Asymptotic Behaviors of The Size of The Largest Cluster in One Dimensional Percolation

Yong-Hua Mao[1], Feng Wang[2] and Xian-Yuan Wu[3*]

[1]School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, China. Email: maoyh@bnu.edu.cn

[2,3]School of Mathematical Sciences, Capital Normal University, Beijing, 100048, China. Email: wangf@mail.cnu.edu.cn; wuxy@mail.cnu.edu.cn

Abstract: This paper focuses on the asymptotic behaviors of the length of the largest 1-cluster in a finite iid Bernoulli sequence. We first reveal a critical phenomenon on the length and then study its limit distribution.

1 Introduction and statement of the results

Percolation is a canonical model on quenched spatial disorder [12], it offers challenging problems in probability theory of relevance to statistical physics [8, 11]. In subcritical percolation, it is widely believed by physicists that the mean size of the largest cluster in a finite system of size N scales like $s_\xi \ln N$. Where s_ξ is called the *crossover size* (since large clusters of size much smaller than s_ξ behave *critically*, while much larger clusters behave *subcritically* [12]). A heuristic theory of the finite size scaling of the largest cluster size in subcritical percolation is presented and supported by numerical simulations in [3]. Note that, besides the prediction on mean size growth, [3] also suggests that as $N \to \infty$ the distribution function of the size of the largest cluster converges to the Fisher-Tippett distribution $e^{-e^{-x}}$ [7].

AMS classification: 60K 35

Key words and phrases: largest cluster; critical phenomenon; skip-free chain; Fibonacci Series.

*Corresponding author

Research supported in part by 973 Project under grant 2011CB808000[1], the Natural Science Foundation of China under grant 10971143[2,3] and the Foundation of Beijing Education Bureau under grant 09224010003[2,3].

Mathematically rigorous results on the size of the largest cluster in subcritical percolation can be found in [4]. Actually, from certain scaling axioms verified for $d = 2$ and believed to hold for $d \le d_c = 6$, Borgs *et al.* have proved in [4] that the mean largest cluster size grows like $s'_\xi \ln(N/s'_\xi)$ as $N/s'_\xi \to \infty$. Where s'_ξ is a corresponding crossover size based on $\xi'(p)$, a correlation length defined in terms of *sponge-crossing probabilities*.

In this paper, we will study the asymptotic behaviors of the size of the largest cluster in one dimension. Usually, people pay little attention to an 1-dimensional percolation problem, because percolation problems in one dimension are always trivial. But for problems on largest cluster, it seems that the situation is quite different. A relatively deeper analysis will reveal that the problem is far from easy or trivial. On the contrary, one will see later that it really offers interesting problems in probability theory.

Now, for $N \ge 1$, write $\mathbb{Z}_N$ as the sublattice $\{1, 2, \cdots, N\}$. For given $p \in (0, 1)$, let us consider the general *site* percolation in $\mathbb{Z}_N$. For any $i \in \mathbb{Z}_N$, independently, i is declared *open* with probability p and is declared *closed* otherwise. Write $C_N(i)$ as the open cluster containing i and $|C_N(i)|$ as the cardinality of $C_N(i)$. Note that in case of i is closed, $C_N(i) = \phi$ and $|C_N(i)| = 0$. Let S_N be the size of the largest open cluster, i.e.

$$S_N := \max_{1 \le i \le N} |C_N(i)|. \tag{1.1}$$

The goal of this paper is to determine the asymptotics of S_N as $N \to \infty$. Without use of the terms in percolation theory, S_N can be stated directly as the longest sequence of only 'heads' in N flips of a coin. S_N bears some resemblance to the longest increasing subsequence of a random permutation of $1, 2, \cdots, N$ [1, 2], although it seems much simpler than the latter.

The original motivation for this work came from the percolation problem of arbitrary words in one dimension studied by Grimmett, Liggett and Richthammer [10]. Let $W = (w_1, w_2, \ldots)$ be an infinite word with $w_i = 0$ or 1, $Y = (Y_1, Y_2, \ldots,)$ be an iid Bernoulli sequence such that $\mathbb{P}(Y_i = 1) = \mathbb{P}(Y_i = 0) = \frac{1}{2}$. For any integer $M \ge 2$, W is called M-seen in Y if there exists a sequence $\{m_i : i \ge\}$ of integers such that $Y_{m_i} = w_i$ and $1 \le m_i - m_{i-1} \le M$ for each $i \ge 1$. (By default, we take $m_0 = 0$.) For any given W, it is believed that the probability that W is M-seen in Y equals zero for any $M \ge 2$, and a positive answer is given for $M = 2$ in [10]. An easier problem should be the corresponding embedding problem of a random word $X = (X_1, X_2, \ldots,)$, an iid Bernoulli sequence (independent to Y) with $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$. Let $X_N = (X_1, X_2, \ldots, X_N)$ and l_N be the length of the largest 0-cluster in X_N. Then it follows from part a) of Theorem 3 in [10] that, if for some $c < \frac{1}{M \ln 2}$

$$\mathbb{P}(l_N \leq c \ln N) \to 1 \quad as\ N \to \infty, \tag{1.2}$$

then

$$\mathbb{P}(X\ is\ M - seen\ in\ Y) = 0.$$

Whereupon, a problem arises: Does l_N really scale as required in (1.2)?

Now, we state the first result of the paper as the following, which exhibits a critical phenomenon on S_N.

Theorem 1.1 *Let $\xi_c = \xi_c(p) = 1/\ln(\frac{1}{p})$. Then*

1. *if $\xi > \xi_c$, then*
$$\lim_{N\to\infty} \mathbb{P}(S_N \geq \xi \ln N) = 0; \tag{1.3}$$
2. *if $\xi < \xi_c$, then*
$$\lim_{N\to\infty} \mathbb{P}(S_N \geq \xi \ln N) = 1; \tag{1.4}$$
3. *if $\xi = \xi_c$, then*
$$1 - e^{-q} = \liminf_{N\to\infty} \mathbb{P}(S_N \geq \xi \ln N) < \limsup_{N\to\infty} \mathbb{P}(S_N \geq \xi \ln N) = 1 - e^{-q/p}, \tag{1.5}$$
where $q = 1 - p$.

Remark 1.1 *For the embedding problem of random word X discussed above, it follows from (1.2) and Theorem 1.1 that*

$$\mathbb{P}(X\ is\ M - seen\ in\ Y) = 0$$

for all $2 \leq M < -\ln(1-p)/\ln 2$.

For the convergence speed in items 1, 2 of Theorem 1.1, we have the following large deviation results.

Corollary 1.2 *For any $x > 0$, we have*

$$\lim_{N\to\infty} \frac{-1}{\ln N} \ln \mathbb{P}(S_N - \xi_c \ln N \geq x \ln N) = x \ln(1/p) \tag{1.6}$$

and

$$\lim_{N\to\infty} \frac{1}{\ln N} \ln\{-\ln \mathbb{P}(S_N - \xi_c \ln N \leq -x \ln N)\} = x \ln(1/p). \tag{1.7}$$

Now we turn to discuss the possible convergence on S_N. For any $\lambda \in [0,1]$, define χ_λ be the integer-valued Fisher-Tippett random variable with distribution function

$$F_\lambda(x) := \mathbb{P}(\chi_\lambda \le x) = \exp\{-qp^{1-\lambda}p^{\lfloor x \rfloor}\}, \quad x \in \mathbb{R}, \tag{1.8}$$

where $\lfloor \cdot \rfloor$ denote the the integer part of number $\cdot$. Note that the distribution function $e^{-ae^{-bx}}$, where $a, b > 0$ are constants, was first discovered by Fisher and Tippett [7]. It is also called the Gumbel distribution in honor of Gumbel's pioneering work on application [9]. Clearly, $F_\lambda(x)$ (resp., $1 - F_\lambda(x)$) goes to zero fast enough as $x \to -\infty$ (resp., $x \to \infty$), this implies that $\mathbb{E}\chi_\lambda$ and $\mathrm{Var}\chi_\lambda$ exist. Furthermore, we have

$$-\infty < M_1 := \inf_{0\le\lambda\le1} \mathbb{E}\chi_\lambda < \sup_{0\le\lambda\le1} \mathbb{E}\chi_\lambda =: M_2 < \infty, \tag{1.9}$$

and

$$0 < \Sigma_1 := \inf_{0\le\lambda\le1} \mathrm{Var}\chi_\lambda < \sup_{0\le\lambda\le1} \mathrm{Var}\chi_\lambda =: \Sigma_2 < \infty. \tag{1.10}$$

For any $N \ge 1$, let $\chi_N = S_N - \lfloor \xi_c \ln N \rfloor$ and $\bar{\chi}_N = S_N - \xi_c \ln N$, denote by F_N, $\bar{F}_N$ the distribution function of χ_N, $\bar{\chi}_N$ respectively.

Our second result concerns the convergence of S_N in distribution after appropriate centering.

Theorem 1.3 *For any $\lambda \in [0,1]$, let $\{N_{\lambda,j}\}_{j\ge1}$ be the subsequence such that $\lim_{j\to\infty} \lambda_j = \lambda$, where $\lambda_j := \xi_c \ln N_{\lambda,j} - \lfloor \xi_c \ln N_{\lambda,j} \rfloor$. Then as $j \to \infty$,*

$$\chi_{\lambda,j} := \chi_{N_{\lambda,j}} \longrightarrow \chi_\lambda \quad \textit{in distribution.} \tag{1.11}$$

Or equivalently, as $j \to \infty$,

$$\bar{\chi}_{\lambda,j} := \bar{\chi}_{N_{\lambda,j}} \longrightarrow \chi_\lambda - \lambda \ \textit{in distribution.} \tag{1.12}$$

In order to show that $\mathbb{E}S_N$ scales as $\xi_c \ln N$, we have to consider the following convergence on moments.

Theorem 1.4 *Suppose $\lambda \in [0,1]$ and $\{N_{\lambda,j}\}$ is a subsequence such that as $j \to \infty$, $\chi_{\lambda,j}$ converges to χ_λ in distribution. Then, for any $m = 1, 2, \ldots$*

$$\lim_{j\to\infty} \mathbb{E}(\chi_{\lambda,j}^m) = \mathbb{E}(\chi_\lambda^m). \tag{1.13}$$

As a consequence, we have

$$\lim_{N\to\infty} \frac{\mathbb{E}S_N}{\ln N} = \xi_c, \tag{1.14}$$

and

$$\Sigma_1 \le \liminf_{N\to\infty} \mathrm{Var} S_N < \limsup_{N\to\infty} \mathrm{Var} S_N = \Sigma_2, \tag{1.15}$$

where ξ_c is given in the statement of Theorem 1.1 and Σ_1, Σ_2 are given in (1.10).

Remark 1.2 *Equation (1.14) indicates that the crossover size in 1-dimensional percolation is $\xi_c(p) = 1/\ln\frac{1}{p}$. Obviously, $\xi_c(p) \sim \frac{p}{1-p} \to \infty$ as $p \to 1$, the critical probability of percolation in one dimension.*

The paper is arranged as follows. In section 2, we transform the problem on S_N to the corresponding problem on a hitting time of a skip-free Markov chain, then we give estimates to the decay rate of the hitting time distribution. In section 3, by using the estimates given in section 2, we prove Theorems 1.1 and 1.3. Finally, we prove Theorem 1.4 in section 4.

2 A skip-free Markov chain and its hitting time

Let us consider $X = \{X_n : n \ge 0\}$, a discrete time skip-free Markov chain on the nonnegative integers with $X_0 = 0$ and transition probability

$$p_{i,j} := \mathbb{P}(X_{n+1} = j \mid X_n = i) = \begin{cases} p, & \text{if } \ j = i+1 \\ \\ q = 1-p, & \text{if } \ j = 0 \end{cases}. \tag{2.1}$$

For any integer $k \ge 1$, define the hitting time of state k as

$$T_k = \inf\{n : X_n = k\}. \tag{2.2}$$

Then, it is straightforward to check that S_N and T_k have the following *dual* relation

$$\mathbb{P}(S_N \ge k) = \mathbb{P}(T_k \le N), \ for\ any\ \ k \ge 1. \tag{2.3}$$

Thus, we have transformed our problem from S_N to T_k. Note that the distribution of T_k is well studied in [5, 6]: let $X' = \{X'_n : n \ge 0\}$ be obtained from X by making k an absorbing state, and let P denote the transition matrix for X', then T_k has probability generating function

$$u \mapsto \prod_{j=0}^{k-1} \left[\frac{(1-\theta_j)u}{1-\theta_j u}\right], \tag{2.4}$$

where $\theta_0, \theta_1, \ldots, \theta_{k-1}$ are the k non-unit eigenvalues of P.

The generating function (2.4) is of course a perfect answer to problems on T_k, but it seems hard to be used directly in our problem. Actually, on the distribution of T_k, we need a concrete estimation (especially in k) rather than a complete theoretic expression.

Given $k \geq 1$, let $P_n = \mathbb{P}(T_k = k+n)$, for $n = 0, 1, 2, \ldots$. Then we have

Lemma 2.1 $\{P_n : n \geq 0\}$ *forms the following generalized Fibonacci series*

$$P_n = \begin{cases} p^k; & n = 0 \\ a_1 P_{n-1} + a_2 P_{n-2} + \cdots + a_n P_0; & 1 \leq n \leq k \\ a_1 P_{n-1} + a_2 P_{n-2} + \cdots + a_k P_{n-k}; & n \geq k+1 \end{cases} \tag{2.5}$$

where $a_i = qp^{i-1}$, $i = 1, 2, \ldots, k$. *In particular* $P_k = P_{k-1} = \cdots = P_2 = P_1 = qp^k$.

Proof. First of all, $T_k = 0$ if and only if $X_1 = 1, X_2 = 2, \ldots, X_k = k$, this implies $P_0 = p^k$. Let $\tau_0^+ := \inf\{n \geq 1 : X_n = 0\}$ be the first returning time of state 0. Then for $n \geq 1$,

$$\begin{aligned} P_n &= \mathbb{P}(\tau_0^+ \leq n, T_k = k+n) = \mathbb{P}(\tau_0^+ \leq n \wedge k, T_k = k+n) \\ &= \sum_{i=1}^{n\wedge k} \mathbb{P}(\tau_0^+ = i, T_k = k+n) = \sum_{i=1}^{n\wedge k} \mathbb{P}(\tau_0^+ = i)\mathbb{P}(T_k = k+n \mid \tau_0^+ = i) \\ &= \sum_{i=1}^{n\wedge k} a_i P_{n-i}, \end{aligned}$$

where $n \wedge k = \min\{n, k\}$. Thus (2.5) follows immediately. Noticing that $a_i = pa_{i-1}$, $i = 2, 3, \ldots, k$, we have

$$\begin{aligned} P_k &= a_1 P_{k-1} + p(a_1 P_{k-2} + \cdots + a_{k-1} P_0) = qP_{k-1} + pP_{k-1} \\ &= P_{k-1} = P_{k-2} = \cdots = P_1 = qP_0 = qp^k. \end{aligned} \tag{2.6}$$

□

Lemma 2.2 *Let* $\alpha_k = (1 - qp^{k-1})\mathbb{I}_{\{p \leq q\}} + (1 - p^k)\mathbb{I}_{\{p > q\}}$ *and* $\beta_k = 1 - qp^k = 1 - P_1$. *Then*

$$\alpha_k \leq \frac{P_{n+1}}{P_n} \leq \beta_k, \quad for\ all \quad n \geq k. \tag{2.7}$$

Where $\mathbb{I}_{\{\cdot\}}$ *be the usual indicator function of set* $\{\cdot\}$.

Proof. For any $n \geq k$ we have

$$P_{n+1} = qP_n + p\sum_{i=1}^{k} a_i P_{n-i} - pa_k P_{n-k} = P_n - P_1 P_{n-k}. \tag{2.8}$$

Together with (2.6), (2.8) implies that P_n decreases in n and then

$$\frac{P_{n+1}}{P_n} = 1 - P_1 \frac{P_{n-k}}{P_n} \le 1 - P_1 = \beta_k, \quad for\ n \ge k.$$

For the lower bound stated in (2.7), let us consider the following two cases: 1), $k \le n < 2k$; and 2), $n \ge 2k$.

Case 1). In case of $n = k$, (2.8) implies

$$\frac{P_{k+1}}{P_k} = 1 - \frac{P_1}{P_k} P_0 = 1 - \frac{P_1}{P_1} P_0 = 1 - p^k \ge \alpha_k.$$

In case of $k < n < 2k$, using (2.8) iteratively, we have

$$\begin{aligned} P_n &= P_{n-1} - P_1 P_{n-k-1} = P_{n-2} - P_1 P_{n-k-2} - P_1 P_{n-k-1} \\ &= \cdots = P_k - P_1[P_0 + P_1 + \cdots + P_{n-k-1}] \\ &= P_1[1 - p^k - (n-k-1)qp^k] \end{aligned}$$

then

$$\frac{P_{n+1}}{P_n} = 1 - \frac{P_1 P_{n-k}}{P_n} = 1 - \frac{qp^k}{1 - p^k - (n-k-1)qp^k}. \tag{2.9}$$

For $p \le q$, we have

$$\frac{p}{q} p^{k-1} + (n-k-1)p^k \le \left(\frac{1}{2}\right)^{k-1} + (k-1)\left(\frac{1}{2}\right)^k \le 1, \quad \forall\ k \ge 1 \tag{2.10}$$

and for $q < p$, noticing that $h(k,q) := [1 + (k-1)q]p^{k-1}$ decreases in q and $h(k,0) = 1$, we have

$$p^{k-1} + (n-k-1)qp^{k-1} \le p^{k-1} + (k-1)qp^{k-1} = [1 + (k-1)q]p^{k-1} \le 1,\ \forall\ k \ge 1. \tag{2.11}$$

It follows from (2.10) and (2.11) that

$$p^k + (n-k-1)qp^k \le p \vee q, \tag{2.12}$$

where $p \vee q = \max\{p, q\}$. Together with (2.9), (2.12) implies $\frac{P_{n+1}}{P_n} \ge \alpha_k$. Thus we finished the proof of Case 1).

Case 2). Using (2.8) iteratively again, we have

$$P_n = P_{n-k} - P_1[P_{n-k-1} + P_{n-k-2} + \cdots + P_{n-2k+1} + P_{n-2k}], \tag{2.13}$$

then

$$\frac{P_{n+1}}{P_n} = 1 - P_1 \frac{P_{n-k}}{P_n} = 1 - P_1 \left\{1 - P_1 \left[\frac{P_{n-k-1} + \cdots + P_{n-2k}}{a_1 P_{n-k-1} + \cdots + a_k P_{n-2k}}\right]\right\}^{-1}.$$

Using the fact that

$$[p \wedge q]p^{k-1}[P_{n-k-1} + \cdots + P_{n-2k}] \leq a_1 P_{n-k-1} + \cdots + a_k P_{n-2k},$$

we obtain $\dfrac{P_{n+1}}{P_n} \geq \alpha_k$ and then finish the proof of Case 2). □

Lemma 2.3 *Let* $\gamma_k = 1 - f(k,p)P_1$*, where* $f(k,p) = \dfrac{1}{1-g(k,p)}$ *and*

$$g(k,p) = \left[p^k \sum_{i=0}^{k-1} \beta_k^i\right] \times \left[\sum_{i=0}^{k-1} p^{k-1-i}\alpha_k^i\right]^{-1}, \tag{2.14}$$

where α_k, β_k *are given in the statement of Lemma 2.2. Then*

$$\frac{P_{n+1}}{P_n} \geq \gamma_k, \quad \textit{for all } n \geq 3k. \tag{2.15}$$

Proof. By (2.13),

$$\begin{aligned} \frac{P_{n-k}}{P_n} &= \left\{1 - \frac{p^k(P_{n-k-1} + \cdots + P_{n-2k})}{P_{n-k-1} + pP_{n-k-2} + \cdots + p^{k-1}P_{n-2k}}\right\}^{-1} \\ &= \left\{1 - \left[p^k \sum_{i=1}^{k} \frac{P_{n-k-i}}{P_{n-2k}}\right] \times \left[\sum_{i=1}^{k} p^{i-1}\frac{P_{n-k-i}}{P_{n-2k}}\right]^{-1}\right\}^{-1} \end{aligned} \tag{2.16}$$

for $n \geq 3k$. By Lemma 2.2, we know that

$$\alpha_k^{k-i} \leq \frac{P_{n-k-i}}{P_{n-2k}} \leq \beta_k^{k-i}, \quad \textit{for all } 1 \leq i \leq k.$$

Then

$$\frac{P_{n-k}}{P_n} \leq \frac{1}{1-g(k,p)} = f(k,p)$$

and

$$\frac{P_{n+1}}{P_n} = 1 - P_1\frac{P_{n-k}}{P_n} \geq 1 - f(k,p)P_1 = \gamma_k$$

for all $n \geq 3k$. □

3 Proofs of Theorems 1.1 and 1.3

In this section, we give proofs to Theorem 1.1, Corollary 1.2 and Theorem 1.3. We first prove items 1, 2 of Theorem 1.1 and Corollary 1.2. Note that item 3 of Theorem 1.1 is a direct consequence of Theorem 1.3.

Proofs of Theorem 1.1 and Corollary 1.2. a) Let $\xi = \xi_c + x$, $x > 0$ and $k = k(N) = \lfloor \xi \ln N \rfloor$. Then, by the monotonicity of P_n and Lemma 2.2, we have

$$\begin{aligned} \mathbb{P}(S_N \geq k) &= \mathbb{P}(T_k \leq N) = \sum_{n=0}^{N-k} P_n = \sum_{n=0}^{k-1} P_n + \sum_{n=k}^{N-k} P_n \\ &\leq kP_0 + \sum_{n=k}^{N-k} P_k \frac{P_n}{P_k} \leq kp^k + P_1 \sum_{n=k}^{N-k} \beta_k^{n-k} \\ &= 1 + kp^k - (1 - qp^k)^{N-2k+1}. \end{aligned} \tag{3.1}$$

Clearly, there exists constant $C_1 > 0$ such that

$$(1 - qp^k)^{N-2k+1} = \left[(1 - qp^{k(N)})^{\frac{1}{qp^{k(N)}}}\right]^{[N-2k(N)+1]qp^{k(N)}} \geq \exp\{-C_1 N^{-x \ln(\frac{1}{p})}\} \tag{3.2}$$

for large enough N. This, together with the fact that $\lim_{N\to\infty} k(N)p^{k(N)} = 0$, implies item 1.

b) Let $\xi = \xi_c - x$, $x > 0$ and $k = k(N) = \lfloor \xi \ln N \rfloor$. Then, by the monotonicity of P_n and Lemma 2.3, we have

$$\begin{aligned} \mathbb{P}(S_N \geq k) &= \mathbb{P}(T_k \leq N) = \sum_{n=0}^{N-k} P_n = \sum_{n=0}^{3k-1} P_n + \sum_{n=3k}^{N-k} P_n \\ &\geq 3kP_{3k} + \sum_{n=3k}^{N-k} P_{3k} \frac{P_n}{P_{3k}} \geq 3kP_{3k} + P_{3k} \sum_{n=3k}^{N-k} \gamma_k^{n-3k} \\ &= 3kP_{3k} + \frac{P_{3k}}{f(k,p)P_1} \left[1 - \gamma_k^{N-4k+1}\right]. \end{aligned} \tag{3.3}$$

Using Lemma 2.2 again, we have

$$\lim_{k\to\infty} 3kP_{3k} \leq \lim_{k\to\infty} 3kP_k\beta_k^{2k} = 0 \tag{3.4}$$

and

$$\lim_{k\to\infty} \frac{P_{3k}}{f(k,p)P_1} \geq \lim_{k\to\infty} \frac{P_k\alpha_k^{2k}}{P_1} = 1. \tag{3.5}$$

Finally,

$$\begin{aligned} \gamma_k^{N-4k+1} &= [1 - f(k,p)qp^k]^{N-4k+1} \\ &= \left[(1 - f(k,p)qp^k)^{\frac{1}{f(k,p)qp^k}}\right]^{[N-4k+1]f(k,p)qp^k} \\ &\leq \exp\{-C_2 N^{x \ln(\frac{1}{p})}\} \end{aligned} \tag{3.6}$$

for some $C_2 > 0$ and large enough N. Note that, in order to obtain (3.5) and (3.6), we have used the fact that

$$\lim_{k\to\infty} f(k,p) = 1.$$

Item 2 follows immediately from (3.3), (3.4), (3.5) and (3.6).

c) Note that the arguments in a), b) also work for $x < 0$, then Corollary 1.2 follows immediately. $\square$

Proof of Theorem 1.3. For any $k \geq 1$, by (3.1) and (3.3), we have

$$3kP_{3k} + \frac{P_{3k}}{f(k,p)P_1}(1 - \gamma_k^{N-4k+1}) \leq \mathbb{P}(S_N \geq k) \leq 1 + kp^k - \beta_k^{N-2k+1}. \tag{3.7}$$

For any given integer l, let $k = k(j,l) = \lfloor \xi_c \ln N_{\lambda,j} \rfloor + l$, then

$$\mathbb{P}(\chi_{\lambda,j} \geq l) = \mathbb{P}(S_{N_{\lambda,j}} \geq k(j,l)).$$

So, by (3.4), (3.5) and (3.7), we have

$$\begin{aligned} 1 - \lim_{j\to\infty} \gamma_k^{N_{\lambda,j}-4k+1} &\leq \liminf_{j\to\infty} \mathbb{P}(\chi_{\lambda,j} \geq l) \\ &\leq \limsup_{j\to\infty} \mathbb{P}(\chi_{\lambda,j} \geq l) \leq 1 - \lim_{j\to\infty} \beta_k^{N_{\lambda,j}-2k+1}. \end{aligned} \tag{3.8}$$

By the property of $N_{\lambda,j}$, we have

$$\begin{aligned} &\lim_{j\to\infty} \beta_{k(j,l)}^{N_{\lambda,j}-2k(j,l)+1} = \exp\left\{-\lim_{j\to\infty}[N_{\lambda,j} - 2k(j,l) + 1]qp^{k(j,l)}\right\} \\ &= \exp\left\{-\lim_{j\to\infty} q[N_{\lambda,j} - 2k(j,l) + 1]p^{(\xi_c \ln N_{\lambda,j} - \lambda_j + l)}\right\} \\ &= e^{-qp^{-\lambda}p^l} \end{aligned}$$

and, similarly,

$$\lim_{j\to\infty} \gamma_{k(j,l)}^{N_{\lambda,j}-4k(j,l)+1} = e^{-qp^{-\lambda}p^l}.$$

Thus (3.8) implies

$$\lim_{j\to\infty} \mathbb{P}(\chi_{\lambda,j} \geq l) = 1 - e^{-qp^{-\lambda}p^l},$$

i.e.

$$\lim_{j\to\infty} \mathbb{P}(\chi_{\lambda,j} \leq l) = e^{-qp^{1-\lambda}p^l} = \mathbb{P}(\chi_\lambda \leq l).$$

We finish the proof of Theorem 1.3. $\square$

4 Proof of Theorem 1.4

In this section, will prove Theorem 1.4. We first prove the corresponding convergence on moments, i.e. (1.13), then we prove (1.14) and (1.15).

Proof of (1.13). For any $\lambda \in [0,1]$, denote by $F_{\lambda,j}(x)$ the distribution function of $\chi_{\lambda,j}$. Intergrating by parts,

$$\mathbb{E}\chi_{\lambda,j}^m = \int_{-\infty}^{\infty} x^m dF_{\lambda,j}(x) = -\int_{-\infty}^{0} mx^{m-1}F_{\lambda,j}(x)dx + \int_{0}^{\infty} mx^{m-1}(1-F_{\lambda,j}(x))dx. \tag{4.1}$$

By convergence in distribution, we have pointwise convergence of $F_{\lambda,j}$ to F_λ, where F_λ is given in (1.8). In order to show the moments of $\chi_{\lambda,j}$ converge to the corresponding moments of χ_λ we need uniform control of $F_{\lambda,j}$ for large j. Actually, we expect that $F_{\lambda,j}(x)$ (resp., $1-F_{\lambda,j}(x)$) go to zero uniformly rapidly as $x \to -\infty$ (resp., $x \to \infty$).

First we try to look for such an uniform control of $\bar{F}_N$ for large N. Recall that $\bar{F}_N$ is the distribution function of $\bar{\chi}_N = S_N - \xi_c \ln N$.

In the case when $x \ge 0$, let $k(x) := \lfloor \xi_c \ln N + x \rfloor$. If x is large enough such that $k(x) > N$, then

$$1 - \bar{F}_N(x) = \mathbb{P}(S_N > \xi_c \ln N + x) \le \mathbb{P}(S_N \ge k(x)) = 0. \tag{4.2}$$

If $x \ge 0$ such that $N/2 \le k(x) \le N$, then by (3.1),

$$1 - \bar{F}_N(x) \le \mathbb{P}(S_N \ge k(x)) \le (k(x)+1)p^{k(x)} \le \frac{N+1}{Np}p^x \le \frac{2}{p}p^x. \tag{4.3}$$

If $x \ge 0$ such that $k(x) < N/2$, then by (3.1),

$$1 - \bar{F}_N(x) \le k(x)p^{k(x)} + 1 - \beta_{k(x)}^{N-2k(x)+1} \le 1 + \frac{1}{2p}p^x - \beta_{k(x)}^{N-2k(x)+1}.$$

Clearly, in the present case, there exists some constant $C' > 0$, which does not depend on N and x, such that

$$\left[1 - qp^{k(x)}\right]^{\frac{1}{qp^{k(x)}}} \ge e^{-C'}.$$

Then

$$\begin{aligned}\beta_{k(x)}^{N-2k(x)+1} &\ge \exp\left\{-C'[N-2k(x)+1]qp^{k(x)}\right\} \ge \exp\left\{-C'q\frac{N-2k(x)+1}{pN}p^x\right\} \\ &\ge \exp\left\{-\frac{qC'}{p}p^x\right\} \ge 1 - \frac{qC'}{p}p^x.\end{aligned}$$

Thus,

$$1 - \bar{F}_N(x) \le 1 + \frac{1}{2p}p^x - \left(1 - \frac{qC'}{p}p^x\right) = C''p^x. \tag{4.4}$$

By (4.2), (4.3) and (4.4), there exists $C_1 > 0$ such that

$$1 - \bar{F}_N(x) \le C_1 p^x, \quad \forall\, x \ge 0 \tag{4.5}$$

for $N \geq 1$.

Now we consider the case when $x \leq 0$. Let $\hbar(x) := \lfloor \xi_c \ln N + x + 1 \rfloor$. Then

$$\bar{F}_N(x) = 1 - \mathbb{P}(S_N > \xi_c \ln N + x) \leq 1 - \mathbb{P}(S_N \geq \hbar(x)). \tag{4.6}$$

If x is small enough such that $\hbar(x) \leq 0$, then

$$\bar{F}_N(x) \leq 1 - \mathbb{P}(S_N \geq \hbar(x)) = 0. \tag{4.7}$$

For $1 \leq k \leq \hbar(0) = \lfloor \xi_c \ln N + 1 \rfloor$, let x be such that $\hbar(x) = k$, i.e.,

$$k - \xi_c \ln N - 1 \leq x < k - \xi_c \ln N, \tag{4.8}$$

then

$$\bar{F}_N(x) \leq 1 - \mathbb{P}(S_N \geq k) = 1 - \mathbb{P}(T_k \leq N) = \mathbb{P}(T_k \geq N+1) = \sum_{n=N-k+1}^{\infty} P_n.$$

Choose $N_0 = N_0(p)$ large enough such that $N - 2\hbar(0) + 1 \geq 1$ for all $N \geq N_0$. Then, for any $N \geq N_0$, $n \geq N - k + 1$ implies $n \geq k$. Thus we can use the upper bound given in Lemma 2.2 and obtain

$$\bar{F}_N(x) \leq \sum_{n=N-k+1}^{\infty} P_n \leq P_{N-k+1} \sum_{n=0}^{\infty} \beta_k^n \leq \frac{P_k}{1-\beta_k} \beta_k^{N-2k+1} = \beta_k^{N-2k+1}. \tag{4.9}$$

Choose $C' > 0$ such that

$$(1 - qp^n)^{\frac{1}{qp^n}} \leq e^{-C'}, \quad for\ all\ n \geq 1.$$

Then

$$\beta_k^{N-2k+1} \leq \exp\{-C'(N-2k+1)qp^k\}.$$

By (4.8), $k \leq \xi_c \ln N + x + 1$, then

$$\bar{F}_N(x) \leq \exp\left\{-\frac{C'[N - 2\hbar(0) + 1]qp}{N} p^x\right\} \leq \exp\{-C'qpp^x\} = e^{-C_2 p^x} \tag{4.10}$$

for all $N \geq N_0$. Combining (4.7) and (4.10), we obtain

$$\bar{F}_N(x) \leq e^{-C_2 p^x}, \quad \forall\ x \leq 0 \tag{4.11}$$

For all $N \geq N_0$. Thus, as in (4.5) and (4.11), we obtain an uniform control of $\bar{F}_N$ for $N \geq N_0$.

By (4.5), (4.11) and the fact that $\bar{\chi}_N \leq \chi_N < \bar{\chi}_N + 1$, we obtain the following uniform control for F_N, the distribution function of $\chi_N = S_N - \lfloor \xi_c \ln N \rfloor$:

$$1 - F_N(x) \leq 1 - \bar{F}_N(x-1) \leq \frac{C_1}{p} p^x, \quad \forall\ x \geq 1 \tag{4.12}$$

and

$$F_N(x) \leq \bar{F}_N(x) \leq e^{-C_2 p^x}, \quad \forall\, x \leq 0 \tag{4.13}$$

for all $N \geq N_0$. Then (1.13) follows from (4.1) and the dominated convergence theorem. As a consequence

$$\lim_{j\to\infty} \mathrm{Var} S_{N_{\lambda,j}} = \lim_{j\to\infty} \mathrm{Var}\chi_{\lambda,j} = \mathrm{Var}\chi_\lambda. \tag{4.14}$$

Proofs of (1.14) and (1.15). To prove (1.14), suppose that $\liminf_{N\to\infty} \dfrac{\mathbb{E}S_N}{\ln N} = D_1 \geq -\infty$. Then, there exists subsequence $\{N_j\}$ such that $\lim_{j\to\infty} \dfrac{\mathbb{E}S_{N_j}}{\ln N_j} = D_1$. Let $\Lambda_j = \xi_c \ln N_j - \lfloor \xi_c \ln N_j \rfloor$. Apparently, the sequence $\{\Lambda_j\}$ is bounded and for some $\lambda \in [0,1]$, there exists a subsubsequence $\{N_{\lambda,j}\}$ of subsequence $\{N_j\}$ such that $\lim_{j\to\infty} \Lambda_{N_{\lambda,j}} = \lambda$. Let $\lambda_j = \Lambda_{N_{\lambda,j}}$. Now, by Theorem 1.3 we know that as $j \to \infty$, $\chi_{\lambda,j} = S_{N_{\lambda,j}} - \lfloor \xi_c \ln N_{\lambda,j} \rfloor$ converges to χ_λ in distribution. By (1.9) and (1.13), we have

$$D_1 = \liminf_{N\to\infty} \frac{\mathbb{E}S_N}{\ln N} = \lim_{j\to\infty} \frac{\mathbb{E}S_{N_{\lambda,j}}}{\ln N_{\lambda,j}} = \lim_{j\to\infty} \frac{\lfloor \xi_c \ln N_{\lambda,j} \rfloor + \mathbb{E}\chi_{\lambda,j}}{\ln N_{\lambda,j}} = \xi_c.$$

Repeating the above argument, we obtain

$$\limsup_{N\to\infty} \frac{\mathbb{E}S_N}{\ln N} = \xi_c$$

and then (1.14) follows.

Finally, (1.15) follows from (1.10), (4.14) and an argument similar to the proof of (1.14).

Thus we finish the proof of Theorem 1.4. □

References

[1] D. Aldous and P. Diaconis (1999) *Longest Increasing Subsequences: From Patience Sorting to the Baik-Deift-Johnsson Theorem.* Bull. Amer. Math. Soc. **36** 413-432.

[2] J. Baik, P. Deift and K. Johansson (1999) *On the Distribution of The Length of The Longest Increasing Subsequence of Random Permutation.* J. Amer. Math. Soc. **12** 1119-1178.

[3] M. Z. Bazant (2000) *The Largest Cluster in Subcritical Percolation.* Physical Review E, 2000 - APS

[4] C. Borgs; J. T. Chayes; H. Kesten and J. Spencer (2001) *The Birth of the Infinite Cluster: Finite-Size Scaling in Percolation.* Communications in Mathematical Physics, Volume 224, Issue 1, pp. 153-204.

[5] M. Brown and Y. S. Shao (1987) *Identifying coefficients in the spectral representation for first passage time distribition.* Probab. Eng. Inform. Sci. **1**, 69-74

[6] J. A. Fill (2009) *On hitting times and fastest strong stationary times for skip-free and more general chains,* Journal of Theoretical Probability, **22**, No. 3, 587-600

[7] R. A. Fisher and L. H. C. Tippett (1928) *Limiting forms of the frequency distribution of the largest or smallest member of a sample.* Mathematical Proceedings of the Cambridge Philosophical Society **24** pp. 180-190

[8] G. R. Grimmett (1999) *Percolation,* 2nd edition, Springer.

[9] E. J. Gumbel (1958) *Statistics of Extremes,* Columbia University Press, New York

[10] G. R. Grimmett, T. M. Liggett and T. Richthammer (2010) *Percolation of arbitrary words in one dimension.* Random Structures & Algorithms, Volume **37**, Issue 1, pp. 85-99

[11] H. Kesten (1982) *Percolation Theory for Mathematicians,* Birkhäuser, Bosten

[12] D. Stauffer and A. Aharony (1992) *Introduction to Percolation Theory,* Taylor & Francis, London