots, N$ .

Let  $T$  be a fixed non-negative integer. We shall study the length of at most  $T$ -contaminated (in other words, at most  $T$ -interrupted) runs of heads. It means that there are at most  $T$  zeros in an  $m$ -length sequence of ones and zeros.

There are several well-known results on the length of the pure head runs. Fair coins were studied in the paper of Erdős and Rényi [1]. Almost sure limit results for the length of the longest runs containing at most  $T$  tails were obtained in [2]. Földes [3] presented asymptotic results for the distribution of the number of  $T$ -contaminated head runs, the first hitting time of a  $T$ -contaminated head run having a fixed length, and the length of the longest  $T$ -contaminated head run. Móri [4] proved an almost sure limit theorem for the longest  $T$ -contaminated head run.

Gordon, Schilling, and Waterman [5] applied extreme value theory to obtain the asymptotic behaviour of the expectation and the variance of the length of the longest  $T$ -contaminated head run. Then, accompanying distributions were obtained for the length of the longest  $T$ -contaminated head run. Ref. [6] proved results on the accuracy of the approximation to the distribution of the length of the longest head run in a Markov chain.

In this paper, we follow the lines of Arratia, Gordon, and Waterman [7], where Poisson approximation was used to find the asymptotic behaviour of the length of the longest at most  $T$ -contaminated head run. We shall use the basic results presented in [7], and give a new approximation for the distribution of the length of the longest at most  $T$ -contaminated head run. We show that for  $T > 0$  the rate of the approximation in our new



Academic Editors: Antónia Földes and Endre Csáki

Received: 1 November 2024

Revised: 29 December 2024

Accepted: 31 December 2024

Published: 3 January 2025

**Citation:** Fazekas, I.; Fazekas, B.; Fórián, L. On the Convergence Rate for the Longest at Most  $T$ -Contaminated Runs of Heads.

*Entropy* **2025**, *27*, 33. <https://doi.org/10.3390/e27010033>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

result is  $O(1/(\log(n))^2)$ , where  $\log$  denotes the logarithm to base  $1/p$ . Here and in what follows,  $f(n) = O(h(n))$  means that  $f(n)/h(n)$  is bounded as  $n \rightarrow \infty$ . We shall see that for  $T > 0$  the rate of the approximation offered by [7] is  $O(\log(\log(n))/\log(n))$ , so our result considerably improves the former result. In our opinion the much better rate  $O(\log(n)/n)$  presented without detailed proof in [7] is just a misprint, that is true only for  $T = 0$ . The main result is Theorem 1. For completeness, we give a proof of the former result, see Proposition 1. In Section 4, we present some simulation results supporting our theorem.

For  $T = 1$  and  $T = 2$ , our result is the same as our former result in [8], where a powerful lemma by Csáki, Földes and Komlós [9] was used in the proof.

## 2. The Approximation of Arratia, Gordon, and Waterman

Using the notation of [7], let  $S_i = X_1 + \dots + X_i$ , and let  $S_{n,t}$  be the largest increment in the sequence  $S_i$  in  $t$  steps; more precisely,  $S_{n,t}$  is the maximal number of heads in a window of length  $t$  starting in the first  $n$  tosses. Let  $R_n(T)$  be the length of the longest at most  $T$ -interrupted runs of heads starting in the first  $n$  tosses. (One can see that  $R_n(T)$  is the length of the longest precisely  $T$ -interrupted runs of heads starting in the first  $n$  tosses.) Then,

$$\{R_n(T) < t\} = \{S_{n,t} < t - T\}.$$

According to Theorem 1 of [7], for the distribution of  $S_{n,t}$ , we have the following approximation. For positive integers  $n, s$ , and  $t$  with  $s \leq t$  and  $s/t > p$ ,

$$|P(S_{n,t} < s) - e^{-EW}| \leq 7tP(X_1 + \dots + X_t = s) + P(X_1 + \dots + X_t > s), \tag{1}$$

$$e^{-n(\frac{s}{t}-p)P(X_1+\dots+X_t=s)} \cdot e^{-2n(1-\frac{s}{t})P(X_1+\dots+X_t=s)P(X_1+\dots+X_t>s)} \leq e^{-EW} \leq e^{-n(\frac{s}{t}-p)P(X_1+\dots+X_t=s)}. \tag{2}$$

In the above inequalities  $EW$  is the expectation of the random variable  $W$  defined in [7]. We shall use inequalities (1) and (2) with  $s = t - T$ . Using notation  $\alpha = n(\frac{s}{t} - p)P(X_1 + \dots + X_t = s)$  and  $\beta = 2n(1 - \frac{s}{t})P(X_1 + \dots + X_t = s)P(X_1 + \dots + X_t > s)$ , the above inequality is of the form

$$e^{-\alpha}e^{-\beta} \leq e^{-EW} \leq e^{-\alpha}. \tag{3}$$

In this paper, the approximation of  $e^{-\alpha}$  will serve as the main term.

Now, we shall analyse that approximation of  $R_n(T)$  which was proposed in [7]. The centering constant in [7] is

$$c_n(T) = \log n + T \log \log n - \log(T!) + \log(q^{T+1}p^{-T}). \tag{4}$$

Let  $x$  be a fixed number so that  $c_n(T) + x = t$  is an integer. We want to estimate  $P(R_n(T) - c_n(T) < x) = P(S_{n,t} < t - T)$ . In the following we shall use both  $\exp(x)$  and  $e^x$  for the usual exponential function.

**Proposition 1.** *Let  $[c_n(T)]$  be the integer part of  $c_n(T)$  and  $\{c_n(T)\} = c_n(T) - [c_n(T)]$  be its fractional part.*

*If  $T = 0$ , then for any integer  $l$ ,*

$$P(R_n(T) - [c_n(0)] < l) = \exp\left(-p^{l-\{c_n(0)\}}\right) \left(1 + O\left(\frac{\log n}{n}\right)\right). \tag{5}$$

If  $T > 0$ , then for any integer  $l$ ,

$$P(R_n(T) - [c_n(T)] < l) = \exp\left(-p^{l - \{c_n(T)\}}\right) \left(1 + O\left(\frac{\log \log n}{\log n}\right)\right). \tag{6}$$

**Remark 1.** In Corollary 3 of [7], the same remainder term  $O\left(\frac{\log n}{n}\right)$  is given for the case  $T > 0$ , too. However, in our opinion, it contains only a part of the remainder terms.

**Proof of Proposition 1.** As our remainder term and the remainder term offered by [7] are different, we give the details of the more or less simple calculation. First, we calculate the right hand side of inequality (1) for  $s = t - T$  and  $t = c_n(T) + x$ , where  $x$  is chosen so that  $t$  is an integer.

$$P(X_1 + \dots + X_t = t - T) = \binom{t}{T} p^t (q/p)^T \leq \kappa \frac{(\log n)^T}{(1/p)^{\log n + T \log \log n}} = O\left(\frac{1}{n}\right).$$

Here and in what follows,  $\kappa$  is an appropriate finite positive constant. Therefore,

$$7tP(X_1 + \dots + X_t = t - T) = O\left(\frac{\log n}{n}\right).$$

For  $T > 0$ , we have

$$\begin{aligned} P(X_1 + \dots + X_t > t - T) &\leq T \binom{t}{t - T + 1} p^{t - T + 1} \leq \kappa t^{T - 1} p^t \\ &\leq \kappa \frac{(\log n)^{T - 1}}{n (\log n)^T} = O\left(\frac{1}{n \log n}\right). \end{aligned}$$

So we obtain

$$|P(S_{n,t} < t - T) - e^{-EW}| = O\left(\frac{\log n}{n}\right). \tag{7}$$

This last formula is valid for  $T = 0$ , too.

Now, we turn to the other parts of the approximation. First, consider  $T = 0$ . Then, the main term of the approximation, i.e.,  $e^{-\alpha}$  in Formula (3) is

$$e^{-\alpha} = e^{-n\left(\frac{t}{t} - p\right)P(X_1 + \dots + X_t = t)} = e^{-p^{-\log(nq) + t}}.$$

We have to approximate  $P(R_n(0) - [c_n(0)] < l)$ , where  $l$  is an integer,  $c_n(0) = \log n + \log q$ , and  $[\cdot]$  denotes the integer part. So, we should apply the previous equality with  $t = [c_n(0)] + l$ , so we obtain

$$e^{-\alpha} = e^{-p^{l - \{c_n(0)\}}},$$

where  $\{\cdot\}$  denotes the fractional part. We see that, if  $T = 0$ , then  $\beta = 0$ , so in inequality (3), we have equality. So, for  $T = 0$ , this part of the approximation is precise, i.e., the main term does not contain a remainder part.

Now, we consider the approximation of the main term for  $T > 0$ .

$$e^{-\alpha} = e^{-n\left(\frac{t-T}{t} - p\right)P(X_1 + \dots + X_t = t - T)} = e^{-n\left(q - \frac{T}{t}\right)\binom{t}{T}q^T p^{t-T}}.$$

Now, denote by  $L$  the base  $1/p$  logarithm of the negative of the exponent, that is,  $L = \log \alpha$ . So,

$$L = \log n + \log\left(q - \frac{T}{t}\right) + \log(t(t - 1) \dots ((t - T + 1))) - \log T! + T \log q + T - t.$$

We shall use  $t = c_n(T) + x$ . Applying Taylor’s expansion of the logarithm function,  $\log(x_0 + y) = \log x_0 + \frac{y}{cx_0} - \frac{y^2}{2cx_0^2}$ , where  $\tilde{x}_0$  is between  $x_0$  and  $x_0 + y$ , and where  $c = \ln(1/p)$ , we obtain

$$L = \log n + \log q - \frac{T}{cqt} - O\left(\frac{1}{t^2}\right) + \log t^T - \frac{t^{T-1}\binom{T}{2}}{ct^T} + O\left(\frac{1}{t^2}\right) - \log T! + T \log q + T - t$$

$$= \log n + T \log t - \frac{1}{ct} \left( \frac{T}{q} + \binom{T}{2} \right) + O\left(\frac{1}{t^2}\right) - \log T! + (T + 1) \log q + T - t.$$

We insert  $t = c_n(T) + x = \log n + T \log \log n + E$ , where  $E$  is defined by the equation at hand so it does not depend on  $n$ . Using again Taylor’s expansions of the logarithm function as  $\log(x_0 + y) = \log x_0 + \frac{y}{cx_0} - \frac{y^2}{2cx_0^2} + \frac{y^3}{3cx_0^3}$ , where  $\tilde{x}_0$  is between  $x_0$  and  $x_0 + y$ , and for the  $1/t$  function, as  $\frac{1}{x_0+y} = \frac{1}{x_0} - \frac{y}{x_0^2} + \frac{y^2}{x_0^3}$ , where  $\tilde{x}_0$  is between  $x_0$  and  $x_0 + y$ , we obtain

$$L = \log n + T \left( \log \log n + \frac{T \log \log n + E}{c \log n} - \frac{(T \log \log n + E)^2}{2c(\log n)^2} + O\left(\frac{(\log \log n)^3}{(\log n)^3}\right) \right) - \frac{1}{c} \left( \frac{T}{q} + \binom{T}{2} \right) \left( \frac{1}{\log n} - \frac{T \log \log n + E}{(\log n)^2} O\left(\frac{(\log \log n)^2}{(\log n)^3}\right) \right) + O\left(\frac{1}{t^2}\right) - \log T! + (T + 1) \log q + T - t.$$

Now, using  $t = c_n(T) + x$  and inserting the value of  $c_n(T)$ , we obtain

$$L = -x + \frac{T^2 \log \log n}{c \log n} + O\left(\frac{1}{\log n}\right),$$

which implies that

$$L = -x + O\left(\frac{\log \log n}{\log n}\right),$$

and this rate is not improvable. We remark that this relation is valid for  $T = 1$ , too.

Therefore, by applying the Taylor series expansion  $e^y = 1 + y + e^{\tilde{y}} \frac{y^2}{2}$  twice, where  $\tilde{y}$  is between 0 and  $y$ , we obtain

$$e^{-\alpha} = e^{-(1/p)^L} = e^{-p^x} \left( 1 - \ln\left(\frac{1}{p}\right) \frac{T^2 \log \log n}{c \log n} + O\left(\frac{1}{\log n}\right) \right) \tag{8}$$

$$= e^{-p^{t-\{c_n(T)\}}} \left( 1 + O\left(\frac{\log \log n}{\log n}\right) \right), \tag{9}$$

and this rate is not improvable.

Now, we consider the  $e^{-\beta}$  part. Here,

$$\beta = 2 \frac{T}{t} \sum_{i=t-T+1}^t \binom{t}{i} p^i q^{t-i} n \binom{t}{T} p^{t-T} q^T$$

with  $t = c_n(T) + x = \log n + T \log \log n + E$ . The largest term in the above sum is the first one, and it is

$$\binom{t}{T-1} p^t \left(\frac{q}{p}\right)^{T-1} = O\left(\frac{1}{n \log n}\right).$$

Then,

$$\binom{t}{T} p^{t-T} q^T = O\left(\frac{1}{n}\right).$$

Using Taylor’s expansion,

$$\frac{T}{t} = O\left(\frac{1}{\log n}\right).$$

So,  $\beta = O(1/n(\log n)^2)$ , and

$$e^{-\beta} = 1 - O\left(\frac{1}{n(\log n)^2}\right).$$

Therefore,

$$\begin{aligned} e^{-\alpha}e^{-\beta} &= e^{-p^{l-\{c_n(T)\}}} \left(1 + O\left(\frac{\log \log n}{\log n}\right)\right) \left(1 - O\left(\frac{1}{n(\log n)^2}\right)\right) \\ &= e^{-p^{l-\{c_n(T)\}}} \left(1 + O\left(\frac{\log \log n}{\log n}\right)\right). \end{aligned} \tag{10}$$

□

### 3. A New Approximation

**Theorem 1.** Let  $T \geq 1$  be an integer. Let

$$\begin{aligned} \tilde{c}_n(T) &= \log(qn) + T \log(\log(qn)) \\ &+ T^2 \frac{\log(\log(qn))}{c \log(qn)} - \frac{T}{cq_0 \log(qn)} - \frac{T^3}{2c} \left(\frac{\log(\log(qn))}{\log(qn)}\right)^2 \\ &+ T^2 \frac{\log(\log(qn))}{cq_0(\log(qn))^2} + T^3 \frac{\log(\log(qn))}{(c \log(qn))^2} \\ &+ \left(T \log\left(\frac{q}{p}\right) - \log(T!)\right) \left(1 + \frac{T}{c \log(qn)} - T^2 \frac{\log(\log(qn))}{c(\log(qn))^2}\right), \end{aligned} \tag{11}$$

where  $\log$  denotes the logarithm to base  $1/p$ ,  $c = \ln(1/p)$ ,  $\ln$  denotes the natural logarithm to base  $e$ , and  $q_0 = \frac{2q}{2+Tq-q}$ . Let  $[\tilde{c}_n(T)]$  denote the integer part of  $\tilde{c}_n(T)$ , while  $\{\tilde{c}_n(T)\}$  denotes the fractional part of  $\tilde{c}_n(T)$ , i.e.  $\{\tilde{c}_n(T)\} = \tilde{c}_n(T) - [\tilde{c}_n(T)]$ .

Then,

$$\begin{aligned} P(R_n(T) - [\tilde{c}_n(T)] < l) & \\ &= \exp\left(-p^{(l-\{\tilde{c}_n(T)\})\left(1 - \frac{T}{c \log(qn)} + T^2 \frac{\log(\log(qn))}{c(\log(qn))^2}\right)}\right) \left(1 + O\left(\frac{1}{(\log n)^2}\right)\right) \end{aligned} \tag{12}$$

for any integer  $l$ , where  $f(n) = O(h(n))$  means that  $f(n)/h(n)$  is bounded as  $n \rightarrow \infty$ .

**Proof.** We use the same approach as in the previous section. First, we calculate the right hand side of inequality (1) for  $s = t - T$  and  $t = \tilde{c}_n(T) + x$ , where  $x$  is chosen so that  $t$  is an integer. As

$$\tilde{c}_n(T) = \log(n) + T \log(\log(n)) + O(1),$$

we obtain

$$P(X_1 + \dots + X_t = t - T) = \binom{t}{T} p^t (q/p)^T \leq \kappa \frac{(\log n)^T}{(1/p)^{\log n + T \log \log n}} = O\left(\frac{1}{n}\right).$$

Therefore,

$$7tP(X_1 + \dots + X_t = t - T) = O\left(\frac{\log n}{n}\right).$$

Similarly,

$$P(X_1 + \dots + X_t > t - T) \leq \kappa t^{T-1} p^t = O\left(\frac{1}{n \log n}\right).$$

So,

$$|P(S_{n,t} < t - T) - e^{-EW}| = O\left(\frac{\log n}{n}\right). \tag{13}$$

Now, we turn to the approximation of the main term  $e^{-\alpha}$ . Denote by  $L$  again the base  $1/p$  logarithm of the negative of the exponent, so

$$\begin{aligned} L &= \log \alpha \\ &= \log n + \log(q - T/t) + \log(t(t-1) \dots ((t-T+1))) - \log T! + T \log q + T - t. \end{aligned}$$

We shall apply it for  $t = \tilde{c}_n(T) + x$ . Therefore,

$$\begin{aligned} L &= \log\left(q - \frac{T}{t}\right) + \log n + \log\left(t^T - \frac{T(T-1)}{2}t^{T-1} + O(t^{T-2})\right) - t \\ &\quad + \log((q/p)^T) - \log(T!) \\ &= \log\left(q - \frac{T}{t}\right) + \log n + \log(t^T) - \frac{T(T-1)t^{T-1}}{2ct^T} + O\left(\frac{1}{t^2}\right) - t \\ &\quad + \log((q/p)^T) - \log(T!) \\ &= \log q - \frac{T}{cqt} + \log n + T \log t - \frac{T(T-1)}{ct} - t \\ &\quad + \log((q/p)^T) - \log(T!) + O\left(\frac{1}{(\log n)^2}\right) \\ &= \log(qn) - \frac{T}{cq_0t} + T \log t - t + \log((q/p)^T) - \log(T!) + O\left(\frac{1}{(\log n)^2}\right), \end{aligned}$$

where we applied Taylor's expansion of the log function up to the second order and used the notation  $q_0 = \frac{2q}{2+Tq-q}$ .

Introduce notation

$$\begin{aligned} D &= -\frac{T^3}{2c} \left(\frac{\log(\log(qn))}{\log(qn)}\right)^2 + T^2 \frac{\log(\log(qn))}{cq_0(\log(qn))^2} + T^3 \frac{\log(\log(qn))}{(c \log(qn))^2} \\ &\quad + \left(T \log\left(\frac{q}{p}\right) - \log(T!)\right) \left(\frac{T}{c \log(qn)} - T^2 \frac{\log(\log(qn))}{c(\log(qn))^2}\right), \end{aligned} \tag{14}$$

$$B = T^2 \frac{\log(\log(qn))}{c \log(qn)} - \frac{T}{cq_0 \log(qn)} + D \tag{15}$$

and

$$A = T \log(\log(qn)) + B.$$

Then,  $t = \tilde{c}_n(T) + x = \tilde{c}_n(T) + l - \{\tilde{c}_n(T)\}$ , where  $l$  is an integer, so

$$t = T \log\left(\frac{q}{p}\right) - \log(T!) + \log(qn) + A + l - \{\tilde{c}_n(T)\}.$$

Inserting this value of  $t$  into the term  $-t$  of  $L$ , we obtain

$$L = -\frac{T}{cq_0t} + T \log t - A - l + \{\tilde{c}_n(T)\} + O\left(\frac{1}{(\log n)^2}\right).$$

Then, use Taylor’s expansion for the function  $1/t$  to obtain

$$L = -\frac{T}{cq_0 \log(qn)} + T^2 \frac{\log(\log(qn))}{cq_0(\log(qn))^2} + T \log(\log(qn)) + T \log(\log(qn)) + B + \log((q/p)^T) - \log(T!) + l - \{\tilde{c}_n(T)\} - A - l + \{\tilde{c}_n(T)\} + O\left(\frac{1}{(\log n)^2}\right).$$

Now, by Taylor’s expansion for the  $\log(x)$  function, we obtain

$$L = -\frac{T}{cq_0 \log(qn)} + T^2 \frac{\log(\log(qn))}{cq_0(\log(qn))^2} + T \log(\log(qn)) + \frac{T(T \log(\log(qn)) + B + \log((q/p)^T) - \log(T!) + l - \{\tilde{c}_n(T)\})}{c \log(qn)} - \frac{1}{2} \frac{T(T \log(\log(qn)) + B + \log((q/p)^T) - \log(T!) + l - \{\tilde{c}_n(T)\})^2}{c(\log(qn))^2} - A - l + \{\tilde{c}_n(T)\} + O\left(\frac{1}{(\log n)^2}\right).$$

Now, we can omit  $B$  from the quadratic term. Then, we apply  $A = T \log(\log(qn)) + B$ , so we obtain

$$L = -\frac{T}{cq_0 \log(qn)} + \frac{T^2 \log(\log(qn))}{cq_0(\log(qn))^2} + \frac{T^2 \log(\log(qn))}{c \log(qn)} + \frac{T(\log((q/p)^T) - \log(T!))}{c \log(qn)} + \frac{T^3 \log(\log(qn))}{(c \log(qn))^2} - \frac{T^2}{q_0(c \log(qn))^2} + \frac{TD}{c \log(qn)} + \frac{T(l - \{\tilde{c}_n(T)\})}{c \log(qn)} - \frac{1}{2} \frac{T^3(\log(\log(qn)))^2}{c(\log(qn))^2} - \frac{1}{2} \frac{T(\log((q/p)^T) - \log(T!) + l - \{\tilde{c}_n(T)\})^2}{c(\log(qn))^2} - \frac{2T}{2} \frac{T \log(\log(qn))(\log((q/p)^T) - \log(T!) + l - \{\tilde{c}_n(T)\})}{c(\log(qn))^2} - B - l + \{\tilde{c}_n(T)\} + O\left(\frac{1}{(\log n)^2}\right) = (l - \{\tilde{c}_n(T)\}) \left(\frac{T}{c \log(qn)} - \frac{T^2 \log(\log(qn))}{c(\log(qn))^2} - 1\right) + O\left(\frac{1}{(\log n)^2}\right).$$

So,

$$e^{-\alpha} = e^{-p}^{(l - \{\tilde{c}_n(T)\}) \left(1 - \frac{T}{c \log(qn)} + \frac{T^2 \log(\log(qn))}{c(\log(qn))^2}\right) + O\left(\frac{1}{(\log n)^2}\right)}.$$

Using Taylor’s expansion again,

$$e^{-\alpha} = e^{-p}^{(l - \{\tilde{c}_n(T)\}) \left(1 - \frac{T}{c \log(qn)} + \frac{T^2 \log(\log(qn))}{c(\log(qn))^2}\right)} \left(1 + O\left(\frac{1}{(\log n)^2}\right)\right).$$

Now, turn to the  $e^{-\beta}$  part, where

$$\beta = 2 \frac{T}{t} \sum_{i=t-T+1}^t \binom{t}{i} p^i q^{t-i} n \binom{t}{T} p^{t-T} q^T$$

and  $t = \tilde{c}_n(T) + x$ . Simple calculations shows that  $\beta \leq \kappa(1/n(\log n)^2)$ , and so

$$e^{-\beta} = 1 + O\left(\frac{1}{n(\log n)^2}\right).$$

Therefore,

$$e^{-\alpha} e^{-\beta} = e^{-p} \left(1 - \{\epsilon_n(T)\}\right) \left(1 - \frac{T}{c \log(qn)} + \frac{T^2 \log(\log(qn))}{c(\log(qn))^2}\right) \left(1 + O\left(\frac{1}{(\log n)^2}\right)\right).$$

□

### 4. Simulation Results

We performed several computer simulation studies for certain fixed values of  $p$  and  $T$ . Here, we present the results of three simulations. The length of each simulated sequence was  $N = 10^6$ , and  $s = 2000$  was the number of repetitions of the  $N$ -length sequences in each case. In each case, the number of contaminations was  $T = 3$ .

Figures 1–3 present the results of the simulations. The left hand side of each figure shows the empirical distribution function of the longest at most  $T$ -contaminated run and its approximation suggested by our Theorem 1. The asterisk (i.e., \*) denotes the result of the simulation, i.e., the empirical distribution of the longest at most  $T$ -contaminated run, and the circle (o) denotes the approximation offered by Theorem 1. The right hand side of each figure shows the approximation by the former result. The asterisk denotes the result of the simulation again, and the circle (o) denotes the approximation offered by Proposition 1. The simulation results support that our new theorem offers a better approximation than the previous one.

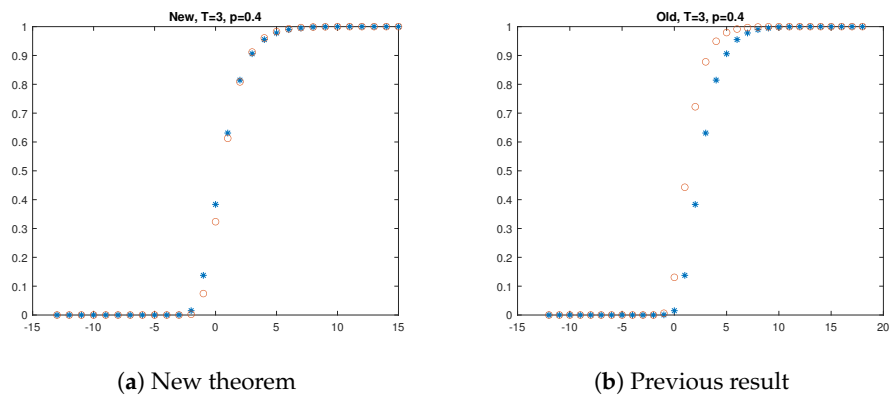


Figure 1. Longest at most  $T = 3$  contaminated run when  $p = 0.4$ .

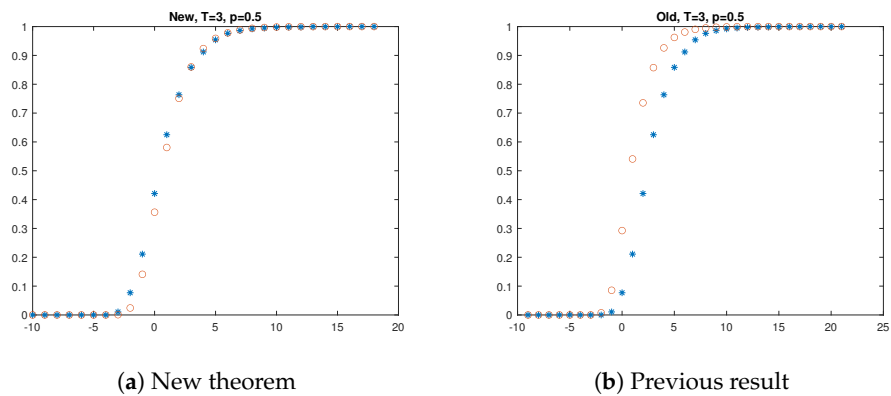
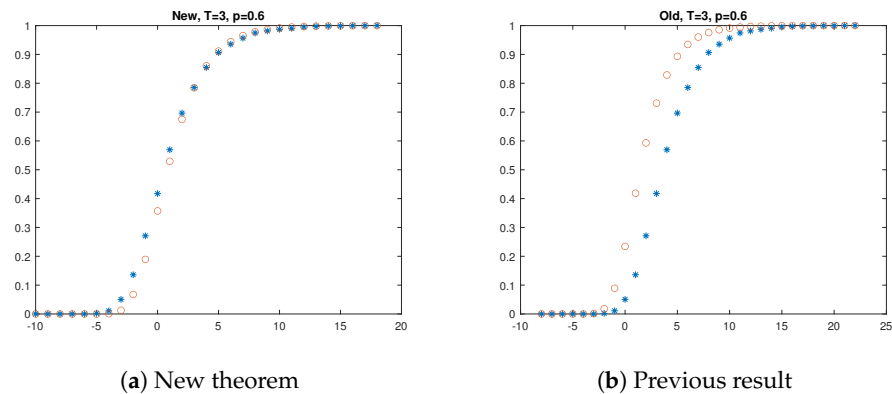


Figure 2. Longest at most  $T = 3$  contaminated run when  $p = 0.5$ .



**Figure 3.** Longest at most  $T = 3$  contaminated run when  $p = 0.6$ .

## 5. Discussion

We were able to obtain a practically applicable approximation for the distribution of the longest at most  $T$ -contaminated head-run. We presented both detailed mathematical proof and simulation evidence.

**Author Contributions:** Conceptualization, I.F.; methodology, I.F.; software, L.F.; validation, B.F.; formal analysis, I.F., B.F. and L.F.; investigation, I.F., B.F. and L.F.; writing—original draft preparation, I.F., B.F. and L.F.; writing—review and editing, I.F.; visualization, L.F.; supervision, I.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Erdős, P.; Rényi, A. On a new law of large numbers. *J. Analyse Math.* **1970**, *23*, 103–111. [[CrossRef](#)]
2. Erdős, P.; Révész, P. On the length of the longest head-run. In *Topics in Information Theory (Second Colloq., Keszthely, 1975)*; Colloq. Math. Soc. János Bolyai; North-Holland: Amsterdam, The Netherlands, 1977; Volume 16, pp. 219–228.
3. Földes, A. The limit distribution of the length of the longest head-run. *Period. Math. Hungar.* **1979**, *10*, 301–310. [[CrossRef](#)]
4. Móri, T.F. The a.s. limit distribution of the longest head run. *Canad. J. Math.* **1993**, *45*, 1245–1262. [[CrossRef](#)]
5. Gordon, L.; Schilling, M.F.; Waterman, M.S. An extreme value theory for long head runs. *Probab. Theory Relat. Fields* **1986**, *72*, 279–287. [[CrossRef](#)]
6. Novak, S.Y. On the length of the longest head run. *Statist. Probab. Lett.* **2017**, *130*, 111–114. [[CrossRef](#)]
7. Arratia, R.; Gordon, L.; Waterman, M.S. The Erdős-Rényi Law in Distribution, for Coin Tossing and Sequence Matching. *Ann. Statist.* **2008**, *18*, 539–570. [[CrossRef](#)]
8. Fazekas, I.; Fazekas, B.; Ochieng Suja, M. Convergence rate for the longest  $T$ -contaminated runs of heads. *Statist. Probab. Lett.* **2024**, *208*, 110059. [[CrossRef](#)]
9. Csáki, E.; Földes, A.; Komlós, J. Limit theorems for Erdős-Rényi type problems. *Studia Sci. Math. Hungar.* **1987**, *22*, 321–332.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.