The Annals of Applied Probability
2003, Vol. 13, No. 4, 1601–1614

LARGE DEVIATIONS IN FIRST-PASSAGE PERCOLATION

BY YUNSHYONG CHOW AND YU ZHANG

Academia Sinica and University of Colorado

Consider the standard first-passage percolation on $\mathbb{Z}^d$, $d \geq 2$. Denote by $\phi_{0,n}$ the face–face first-passage time in $[0, n]^d$. It is well known that

$$\lim_{n\to\infty} \frac{\phi_{0,n}}{n} = \mu(F) \qquad \text{a.s. and in } L_1,$$

where F is the common distribution on each edge. In this paper we show that the upper and lower tails of $\phi_{0,n}$ are quite different when $\mu(F) > 0$. More precisely, we can show that for small $\varepsilon > 0$, there exist constants $\alpha(\varepsilon, F)$ and $\beta(\varepsilon, F)$ such that

$$\lim_{n\to\infty} \frac{-1}{n} \log P\left(\phi_{0,n} \leq n(\mu - \varepsilon)\right) = \alpha(\varepsilon, F)$$

and

$$\lim_{n\to\infty} \frac{-1}{n^d} \log P\left(\phi_{0,n} \geq n(\mu + \varepsilon)\right) = \beta(\varepsilon, F).$$

1. Introduction and main results. We consider $\mathbb{Z}^d$, $d \geq 2$, as a graph with edges connecting each pair of vertices $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ with $\|x - y\| = 1$, where the norm is defined by

$$\|x - y\| = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_d - y_d|.$$

To each edge e we attach a nonnegative random variable $t(e)$. The basic assumption is that the random variables $\{t(e) : e \in \mathbb{Z}^d\}$ are i.i.d. with the common distribution F satisfying

$$F(0^-) = 0, \qquad t(e) \text{ is not a constant} \quad \text{and} \quad E(t(e)) < \infty. \tag{1.1}$$

More formally, we consider the following probability space. As sample space we take $\Omega = \prod_{e\in\mathbb{Z}^d}[0, \infty)$, points of which are represented as *configurations*. Let P be the corresponding product measure on Ω and the expectation with respect to P is denoted by E. For any two vertices u and v, a path γ from u to v is an alternating sequence $(v_0, e_1, v_1, \ldots, e_n, v_n)$ of vertices and edges in $\mathbb{Z}^d$ with $v_0 = u$ and $v_n = v$. The path γ is said running from A to B if $u \in A$ and $v \in B$, and $\gamma \subset C$ means all its vertices v_i are contained in C. Here A, B and C are

Received September 2002; revised January 2003.
AMS 2000 subject classification. 60K35.
Key words and phrases. First-passage percolation, large deviations.

subsets of $\mathbb{Z}^d$. Define *the passage time* of γ as $t(\gamma) = \sum_{i=1}^{n} t(e_i)$. Two popular first-passage times, face–face and point–point, are defined by

$$\phi\left(\prod_{1\le i\le d}[x_i, y_i]\right) = \inf\left\{t(\gamma) : \gamma \text{ is a path from } A \text{ to } B \text{ and } \gamma \subset \prod_{1\le i\le d}[x_i, y_i]\right\}, \tag{1.2}$$

where $A = \{x_1\} \times \prod_{2\le i\le d}[x_i, y_i]$ and $B = \{y_1\} \times \prod_{2\le i\le d}[x_i, y_i]$, and

$$a_{m,n} = \inf\{t(\gamma) : \gamma \text{ is a path from } (m, 0, \dots, 0) \text{ to } (n, 0, \dots, 0)\}.$$

For short we denote $\phi([m,n]^d)$ by $\phi_{m,n}$. It is well known [Grimmett and Kesten (1984)] that

$$\lim_{n\to\infty} \frac{1}{n} a_{0n} = \lim_{n\to\infty} \frac{1}{n}\phi_{0,n} = \mu(F) \qquad \text{a.s. and in } L_1,$$

and $\mu(F) = 0$ iff $F(0) \ge p_c$, where $p_c = p_c(d)$ is the critical probability for Bernoulli (bond) percolation on $\mathbb{Z}^d$ and the nonrandom constant $\mu(F)$ is the so called *time constant*.

In this paper we shall mainly investigate large deviations of $\theta_{0,n}$ for $\theta = a, \phi$. In fact, it is known [see Zhang (1995)] that $\theta_{0,n}$ cannot be very large if $F(0) > p_c$. The so called critical case, $F(0) = p_c$, is more complicated. Few results are known in this direction. So we focus on the case that $F(0) < p_c$. A subadditive argument shows [Kesten (1986)] that

$$\lim_{n\to\infty} \frac{-1}{n} \log P\big(a_{0,n} \le n(\mu - \varepsilon)\big) = \alpha(\varepsilon, F) \tag{1.3}$$

and the constant $\alpha(\varepsilon, F)$ is positive if ε is small. Regarding to ϕ we have the following

THEOREM 1. *Assume* (1.1), $F(0) < p_c$ *and* $0 < \varepsilon$. *Then*

$$\lim_{n\to\infty} \frac{-1}{n} \log P\big(\phi_{0,n} \le n(\mu - \varepsilon)\big) = \alpha(\varepsilon, F). \tag{1.4}$$

Note that if $\mu < \varepsilon$ then $P(\phi_{0,n} \le n(\mu - \varepsilon)) = 0$ and thus $\alpha(\varepsilon, F) = \infty$. However, it seems not trivial to decide the upper tail behavior of $\theta_{0,n}$ for $\theta = a$ or ϕ. It is proved [see Chapter 5 in Kesten (1986)] that

$$\lim_{n\to\infty} \frac{-1}{n} \log P\big(a_{0,n} \ge n(\mu + \varepsilon)\big) = \infty$$

and under the condition that $t(e)$ is bounded with $F(0) < p_c$,

$$0 < \liminf_{n\to\infty} \frac{-1}{n^d} \log P\big(a_{0,n} \ge n(\mu + \varepsilon)\big). \tag{1.5}$$

Note that the bounded condition on $t(e)$ is important. Otherwise, (1.5) could fail. For example, let $t(e)$ be exponentially distributed with a parameter $\lambda < p_c$. We know that if all $2d$ bonds connected to the origin take values larger than $n(\mu+\varepsilon)$, then $a_{0,n} \geq n(\mu+\varepsilon)$. Hence

$$P\big(a_{0,n} \geq n(\mu+\varepsilon)\big) \geq P\big(t(e) \geq n(\mu+\varepsilon)\big)^{2d} = \exp\big(-2d\lambda n(\mu+\varepsilon)\big).$$

Therefore, (1.5) does not hold. However, we do not need this bounded condition for face–face first-passage time. More precisely, we will show:

THEOREM 2. *Assume* (1.1), $F(0) < p_c$ *and* $0 < \varepsilon$. *Then*

$$\limsup_{n\to\infty} \frac{-1}{n^d} \log P\big(\phi_{0,n} \geq n(\mu+\varepsilon)\big) < \infty \text{ holds for } 0 < \varepsilon \leq Et(e) - \mu. \tag{1.6}$$

Furthermore, if $E(\exp \theta t(e)) < \infty$ *for some positive* $\theta < \infty$, *then*

$$0 < \liminf_{n\to\infty} \frac{-1}{n^d} \log P\big(\phi_{0,n} \geq n(\mu+\varepsilon)\big). \tag{1.7}$$

With (1.6) and (1.7) it is natural to study the limit behavior of the upper tail of $\phi_{0,n}$ as we did in Theorem 1. However, it is difficult to show the existence of a limit regarding to the upper tail of $\phi_{0,n}$. Indeed, most limit behaviors in first passage percolation are obtained by using subadditive arguments. More precisely, one uses two paths to construct a longer path to set a subadditive inequality. This argument will not work on the upper tail of $\phi_{0,n}$, since the tail is extremely small with an order $\exp(-Cn^d)$ indicated in Theorem 2. To overcome the difficulty, we borrow ideas from the min-cut and max-flow theorem. Take as an example the Bernoulli first passage percolation on $\mathbb{Z}^3$, where $t(e)$ can only take zero or one. Then $\{\phi_{0,n} = k\}$ is equivalent to the event that there are only k disjoint dual surfaces piling from the top to the bottom of $[0,n]^3$ such that each of their plaquettes takes value one. These k surfaces in $[0,n]^3$ could be connected not only vertically with those corresponding surfaces in $[0,n]^2 \times [n+1, 2n]$, but also horizontally with those corresponding surfaces in $[0,n] \times [n+1, 2n] \times [0,n]$. As a result k disjoint dual surfaces could be constructed piling from the top to the bottom of $[0,n] \times [0, \ell n]^2$. This sets up a multi-subadditive argument.

THEOREM 3. *Assume* (1.1), $F(0) < p_c$ *and* $0 < \varepsilon$. *Then there exists a constant* $\beta(\varepsilon, F) \geq 0$ *such that*

$$\lim_{n\to\infty} \frac{-1}{n^d} \log P\big(\phi_{0,n} \geq n(\mu+\varepsilon)\big) = \beta(\varepsilon, F). \tag{1.8}$$

Note that Theorem 2 above implies $0 < \beta(\varepsilon, F) < \infty$ if $0 < \varepsilon \leq Et(e) - \mu$ and $E(\exp \theta t(e)) < \infty$ for some positive $\theta < \infty$.

Theorems 1, 2 and 3 will be proved in Sections 2, 3 and 4, respectively. Finally, we remark on the following:

1. We are unable to show a result similar to (1.8) for $a_{0,n}$.
2. It can be shown that $\alpha(\varepsilon, F)$ and $\beta(\varepsilon, F)$ are convex in $\varepsilon > 0$. In fact the former is done in (2.2). This implies that $\alpha(\varepsilon, F)$ and $\beta(\varepsilon, F)$ are continuous in $\varepsilon > 0$. We conjecture that $\alpha(\varepsilon, F)$ and $\beta(\varepsilon, F)$ are also continuous in F.

2. Proof of Theorem 1. In this section we will show that the limit in (1.4) exists and equals to $\alpha(\varepsilon, F)$ defined in (1.3). Unlike $a_{0,n}$, $\log P(\phi_{0,n} \le n(\mu - \varepsilon))$ is not a subadditive sequence any more. To avoid the problem, we use the idea in Hammersley and Welsh (1965).

First we show the continuity of $\alpha(\varepsilon, F)$ in $\varepsilon > 0$. It is clear that for any $0 < \delta < \varepsilon$,

$$P\big(a_{0,n} \le n(\mu - \varepsilon - \delta), a_{n,2n} \le n(\mu - \varepsilon + \delta)\big) \le P\big(a_{0,2n} \le 2n(\mu - \varepsilon)\big). \tag{2.1}$$

By the FKG inequality and translation invariance, we obtain from (2.1) and (1.3) that

$$\alpha(\varepsilon + \delta, F) + \alpha(\varepsilon - \delta, F) \ge 2\alpha(\varepsilon, F), \tag{2.2}$$

which means $\alpha(\varepsilon, F)$ is locally convex in ε. Hence $\alpha(\varepsilon, F)$ is continuous in ε.

Let $\overrightarrow{e_1} = (1, 0, \dots, 0)$. Now we investigate the relationship between the point–point first-passage time $a_{m,n}$ and the cylinder point–point first-passage times defined by

$$t_{m,n}(k) = \inf\{t(\gamma) : \gamma \text{ is a path from } m\overrightarrow{e_1} \text{ to } n\overrightarrow{e_1} \text{ and } \gamma \subset [-k+m, n+k] \times \mathbb{Z}^{d-1}\},$$

where $k \ge 0$. Denote $t_{m,n}(0)$ by $t_{m,n}$ for short. A standard subadditive argument yields

$$\begin{aligned} &\lim_{n\to\infty} \frac{-1}{n} \log P\big(t_{0,n}(k) \le n(\mu - \varepsilon)\big) \\ &\qquad = \inf_n \frac{-1}{n} \log P\big(t_{0,n}(k) \le n(\mu - \varepsilon)\big) \\ &\qquad \equiv \tau_k(\varepsilon, F). \end{aligned} \tag{2.3}$$

The same argument used in (2.2) shows that $\tau_k(\varepsilon, F)$ is locally convex and thus continuous in $\varepsilon > 0$ for each $k \ge 0$. We claim that

$$\tau_0(\varepsilon, F) = \tau_k(\varepsilon, F) = \alpha(\varepsilon, F) \qquad \text{for all } k \ge 0. \tag{2.4}$$

It is clear from the definitions that $a_{0,n} \le t_{0,n}(k) \le t_{0,n}$, which implies

$$\tau_0(\varepsilon, F) \ge \tau_k(\varepsilon, F) \ge \alpha(\varepsilon, F) \qquad \text{for } k \ge 0. \tag{2.5}$$

It remains to verify the other direction. Since $t(e) \ge 0$, we may choose M such that

$$P\big(0 \le t(e) < M\big) \equiv \delta > 0.$$

For each $k \geq 0$ we denote by $A(k)$ the event that all those $2k$ edges e from $-k\overrightarrow{e_1}$ to the origin and $n\overrightarrow{e_1}$ to $(n+k)\overrightarrow{e_1}$ along the first coordinate take values less than M. Then

$$P(A(k)) \geq \delta^{2k} \tag{2.6}$$

and $t_{-k,n+k} \leq t_{0,n}(k) + 2kM$ on $A(k)$. As a consequence,

$$P\big(t_{0,n}(k) \leq n(\mu - \varepsilon), A(k)\big) \leq P\big(t_{-k,n+k} \leq n(\mu - \varepsilon) + 2kM\big).$$

By translation invariance the following holds for $\varepsilon' < \varepsilon$ and n large:

$$\begin{aligned} P\big(t_{-k,n+k} \leq n(\mu - \varepsilon) + 2kM\big) &= P\big(t_{0,n+2k} \leq n(\mu - \varepsilon) + 2kM\big) \\ &\leq P\big(t_{0,n+2k} \leq (n+2k)(\mu - \varepsilon')\big). \end{aligned}$$

Using the FKG inequality, (2.6) and the continuity of $\tau_0(\varepsilon, F)$ in ε, we obtain from the previous two equations that

$$\tau_k(\varepsilon, F) = \lim_{n\to\infty} \frac{-1}{n} \log P\big(t_{0,n}(k) \leq n(\mu - \varepsilon)\big) \geq \tau_0(\varepsilon, F). \tag{2.7}$$

We explore the relationship between $a_{0,n}$ and $t_{0,n}(k)$ for large k. A route for $a_{0,n}$ is defined as a path γ from the origin to $n\overrightarrow{e_1}$ with $t(\gamma) = a_{0,n}$. It is proved [see page 258 in Kesten (1986)] that such a route exists for $a_{0,n}$ when $F(0) < p_c$. Such existence implies that for each configuration ω and a fixed n, $a_{0,n}(\omega) = t_{0,n}(k, \omega)$ for all large k. That is

$$\{t_{0,n}(k) \leq n(\mu - \varepsilon)\} \uparrow \{a_{0,n} \leq n(\mu - \varepsilon)\} \qquad \text{as } k \text{ increases to } \infty. \tag{2.8}$$

Hence for fixed n, we have from (2.3) that for k large,

$$\begin{aligned} \tau_k(\varepsilon, F) &\leq \frac{-1}{n} \log P\big(t_{0,n}(k) \leq n(\mu - \varepsilon)\big) \\ &\leq \frac{-1}{n} \log P\big(a_{0,n} \leq n(\mu - \varepsilon)\big) + \frac{1}{n}. \end{aligned} \tag{2.9}$$

It follows from (2.7) and (1.3) that $\tau_0(\varepsilon, F) \leq \alpha(\varepsilon, F)$. This verifies (2.4) in view of (2.5).

Knowing the relationship between $a_{0,n}$ and $t_{0,n}(k)$ we now connect ϕ with τ. Define

$$h_{m,n}(k) = \inf\big\{t(\gamma) : \gamma \text{ is a path from } m\overrightarrow{e_1} \text{ to } n\overrightarrow{e_1} \text{ and } \gamma \subset [m,n] \times [-k,k]^{d-1}\big\}.$$

Again a subadditive argument shows the existence of the following limit for fixed $k \geq 0$:

$$\begin{aligned} &\lim_{n\to\infty} \frac{-1}{n} \log P\big(h_{0,n}(k) \leq n(\mu - \varepsilon)\big) \\ &= \inf_n \frac{-1}{n} \log P\big(h_{0,n}(k) \leq n(\mu - \varepsilon)\big) \\ &\equiv \eta_k(\varepsilon, F). \end{aligned} \tag{2.10}$$

Since $h_{0,n}(k) \geq t_{0,n}$, $\eta_k(\varepsilon, F) \geq \tau_0(\varepsilon, F)$ holds trivially. We will show that

$$\limsup \eta_k(\varepsilon, F) \leq \tau_0(\varepsilon, F) \quad \text{and thus} \quad \lim_{k\to\infty} \eta_k(\varepsilon, F) = \tau_0(\varepsilon, F). \tag{2.11}$$

Similar to (2.8) and (2.9) we have that for fixed n,

$$\{h_{0,n}(k) \leq n(\mu - \varepsilon)\} \uparrow \{t_{0,n} \leq n(\mu - \varepsilon)\} \qquad \text{as } k \text{ increases to } \infty.$$

Hence for fixed n, we get from (2.10) that for k large

$$\begin{aligned} \eta_k(\varepsilon, F) &\leq \frac{-1}{n} \log P\big(h_{0,n}(k) \leq n(\mu - \varepsilon)\big) \\ &\leq \frac{-1}{n} \log P\big(t_{0,n} \leq n(\mu - \varepsilon)\big) + \frac{1}{n}, \end{aligned} \tag{2.12}$$

from which (2.11) follows. Because $\phi([0, n] \times [-n/2, n/2]^{d-1}) \leq h_{0,n}(k)$ for $2k \leq n$,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{-1}{n} \log P\big(\phi_{0,n} \leq n(\mu - \varepsilon)\big) \\ &\qquad \leq \limsup_{n\to\infty} \frac{-1}{n} \log P\big(h_{0,n}(k) \leq n(\mu - \varepsilon)\big) = \eta_k(\varepsilon, F) \end{aligned}$$

by (2.10) and translation invariance. It follows then from (2.11) and (2.4) that

$$\limsup_{n\to\infty} \frac{-1}{n} \log P\big(\phi_{0,n} \leq n(\mu - \varepsilon)\big) \leq \tau_0(\varepsilon, F) = \alpha(\varepsilon, F). \tag{2.13}$$

For the opposite direction of (2.13), we introduce

$$t_n(a, b) = \inf\big\{t(\gamma) : \gamma \text{ is a path from } a \text{ to } b \text{ and } \gamma \subset [0, n] \times [0, n]^{d-1}\big\},$$

where $a \in \{0\} \times [0, n]^{d-1}$ and $b \in \{n\} \times [0, n]^{d-1}$. It is clear by counting that

$$P\big(\phi_{0,n} \leq n(\mu - \varepsilon)\big) \leq 2(n + 1)^{d-1} \max_{a,b} P\big(t_n(a, b) \leq n(\mu - \varepsilon)\big). \tag{2.14}$$

On the other hand, by symmetry and translation invariance,

$$\begin{aligned} \big\{P\big(t_n(a, b) \leq n(\mu - \varepsilon)\big)\big\}^2 &\leq P\big(t_{2n}(a, a') \leq 2n(\mu - \varepsilon)\big) \\ &\leq P\big(t_{0,2n} \leq 2n(\mu - \varepsilon)\big), \end{aligned} \tag{2.15}$$

where $a' \in \{2n\} \times [0, n]^{d-1}$ and has the same coordinates as a except the first one. Combining (2.14) and (2.15),

$$\begin{aligned} \tau_0(\varepsilon, F) &= \lim_{n\to\infty} \frac{-1}{2n} \log P\big(t_{0,2n} \leq 2n(\mu - \varepsilon)\big) \\ &\leq \liminf_{n} \frac{-1}{n} \log P\big(\phi_{0,n} \leq n(\mu - \varepsilon)\big). \end{aligned}$$

This, together with (2.13), implies (1.4). The proof of Theorem 1 is complete.

3. Proof of Theorem 2. We only prove Theorem 2 for $d = 3$. The same argument can be adopted directly for any $d > 1$. If all bonds in $[0, n]^3$ take values no less than $Et(e)$ and $Et(e) \geq \mu + \varepsilon$, then $\phi_{0,n} \geq nEt(e) \geq n(\mu + \varepsilon)$. Hence $P(\phi_{0,n} \geq n(\mu + \varepsilon)) \geq P(t(e) \geq Et(e))^{n^3}$ and (1.6) follows easily. Note that $Et(e) > \mu$ under assumption (1.1). It remains to show (1.7). Define

$$T_{l,k,m} = \inf\{t(\gamma) : \gamma \text{ is a path from } (l, 0, 0) \text{ to } (k, 0, 0) \text{ and } \gamma \subset [l, k] \times [-m, m]^2\}.$$

We first show the following lemma under an extra exponential moment assumption.

LEMMA 3.1. *Assume* (1.1), $F(0) < p_c, 0 < \varepsilon$ *and* $E(\exp\theta t(e)) < \infty$ *for some positive* $\theta < \infty$. *Then there exists a constant* $\eta > 0$ *such that*

$$P\big(T_{0,k,m} \geq k(\mu + \varepsilon)\big) \leq \exp(-\eta k) \qquad \text{for any large } k = nm.$$

PROOF. Let γ be a route for $t_{0,n}$, where $t_{0,n}$ is the cylinder point–point first-passage time defined in Section 2. Define

$$h_n(\gamma) = \max_{2 \leq i \leq 3} \{|m_i| : (m_1, m_2, m_3) \in \gamma\}$$

and

$$h_n = \max\{h_n(\gamma) : \gamma \text{ is a route for } t_{0,n}\}.$$

It is known [see Theorem 8.15 in Smythe and Wierman (1978)] that

$$\limsup h_n/n \leq 1 \qquad \text{almost surely.} \tag{3.1}$$

In fact, they only proved (3.1) for $d = 2$, but extension to any $d > 2$ is straightforward.

We claim that for $k = m$ and m large,

$$ET_{0,m,m} \leq m(\mu + \varepsilon). \tag{3.2}$$

To see this, define $H_n = \{h_n/n \leq 1\}$. Since $t_{0,m} = T_{0,m,m}$ on H_m, we have

$$Et_{0,m} \geq E(t_{0,m}; H_m) = E(T_{0,m,m}; H_m) = E(T_{0,m,m}) - E(T_{0,m,m}; H_m^C). \tag{3.3}$$

By Theorem 2.18 in Kesten (1986), $\lim_{n\to\infty} Et_{0,n}/n = \mu$. Hence,

$$Et_{0,m}/m \leq \mu + \varepsilon/2 \qquad \text{holds for } m \text{ large.} \tag{3.4}$$

In view of (3.3) and (3.4), it suffices to show $\lim_{m\to\infty} E(T_{0,m,m}/m; H_m^C) = 0$. Since $T_{0,m,m} \leq \sum_{e\in\gamma} t(e)$, where γ is the path from $(0, \ldots, 0)$ to $(m, 0, \ldots, 0)$

along the first coordinate direction. By using the Cauchy–Schwarz inequality twice,

$$\{E(T_{0,m,m}/m;\, H_m^C)\}^2 \le E\left\{\left(\sum_{e\in\gamma} t(e)/m\right)^2\right\} \cdot P(H_m^C) \le E\{t(e)^2\} \cdot P(H_m^C).$$

This, together with (3.1), verifies (3.2).

It is clear from the definition that $T_{0,nm,m} \le \sum_{i=0}^{n-1} T_{im,(i+1)m,m}$ and thus

$$P\big(T_{0,nm,m} \ge nm(\mu+2\varepsilon)\big) \le P\left(\sum_{i=0}^{n-1} T_{im,(i+1)m,m} \ge nm(\mu+2\varepsilon)\right).$$

Since $\{T_{im,(i+1)m,m}\}$ are i.i.d. with a finite common mean $ET_{0,m,m}$, a standard large deviation argument [see Durrett (1996)] and (3.2) imply that for n large,

$$P\big(T_{0,nm,m} \ge nm(\mu+2\varepsilon)\big) \le \exp(-Cn)$$

holds for some constant $C > 0$. Lemma 3.1 follows by setting $\eta = C/m$. □

Now we prove (1.7). Take $k = mn$ and divide $[0,k]^2$ into $(k/m)^2 = n^2$ equal subsquares of size $m \times m$,

$$S_1, S_2, \dots, S_{n^2}.$$

Since $\{\phi([0,k]\times S_i) \ge k(\mu+\varepsilon)\}$, $\{\phi([0,k]\times S_j) \ge k(\mu+\varepsilon)\}$ are independent for $i \ne j$ and $\{\phi_{0,k} \ge k(\mu+\varepsilon)\} \subseteq \bigcap_{i=1}^{n^2}\{\phi([0,k]\times S_i) \ge k(\mu+\varepsilon)\}$, we have

$$P\big(\phi_{0,k} \ge k(\mu+\varepsilon)\big) \le \big\{P\big(\phi([0,k]\times[0,m]^2) \ge k(\mu+\varepsilon)\big)\big\}^{n^2}. \tag{3.5}$$

By Lemma 3.1 and translation invariance,

$$P\big(\phi([0,k]\times[0,m]^2) \ge k(\mu+\varepsilon)\big) \le P\big(T_{0,k,m/2} \ge k(\mu+\varepsilon)\big) \le \exp(-\eta k). \tag{3.6}$$

Combining (3.5) and (3.6), we get that for k large and $m|k$,

$$P\big(\phi_{0,k} \ge k(\mu+\varepsilon)\big) \le \exp(-\eta k n^2) = \exp(-C' k^3),$$

where $C' = \eta/m^2$. This verifies (1.7) and thus Theorem 2.

4. Proof of Theorem 3. For convenience we work on $\phi_{1,n}$ instead of $\phi_{0,n}$. The basic idea is to show that $-\log P(\phi_{1,n} \ge n(\mu+\varepsilon))$ is sort of a multi-subadditive sequence. The proof is done by verifying $\kappa' = \kappa$, where

$$\begin{aligned} \kappa' &= \liminf \frac{1}{n^d} \log P\big(\phi_{1,n} \ge n(\mu+\varepsilon)\big) \quad \text{and} \\ \kappa &= \limsup \frac{1}{n^d} \log P\big(\phi_{1,n} \ge n(\mu+\varepsilon)\big). \end{aligned} \tag{4.1}$$

Note that $\kappa' \le \kappa \le 0$ as $P(\phi_{1,n} \ge n(\mu+\varepsilon)) \le 1$. If $\kappa = -\infty$, then $\kappa' = \kappa = -\infty$ and we are done. So we may assume κ is finite hereafter.

Denote the front and rear faces of the box $[1,n]^d$ in the ith coordinate by

$$F_i(n) = \{v = (v_1, \dots, v_d) \in [1,n]^d : v_i = 1\}$$

and

$$R_i(n) = \{v = (v_1, \dots, v_d) \in [1,n]^d : v_i = n\}.$$

Choose $\delta > 0$ such that

$$r = P(t(e) > 2\delta) > 0. \tag{4.2}$$

The number δ will be used as the mesh size in measuring the first-passage time. Define

$$G_k = \{v \in [1,n]^d : \inf\{t(\gamma) : \gamma \text{ is a path from } F_1(n) \text{ to } v \text{ and } \gamma \subset [1,n]^d\} < k\delta\}.$$

In general, the geometry of G_k could be very complicated. In any case G_k is connected as any vertex $v \in G_k$ could be reached from $F_1(n)$ along some path γ with $t(\gamma) < k\delta$. Let

$$\ell_n = \lfloor n(\mu+\varepsilon)/\delta \rfloor \qquad \text{so that } \ell_n\delta \le n(\mu+\varepsilon) \le (l_n+1)\delta. \tag{4.3}$$

On the event $\{\phi_{1,n} \ge n(\mu+\varepsilon)\}$ the goal face $R_1(n)$ is contained in a certain connected component, say O_k, of $[1,n]^d \backslash G_k$ for each $1 \le k \le \ell_n$. Define

$$S_k = \{v \in G_k : \text{ there is a vertex } u \in O_k \text{ with } \|u - v\| = 1\}. \tag{4.4}$$

Roughly speaking, these random sets $\{S_k : 1 \le k \le \ell_n\}$ propagate from $F_1(n)$ toward $R_1(n)$. A key feature about S_k that will be used later is that any path γ running from $F_1(n)$ to $R_1(n)$ must cross S_k somewhere. Hence there is an edge $e = \langle v, u\rangle \in \gamma$ with $v \in S_k \subseteq G_k$ and $u \in O_k$. In particular,

$$\inf\{t(\gamma) : \gamma \text{ is a path from } F_1(n) \text{ to } v \text{ and } \gamma \subset [1,n]^d\} < k\delta \qquad \text{for } v \in S_k. \tag{4.5}$$

The event $\{\phi_{1,n} \ge n(\mu+\varepsilon)\}$ is classified by the intersections of $\{S_k\}$ with the faces of $[1,n]^d$ as follows:

$$\begin{aligned} &\{\phi_{1,n} \ge n(\mu+\varepsilon)\} \\ &\qquad \subseteq \bigcup \{U_{i,k}(w) = C_{i,k}, L_{i,k}(w) = D_{i,k} \text{ for } 2 \le i \le d \text{ and } 1 \le k \le \ell_n\}, \end{aligned} \tag{4.6}$$

where $U_{i,k}(w) = S_k(w) \cap R_i(n)$, $L_{i,k}(w) = S_k(w) \cap F_i(n)$ and the union is taken over all feasible nonempty deterministic subsets $C_{i,k}$ in $R_i(n)$ and $D_{i,k}$ in $F_i(n)$. Let N_n be the number of terms in the union of (4.6). Among these N_n terms let us assume that

$$Z_n = \{U_{i,k}(w) = \tilde{C}_{i,k}, L_{i,k}(w) = \tilde{D}_{i,k} \text{ for } 2 \le i \le d \text{ and } 1 \le k \le \ell_n\} \tag{4.7}$$

has the maximum probability. It is clear from (4.6) that

$$P(\phi_{1,n} \geq n(\mu + \varepsilon)) \leq N_n \cdot P(Z_n). \tag{4.8}$$

We claim the following lemma.

LEMMA 4.1. $\limsup(\log P(Z_n))/n^d \geq \kappa$.

PROOF. By (4.1) and (4.8) it suffices to show via a combinatorial argument that

$$\lim(\log N_n)/n^d = 0. \tag{4.9}$$

Those prescribed sets $\{C_{i,k}\}$ and $\{D_{i,k}\}$ on the faces $R_i(n)$ and $F_i(n)$, respectively, could be very complicated. Fortunately here we need not to know such details. Since there are $2(d-1)$ faces involved, we have from symmetry that

$$N_n \leq \{\# \text{ of choosing } C_{2,1}, C_{2,2}, \dots, C_{2,\ell_n} \text{ on } R_2(n)\}^{2(d-1)}. \tag{4.10}$$

The sets $\{C_{2,i}; 1 \leq i \leq \ell_n\}$ are determined once we know the following mapping which associates to each lattice point v on the face $R_2(n)$ a pair of integers (t, s), where $t = \inf\{k \leq \ell_n : v \in G_k\} \wedge (\ell_n + 1)$ and $s = \#\{k \leq \ell_n : v \in S_k\}$. Hence,

$$\# \text{ of choosing } C_{2,1}, C_{2,2}, \dots, C_{2,\ell_n} \text{ on } R_2(n) \leq (n^{d-1})^{(\ell_n+1)^2}.$$

Equation (4.9) is easily verified by using (4.3) and (4.10). □

Now we proceed to show $\kappa = \kappa'$ defined in (4.1). Fix $\eta > 0$. By Lemma 4.1 there exist m, which can be made arbitrarily large, and $\{\tilde{C}_{2,k}, \tilde{D}_{2,k} : 1 \leq j \leq l_m\}$ such that

$$\log P(Z_m) \geq (\kappa - \eta)m^d, \tag{4.11}$$

where Z_m is given in (4.7) with n there replaced by m. We will use the event Z_m on the basic block $[1, m]^d$ and its independent copies to build up events on a larger block. The enlargement is done sequentially along the dth, $(d-1)$st, ... and finally the first coordinate. For brevity we will consider only the case $d = 3$ in the following. The other cases can be treated similarly.

We start with $[1, m]^2 \times [1, 2m]$. We set the event Z_m on $A_1 = [1, m]^3$ and event $\bar{Z}_m$ on $A_2 = [1, m]^2 \times [m + 1, 2m]$, which is obtained by reflecting an independent copy of Z_m on A_1 with respect to the hyperplane $x_3 = m + 1/2$. We require $t(e) > 2\delta$ for any edge e in

$$E = \{\langle (k, j, m), (k, j, m + 1)\rangle : 2 \leq k \leq m \text{ and } 1 \leq j \leq m\},$$

which connects A_1 and A_2. Denote by V_3 the event thus obtained on $[1, m]^2 \times [1, 2m]$. By symmetry,

$$P(\bar{Z}_m) = P(Z_m). \tag{4.12}$$

Moreover, the construction shows both $L_{3,k}(\omega)$ in event $\bar{Z}_m$ and $U_{3,k}(\omega)$ in event Z_m equal to $\tilde{C}_{3,k}$ except being situated respectively on hyperplanes $x_3 = m+1$ and $x_3 = m$, which can be connected vertically by edges in E above. As a result, hyperface $S_k(\omega)$ in Z_m could be joined vertically with $S_k(\omega)$ in $\bar{Z}_m$ for $1 \le k \le l_m$. We claim that

$$\phi([1,m] \times [1,m] \times [1,2m]) \ge l_m \delta \qquad \text{on event } V_3 \tag{4.13}$$

by showing that any optimal path π in $[1,m]^2 \times [1,2m]$ will stay either in A_1 or in A_2 before crossing S_{l_m} in A_1 or that in A_2. [See (1.2) for definition of ϕ.] Otherwise, some connecting edge in E appears in π. Let $\bar{e} = \langle v, u\rangle$ be the first such edge. (See Figure 1 for a two-dimensional version.) We may assume $v \in A_1$ without loss of generality. Let $k = \max\{1 \le j \le \ell_m : v \in O_j\}$, which is assumed nonempty temporarily. It is clear from the definition that v lies between hyperfaces

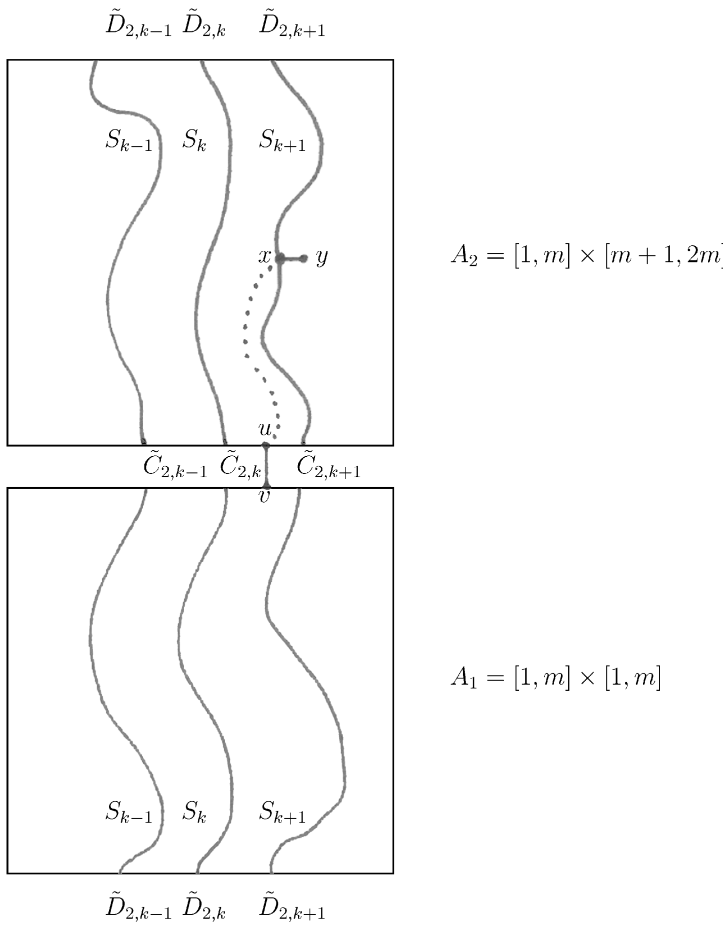

FIG. 1.

S_k and S_{k+1} in A_1. Then u lies between S_k and S_{k+1} in A_2. The path π must cross S_{k+1} in A_2 or that in A_1 in order to reach the goal face $\{m\} \times [1, m] \times [1, 2m]$. Suppose it happens in A_2 with $\tilde{e} = \langle x, y \rangle$ as the crossing edge (see Figure 1). By (4.5) there is a path γ from $\{1\} \times [1, m] \times [m + 1, 2m]$ to x with

$$t(\gamma) < (k + 1)\delta. \tag{4.14}$$

Since v lies outside S_k in A_1 and the subpath of π from its starting point, say, w to v lies in A_1, the passage time of such subpath from w to x is no less than $k\delta + t(\bar{e}) \geq (k + 2)\delta$, which is worse than $t(\gamma)$ in (4.14). Similarly, the crossing of π through S_{k+1} cannot happen in A_1 as it would cost the path π at least 2δ more for using another connecting edge in E in order to get back to A_1. In case $\{1 \leq j \leq \ell_m : v \in O_j\} = \varnothing$, it means all $\{U_{3,k}(\omega) : 1 \leq k \leq \ell_m\}$ in A_1 coincide at v and thus all $\{L_{3,k}(\omega) : 1 \leq k \leq \ell_m\}$ in A_2 coincide at u. Since $L_{i,k}(w) = S_k(w) \cap F_i(n)$, $u \in S_1$ in A_2 in this case. Because $t(\bar{e}) > 2\delta$, the subpath of π from w to u cannot be optimal in view of (4.5).

Let $n = Mm + r$, where $M = \lfloor n/m \rfloor$ and $0 \leq r < m$. Hence

$$n - m \leq Mm. \tag{4.15}$$

We may add, along the x_3 coordinate, alternately more independent copies of Z_m and $\bar{Z}_m$ on top of the previous $[1, m]^2 \times [1, 2m]$ until the block $H_1 = [1, m]^2 \times [1, (M + 1)m]$ is reached. As before we require $t(e) > 2\delta$ for all edges connecting these $(M + 1)$ blocks each of size $[1, m]^3$. Denote by W_3 the resulted event on H_1. We can show as we did in (4.13) that

$$\phi(H_1) \geq l_m\delta \qquad \text{on event } W_3 \tag{4.16}$$

and thus $W_3 \subseteq \{\phi(H_1) \geq l_m\delta\}$. Since there are $(m - 1)mM$ connecting edges in block H_1, we get from (4.11), (4.12) and (4.2) that

$$\begin{aligned} \log P\big(\phi(H_1) \geq l_m\delta\big) &\geq \log P(W_3) \\ &\geq (\kappa - \eta)m^3(M + 1) + (m - 1)mM \log r. \end{aligned} \tag{4.17}$$

Now we will enlarge the block H_1 along the second coordinate. Let $\bar{W}_3$ be the event on $H_2 = [1, m] \times [m + 1, 2m] \times [1, (M + 1)m]$ obtained from reflecting an independent copy of W_3 on H_1 with respect to the hyperplane $x_2 = m + 1/2$. Write

$$H_i = \bigcup_{j=1}^{M+1} \{[1, m] \times [(i - 1)m + 1, im] \times [(j - 1)m + 1, jm]\} \equiv \bigcup_{j=1}^{M+1} H_{i,j}$$

so that each $H_{i,j}$ is identical to the basic block $[1, m]^3$ and $H_{1,j}$, $H_{2,j}$ could be connected by edges in

$$E_j = \{\langle (s, m, k), (s, m + 1, k) \rangle : 2 \leq s \leq m \text{ and } (j - 1)m + 1 \leq k \leq jm\}.$$

As before we require $t(e) > 2\delta$ for any edge in $\bigcup_{j=1}^{M+1} E_j$ and denote by V_2 the event thus obtained on $[1,m] \times [1,2m] \times [1,(M+1)m]$. Because any edge connecting H_1 and H_2 must belong to a certain E_j and a result similar to (4.13) holds on each $H_{1,j} \cup H_{2,j}$, the same argument shows

$$\phi\big([1,m] \times [1,2m] \times [1,(M+1)m]\big) \geq l_m\delta \qquad \text{on event } V_2.$$

As in (4.16) we then add, along the x_2 coordinate, alternately more independent copies of W_3 and $\bar{W}_3$ on top of the previous $[1,m] \times [1,2m] \times [1,(M+1)m]$ until the block $I = [1,m] \times [1,(M+1)m]^2$ is reached. We still require $t(e) > 2\delta$ for any edge connecting these $(M+1)$ blocks each of the same size as H_1. Denote by W_2 the resulted event on I. Similar to (4.16) and (4.17), we will have

$$\phi(I) \geq l_m\delta \qquad \text{on event } W_2 \tag{4.18}$$

and thus $W_2 \subseteq \{\phi(I) \geq l_m\delta\}$. Using (4.17) and the fact that there are $(m-1) \times m(M+1)$ connecting edges between blocks H_1 and H_2, we get

$$\begin{aligned} \log P\big(\phi(I) \geq l_m\delta\big) &\geq \log P(W_2) \\ &\geq (\kappa-\eta)m^3(M+1)^2 + 2(m-1)mM(M+1)\log r. \end{aligned} \tag{4.19}$$

Now taking M independent copies of the event W_2 in I and putting these M blocks side by side along the first coordinate, we get a big block $J = [1,Mm] \times [1,(M+1)m]^2$. As before we require $t(e) > 2\delta$ for all the $(M-1)(M+1)^2m^2$ horizontal edges connecting these M blocks. Denote by W_1 the resulted event on J. By (4.18), (4.3) and (4.15), the face–face first-passage time $\phi(J)$ on the event W_1 is at least

$$\begin{aligned} Ml_m\delta + (M-1)2\delta &\geq M\big(m(\mu+\varepsilon) - \delta\big) + (M-1)2\delta \\ &\geq n(\mu+\varepsilon) - m(\mu+\varepsilon) + \lfloor n/m \rfloor\delta - 2\delta \\ &\geq n(\mu+\varepsilon) \end{aligned} \tag{4.20}$$

if n is large. As J is thinner in the x_1 coordinate but thicker in both the x_2 and x_3 coordinates than $[1,n]^3$, the face–face first-passage time $\phi_{1,n}$ in $[1,n]^3$ is no less than that in J. By (4.2) and (4.20),

$$P\big(\phi_{1,n} \geq n(\mu+\varepsilon)\big) \geq P(W_1) \geq \{P(W_2)\}^M \cdot r^{(M-1)(M+1)^2m^2}.$$

Since $M = \lfloor n/m \rfloor$, (4.19) and (4.1) imply

$$\kappa \geq \kappa' \geq (\kappa-\eta) + \big\{2(m-1)m^{-2} + m^{-1}\big\}\log r.$$

Letting first $m \to \infty$ and then $\eta \to 0$, we get $\kappa' = \kappa$. Theorem 3 is proved.

Acknowledgments. Thanks to an anonymous referee for tightening up the proof of Lemma 4.1. The second author would like to thank the hospitality of the Institute of Mathematics, Academia Sinica. Part of the work was done when he visited there.

REFERENCES

DURRETT, R. (1996). *Probability*: *Theory and Examples*, 2nd ed. Wadsworth, Belmont, CA.

GRIMMETT, G. and KESTEN, H. (1984). First-passage percolation, network flows and electrical resistances. *Z. Wahrsch. Verw. Gebiete* **66** 335–366.

HAMMERSLEY, J. M. and WELSH, D. J. A. (1965). First-passage percolation, subadditive processes, stochastic networks and generalized renewal theory. In *Bernoulli*, *Bayes*, *Laplace Anniversary Volume* (J. Neyman and L. Le Cam, eds.) 61–110. Springer, Berlin.

KESTEN, H. (1986). *Aspects of First-Passage Percolation*. *Lecture Notes in Math.* **1180** 125–264. Springer, Berlin.

SMYTHE, R. T. and WIERMAN, J. C. (1978). *First Passage Percolation on the Square Lattice*. *Lecture Notes in Math.* **671**. Springer, Berlin.

ZHANG, Y. (1995). Supercritical behaviors in first-passage percolation. *Stochastic Process. Appl.* **59** 251–266.

INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
TAIPEI
TAIWAN 11529
E-MAIL: chow@math.sinica.edu.tw

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF COLORADO
COLORADO SPRINGS, COLORADO 80933
E-MAIL: yzhang@math.uccs.edu