The Annals of Applied Probability
1993, Vol. 3, No. 1, 56–80

INEQUALITIES FOR THE TIME CONSTANT IN FIRST-PASSAGE PERCOLATION

BY J. VAN DEN BERG[1] AND H. KESTEN[2]

CWI and Cornell University

Consider first-passage percolation on $\mathbb{Z}^d$. A classical result says, roughly speaking, that the shortest travel time from $(0, 0, \dots, 0)$ to $(n, 0, \dots, 0)$ is asymptotically equal to $n\mu$, for some constant μ, which is called the *time constant*, and which depends on the distribution of the time coordinates. Except for very special cases, the value of μ is not known. We show that certain changes of the time coordinate distribution lead to a decrease of μ; usually μ will strictly decrease. Two examples of our results are:

(i) If F and G are distribution functions with $F \leq G$, $F \not\equiv G$, then, under mild conditions, the time constant for G is *strictly* smaller than that for F.

(ii) For $0 < \varepsilon_1 < \varepsilon_2 \leq a < b$, the time constant for the uniform distribution on $[a - \varepsilon_2, b + \varepsilon_1]$ is strictly smaller than for the uniform distribution on $[a, b]$.

We assume throughout that all our distributions have finite first moments.

1. Introduction.

First-passage percolation was started by Hammersley and Welsh (1965). [See Smythe and Wierman (1978) and Kesten (1986, 1987) for more information.]

In this paper we restrict ourselves to the d-dimensional cubic lattice with $d \geq 2$. The vertices of this lattice are the elements of $\mathbb{Z}^d$ and will typically be denoted by v or w. The special vertices $(0, \dots, 0)$ and $(n, 0, \dots, 0)$ will be denoted by $\mathbf{0}$ and $\mathbf{n}$, respectively.

If $v = (v_1, \dots, v_d)$ and $w = (w_1, \dots, w_d)$ we denote

$$\|v - w\| = \sum_{i=1}^{d} |v_i - w_i|. \tag{1.1}$$

By the *distance* between v and w we mean $\|v - w\|$.

Two vertices are said to be *adjacent* or *neighbors* if their distance is 1. The edges in the lattice are the line segments between adjacent vertices. They will typically be denoted by e, possibly with a subscript or superscript. The set of all edges is denoted by E.

Received September 1991; revised March 1992.

[1]Supported by the Army Research Office through the Mathematical Sciences Institute (MSI) at Cornell University.

[2]Partially supported by the NSF through a grant to Cornell University.

AMS 1991 *subject classifications*. 60K35, 82A42.

Key words and phrases. First-passage percolation, time constant, inequality.

Paths will typically be denoted by π. If e is an edge on π, we write, with some abuse of notation, $e \in \pi$. The *length* of a path π is the number of edges in π and denoted by $|\pi|$.

Now we assign to each edge e a random variable $t(e)$, which is called the *time coordinate* of e, and which can be interpreted as the time it takes to traverse e. We assume that the $t(e)$, $e \in E$, are i.i.d. nonnegative random variables and denote their common distribution (*the time coordinate distribution*) by F. *Throughout this paper we assume that the time coordinate distribution has finite first moment.*

The *travel time* $t(\pi)$ of a path π is defined by

$$t(\pi) = \sum_{e \in \pi} t(e). \tag{1.2}$$

The *shortest travel time* from vertex v to vertex w, $t(v, w)$, is defined as

$$t(v, w) = \inf\{t(\pi): \pi \text{ is a path from } v \text{ to } w\}. \tag{1.3}$$

The shortest travel time from $\mathbf{0}$ to $\mathbf{n}$ will be denoted by a_n, that is,

$$a_n = t(\mathbf{0}, \mathbf{n}). \tag{1.4}$$

A classical result, based on the subadditive ergodic theorem, says that there exists a constant $\mu = \mu(F) < \infty$ such that

$$\frac{a_n}{n} \to \mu \ (n \to \infty) \quad \text{a.s. and in } L_1. \tag{1.5}$$

μ is called the *time constant*. Since μ has an obvious interpretation in terms of optimization problems, its determination as a functional of F (and d) is a basic problem in first-passage percolation. Unfortunately not much progress has been made in this direction. It appears even to be a very hard problem to give accurate rigorous estimates of $\mu(F)$ [see Hammersley and Welsh (1965), Section 6, Smythe and Wierman (1978), Section 7.2, Smythe (1980), Janson (1981) and Ahlberg and Janson (1984)].

Our results here, which give an inequality between $\mu(\tilde{F})$ and $\mu(F)$ for certain pairs $(\tilde{F}, F)$ can be some help in carrying over estimates for $\mu(\tilde{F})$ to $\mu(F)$ and vice versa.

Some further motivation for our inequalities came from the result of Cox (1980) and Cox and Kesten (1981) that $\mu(F)$ is continuous in F, that is, if F_n converges to F weakly, then $\mu(F_n) \to \mu(F)$. In particular this implies that if F has unbounded support, and F_x is the distribution obtained from F by truncating at x, then $\mu(F_x) \to \mu(F)$ as $x \to \infty$. This raises the question whether one could even have

$$\mu(F_x) = \mu(F) \quad \text{for sufficiently large } x.$$

Our Theorem 2.13 shows that, under mild conditions, this is not the case, that is, for every x, $\mu(F_x)$ is strictly smaller than $\mu(F)$.

Another natural question is what happens to the time constant if the time coordinate distribution is "stretched out" by a certain factor, as in the following case: Let $0 < \varepsilon < a < b$, F the uniform distribution on $[a, b]$ and $\tilde{F}$

the uniform distribution on $[a - \varepsilon, b + \varepsilon]$. It is not so difficult to show that $\mu(\tilde{F}) \le \mu(F)$. The fact that even strict inequality holds is more difficult to prove, but follows from Theorem 2.9(b).

Finally, some more notational remarks:

We use the term increasing where some people use nondecreasing. For example, g is increasing on A means $g(y_2) \ge g(y_1)$ for $y_1 < y_2$, $y_i \in A$. A similar convention is adopted for "decreasing."

Important "universal" constants (which only depend on the dimension d, e.g., the number of vertices at distance 10 from 0) will mostly be denoted by C_1, C_2 and so on. Symbols $D_1, D_2, \dots$ denote important constants which may depend on various quantities in our problem (such as the time coordinate distribution) but not on n. Such dependence will be indicated in the notation when the constant D_i is introduced.

If V is a set of vertices then ∂V denotes the *interior boundary* of V, that is, the set of all vertices in V which have a neighbor outside V. The cardinality of V is denoted by $|V|$. The *diameter* of V, that is, $\max\{\|v - w\|, v, w \in V\}$, is denoted by $\operatorname{diam}(V)$.

In several places in this paper we have to deal with one or more families of time coordinates in addition to $\{t(e)\colon e \in E\}$. These will usually be denoted by $\{\tilde{t}(e)\colon e \in E\}$, $\{t^*(e)\colon e \in E\}$ and so on. In these cases we define $\tilde{t}(v, w)$, $\tilde{a}_n$, $t^*(v, w)$, a_n^* and so on as the obvious analogues of $t(v, w)$, a_n and so on. For instance if π is a path, then $\tilde{t}(\pi) = \Sigma\tilde{t}(e)$.

2. Statement of results.

2. Statement of results. Our principal result is Theorem 2.9. Its basic condition is phrased in terms of the following partial order between distribution functions on $\mathbb{R}$.

(2.1) Definition. For the distribution functions F and $\tilde{F}$ we say that $\tilde{F}$ is *more variable* than F if

$$\int \varphi(x)\, d\tilde{F}(x) \le \int \varphi(x)\, dF(x) \tag{2.2}$$

for every concave increasing function $\varphi\colon \mathbb{R} \to \mathbb{R}$ for which the two integrals in (2.2) converge absolutely.

This partial order between distribution functions has been used in reliability and queueing theory and has a long history. [See Stoyan and Daley (1983), Section 1.4 or Ross (1983), Section 8.5 and their references. Note that our terminology does not quite agree with that of Ross (1983). Ross' order is the same as the convex order of Stoyan and Daley (1983), while we use the concave order of the latter reference. The two orders are equivalent when F and $\tilde{F}$ have the same mean.] These references also give quite a number of examples of pairs $F, \tilde{F}$ with $\tilde{F}$ more variable than F. A simple criterion for $\tilde{F}$ to be more variable than F, when they both have finite mean, is that

$$\int_{-\infty}^{x} F(y)\,dy \le \int_{-\infty}^{x} \tilde{F}(y)\,dy$$

for all x [cf. Stoyan and Daley (1983), Section 1.4]. Another way to make such examples is by taking a distribution F, and then constructing $\tilde{F}$ by pushing mass away from some point ξ. This is seen in the following sufficient condition.

(2.3) CUT CRITERION OF KARLIN AND NOVIKOFF [see Stoyan and Daley (1983), Section 1.5]. Let F and $\tilde{F}$ be two distributions with finite mean. Assume that

$$\int x\,d\tilde{F}(x) \le \int x\,dF(x), \tag{2.4}$$

and that for some number ξ,

$$\begin{aligned} F(x) &\le \tilde{F}(x) \quad \text{when } x < \xi, \\ F(x) &\ge \tilde{F}(x) \quad \text{when } x > \xi. \end{aligned} \tag{2.5}$$

Then $\tilde{F}$ is more variable than F.

Examples (2.17)–(2.19) are based on this criterion.

It will be crucial to our proof that $\tilde{F}$ more variable than F is equivalent to a condition on a coupled pair of random variables $t, \tilde{t}$ with marginal distributions F and $\tilde{F}$. Specifically, the following result has been known for quite some time [see Strassen (1965), Theorem 9 and Whitt (1980)].

(2.6) THEOREM. *Let F and $\tilde{F}$ be two distributions with finite first moment. $\tilde{F}$ is more variable than F if and only if there exists a pair of random variables $t, \tilde{t}$ on one probability space, with marginal distributions F and $\tilde{F}$, respectively, and satisfying*

$$f(y) := E\{\tilde{t} \mid t = y\} \le y \tag{2.7}$$

for almost all y $[F]$.

We need one more definition before we state our principal result.

(2.8) DEFINITION. Let F be a distribution with supp(F) (= support of F) $\subset [0, \infty)$. Let $r = \min(\operatorname{supp}(F))$. We say that F is *useful* if

$$\begin{aligned} F(r) &< p_c \quad \text{in case } r = 0, \\ F(r) &< \vec{p}_c \quad \text{in case } r > 0. \end{aligned}$$

Here p_c and $\vec{p}_c$ stand for the critical probability for bond percolation and oriented bond percolation on $\mathbb{Z}^d$, respectively.

(2.9) THEOREM. (a) *Let F and $\tilde{F}$ be two distributions on $[0,\infty)$ with finite mean such that $\tilde{F}$ is more variable than F. Then*

$$\mu(\tilde{F}) \leq \mu(F). \tag{2.10}$$

(b) *If in addition $d \geq 2$, F is useful and $\tilde{F} \not\equiv F$, then even*

$$\mu(\tilde{F}) < \mu(F). \tag{2.11}$$

The next result is a special case of Theorem 2.9(b).

(2.12) DEFINITION. For the probability distributions F and G, we say that *F strictly dominates G* if $F(x) \leq G(x)$ for all x, but $F \not\equiv G$.

(2.13) THEOREM. *Let F and $\tilde{F}$ have finite mean. If F is useful and strictly dominates $\tilde{F}$, then*

$$\mu(\tilde{F}) < \mu(F).$$

(2.14) REMARK. The requirement $F(0) < p_c$ for a useful F cannot be dropped, because if $F(0) \geq p_c$, then $\mu(F) = 0$ [cf. Kesten (1986), Theorem 6.1]. We do not know whether (2.11) remains valid if $r = \min(\text{supp}(F)) > 0$ and $F(r) \geq \vec{p}_c$.

(2.15) REMARK. It can be shown fairly easily that the validity of Theorem 2.13 for every pair $F, \tilde{F}$ in which F is useful and strictly dominates $\tilde{F}$ is equivalent to the following result: Let F be useful and have finite mean. Let A be a Borel set with $F(A) > 0$. If $\pi(n)$ is an optimal path on which a_n is achieved (see Section 4 for more precise definition), then

$$\liminf \frac{1}{n} E\{\text{number of } e \text{ in } \pi(n) \text{ with } t(e) \in A\} > 0. \tag{2.16}$$

(2.16) can also be expressed heuristically as "F is absolutely continuous with respect to the limit of the expected empirical distribution of the $t(e)$ along the optimal path."

The remainder of this section consists of examples. Section 3 contains the proof of Theorems 2.9(a), and the reduction of Theorem 2.13 to 2.9(b). Section 4 gives some auxiliary results for Theorem 2.9(b). Finally, Section 5 gives the main step in the proof of Theorem 2.9(b), namely a "geometrical" construction.

(2.17) EXAMPLE. Consider the following case mentioned in the abstract. F is the uniform distribution on $[a, b]$ and $\tilde{F}$ the uniform distribution on $[a - \varepsilon_2, b + \varepsilon_1]$ with $0 < \varepsilon_1 \leq \varepsilon_2$ so that (2.4) holds. $\tilde{F}$ is more variable than F by the cut criterion (2.3). Indeed (2.5) is obvious from the fact that F and $\tilde{F}$ are just linear functions on $[a, b]$ and $[a - \varepsilon_2, b + \varepsilon_1]$, respectively. Theorem

2.9(b) applies. Thus if $d \geq 2$, then

$$\mu(\tilde{F}) < \mu(F)$$

in this example.

(2.18) EXAMPLE. Let $0 \leq a < b < \infty$ and F, $\tilde{F}$ two distributions with

$$\operatorname{supp}(F) \subset [a, b] \quad \text{and} \quad \operatorname{supp}(\tilde{F}) \subset \{a, b\}.$$

If (2.4) holds, then $\tilde{F}$ is more variable than F, again by the cut criterion (2.3) [again (2.5) is easy to check in this case]. Thus, (2.10) holds. If F is useful, $d \geq 2$, and F is not concentrated on a and b only, then even

$$\mu(\tilde{F}) < \mu(F).$$

(2.19) EXAMPLE. For $a < b$, let $U[a, b]$ be the uniform distribution on the interval $[a, b]$. For integers $0 \leq l \leq m$, let $U\{l, \dots, m\}$ be the uniform distribution on the set of integers $\{l, \dots, m\}$. We claim that, for $d = 2$ and $1 \leq l < m$,

$$\mu\big(U\big[l - \tfrac{1}{2}, m + \tfrac{1}{2}\big]\big) < \mu(U\{l, \dots, m\}) < \mu(U[l, m]). \tag{2.20}$$

The left-hand inequality follows because $U[l - 1/2, m + 1/2]$ is the convolution of $U\{l, \dots, m\}$ and $U[-1/2, 1/2]$. It is now easy to see from the definition and Jensen's inequality that $U[l - 1/2, m + 1/2]$ is more variable than $U\{l, \dots, m\}$. Alternatively we can apply Theorem 2.6 with $\tilde{t} = t + u$, where t and u are independent random variables with distributions $U\{l, \dots, m\}$ and $U[-1/2, 1/2]$, respectively. Moreover, for $d = 2$ and $m > l \geq 1$, $U\{l, \dots, m\}$ is useful since it puts mass $(m - l + 1)^{-1} \leq 1/2 < \vec{p}_c$ (for $d = 2$) on $\{l\}$.

For the right-hand inequality in (2.20) we introduce the intermediate distribution G which puts mass $[2(m - l)]^{-1}$ on each of the points l and m, and mass $(m - l)^{-1}$ on each integer j, $l < j < m$. G is obtained from $U[l, m]$ by moving one half of the mass in $[j, j + 1]$ to j and half to $(j + 1)$. For the same reasons as in the preceding example, G is more variable than $U[l, m]$. Once again the cut criterion of Karlin and Novikoff can now be used to show that $U\{l, \dots, m\}$ is more variable than G.

(2.21) REMARK. It follows from (2.20) and the continuity of μ that for some $c \in (0, 1/2)$, $\mu(U[l - c, m + c]) = \mu(U\{l, \dots, m\})$.

(2.22) REMARK. It is easy from the above to make examples of pairs of distribution functions F and $\tilde{F}$ with

$$0 < \mu(\tilde{F}) < \mu(F) \quad \text{even though} \int_{x \geq 0} x\, d\tilde{F}(x) > \int_{x \geq 0} x\, dF(x). \tag{2.23}$$

For instance, by (2.20) this holds for $\tilde{F} = U([l - 1/2 + \varepsilon, m + 1/2 + \varepsilon])$, $F = U\{l, \dots, m\}$ when $\varepsilon > 0$ is sufficiently small. Other examples of (2.23) appear already in Smythe and Wierman [(1978), Section 7.3].

3. Proofs of Theorems 2.9(a) and 2.13.

PROOF OF THEOREM 2.9(a). This is almost immediate from the definitions. It is easy to see [compare Ross (1983), Proposition 8.54 and Stoyan and Daley (1983), equation (1.10.5)] that if $\tilde{F}$ is more variable than F, and φ is an increasing concave function from $\mathbb{R}^k \to \mathbb{R}$, then

$$E\varphi(\tilde{t}_1, \dots, \tilde{t}_k) \le E\varphi(t_1, \dots, t_k) \tag{3.1}$$

whenever $\tilde{t}_1, \dots, \tilde{t}_k$ are i.i.d. with distribution $\tilde{F}$, and $t_1, \dots, t_k$ are i.i.d. with distribution F. We apply this with

$$\varphi = a_{n,N} = \min\{t(\pi): \pi \text{ a path from } \mathbf{0} \text{ to } \mathbf{n} \text{ in the cube } [-N, N]^d\}.$$

Clearly $a_{n,N}$ is an increasing function of the variables $t(e)$, $e \in [-N, N]^d$. Since $a_{n,N}$ is the min of the linear functions $t(\pi) = \Sigma_{e \in \pi} t(e)$, it is also a concave function of the $t(e)$. Thus, by (3.1),

$$E\tilde{a}_{n,N} \le Ea_{a,N}$$

(see end of Section 1 for $\tilde{a}$). It follows, by letting $N \to \infty$, that

$$E\tilde{a}_n \le Ea_n. \tag{3.2}$$

This, together with (1.5), implies (2.10). □

PROOF OF THEOREM 2.13 FROM THEOREM 2.9(b). It is a standard fact [cf. Stoyan and Daley (1983), Section 1.2] that $F(x) \le \tilde{F}(x)$ for all x implies (2.2) for all increasing functions φ. Hence if F strictly dominates $\tilde{F}$, then $\tilde{F}$ is more variable than F. Thus, once we prove Theorem 2.9(b), (2.11) will follow. [Note that the case $d = 1$ is trivial in this case, for then $\mu(\tilde{F}) = \int x\, d\tilde{F}(x) < \int x\, dF(x) = \mu(F)$.] □

4. Preliminaries for Theorem 2.9(b). First we show that we may restrict ourselves in Theorem 2.9(b) to the case where

$$F(0) \le \tilde{F}(0) < p_c. \tag{4.1}$$

To see this, note first that if $\tilde{F}$ is more variable than F, and F and $\tilde{F}$ are concentrated on $[0, \infty)$, then $\tilde{F}(0) \ge F(0)$. This follows from Theorem 2.6. Indeed, let $(t, \tilde{t})$ be a pair of random variables with marginal distributions F and $\tilde{F}$ and satisfying (2.7). Then (2.7) for $y = 0$ implies $\tilde{t} = 0$ almost everywhere on $\{t = 0\}$, or $\tilde{F}(0) \ge F(0)$.

Thus, the first inequality in (4.1) holds. The second inequality can fail for a useful F only if $F(0) < p_c \le \tilde{F}(0)$. But in this case $\mu(F) > 0$ and $\mu(\tilde{F}) = 0$, by Kesten [(1986), Theorem 6.1 and the equality $p_T = p_c$ which was proved by Menshikov (1986) and Aizenman and Barsky (1987)]. Thus, in this case (2.11) is trivial.

$F(0) < p_c$ guarantees that the infimum in (1.3) is taken on. Similarly $\tilde{F}(0) < p_c$ implies that

$$\tilde{a}_n = \min\{\tilde{t}(\pi)\text{: is a path from } \mathbf{0} \text{ to } \mathbf{n}\}$$

[see Kesten (1986), Remark 9.23]. We may therefore define the *optimal path* from v to w as a path π with $t(\pi) = t(v, w)$, when we are using time coordinates with distribution F; similarly we can find $\tilde{\pi}$ such that $\tilde{t}(\tilde{\pi}) = \tilde{t}(v, w)$. To define the optimal path uniquely we order all paths on $\mathbb{Z}^d$ in some arbitrary way, and if several paths from v to w have the minimal travel time $t(v, w)$, then we choose the first one in our ordering as the optimal path. The optimal paths from $\mathbf{0}$ to $\mathbf{n}$ will be denoted by $\pi(n)$ and $\tilde{\pi}(n)$, respectively. Thus [cf. (1.4)]

$$\begin{aligned} a_n &= t(\mathbf{0}, \mathbf{n}) = t(\pi(n)), \\ \tilde{a}_n &= \tilde{t}(\mathbf{0}, \mathbf{n}) = \tilde{t}(\tilde{\pi}(n)). \end{aligned} \tag{4.2}$$

Since $\tilde{\pi}(n)$ is the optimal path for $\tilde{a}_n$ this gives

$$E\{\tilde{a}_n\} = E\{\tilde{t}(\tilde{\pi}(n))\} \leq E\{\tilde{t}(\pi(n))\}. \tag{4.3}$$

Our approach is to show that in turn

$$E\{\tilde{t}(\pi(n))\} \leq E\{t(\pi(n))\} - \eta n = E\{a_n\} - \eta n \tag{4.4}$$

for some $\eta > 0$. This will imply (2.11) by means of (1.5).

We next prove two technical lemmas which will be useful in the next section. In the sequel $(t, \tilde{t})$ is a pair of random variables with marginal distributions F and $\tilde{F}$, for which (2.7) holds. Such a pair exists under the hypotheses of Theorem 2.9.

(4.5) LEMMA. *It suffices to prove Theorem* 2.9(b) *under the additional assumption*

$$P\{\tilde{t} > t\} > 0. \tag{4.6}$$

PROOF. Assume that the conditions of Theorem 2.9(b) hold, but that (4.6) fails. Let ρ be a random variable independent of $(t, \tilde{t})$ and with distribution

$$P\{\rho = +1\} = P\{\rho = -1\} = \tfrac{1}{2}.$$

Then define

$$\bar{t} = t + \rho(t - \tilde{t}).$$

If (4.6) fails, then

$$P\{\bar{t} \geq \tilde{t}\} = 1.$$

Consequently, if $\bar{F}$ denotes the distribution of $\bar{t}$, then $\mu(\bar{F}) \geq \mu(\tilde{F})$. It therefore suffices to prove

$$\mu(\bar{F}) < \mu(F). \tag{4.7}$$

However, by the definition of $\bar{t}$ we have

$$E\{\bar{t}|t\} = t,$$

so that $\bar{F}$ is more variable than F. Also $\tilde{F} \not\equiv F$ shows that if (4.6) fails, then $P\{\tilde{t} < t\} > 0$ and hence

$$P\{\bar{t} > t\} > 0.$$

In addition, if φ_0 is strictly concave, then on $\{t \neq \tilde{t}\}$,

$$E\{\varphi_0(\bar{t})|t, \tilde{t}\} = \tfrac{1}{2}\left[\varphi_0(t + (t - \tilde{t})) + \varphi_0(t - (t - \tilde{t}))\right] < \varphi_0(t),$$

so that $E\varphi_0(\bar{t}) < E\varphi_0(t)$ and $\bar{F} \not\equiv F$.

Thus, if (4.6) fails for our original $(t, \tilde{t})$, then it does hold for $(t, \bar{t})$, and (4.7) implies (2.11). We may therefore add (4.6) to our hypotheses. □

From now on we assume that (4.6) holds. We remind the reader of the notation

$$r = \min(\operatorname{supp}(F)),$$

which will be used throughout the remainder of this paper.

(4.8) LEMMA. *Under the conditions of Theorem* 2.9(b) *and* (4.6) *there exist* $k > 0$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$ *and a bounded Borel set* $I_0 \subset [r, \infty)$ *such that*:

(a) $F([r, \inf(I_0)]) > 0$, $F(I_0) > 0$, *and* $F([\sup(I_0), \infty)) > 0$.

(b) *For all* $y \in I_0$,

$$P\{\tilde{t} > y + \alpha | t = y\} \geq \beta. \tag{4.9}$$

(c) *For all* $y_1, \dots, y_k, y'_1, \dots, y'_{k+2} \in I_0$,

$$\sum_1^{k-2} y'_i < \sum_1^{k} y_i < \sum_1^{k+2} y'_i \tag{4.10}$$

and

$$\sum_1^{k} (y_i + \alpha) > \gamma + \sum_1^{k+2} y'_i. \tag{4.11}$$

PROOF. By (4.6) we can find a Borel set $B \subset [r, \infty)$ such that $F(B) > 0$ and such that, for all $y \in B$,

$$P\{\tilde{t} > y | t = y\} > 0. \tag{4.12}$$

By decreasing B, if necessary, and taking $\alpha, \beta > 0$ sufficiently small we may assume that (4.9) holds for all $y \in B$. If F has an atom $y_0 \in B$, take $I_0 = \{y_0\}$, and take k so large that $k(y_0 + \alpha) > (k + 2)y_0$. Then (4.11) holds for any choice of γ in $(0, k(y_0 + \alpha) - (k + 2)y_0)$. Note further that y_0 cannot be 0, since (4.12) for $y_0 = 0$, together with $\tilde{t} \geq 0$, would give

$$f(0) = E\{\tilde{t}|t = 0\} > 0,$$

in contradiction to (2.7). Thus $y_0 > 0$ and (4.10) is automatic. Also (a) is clear, so that the lemma is proven if F has an atom in B.

If F has no atom in B, let $y_0 > r$ be a point of increase of the conditional distribution of t, given $t \in B$. Then

$$P\{t \in (y_0 - \delta, y_0 + \delta) \cap B\} > 0 \quad \text{for all } \delta > 0.$$

Since the conditional distribution of t, given $t \in B$, has no atoms, its support contains infinitely many points and we can therefore choose $y_0 > r$ such that also

$$F((y_0, \infty)) > 0.$$

Now take k as before, and choose $\delta_0 > 0$ so small that

$$k(y_0 - \delta_0) > (k-2)(y_0 + \delta_0), \qquad (k+2)(y_0 - \delta_0) > k(y_0 + \delta_0),$$
$$k(y_0 - \delta_0 + \alpha) > (k+2)(y_0 + \delta_0),$$
$$F([y_0 + \delta_0, \infty)) > 0, \qquad F([r, y_0 - \delta_0]) > 0.$$

It is trivial to check that (a)–(c) hold for $I_0 = (y_0 - \delta_0, y_0 + \delta_0) \cap B$, $0 < \gamma < k(y_0 - \delta_0 + \alpha) - (k+2)(y_0 + \delta_0)$. □

5. Proof of Theorem 2.9(b): A block construction. Throughout this section the hypotheses of Theorem 2.9(b) are in force. $\{(t(e), \tilde{t}(e)) \colon e \in E\}$ will be an i.i.d. family such that each $(t(e), \tilde{t}(e))$ has the same distribution as a pair $(t, \tilde{t})$ with marginal distributions F and $\tilde{F}$, and satisfying (2.7).

The proof of (4.4) will use a block-rescaling technique. This technique has become rather standard in percolation and related fields. To start we introduce the cubes which will play the role of "renormalized sites."

For each $N > 0$, define the hypercubes $S(l; N)$, $l \in \mathbb{Z}^d$, by

$$S(l;N) = \{v \in \mathbb{Z}^d \colon Nl \le v < N(l+1)\}. \tag{5.1}$$

(Note: $v \le w$ means: $v_i \le w_i$ for $1 \le i \le d$. $v < w$ means $v_i < w_i$ for $1 \le i \le d$.)

For fixed N we call these hypercubes *N-cubes*. Note that for each N the N-cubes form a partition of $\mathbb{Z}^d$. The *distance* between two N-cubes $S(l;N)$ and $S(l';N)$ is defined as $\|l - l'\|$.

The first lemma follows from a standard Peierls argument [compare with Grimmett and Kesten (1984), proof of (3.12)]. We skip its proof.

(5.2) LEMMA. *Suppose that, for each N, all N-cubes are randomly colored black or white in such a way that the process [colors of $S(l;N)$, $l \in \mathbb{Z}^d$] is translation invariant. Moreover, suppose that there exists a constant C_0 (independent of N) such that, for each l and N, the color of $S(l;N)$ is completely determined by the time coordinates of the edges in $\bigcup\{S(l';N) \colon \|l' - l\| \le C_0\}$. Finally, suppose that $\lim_{N\to\infty} P(S(0;N)$ is black$) = 1$. Then, for all sufficiently large N there exist $\varepsilon = \varepsilon(N) > 0$ and $D = D(N) > 0$ such that, for all $v, w \in \mathbb{Z}^d$,*

$$P\{\exists \text{ path from } v \text{ to } w \text{ which visits at most } \varepsilon\|v - w\| \text{ distinct black } N\text{-cubes}\} \le e^{-D\|v-w\|}. \tag{5.3}$$

For the following lemma we need a further family of hypercubes. For each $l \in \mathbb{Z}^d$ and natural number N we define

$$(5.4) \qquad T(l;N) = \{v \in \mathbb{Z}^d : Nl - N \le v \le Nl + 2N\}.$$

Each $T(l;N)$ is a hypercube of size $3N$, containing $S(l;N)$ in its center. Note that both values $Nl_i - N$ and $N(l_i + 1) + N$ are permitted for v_i in (5.4) [this in distinction to (5.1), where a strong inequality was used to delimit v_i on the right]. We shall call these new hypercubes *large N-cubes*.

(5.5) LEMMA. *If the distribution F of the time coordinates is useful, then there exist $\delta = \delta(F) > 0$, and $D_0 = D_0(F) > 0$ such that, for all vertices v, w,*

$$(5.6) \qquad P\{t(v,w) < (r+\delta)\|v-w\|\} \le e^{-D_0\|v-w\|},$$

where, as before, $r = \min(\mathrm{supp}(F))$.

PROOF. The proof is similar in spirit to Lemmas 3.5–3.7 in Grimmett and Kesten (1984), but fortunately somewhat simpler. Since F is useful, we can find $r' > r$ such that

$$(5.7) \qquad P\{t(e) < r'\} < p_c \quad \text{if } r = 0,$$

respectively,

$$(5.8) \qquad r' - r < r \quad \text{and } P\{t(e) < r'\} < \vec{p}_c \quad \text{if } r > 0.$$

We color the N-cubes randomly black or white as follows: In the case $r = 0$, each $S(l;N)$ is colored white if and only if there is a path from some $u_1 \in \partial S(l;N)$ to some $u_2 \in \partial T(l;N)$ with all edges having time coordinate smaller than r'. In the case $r > 0$, $S(l;N)$ is colored white if and only if there is a path from some $u_1 \in \partial S(l;N)$ to some $u_2 \in \partial T(l;N)$ which has length $\|u_1 - u_2\|$ and all of whose edges have time coordinate smaller than r'.

If we call an edge *open* if its time coordinate is smaller than r', then, for given u_1 and u_2, the event above corresponds with the existence of an open path from u_1 to u_2 in the case $r = 0$, and with the existence of an oriented open path (with properly chosen orientation) from u_1 to u_2 in the case $r > 0$. [The orientation should be chosen such that one may pass from a vertex a only to the vertices $a + (\mathrm{sgn}(u_{2,i} - u_{1,i}))e_i$, $1 \le i \le d$, where e_i is the ith coordinate vector.] From the exponential bounds in Hammersley (1957) and van den Berg and Kesten (1985) and the fact that the p_T of some of the early references in percolation equals p_c [Menshikov (1986) and Aizenman and Barsky (1987), especially the latter one in the oriented case; see also Grimmett (1989), Chapter 3] it follows that the probability of the existence of such an open path from u_1 to u_2 decreases exponentially in $\|u_2 - u_1\|$. Since $|\partial S(l;N)|$ and $|(\partial T(l;N)|$ are bounded by powers of N it is clear that

$$(5.9) \qquad \lim_{N\to\infty} P\{S(\mathbf{0},N) \text{ is black}\} = 1.$$

It is also clear that the color of each $S(l;N)$ is determined by the time coordinates in $T(l,N)$. Hence we can apply the preceding lemma. So fix N,

$\varepsilon > 0$ and $D = D(F, N) > 0$ such that, for all vertices v and w:

$$P\{\exists \text{ path from } v \text{ to } w \text{ which visits at most } \varepsilon\|v - w\| \text{ distinct black } N\text{-cubes}\} \le e^{-D\|v-w\|}. \tag{5.10}$$

Now let v and w be given. Let $S(l_v; N)$ be the unique N-cube containing v and assume $w \notin T(l_v; N)$. Take any $\delta > 0$. Suppose $t(v, w) < (r + \delta)\|v - w\|$. Hence there exists a path π from v to w with

$$t(\pi) < (r + \delta)\|v - w\|. \tag{5.11}$$

Let n_B be the number of distinct black N cubes visited by π. Since each cube $T(l; N)$ intersects at most 7^d cubes $T(k; N)$, there exists a sequence of black N-cubes $S(l_1; N), \dots, S(l_j; N)$, each one visited by π, and such that:

(a) $j \ge 7^{-d} n_B - 2$.
(b) For all i and k with $1 \le k \le j$,

$$T(l_i; N) \cap T(l_k; N) = \phi.$$

(c) $w \notin T(l_1; N) \cup \cdots \cup T(l_j; N)$.

We are now ready to complete the proof. It is clear that the travel time from v to w cannot be less than $r\|v - w\|$. In case $r > 0$ this value can only be achieved along a path from v to w which has the minimal length $\|v - w\|$, and all of whose edges have time coordinate as small as possible, namely r. Our coloring procedure is such that an extra amount of at least $(r' - r)$ is needed beyond the minimal time for each segment of π from a point $u_1 \in \partial S(l; N)$ to a point $u_2 \in \partial T(l; N)$ with $S(l; N)$ black (recall that $r' - r < r$ if $r > 0$). This statement remains valid even when $r = 0$.

Since for each l_i, $1 \le i \le j$, the path π visits $S(l_i; N)$ and ends at $w \notin T(l_i; N)$, π must contain a segment from some $u_1 \in \partial S(l_i; N)$ to some $u_2 \in \partial T(l_i: N)$. Moreover these segments can be taken disjoint for different i by (b). Therefore [see (a)]

$$t(\pi) \ge r\|v - w\| + (r' - r)\left(7^{-d} n_B - 2\right). \tag{5.12}$$

From (5.11) and (5.12) we get the following for $\|v - w\| \ge D_1 = D_1(\delta, F)$:

$$n_B \le \frac{C_1}{(r' - r)} \delta \|v - w\|. \tag{5.13}$$

In short, we have for every $\delta > 0$, that the event $\{t(v, w) < (r + \delta)\|v - w\|\}$ implies the existence of a path from v to w which visits at most $C_1(r' - r)^{-1}\delta\|v - w\|$ distinct black N-cubes. Hence, if we take $\delta > 0$ so small that $C_1 \delta/(r' - r) \le \varepsilon$, then by (5.10) for $\|v - w\| \ge D_1$,

$$P\{t(v, w) < (r + \delta)\|v - w\|\} \le e^{-D\|v-w\|}. \tag{5.14}$$

So far we have proved the required result for all v and w with $w \notin T(l_v; N)$ and $\|v - w\| \ge D_1$. The result can be extended to all v and w by adjusting δ and D_0. □

We need also a collection of rectangular boxes. For each natural number N, $l \in \mathbb{Z}^d$ and $-d \le j \le d$, $j \ne 0$, define

$$B^j(l;N) = T(l;N) \cap T\big(l + 2\,\mathrm{sgn}(j)e_{|j|};N\big), \tag{5.15}$$

where again e_j is the j-th coordinate vector. We shall call these $B^j(l;N)$ *N-boxes*. In the case $d = 2$, $B^1(l;N)$ is the $N \times 3N$ closed rectangle $[(l_1+1)N, (l_1+2)N] \times [(l_2-1)N, (l_2+2)N]$ which lies "east" of $S(l;N)$ and has its long side in the "north–south" direction. $B^{-1}(l;N)$ is a similar rectangle "west" of $S(l;N)$ and so on. In general, each N-box is a closed box of size $3N \times \cdots \times 3N \times N \times 3N \times \cdots \times 3N$, and each $S(l;N)$ is "surrounded" by $2d$ N-boxes. The importance of the N-boxes (and the reason why they have been used many times before in percolation) comes from the fact that a path which starts in $S(l;N)$ and ends outside or on the boundary of $T(l;N)$ must have a segment which lies entirely in one of the surrounding N-boxes, and which connects the two opposite large faces of that N-box (i.e., which crosses the N-box "in the short direction"). This fact will be used below. By "crossing an N-box" we will always mean "crossing that N-box in the short direction." Finally we define, for any box B,

$$\overline{B} = \{v \in \mathbb{Z}^d \colon \exists\, w \in B \text{ with } \|v - w\| \le \mathrm{diam}(B)\}. \tag{5.16}$$

Clearly $B \subset \overline{B}$.

We now wish to prove (4.4) for some $\eta > 0$. The idea of the proof is to find sufficiently many N-boxes along $\pi(n)$ where the $\tilde{t}$-values give a certain reduction in the travel time when compared to the t-values. Actually, it turns out to be helpful to introduce an extra randomization. Let $\{\xi(e) \colon e \in E\}$ be another i.i.d. family of random variables, which is independent of $\{(t(e), \tilde{t}(e) \colon e \in E\}$ and with

$$P\{\xi(e) = 0\} = P\{\xi(e) = 1\} = \tfrac{1}{2}. \tag{5.17}$$

Define

$$\hat{t}(e) = \xi(e)\tilde{t}(e) + (1 - \xi(e))t(e) \tag{5.18}$$

and in accordance with our convention at the end of Section 1 use $\hat{a}_n$ to denote the passage time from $\mathbf{0}$ to $\mathbf{n}$ for the $\hat{t}$ values.

(5.19) Lemma.

$$E\{\tilde{a}_n\} \le E\{\hat{a}_n\} \le E\{\hat{t}(\pi(n))\} \le E\{t(\pi(n))\} = E\{a_n\}.$$

Proof. The distribution of $\hat{t}(e)$ is $\hat{F} = (1/2)(F + \tilde{F})$. Since $\tilde{F}$ is more variable than F, it is also more variable than $\hat{F}$. Therefore the first inequality is just (3.2) with F replaced by $\hat{F}$. The second inequality is immediate from the definition of $\hat{a}_n$ as an infimum over all paths from $\mathbf{0}$ to $\mathbf{n}$. Also the equality at the end is just (4.2).

For the third inequality, let $\mathscr{F}$ be the σ-field generated by $\{t(e): e \in E\}$. Then $\pi(n)$ is $\mathscr{F}$-measurable and

$$\begin{aligned} E\{\hat{t}(\pi(n))\} &= E\Big\{\sum_e \hat{t}(e) I[e \in \pi(n)]\Big\} \\ &= E\Big\{\sum_e I[e \in \pi(n)] E\{\hat{t}(e) | \mathscr{F}\}\Big\} \\ &= E\Big\{\sum_e I[e \in \pi(n)] \tfrac{1}{2}\big(E\{\tilde{t}(e) | \mathscr{F}\} + t(e)\big)\Big\} \\ &\qquad\qquad \text{[by (5.18) and (5.17)]} \\ &\le E\Big\{\sum_e I[e \in \pi(n)] t(e)\Big\} \qquad \text{[by (2.7)]} \\ &= E\{t(\pi(n))\}. \end{aligned}$$

□

By (1.5), the inequalities in (5.19) imply

$$\mu(\tilde{F}) \le \mu(\hat{F})$$

and

$$\mu(\hat{F}) \le \mu(F). \tag{5.20}$$

If we can replace (5.20) by a strict inequality, then Theorem 2.9(b) follows. This is roughly done as follows. The third inequality in (5.19) tells us that the $\hat{t}$-travel time *along* $\pi(n)$ is, on the average, no more than the minimal t-travel time from $\mathbf{0}$ to $\mathbf{n}$. The idea is to show that the $\hat{t}$-travel time can be further improved by making "bypasses," that is, by replacing some stretches of $\pi(n)$ by other paths with a shorter $\hat{t}$-travel time. We have to find of the order of n such bypasses such that the total savings in travel time is of order n. This will then result in (4.4) (with $\tilde{t}$ replaced by $\hat{t}$) for some $\eta > 0$ and a strict inequality in (5.20). The bypasses will be constructed by modifying the configuration in certain N-boxes crossed by $\pi(n)$. The "probability cost" of such modifications can be controlled well enough [see (5.47), (5.48) and (5.51)] to obtain a proper lower bound on the expected number of boxes with a bypass.

In a way this method of modification shows that any configuration that *can* occur along the optimal path, actually occurs along the optimal path with a positive frequency (at least in expectation). Modification techniques have been used before in percolation [e.g., Campanino and Russo (1985) Lemma 4.5, Menshikov (1986) and Aizenman and Grimmett (1991)].

We turn to the construction of bypasses. Fix I_0, k and $\alpha - \gamma$ as in Lemma 4.8. If π is a path, E_π will denote the set of edges in π.

(5.21) DEFINITION. A pair of paths π and π' is called *feasible* if it has the following properties:

(a) π and π' have the same initial points and the same endpoints,
(b) π and π' are edge disjoint,
(c) $|\pi| = k$, $|\pi'| \le k + 2$,
(d) $t(e) \in I_0$ for all $e \in \pi$, and $t(e) \le \sup(I_0)$ for all $e \in \pi'$,
(e) π is a segment of $\pi(n)$.

Note that only properties (d) and (e) here involve the random travel times. (a)–(c) are geometric requirements.

Most of the remainder of this section is devoted to proving the following proposition.

(5.22) PROPOSITION. *Assume that the conditions of Theorem* 2.9(b) *and* (4.6) *hold. Then there exists a* $D_2 = D_2(F, \tilde{F})$ *such that for all sufficiently large* n *there is a sequence of pairs* (π_i, π_i'), *such that each pair satisfies conditions* (5.21)(*a*)–(*c*), *and*

$$(E_{\pi_i} \cup E_{\pi_i'}) \cap (E_{\pi_j} \cup E_{\pi_j'}) = \phi \quad \textit{for all } i \neq j \tag{5.23}$$

and

$$\sum_i P\{(\pi_i, \pi_i') \textit{ is feasible}\} \ge D_2 n. \tag{5.24}$$

Before proving the proposition we show that it implies Theorem 2.9(b). We shall call a pair of paths (π, π') *advantageous* if it is feasible, and if in addition

$$\tilde{t}(e) > t(e) + \alpha \quad \text{and} \quad \xi(e) = 1 \quad \text{for all } e \in \pi,$$
$$\xi(e) = 0 \quad \text{for all } e \in \pi'.$$

Roughly speaking, for each advantageous pair (π_i, π_i') we modify $\pi(n)$ by replacing π_i by π_i'. It turns out that this leads to an amount of saved $\hat{t}$-travel time of at least γ. Moreover, the expected number of advantageous pairs will be at least $D_3 n$. This leads to a total expected saving of $\gamma D_3 n$ and a decrease in the time constant of at least γD_3. We now show how to do this more precisely.

PROOF OF THEOREM 2.9(b) FROM PROPOSITION 5.22. Since feasibility is defined in terms of the $t(e)$ only, we have from (4.9) and the independence of the ξ's from the $(t(e), \tilde{t}(e))$, that

$$P\{(\pi, \pi') \text{ is advantageous}\} \ge \beta^k 2^{-2k-2} P\{(\pi, \pi') \text{ is feasible}\}. \tag{5.25}$$

Now suppose that Proposition 5.22 holds and that n is sufficiently large. For brevity call i feasible (advantageous) if (π_i, π_i') is feasible (advantageous). Let $\pi'(n)$ be the path obtained from $\pi(n)$ by replacing π_i by π_i' for every advantageous i. By the disjointness condition (5.23) and the fact that π_i and π_i' have the same endpoints, $\pi'(n)$ is well defined. $\pi'(n)$ is not necessarily self avoiding but that will not influence our argument. From the definition of advantageous and the properties of I_0, it follows immediately that

$$\begin{aligned}
\hat{t}(\pi(n)) - \hat{t}(\pi'(n)) &= \sum_i \left[\sum_{e \in \pi_i} \hat{t}(e) - \sum_{e \in \pi_i'} \hat{t}(e) \right] I(i \text{ is advantageous}) \\
&= \sum_i \left[\sum_{e \in \pi_i} \tilde{t}(e) - \sum_{e \in \pi_i'} t(e) \right] I(i \text{ is advantageous}) \\
&\geq \sum_i \left[\sum_{e \in \pi_i} (t(e) + \alpha) - \sum_{e \in \pi_i'} t(e) \right] I(i \text{ is advantageous}) \\
&\geq \sum_i \gamma I(i \text{ is advantageous}).
\end{aligned}$$

Taking expectations gives

$$\begin{aligned}
E\{\hat{a}_n\} &\leq E\{\hat{t}(\pi'(n))\} \\
&\leq E\{\hat{t}(\pi(n))\} - \gamma \sum_i P\{i \text{ is advantageous}\} \\
&\leq E\{a_n\} - \gamma \beta^k 2^{-2k-2} \sum_i P\{i \text{ is feasible}\} \qquad [\text{by } (5.19) \text{ and } (5.25)] \\
&\leq E\{a_n\} - \beta^k 2^{-2k-2} \gamma D_2 n \qquad [\text{by } (5.24)].
\end{aligned}$$

Thus Proposition 5.22 indeed implies a strict inequality in (5.20), and hence (2.11). □

It remains to prove Proposition 5.22.

Proof of Proposition 5.22. We will distinguish between the cases of bounded and unbounded support of F. In both cases we show that the configurations in certain of the blocks which are crossed by $\pi(n)$ can be modified so as to contain a feasible pair of paths. We first treat the bounded case in detail, and then the unbounded case (which is easier) in less detail.

Proof for the bounded case. Let $M = \sup \operatorname{supp}(F)$. Recall the N-boxes defined in (5.15) and the definition of $\overline{B}$ in (5.16). Let, for each $N > 0$, $B(N)$ be a generic N-box. We need a sequence of numbers $\nu(N)$, $N = 1, 2, \dots$ satisfying the following conditions:

(5.26) $$\nu(N) \geq \sup(I_0),$$

(5.27) $$F([\nu(N), \infty)) > 0,$$

(5.28) $$\lim_{N \to \infty} P\{\text{there exists an } e \in B(N) \text{ with } t(e) > \nu(N)\} = 0.$$

Such $\{\nu(N)\}$ exist by property (a) in Lemma 4.8. [Note the strict inequality in (5.28); if F has an atom at the right end of its support, then we can take $\nu(N)$ equal to this right endpoint of supp(F).] Let δ and D_0 be such that (5.6) holds. Now color an N-box B black if (i) and (ii) below occur.

(i) For each path π which lies entirely in $\overline{B}$ and whose endpoints v_π and w_π satisfy $\|v_\pi - w_\pi\| \geq N$,

$$t(\pi) \geq (r+\delta)\|v_\pi - w_\pi\|. \tag{5.29}$$

(ii) For all $e \in B$, $t(e) \leq \nu(N)$.

We claim that (uniformly in the location of the N-box),

$$\lim_{N\to\infty} P\{B(N) \text{ is black}\} = 1. \tag{5.30}$$

Clearly the probability that (ii) holds tends to 1 [by (5.28)]. The same holds for property (i), since by (5.6),

$$P\{(5.29) \text{ fails for some } \pi \in \overline{B}\} \leq \sum_{\substack{v,w\in\overline{B}\\ \|v-w\|\geq N}} e^{-D_0\|v-w\|}$$

$$\leq |\overline{B}|^2 e^{-D_0 N} \to 0 \quad \text{as } N \to \infty$$

(note that $|\overline{B}|$ is bounded by a polynomial in N). Thus (5.30) holds.

We claim further that (5.29) implies

$$t(v,w) \geq (r+\delta)\|v-w\| \quad \text{for all } v,w \in B \text{ with } \|v-w\| \geq N. \tag{5.31}$$

This is true because any path π from v to w is either contained in $\overline{B}$ [in which case (5.31) is immediate from (5.29)] or has a segment contained in $\overline{B}$ from v to some $w' \in \partial\overline{B}$. By (5.16) $\|v - w'\| \geq \text{diam}(B) \geq N$. Applying (5.29) to this segment gives

$$t(v,w) \geq t(v,w') \geq (r+\delta)\text{diam}(B) \geq (r+\delta)\|v-w\|,$$

as desired.

Next we color the N-cubes. $S(l;N)$ is colored black if each of its surrounding N-boxes is black. It follows from (5.30) that

$$\lim_{N\to\infty} P\{S(0;N) \text{ is black}\} = 1. \tag{5.32}$$

It is also clear that the other conditions of Lemma 5.2 hold. So we apply that lemma and take N, E and D so that for all $v, w \in \mathbb{Z}^d$, (5.3) holds. N will also have to satisfy (5.38) and (5.39) below, but it is easily seen that all these requirements are met for large N. From (5.3) we immediately get that, for some $D_4 = D_4(F, \tilde{F}, N)$ and all large n,

$$E\{\text{the number of distinct black } N\text{-cubes visited by } \pi(n)\} > D_4 n. \tag{5.33}$$

As we noted in the beginning of this section, any path going from some vertex of an N-cube $S(l;N)$ to some vertex outside $T(l;N)$ must cross at least one of the $2d$ N-boxes surrounding $S(l;N)$. Hence (5.33) implies that there exists a $D_5 = D_5(F, \tilde{F}, N) > 0$ such that, for all large n,

$$E\{\text{number of black } N\text{-boxes crossed by } \pi(n)\} \geq D_5 n. \tag{5.34}$$

Since each N-box intersects at most C_3 other N-boxes for some $C_3 < \infty$, we can even find a collection $\mathscr{B}$ of *disjoint* N-boxes such that, for all large n,

$$\sum_{B \in \mathscr{B}} P\{B \text{ is black and } \pi(n) \text{ crosses } B\} \geq \frac{D_5}{(C_3+1)} n. \tag{5.35}$$

Our next step is a modification argument, showing that the probability of an N-box B containing a feasible pair of paths is at least D_6 times the probability that B is black and crossed by $\pi(n)$. First take two integers l_1 and l_2 such that

$$\frac{\delta l_1}{3d(r+\delta/2)} > k \tag{5.36}$$

and

$$l_2 \geq k+2. \tag{5.37}$$

Now fix N so large that (5.33) holds for all large n, and such that

$$N > 2(l_1 + l_2 + k), \tag{5.38}$$

$$\frac{[(\delta/2)N - 4dl_1M - 4dl_2 \sup(I_0)]2l_1}{[4dl_1M + 4dl_2\sup(I_0) + 3dN(r+\delta/2)]} > k \tag{5.39}$$

[this is possible by (5.36)]. We have now fixed N and will use ν as an abbreviation for $\nu(N)$.

We are now ready for the modification argument. Let B be an N-box, which contains neither **0** nor **n**. Let U (respectively, V and W) be the set of all vertices in B at distance $\leq l_1$ from ∂B [respectively, with distance in $(l_1, l_1 + l_2)$ and $\geq l_1 + l_2$]; see Figure 1. Next we choose a new set $\{t^*(e)\}$ of time coordinates. We take $t^*(e) = t(e)$ for e not contained in B, and for the remaining e's we take the $t^*(e)$'s as independent copies of the $t(e)$ (and independent of each other). The remaining considerations are carried out only

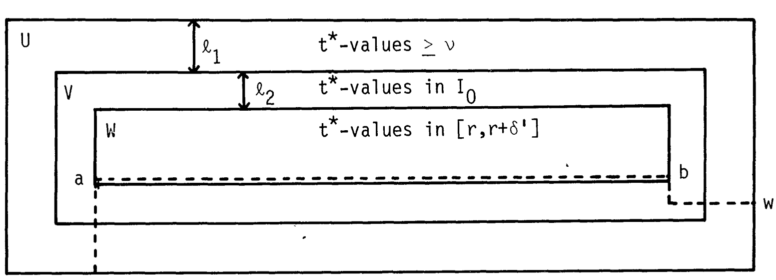

FIG. 1. *Illustration of the box B with the sets U, V, W and the path π_2' from v to w (dashed).*

on the event Γ_1, defined as

$$\Gamma_1 := \{B \text{ is black and } \pi(n) \text{ crosses } B\}$$
$$\cap \{t^*(e) \geq \nu \text{ for all } e \text{ with at least one endpoint in } U\}$$
$$\cap \{t^*(e) \in I_0 \text{ for all } e \text{ with both endpoints in } V\}$$
$$\cap \{t^*(e) \in [r, r+\delta'] \text{ for all } e \text{ with at least one endpoint in } W\},$$

where

$$\delta' = \min\left\{\frac{\delta}{2}, \inf(I_0) - r\right\}. \tag{5.40}$$

Note that property (a) in Lemma 4.8 guarantees $F([r, r+\delta']) > 0$, while $F([\nu, \infty)) > 0$ by (5.27).

We are going to show that on Γ_1 the t^*-travel time from $\mathbf{0}$ to $\mathbf{n}$ is strictly smaller than the t-travel time a_n. Let $\pi(n)$ be the optimal t-path for a_n, as usual. We think of $\mathbf{0}$ as its initial point and $\mathbf{n}$ as its endpoint, and define v and w as the first and last point of $\pi(n)$ in B. We denote the segments of $\pi(n)$ from $\mathbf{0}$ to v and from v to w by π_1 and π_2, respectively. We next choose a path π_2' in B from v to w for which (on Γ_1)

$$t^*(\pi_2') \leq 4dl_1M + 4dl_2 \sup(I_0) + (r+\delta')\|v - w\|. \tag{5.41}$$

One way to construct such a path is as follows. Choose a path λ_1 from v to some point a in W of at most dl_1 steps in U and at most dl_2 steps in V (see Figure 1). Similarly we can connect w to a point b in W by a path λ_3 of at most dl_1 steps in U and at most dl_2 steps in V. Clearly $\|a - b\| \leq \|v - w\| + 2d(l_1 + l_2)$, so that a can be connected to b by a path λ_2 in W of at most $\|v - w\| + 2d(l_1 + l_2)$ steps. Now take for π_2' the concatenation of λ_1, λ_2 and the reverse of λ_3. Next we define π^* as the path obtained from $\pi(n)$ by replacing π_2 by π_2'. Our next task is to show that on Γ_1,

$$t^*(\pi^*) < t(\pi(n)). \tag{5.42}$$

Since $t^*(e) = t(e)$ for e not contained in B,

$$t(\pi(n)) - t^*(\pi^*) = t(\pi_2) - t^*(\pi_2').$$

Moreover, if $\|v - w\| \geq N$, then because B is black, $t(\pi_2) \geq (r+\delta)\|v - w\|$ [see (5.31)]. If $\|v - w\| < N$, then π_2 still crosses B and hence contains a segment, $\bar{\pi}$ say, between opposite faces of B and we can apply (5.31) to the endpoints of $\bar{\pi}$. Then $t(\pi_2) \geq t(\bar{\pi}) \geq N(r+\delta)$. Thus in any case

$$t(\pi_2) \geq (r+\delta)\max(\|v - w\|, N). \tag{5.43}$$

Together with (5.41) this yields

$$\begin{aligned} t(\pi(n)) - t^*(\pi^*) &\geq (r+\delta)\max(\|v - w\|, N) - 4dl_1M - 4dl_2\sup(I_0) - (r+\delta')\|v - w\| \\ &\geq \frac{\delta}{2}N - 4dl_1M - 4dl_2\sup(I_0) > 0 \qquad [\text{by } (5.39)], \end{aligned}$$

which proves (5.42). We now define $\pi^*(n)$ as the optimal path from **0** to **n** for the t^* values. Then

$$
\begin{aligned}
t(\pi^*(n)) &\geq t(\pi(n)) \qquad [\text{by definition of } \pi(n)] \\
&\geq t^*(\pi^*) + \frac{\delta}{2}N - 4dl_1M - 4dl_2\sup(I_0) \\
&\geq t^*(\pi^*(n)) + \frac{\delta}{2}N - 4dl_1M - 4dl_2\sup(I_0) \\
&> t^*(\pi^*(n)) \qquad [\text{by definition of } \pi^*(n)].
\end{aligned} \tag{5.44}
$$

Since $t^*(e) = t(e)$ for e not in B and $t^*(e) \geq t(e)$ for $e \subset B$ and with one endpoint in U (because B is black and we are on Γ_1), the fact that $t^*(\pi^*(n))$ is strictly less than $t(\pi^*(n))$ must be due to edges on $\pi^*(n)$ which have both endpoints in $V \cup W$. In particular, some such edges must exist. We claim that on the event Γ_1,

(5.45) $\pi^*(n)$ has a segment of length k which lies entirely in V.

To prove (5.45) consider first the case that $\pi^*(n)$ contains a vertex in W. Then (5.45) is obvious, for then $\pi^*(n)$ must have crossed V, which has width $l_2 - 2 \geq k$ [cf. (5.37)]. So we may assume that $\pi^*(n)$ has no vertex in W. Now let c and d be two distinct vertices of $\pi^*(n)$ in V, and let $\overline{\pi}$ be the segment of $\pi^*(n)$ from c to d. It is not difficult to see [from the fact that $V \cup W$ is a "rectangular box," with all $t^*(e) \leq \nu$ in this box, while $t^*(e) \geq \nu$ for $e \subset B \setminus V \cup W$ on the event Γ_1] that any path from c to d which lies entirely in B, and which has minimal t^*-value among such paths, must lie entirely in $V \cup W$. Hence if $\overline{\pi}$ contains a vertex outside V it must contain a vertex outside B. From this it follows that $\pi^*(n)$ is a concatenation of "excursions," each of which consists of three segments: one which starts outside B and which lies, except for its last vertex, outside V; one which lies entirely in V; and one which starts in V, ends outside B, and lies, except for its first vertex, entirely outside V. Let σ be the number of excursions of $\pi^*(n)$. Each excursion crosses the set U twice and hence has t^*-travel time $\geq 2\nu l_1$. Hence the t^*-travel time of the segment of $\pi^*(n)$ between the first vertex where it enters B and the last vertex where it exits B is at least $\sigma 2\nu l_1$. However, analogously to (5.41), we can connect any two vertices in B by a path with t^*-travel time at most $4dl_1M + 4dl_2\sup(I_0) + d3N(r + \delta/2)$. Hence, since $\pi^*(n)$ is optimal,

$$\sigma 2\nu l_1 \leq 4dl_1M + 4dl_2\sup(I_0) + d3N(r + \delta/2),$$

that is,

$$\sigma \leq \frac{4dl_1M + 4dl_2\sup(I_0) + 3dN(r + \delta/2)}{2\nu l_1}.$$

From (5.44) we obtain a lower bound for the total savings due to the excursions, that is, the difference between the t- and the t^*-travel time of $\pi^*(n)$.

Thus there is at least one excursion which contributes a saving of

$$\sigma^{-1}\left[\frac{\delta}{2}N - 4dl_1M - 4dl_2\sup(I_0)\right]$$
$$\geq \frac{[(\delta/2)N - 4dl_1M - 4dl_2\sup(I_0)]2\nu l_1}{[4dl_1M + 4dl_2\sup(I_0) + 3dN(r+\delta/2)]}.$$

Take such an excursion. Clearly the only contribution to the saving in travel time by this excursion comes from its segment which lies entirely in V. Since each edge e in V contributes at most $t(e) - t^*(e) \leq t(e) \leq \nu$ to the saving, the segment in V of our excursion has to have length at least

$$\frac{[(\delta/2)N - 4dl_1M - 4dl_2\sup(I_0)]2l_1}{[4dl_1M + 4dl_2\sup(I_0) + 3dN(r+\delta/2)]} > k$$

[by (5.39)]. This proves (5.45).

Now let π_+ be a segment of $\pi^*(n)$ of length k which lies entirely in V. Call its endpoints c and d again. Note that (4.10) and the optimality of $\pi^*(n)$ imply that also

$$\|c - d\| \geq k. \tag{5.46}$$

Indeed, if (5.46) would fail, then there would exist a path from c to d in the rectangular box $V \cup W$ of length at most $k-2$. The sum of the $t^*(e)$ along this alternative path would be at most $(k-2)\sup(I_0) < t^*(\pi_+)$ [by (4.10)]. Thus (5.46) holds and in fact $\|c - d\| = k$. From this it is easy to see that there exists also another path π'_+ from c to d of length at most $(k+2)$, edge disjoint from π_+ and lying entirely in $V \cup W$. [See Figure 2 for some typical choices of π'_+ in dimension 2; note that in Figure 2(b) and (c) at least one of the dashed paths will lie in $V \cup W$, because π_2 cannot be adjacent to *both* faces of $V \cup W$ which are parallel to the x_1 axis, since these faces are distance at least $N - 2l_1 - 2l_2 > 2k$ apart.] We have therefore shown that if Γ_1 occurs, then the pair (π_+, π'_+) is a feasible pair in the t^*-configuration. By our choice of $\nu = \nu(N)$, I_0 and δ',

$$\begin{aligned} D_6 &= D_6(F, \tilde{F}, N, l_1, l_2) \\ &:= P\{t^*(e) \geq \nu \text{ for all } e \text{ with at least one endpoint in } U\} \\ &\quad \times P\{t^*(e) \in I_0 \text{ for all } e \text{ with both endpoints in } V\} \\ &\quad \times P\{t^*(e) \in [r, r+\delta'] \text{ for all } e \text{ with at least one endpoint in } W\} \\ &> 0 \end{aligned} \tag{5.47}$$

and is independent of n. We obtain

$$\begin{aligned} &P\{B \text{ contains a feasible pair in the } t\text{-configuration}\} \\ &\quad = P\{B \text{ contains a feasible pair in the } t^*\text{-configuration}\} \\ &\quad \geq P\{\Gamma_1\} = D_6 P\{B \text{ is black and } \pi(n) \text{ crosses } B\}. \end{aligned} \tag{5.48}$$

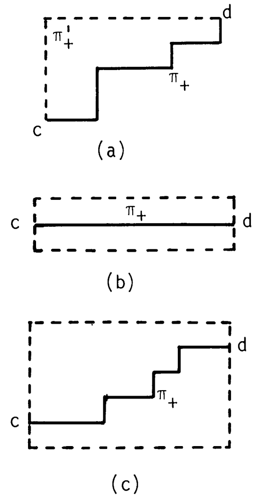

FIG. 2. *Some examples of π_+ (solidly drawn) and possible choices for the bypass π'_+ (dashed). In each case at least one of the dashed paths lies in $V \cup W$. The length of the dashed paths is $k = |\pi_+|$ in case (a) and $(k + 2)$ in cases (b) and (c).*

Proposition 5.22 now follows from (5.35) and the obvious fact that each N-box contains at most $D_7 = D_7(N)$ pairs (π, π') which satisfy conditions (a)–(c) in Definition (5.21). □

PROOF OF PROPOSITION 5.22 IN THE UNBOUNDED CASE. This proof is simpler than that of the bounded case. Follow the proof of the bounded case. This time the sequence $\nu(N)$, $N = 1, 2, \dots$ is taken in such a way that it satisfies (5.26) and

$$\lim_{N \to \infty} P\bigg(\sum_{e \in B(N)} t(e) \geq \nu(N) \bigg) = 0. \tag{5.49}$$

Equation (5.27) will hold automatically since supp(F) is unbounded. We color an N-box B black if (i) and (ii) below hold:

(i) For each path π which lies entirely in $\overline{B}$ and whose endpoints v_π and w_π satisfy $\|v_\pi - w_\pi\| \geq N$,

$$t(\pi) \geq (r + \delta)\|v_\pi - w_\pi\|,$$

(ii) $\sum_{i \in B} t(e) < \nu(N)$

[only (ii) is different from before]. As in (5.34) we conclude that for all large N there exists a $D_8 = D_8(F, \tilde{F}, N) > 0$ such that for all n with $\mathbf{n} \notin T(\mathbf{0}; N)$,

$$E\{\text{number of distinct black } N\text{-boxes crossed by } \pi(n)\} \geq D_8 n. \tag{5.50}$$

The modification argument for the unbounded case is quite different from that for the bounded case. Let B be a black N-box crossed by $\pi(n)$. Let v and w be

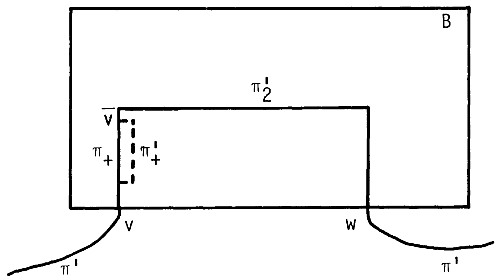

FIG. 3. *Illustration of the paths $\pi_+, \pi'_+, \pi'_2, \pi'$. π'_+ is dashed.*

the first and last vertex, respectively, on $\pi(n) \cap \partial B$ which have a neighbor in the interior of B. Let v' be the neighbor of v in the interior of B; assume $v' = v + \zeta e_l$, $\zeta = +1$ or -1. Let π_+ be the straight line segment of length k which runs from $v + 2\zeta e_l$ to $\bar{v} := v + (k+2)\zeta e_l$. Let π'_+ be a "bypass of π_+," that is, a path from $v + 2\zeta e_l$ to $\bar{v}$ of $(k+2)$ steps, which also lies in B and is edge-disjoint form π_+ [compare Figure 2(b)]. Next, let π'_2 be a path in B which starts at v, then follows the line segment from v to $v + (k+3)\zeta e_l$, and then from $v + (k+3)\zeta e_l$ to w inside B in at most $\|v - w\| + (2k+9)$ steps, without intersecting $\pi_+ \cup \pi'_+$ (see Figure 3). Note that π'_2 contains π_+. Finally, π' will be the path which coincides with $\pi(n)$ from $\mathbf{0}$ to v, then follows π'_2 to w, and then again coincides with $\pi(n)$ from n to $\mathbf{n}$.

Now choose $t^*(e)$. Again $t^*(e) = t(e)$ for e not contained in B. For the edges in B, the $t^*(e)$ are independent copies of the $t(e)$. Finally, take $\Gamma_2 = \bigcap_{i=1}^4 \Gamma_{2,i}$, where the $\Gamma_{2,i}$ are the following events:

$$\Gamma_{2,1} = \{B \text{ is black and } \pi(n) \text{ crosses } B\},$$

$$\Gamma_{2,2} = \{t^*(e) \in I_0 \text{ for all } e \text{ in } \pi_+ \cup \pi'_+\}$$

$$\Gamma_{2,3} = \{t^*(e) \in [r, r + \delta/2) \text{ for all } e \in \pi'_2 \setminus \pi_+\},$$

$$\Gamma_{2,4} = \{t^*(e) > \nu \text{ for all } e \text{ in } B \setminus (\pi'_2 \cup \pi'_+)\}.$$

Note that for some D_9 depending on F, $\tilde{F}$ and N only

$$P\left\{ \bigcap_2^4 \Gamma_{2,i} \,\middle|\, \text{all } t \text{ values} \right\} \geq D_9 > 0. \tag{5.51}$$

Let $\pi_2(n)$ be the part of $\pi(n)$ from v to w. Then, analogously to (5.41)–(5.44) we obtain that, on Γ_2,

$$t(\pi(n)) - t^*(\pi') = t(\pi_2(n)) - t^*(\pi'_2) > 0,$$

provided N is large enough [since $t^*(\pi'_2) \leq (r + \delta/2)(\|v - w\| + 2k + 9) + k \sup(I_0)$]. It follows that $a_n^* < a_n$ and that $\pi^*(n)$ must contain some edge in

B. However, on $\Gamma_{2,4} \cap \Gamma_{2,1}$ we have for *any* path π containing an edge f in $B \setminus (\pi'_2 \cup \pi'_+)$ that $t^*(\pi) > t(\pi)$. Indeed, $t^*(f)$ alone exceeds $\sum_{e \in B} t(e) \geq \sum_{e \in \pi \cap B} t(e)$. Thus $\pi^*(n)$ cannot contain any edge in $B \setminus (\pi'_2 \cup \pi'_+)$. Thus $\pi^*(n)$ must enter B at v and leave at w, or vice versa, and can only use edges of $\pi'_2 \cup \pi'_+$. By property (4.10) of I_0 we have

$$\sum_{e \in \pi_+} t^*(e) < \sum_{e \in \pi'_+} t^*(e)$$

on $\Gamma_{2,2}$. It follows that $\pi^*(n)$ must contain all of π_+ when Γ_2 occurs. Therefore, on the event Γ_2, (π_+, π'_+) is a feasible pair in the t^*-configuration. The rest is as before. □

Acknowledgment. We are grateful to Peter Donnelly for pointing out to us the usefulness of the partial ordering "$\tilde{F}$ is more variable than F." Our original conditions for Theorem 2.9 were seemingly more restrictive and complicated. Also, Example 2.18 was suggested to us by Peter Donnelly.

REFERENCES

AHLBERG, R. and JANSON, S. (1984). Upper bounds for the connectivity constant. Preprint. Uppsala Univ.

AIZENMAN, M. and BARSKY, D. J. (1987). Sharpness of the phase transition in percolation models. *Comm. Math. Phys.* **108** 489–526.

AIZENMAN, M. and GRIMMETT, G. (1991). Strict monotonicity of critical points in percolation and ferromagnetic models. *J. Statist. Phys.* **63** 817–835.

CAMPANINO, M. and RUSSO, L. (1985). An upper bound for the critical probability for the three-dimensional cubic lattice. *Ann. Probab.* **13** 478–491.

COX, J. T. (1980). The time constant of first-passage percolation on the square lattice. *Adv. in Appl. Probab.* **12** 864–879.

COX, J. T. and KESTEN, H. (1981). On the continuity of the time constant of first-passage percolation. *J. Appl. Probab.* **18** 809–819.

DURRETT, R. (1988). *Lectures Notes on Particle Systems and Percolation.* Wadsworth and Brooks/Cole, Pacific Grove, CA.

GRIMMETT, G. (1989). *Percolation.* Springer, New York.

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

HAMMERSLEY, J. M. (1957). Percolation processes. Lower bounds for the critical probability. *Ann. Math. Statist.* **28** 790–795.

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. M. Le Cam, eds.) 61–110. Springer, New York.

JANSON, S. (1981). An upper bound for the velocity of first-passage percolation. *J. Appl. Probab.* **18** 256–262.

KESTEN, H. (1986). Aspects of first passage percolation. *Ecole d'Eté de Probabilités de Saint-Flour XIV. Lecture Notes in Math.* **1180** 125–264. Springer, New York.

KESTEN, H. (1987). Percolation theory and first-passage percolation. *Ann. Probab.* **15** 1231–1271.

MENSHIKOV, M. V. (1986). Coincidence of critical points in percolation problems. *Soviet Math. Dokl.* **33** 856–859.

ROSS, S. M. (1983). *Stochastic Processes.* Wiley, New York.

SMYTHE, R. T. (1980). Percolation models in two and three dimension. In *Biological Growth and Spread. Lecture Notes in Biomath.* **38** 504–511. Springer, New York.

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

STOYAN, D. and DALEY, D. J. (1983). Comparison methods for queues and other stochastic models. Wiley, New York.

STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

VAN DEN BERG, J. and KESTEN, H. (1985). Inequalities with applications to percolation and reliability. *J. Appl. Probab.* **22** 556–569.

WHITT, W. (1980). The effect of variability in the GI/G/s queue. *J. Appl. Probab.* **17** 1062–1071.

CENTRUM VOOR WISKUNDE EN INFORMATICA
KRUISLAAN 413
1098 SJ AMSTERDAM
THE NETHERLANDS

DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853