The Annals of Probability
2015, Vol. 43, No. 4, 1992–2025
DOI: 10.1214/14-AOP926

A HSU–ROBBINS–ERDŐS STRONG LAW IN FIRST-PASSAGE PERCOLATION

BY DANIEL AHLBERG

Instituto Nacional de Matemática Pura e Aplicada

Large deviations in the context of first-passage percolation was first studied in the early 1980s by Grimmett and Kesten, and has since been revisited in a variety of studies. However, none of these studies provides a precise relation between the existence of moments of polynomial order and the decay of probability tails. Such a relation is derived in this paper, and is used to strengthen the conclusion of the shape theorem. In contrast to its one-dimensional counterpart—the Hsu–Robbins–Erdős strong law—this strengthening is obtained without imposing a higher-order moment condition.

1. Introduction. The study of large deviations in first-passage percolation was pioneered by Grimmett and Kesten [13]. In their work, they investigate the rate of convergence of travel times toward the so-called time constant, by providing some necessary and sufficient conditions for exponential decay of the probability of linear order deviations. Although the rate of convergence toward the time constant has received considerable attention in the literature, there is no systematic study of the regime for polynomial decay of the probability tails. This is remarkable since it is precisely in this regime that strong laws such as the celebrated shape theorem are obtained. In this paper, we derive a precise characterization of the regime of polynomial decay in terms of a moment condition. As a consequence, we improve upon the statement of the shape theorem without strengthening its hypothesis.

Consider the $\mathbf{Z}^d$ nearest-neighbor lattice for $d \geq 2$, with nonnegative i.i.d. random weights assigned to its edges. The random weights induce a random pseudo-metric on $\mathbf{Z}^d$, known as first-passage percolation, in which distance between points are given by the minimal weight sum among possible paths. An important subadditive nature of these distances, sometimes referred to as *travel times*, was in the 1960s identified and studied by Hammersley and Welsh [14] and Kingman [18]. A particular fact dating back to these early studies is that, under weak conditions, travel times grow linearly with respect to Euclidean distance between points, at a rate depending on the direction, and with sublinear corrections. This asymptotic rate is referred to as the *time constant*.

Received June 2013; revised January 2014.
MSC2010 subject classifications. Primary 60K35; secondary 60F10, 60F15.
Key words and phrases. First-passage percolation, shape theorem, large deviations.

The work of Grimmett and Kesten on the rate of convergence toward the time constant was later continued by Kesten himself in [16], and more recently refined in [7] and [9]. These studies, just like the present study, investigates deviations of linear order. Other studies have pursued stronger concentration inequalities, describing deviations of sublinear order, notably Kesten [17], Talagrand [20] and Benjamini, Kalai and Schramm [5]. Other authors have considered deviations of linear order for the related concept of so-called chemical distance in Bernoulli percolation [4, 12]. A common feature among these studies is that they aim to derive exponential decay of the probability tails for deviations in coordinate directions, and generally require exponential decay of the tails of the weights in order to get there. There is no previous study characterizing the regime of polynomial rate of decay on the probability tails of linear order deviations. This study attends to this matter and provides necessary and sufficient conditions for polynomial rate of decay in terms of a moment condition, valid for all directions simultaneously. The results obtained strengthens earlier strong laws in first-passage percolation, in particular the shape theorem, due to Richardson [19] and Cox and Durrett [8], which states the precise conditions under which the set of points in $\mathbf{Z}^d$ within distance t from the origin (in the random metric), rescaled by t, converges to a deterministic compact and convex set.

1.1. *The shape theorem.* We will throughout this paper assume that $d \geq 2$, as the one-dimensional case coincides with the study of i.i.d. sequences. Let $\mathcal{E}$ denote the set of edges of the $\mathbf{Z}^d$ lattice, and let τ_e denote the random weight associated with the edge e in $\mathcal{E}$. The collection $\{\tau_e\}_{e\in\mathcal{E}}$ of weights, commonly referred to as *passage times*, will throughout be assumed to be nonnegative and i.i.d. The distance, or *travel time*, $T(y,z)$ between two points y and z of $\mathbf{Z}^d$ is defined as the minimal path weight, as induced by the random environment $\{\tau_e\}_{e\in\mathcal{E}}$, among paths connecting y and z. That is, given a path Γ, let $T(\Gamma) := \sum_{e\in\Gamma} \tau_e$ and define

$$T(y,z) := \inf\{T(\Gamma) : \Gamma \text{ is a path connecting } y \text{ and } z\}.$$

As mentioned above, travel times grow linearly in comparison with Euclidean or ℓ^1-distance on $\mathbf{Z}^d$, denoted below by $|\cdot|$ and $\|\cdot\|$, respectively. The precise meaning of this informal statement refers to the existence of the limit

$$\mu(z) := \lim_{n\to\infty} \frac{T(0,nz)}{n} \qquad \text{in probability,} \tag{1}$$

which we refer to as the *time constant*, and which may depend on the direction. Indeed, the growth is only linear in the case that $\mu(z) > 0$, which is known to be the case if and only if $\mathbf{P}(\tau_e = 0) < p_c(d)$, where $p_c(d)$ denotes the critical probability for bond percolation on $\mathbf{Z}^d$ (see [16]).

Existence of the limit in (1) was first obtained with a moment condition in [14], but indeed exists finitely without any restriction on the passage time distribution (see [8, 16], and the discussion in Appendix A below). Moreover, the convergence in (1) holds almost surely and in L^1 if and only if $\mathbf{E}[Y]<\infty$, where Y denotes the minimum of $2d$ random variables distributed as τ_e, as a consequence of the subadditive ergodic theorem [18]. A more comprehensive result, the shape theorem, provides simultaneous convergence in all directions. A weak form thereof can be concisely stated as

$$\limsup_{z\in\mathbf{Z}^d:\,\|z\|\to\infty} \mathbf{P}\big(\big|T(0,z)-\mu(z)\big|>\varepsilon\|z\|\big)=0 \qquad \text{for every } \varepsilon>0. \tag{2}$$

Also the convergence in (2) holds without restrictions to the passage time distribution (see Section 2.2). However, under the assumption of a moment condition, it is possible to obtain an estimate on the rate of decay in (2). This is precisely the aim of this study, and the content of Theorems 3 and 4 below. Based thereon, we will derive the following Hsu–Robbins–Erdős type of strong law, which characterizes the summability of tail probabilities of the above type.

THEOREM 1. *For every $\alpha>0$, $\varepsilon>0$ and $d\geq 2$,*

$$\mathbf{E}\big[Y^\alpha\big]<\infty \quad\Longleftrightarrow\quad \sum_{z\in\mathbf{Z}^d}\|z\|^{\alpha-d}\mathbf{P}\big(\big|T(0,z)-\mu(z)\big|>\varepsilon\|z\|\big)<\infty.$$

Apart from characterizing the summability of probabilities of large deviations away from the time constant, Theorem 1 has several implications for the shape theorem, which we will discuss next. Cox and Durrett's version of the shape theorem is a strengthening of (2), and can be stated as follows: if $\mathbf{E}[Y^d]<\infty$, then

$$\limsup_{z\in\mathbf{Z}^d:\,\|z\|\to\infty} \frac{|T(0,z)-\mu(z)|}{\|z\|}=0 \qquad \text{almost surely.} \tag{3}$$

(A more popular way to phrase the shape theorem is offered below.) It is well known that $\mathbf{E}[Y^d]<\infty$ is also necessary for the convergence in (3). Let

$$\mathscr{Z}_\varepsilon:=\big\{z\in\mathbf{Z}^d:\big|T(0,z)-\mu(z)\big|>\varepsilon\|z\|\big\},$$

and note that (3) is equivalent to saying that the cardinality of the set $\mathscr{Z}_\varepsilon$, denoted by $|\mathscr{Z}_\varepsilon|$, is finite for every $\varepsilon>0$ with probability one. This statement is by Theorem 1, under the same assumption of $\mathbf{E}[Y^d]<\infty$, strengthened to say that $\mathbf{E}|\mathscr{Z}_\varepsilon|<\infty$ for every $\varepsilon>0$.

The shape theorem is commonly presented as a comparison between the random set

$$\mathscr{B}_t:=\big\{z\in\mathbf{Z}^d:T(0,z)\leq t\big\}$$

and the discrete "ball" $\mathscr{B}_t^\mu:=\{z\in\mathbf{Z}^d:\mu(z)\leq t\}$. The growth of the first-passage process may be divided into two regimes characterized by the time constant: $\mu\equiv 0$, and $\mu(z) > 0$ for all $z \neq 0$. In the more interesting regime $\mu \not\equiv 0$, Cox and Durrett's shape theorem can be phrased: if $\mathbf{E}[Y^d] < \infty$, then for every $\varepsilon > 0$ the two inclusions

$$\mathscr{B}^{\mu}_{(1-\varepsilon)t} \subset \mathscr{B}_t \subset \mathscr{B}^{\mu}_{(1+\varepsilon)t} \tag{4}$$

hold for all t large enough, with probability one. A simple inversion argument shows that this formulation is equivalent to the one in (3).

The time constant $\mu(\cdot)$ extends continuously to $\mathbf{R}^d$ and inherits the properties of a semi-norm. In other words, one may interpret (4) as $\frac{1}{t}\mathscr{B}_t$ being asymptotic to the unit ball $\{x \in \mathbf{R}^d : \mu(x) \leq 1\}$ expressed in this norm, and failure of either inclusion in (4) indicates a linear order deviation of $\mathscr{B}_t$ from this asymptotic shape. Inspired by Theorem 1, one may wonder whether the size, that is Lebesgue measure, of the set of times for which (4) fails behaves similarly as the size of $\mathscr{Z}_\varepsilon$. Assume that $\mu \not\equiv 0$ and let

$$\mathscr{T}_\varepsilon := \{t \geq 0 : \text{either inclusion in (4) fails}\}.$$

The shape theorem says that $\mathbf{E}[Y^d] < \infty$ is sufficient for the supremum of $\mathscr{T}_\varepsilon$, and hence the Lebesgue measure $|\mathscr{T}_\varepsilon|$ of the set $\mathscr{T}_\varepsilon$, to be finite almost surely, for every $\varepsilon > 0$. As it turns out, the same condition is not sufficient to obtain finite expectation, as our next result shows.

THEOREM 2. *Assume that $\mu \not\equiv 0$, and let $\alpha > 0$, $\varepsilon > 0$ and $d \geq 2$. Then*

$$\mathbf{E}\big[Y^{d+\alpha}\big] < \infty \quad \Longleftrightarrow \quad \mathbf{E}\big[|\mathscr{T}_\varepsilon|^\alpha\big] < \infty \quad \Longleftrightarrow \quad \mathbf{E}\big[(\sup \mathscr{T}_\varepsilon)^\alpha\big] < \infty.$$

The title of the paper refers to Theorem 1, and is motivated by the following comparison with its one-dimensional analogue. Let S_n denote the sum of n i.i.d. random variables with mean m. The strong law of large numbers states that S_n/n converges almost surely to its mean as n tends to infinity. This is the one-dimensional analogue of the shape theorem. Equivalently put, the number of n for which $|S_n - nm| > \varepsilon n$ is almost surely finite for every $\varepsilon > 0$. This is true if and only if the mean m is finite. Hsu and Robbins [15] proved that if also second moments are finite, then

$$\sum_{n=1}^{\infty} \mathbf{P}\big(|S_n - nm| > \varepsilon n\big) < \infty.$$

Erdős [10, 11] showed that finite second moment is in fact also necessary for this stronger conclusion to hold. In particular, the stronger conclusion requires a stronger hypothesis. The analogous strengthening of the shape theorem (Theorem 1) holds without the need of a stronger hypothesis.

REMARK. A sequence $X_1, X_2, \ldots$ of random variables which for all $\varepsilon > 0$ and some random variable X satisfies $\sum_{n=1}^{\infty} \mathbf{P}(|X_n - X| > \varepsilon) < \infty$ is necessarily convergent to X almost surely. As introduced by Hsu and Robbins, this stronger mode of convergence was, "for want of a better name," by them called *complete*. In their language, Theorem 1 implies that $\mathbf{E}[Y^d] < \infty$ is necessary and sufficient not only for the almost sure, but also for complete convergence in the shape theorem.

REMARK. In the statement of the shape theorem, (4) is often exchanged for $(1-\varepsilon)\widetilde{\mathscr{B}}^{\mu} \subset \frac{1}{t}\widetilde{\mathscr{B}}_t \subset (1+\varepsilon)\widetilde{\mathscr{B}}^{\mu}$, where $\widetilde{\mathscr{B}}_t$ is the "fattened" set obtained by replacing each site in $\mathscr{B}_t$ by a unit cube centered around it, and where $\widetilde{\mathscr{B}}^{\mu} = \{x \in \mathbf{R}^d : \mu(x) \leq 1\}$. This formulation is equivalent to the one based on (4). In Section 7, where Theorem 2 is proved, it will be clear why it is more convenient to work with the discrete sets in (4).

1.2. *Large deviation estimates.* Grimmett and Kesten were concerned with large deviation estimates and large deviation principles for travel times in coordinate directions, such as the family $\{T(0, n\mathbf{e}_1) - n\mu(\mathbf{e}_1)\}_{n\geq 1}$. Their results in [13] were soon improved upon in [16]. Deviations above and below the time constants behave quite differently. The first observation in this direction is that the probability of large deviations below the time constant decay at an exponential rate without restrictions to the passage time distribution. This was proved for travel times in coordinate direction in [13], $d = 2$, and [16], $d \geq 2$. A further indication is that the probability of deviations above the time constant decays superexponentially, subject to a sufficiently strong moment condition on Y (at least exponential). This observation was first made by Kesten [16], Theorem 5.9, but more recently refined by Chow and Zhang [7] and Cranston, Gauthier and Mountford [9].

The main contribution of this study is an estimate on linear order deviations above the time constant under the assumption that $\mathbf{E}[Y^{\alpha}] < \infty$ for some $\alpha > 0$, presented in Theorem 4 below. The proof of this result will make crucial use of the fact that what determines the rate of decay of the probability tail of the travel time $T(0, z)$ is the distribution of the weights of the $2d$ edges reaching out of the origin and the $2d$ edges leading in to the point z. This idea is in itself not new. A similar idea was used by Cox and Durrett to prove existence of the limit in (1) without any moment condition (see [8] for $d = 2$, and [16] for higher dimensions). Also Zhang [21] builds on their work and obtains concentration inequalities based on a moment condition for the travel time between sets whose size is growing logarithmically in their distance.

We provide two estimates for deviations of linear order, considering deviations below and above the time constant separately. The first extends the result of Grimmett and Kesten [13, 16] from the coordinate axis to all of $\mathbf{Z}^d$.

THEOREM 3. *For every $\varepsilon > 0$, there are $M = M(\varepsilon)$ and $\gamma = \gamma(\varepsilon) > 0$ such that for every $z \in \mathbf{Z}^d$ and $x \geq \|z\|$*

$$\mathbf{P}\big(T(0, z) - \mu(z) < -\varepsilon x\big) \leq M e^{-\gamma x}.$$

A similar exponential rate of decay cannot hold in general for deviations above the time constant, since $\mathbf{P}(T(0,z) - \mu(z) > \varepsilon x)$ is bounded from below by $\mathbf{P}(Y > Mx)$ for any sufficiently large M. One may instead wonder whether the decay of $\mathbf{P}(T(0,z) - \mu(z) > \varepsilon x)$ in fact is determined by the probability tails of Y. Interestingly, in the regime of polynomial decay on the probability tails, this is indeed the case.

THEOREM 4. *Assume that* $\mathbf{E}[Y^\alpha] < \infty$ *for some* $\alpha > 0$. *For every* $\varepsilon > 0$ *and* $q \geq 1$ *there exists* $M = M(\alpha, \varepsilon, q)$ *such that for every* $z \in \mathbf{Z}^d$ *and* $x \geq \|z\|$

$$\mathbf{P}\big(T(0,z) - \mu(z) > \varepsilon x\big) \leq M\mathbf{P}(Y > x/M) + \frac{M}{x^q}.$$

The strength in Theorem 3 is that it gives exponential decay without the need of a moment condition, which is best possible; see [16], Theorem 5.2. Moreover, the upper bound is independent of the direction. This fact is a consequence of the equivalence between μ and the usual ℓ^p-distances.

The strength in Theorem 4 is that under a minimal moment assumption, it relates the probability tail of $T(0,z) - \mu(z)$ directly with that of Y, together with an additional error term. In the case of polynomial decay of the tails of Y, this result is essentially sharp. It is not clear whether the polynomially decaying error term that appears in Theorem 4 could be improved or not. However, in view of the exponential decay obtained in the case of a moment condition of exponential order, and the superexponential decay for bounded passage times (see [16], Theorem 5.9, or [7] and [9]), it seems possible that the error in fact may decay at least exponentially fast. This is the most interesting question left open in this study, together with the question whether it is possible to remove the moment condition in Theorem 4 completely.

Theorem 1 is easily derived from Theorem 3 and Theorem 4. The proof of Theorem 3 will follow the steps of [16], whereas the proof of Theorem 4 will be derived from first principles, via a regeneration argument similar to that used in [1]. A similar characterization of deviations away from the time constant as the one presented here has in parallel been derived for first-passage percolation of cone-like subgraphs of the $\mathbf{Z}^d$ lattice by the same author in [2], however, detailed proofs appear only here. Complementary to Theorem 1, we may also obtain necessary and sufficient conditions for summability of tails in radial directions from Theorems 3 and 4.

COROLLARY 5. *For any* $\alpha > 0$, $\varepsilon > 0$ *and* $z \in \mathbf{Z}^d$,

$$\mathbf{E}\big[Y^\alpha\big] < \infty \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} n^{\alpha-1}\mathbf{P}\big(\big|T(0,nz) - n\mu(z)\big| > \varepsilon n\big) < \infty.$$

Another consequence of Theorems 3 and 4 is the following characterization of L^p-convergence, of which a proof may be found in [2].

COROLLARY 6. *For every* $p > 0$,

$$\mathbf{E}[Y^p] < \infty \quad \Longleftrightarrow \quad \limsup_{z \in \mathbf{Z}^d : \|z\| \to \infty} \mathbf{E}\left|\frac{T(0,z) - \mu(z)}{\|z\|}\right|^p = 0.$$

Constants given above and also later on in this paper generally depend on the dimension d and on the actual passage time distribution. However, this will not always be stressed in the notation. We would also like to remind the reader that above and for the rest of this paper we will let $|\cdot|$ denote Euclidean distance, and let $\|\cdot\|$ denote ℓ^1-distance. Although the former notation will also be used to denote cardinality for discrete sets, and Lebesgue measure for (measurable) subsets of $\mathbf{R}^d$, we believe that what is referred to will always be clear form the context. Finally, we will denote the d coordinate directions by $\mathbf{e}_i$ for $i = 1, 2, \dots, d$, and recall that Y denotes the minimum of $2d$ independent random variables distributed as τ_e.

We continue this paper with a discussion of some preliminary results and observations in Section 2. In Section 3, we prove Theorem 3, and in Section 4 we describe a regenerative approach that in Section 5 will be used to prove Theorem 4. Finally, Theorem 1 is derived in Section 6, and Theorem 2 in the ending Section 7.

2. Convergence toward the asymptotic shape. Before moving on to the core of this paper, we will first discuss some preliminary observations and results for later reference. We will begin with a few properties of the time constant, and their consequences for the asymptotic shape $\{x \in \mathbf{R}^d : \mu(x) \le 1\}$. We thereafter describe Cox, Durrett and Kesten's approach to convergence without moment condition, in order to state Kesten's version of the shape theorem. Kesten's theorem will be required in order to prove Theorem 3 without moment condition; a first application is found in Proposition 7 below.

2.1. *The time constant and asymptotic shape.* The foremost characteristic of first-passage percolation is its subadditive property, inherited from its interpretation as a pseudo-metric on $\mathbf{Z}^d$. This property takes the expression

$$T(x,z) \le T(x,y) + T(y,z) \qquad \text{for all } x, y, z \in \mathbf{Z}^d,$$

and will be used repeatedly throughout this study. Subadditivity also carries over in the limit. The time constant μ was defined in (1) on $\mathbf{Z}^d$, but extends in fact continuously to all of $\mathbf{R}^d$. The extension is unique with respect to preservation of the following properties:

$$\mu(ax) = |a|\mu(x) \qquad \text{for } a \in \mathbf{R} \text{ and } x \in \mathbf{R}^d,$$
$$\mu(x+y) \le \mu(x) + \mu(y) \qquad \text{for } x, y \in \mathbf{R}^d,$$
$$\big|\mu(x) - \mu(y)\big| \le \mu(\mathbf{e}_1)\|x - y\| \qquad \text{for } x, y \in \mathbf{R}^d.$$

The third of the above properties is easily obtained from the previous two, and shows that $\mu : \mathbf{R}^d \to [0, \infty)$ is Lipschitz continuous.

As mentioned above, there are two regimes separating the behavior of μ. Either $\mu \equiv 0$, or $\mu(x) \neq 0$ for all $x \neq 0$. The separating factor is, as mentioned, whether $\mathbf{P}(\tau_e = 0) \geq p_c(d)$ or not, where $p_c(d)$ denotes the critical probability for bond percolation on $\mathbf{Z}^d$. In the latter regime μ satisfies all the properties of a norm on $\mathbf{R}^d$, and the unit ball $\{x \in \mathbf{R}^d : \mu(x) \leq 1\}$ in this norm can be shown to be compact convex and to have nonempty interior. Consequently, μ is bounded away from 0 and infinity on any compact set not including the origin. In particular,

$$0 < \inf_{\|x\|=1} \mu(x) \leq \sup_{\|x\|=1} \mu(x) < \infty.$$

A careful account for the above statements is found in [16]. (See also Appendix A below.)

2.2. *A shape theorem without moment condition*. Cox and Durrett [8] found a way to prove existence of the limit in (1) without restrictions to the passage time distribution. Their argument was presented for $d = 2$, and later extended to higher dimensions by Kesten [16]. As a consequence, Kesten showed that the moment condition in the shape theorem can be removed to the cost of a weakening of its conclusion. Since Kesten's result will be important in order to derive an estimate on large deviations below the time constant (Theorem 3), we will recall the result here. To reproduce the result in a fair amount of detail requires that some notation is introduced. However, a bit loosely put, Kesten's result states that if $\mathscr{B}_t$ is replaced by the set $\overline{\mathscr{B}_t}$ containing $\mathscr{B}_t$ and each other point in $\mathbf{Z}^d$ "surrounded" by $\mathscr{B}_t$, then (4) holds for all large enough t almost surely also without the moment condition. That is, $\overline{\mathscr{B}_t}$ should be thought of as containing all points from which there is no infinite self-avoiding path disjoint with $\mathscr{B}_t$.

Given $\delta > 0$, pick $\bar{t} = \bar{t}(\delta)$ such that $\mathbf{P}(\tau_e \leq \bar{t}) \geq 1 - \delta$. Next, color each vertex in $\mathbf{Z}^d$ either black or white depending on whether at least one of the edges adjacent to it has weight larger than $\bar{t}$ or not. The moral here is that if δ is small, then an infinite connected component of white vertices will exist with probability one, and that travels within this white component are never "slow." Based on this idea, we go on and define "shells" of white vertices around each point in $\mathbf{Z}^d$. If we make sure that these shells exist, are not too large, but intersect the infinite white component, then the shells may be used to "surround" points in $\mathbf{Z}^d$.

Without reproducing all the details, Kesten shows that it is possible to define a set $\Delta_z \subset \mathbf{Z}^d$ consisting of white vertices and which, given that $\delta = \delta(d) > 0$ is sufficiently small, almost surely satisfies the following properties (see Appendix B below):

(1) Δ_z is a finite connected subset of $\mathbf{Z}^d$,
(2) every path connecting z to infinity has to intersect Δ_z,

(3) there is a point in Δ_z which is connected to infinity by a path of white vertices,

(4) either every path between y and z in $\mathbf{Z}^d$ intersects both Δ_y and Δ_z, or $\Delta_y \cap \Delta_z \neq \varnothing$.

Moreover, the shells may be chosen so that their diameter, defined as the maximal ℓ^1-distance between a pair of its elements, for some $M < \infty$ and $\gamma > 0$ satisfies

$$\mathbf{P}\big(\mathrm{diam}(\Delta_z) > n\big) \le Me^{-\gamma n} \qquad \text{for all } n \ge 1. \tag{5}$$

The advantage of the construction of shells is that although travel times between points may be too heavy-tailed to obey a strong law, the travel time between shells of two points in $\mathbf{Z}^d$ have finite moments of all orders. That $T(\Delta_y, \Delta_z)$ is a lower bound for $T(y, z)$ is a consequence of the fourth property. A complementary upper bound is obtained by summing the weights of paths connecting y and Δ_y, Δ_y and Δ_z, and Δ_z and z, respectively. These paths may not intersect and form a path between y and z, so in order to obtain an upper bound, we also have to consider the maximal weight of a path between two points in Δ_y and Δ_z, respectively. Since each shell is white and connected, and the $2d$ edges adjacent to a white vertex have weight at most $\bar{t}$, we arrive at the following inequality:

$$0 \le T(y, z) - T(\Delta_y, \Delta_z) \le T(y, \Delta_y) + T(\Delta_z, z) + 2d\bar{t}\big(|\Delta_y| + |\Delta_z|\big). \tag{6}$$

Without the need of a moment condition $(T(\Delta_0, \Delta_{nz}) + 2d\bar{t}|\Delta_{nz}|)_{n\ge 1}$ is found to satisfy the conditions of the subadditive ergodic theorem; see [16], Theorem 2.26. Consequently, the limit of $\frac{1}{n}T(\Delta_0, \Delta_{nz})$ as $n \to \infty$ exists almost surely and in L^1, and together with (6), existence of the limit in (1) is obtained.

Let us now move on to state Kesten's version of the shape theorem. For our purposes, it will be practical to present the statement on the form of a limit, in analogy to (3). On this form, Kesten's theorem [16], Theorem 3.1, simply states that

$$\limsup_{z\in\mathbf{Z}^d : \|z\|\to\infty} \frac{|T(0, \Delta_z) - \mu(z)|}{\|z\|} = 0 \qquad \text{almost surely.} \tag{7}$$

The weak version of the shape theorem stated in (2) is now easily obtained from (7) together with (6). As a comparison, recall that Cox and Durrett's version of the shape theorem states that if $\mathbf{E}[Y^d] < \infty$, then (7) holds also if Δ_z is replaced by z.

2.3. *Point-to-shape travel times.* In view of the convergence of the set $\mathscr{B}_t$ toward a convex compact set described in terms of $\mu(\cdot)$, it is reasonable to study the travel time to points at a large distance with respect to this norm. That is, introduce what could be referred to as *point-to-shape* travel times as $T(0, \neg\mathscr{B}_n^{\mu})$, where $\neg\mathscr{B}_n^{\mu} := \mathbf{Z}^d \setminus \mathscr{B}_n^{\mu} = \{z \in \mathbf{Z}^d : \mu(z) > n\}$. This definition only makes sense in the case that $\mu \not\equiv 0$, and in this case a strong law for the point-to-shape travel times holds without restriction on the passage time distribution.

PROPOSITION 7. *Assume that $\mu \not\equiv 0$. Then*

$$\lim_{n\to\infty} \frac{T(0, \neg\mathscr{B}_n^\mu)}{n} = 1 \qquad \textit{almost surely.}$$

PROOF. Let m_n denote the least integer for which $m_n\mu(\mathbf{e}_1) > n$. By definition, we have

$$\frac{T(0, \neg\mathscr{B}_n^\mu)}{n} \le \frac{T(0, \Delta_{m_n\mathbf{e}_1}) + \bar{t}|\Delta_{m_n\mathbf{e}_1}|}{n} \to 1 \qquad \text{almost surely.}$$

So, it is sufficient to show that the event

$$A_\delta = \left\{\liminf_{n\to\infty} \frac{T(0, \neg\mathscr{B}_n^\mu)}{n} \le 1 - \delta\right\}$$

has probability 0 to occur for every $\delta > 0$. On the event A_δ there is an increasing sequence $(n_k)_{k\ge1}$ of integers for which $T(0, \neg\mathscr{B}_{n_k}^\mu) \le (1 - \delta/2)n_k$. For each such n_k, there is a site v_k such that $\mu(v_k) > n_k$, but $T(0, v_k) \le (1-\delta/2)n_k$. When n_k is large we may further assume that $\mu(v_k) \le 2n_k$. Consequently, we conclude that for large k

$$T(0, v_k) - \mu(v_k) \le -\delta n_k/2 \le -\delta\mu(v_k)/4 \le -\varepsilon\|v_k\|, \tag{8}$$

for some $\varepsilon > 0$. However, $T(0, \Delta_v) \le T(0, v)$, so the occurrence of (8) for infinitely many k is contradicted by (7), almost surely. That is, $\mathbf{P}(A_\delta) = 0$ for every $\delta > 0$, as required. □

3. Large deviations below the time constant. We will follow the approach of Kesten [16], Theorem 5.2, on our way to a proof of Theorem 3. If $\mu \equiv 0$, then there is nothing to prove. So, we may assume that $\mu \not\equiv 0$. Unlike Kesten, we will work with the point-to-shape travel times introduced above in order to obtain a bound on deviations in all directions simultaneously, and not only for coordinate directions. The first and foremost step is this next lemma.

LEMMA 8. *Let $X_{\ell,\ell+m}^{(q)}$ for $q = 1, 2, \ldots$ denote independent random variables distributed as $T(\mathscr{B}_\ell^\mu, \neg\mathscr{B}_{\ell+m}^\mu)$. There exists $C < \infty$ such that for every $n \ge m \ge \ell \ge 1$ and $x > 0$ we have*

$$\mathbf{P}\big(T\big(0, \neg\mathscr{B}_n^\mu\big) < x\big) \le \sum_{Q+1\ge n/(m+C\ell)} n^{d-1}\left(C\frac{m}{\ell}\right)^{d(Q-1)} \mathbf{P}\left(\sum_{q=1}^{Q} X_{\ell,\ell+m}^{(q)} < x\right).$$

PROOF. Pick $z \in \mathbf{Z}^d$ such that $\mu(z) > n$. Let $\Gamma = \Gamma(z)$ be a self-avoiding path connecting the origin to z. Choose a subsequence $v_0, v_1, \ldots, v_Q$ of the vertices in Γ as follows. Set $v_0 = 0$. Given v_q, choose v_{q+1} to be the first vertex in Γ succeeding v_q such that

$$\mu(v_{q+1} - v_q) > m + 2\ell.$$

When no such vertex exists, stop and set $Q = q$. To find a lower bound on Q, note that

$$n < \mu(z) \le \mu(z - v_Q) + \mu(v_Q) \le \mu(z - v_Q) + \sum_{q=0}^{Q-1} \mu(v_{q+1} - v_q).$$

Since $\mu(v_{q+1} - v_q) \le m + 2\ell + \mu(\mathbf{e}_1)$ and $\mu(z - v_Q) \le m + 2\ell$, we see that Q must satisfy

$$n \le (Q+1)\big(m + \big(2 + \mu(\mathbf{e}_1)\big)\ell\big). \tag{9}$$

Next, pick $r > 0$ such that $[-r, r]^d \subseteq \mathscr{B}_1^\mu$ and tile $\mathbf{Z}^d$ with copies of $(-r\ell, r\ell]^d$ such that each box is centered at a point in $\mathbf{Z}^d$, and each point in $\mathbf{Z}^d$ is contained in precisely one box. Let Λ_q denote the box that contains v_q, and let w_q denote the center of Λ_q. Of course, the tiling can be assumed chosen such that $w_0 = v_0 = 0$. Denote by Γ_q the part of the path Γ that connects v_q and v_{q+1}. Note that for $q_1 \neq q_2$ the two pieces Γ_{q_1} and Γ_{q_2} are edge disjoint. By construction, v_q is contained in the copy of $\mathscr{B}_\ell^\mu$ centered at w_q, while v_{q+1} is not contained in the copy of $\mathscr{B}_{\ell+m}^\mu$ centered at w_q (see Figure 1). That is,

$$\mu(v_q - w_q) \le \ell \quad \text{and} \quad \mu(v_{q+1} - w_q) > \ell + m. \tag{10}$$

Moreover, the points $w_0, w_1, \dots, w_{Q-1}$ have to satisfy

$$\mu(w_{q+1} - w_q) \le m + 4\ell + \mu(\mathbf{e}_1). \tag{11}$$

Let W_Q denote the set of all sequences $(w_0, w_1, \dots, w_{Q-1})$ such that $w_0 = 0$, each w_q is the center of some box Λ_q, and w_q and w_{q+1} satisfies (11) for each $q = 0, 1, \dots, Q-2$.

Given $x > 0$, $Q \in \mathbf{Z}_+$ and $w = (w_0, w_1, \dots, w_{Q-1}) \in W_Q$, let $A(x, w)$ denote the event that there exists a path Γ from the origin to z with edge disjoint segments $\Gamma_0, \Gamma_1, \dots, \Gamma_{Q-1}$ such that:

(1) $\sum_{q=0}^{Q-1} T(\Gamma_q) < x$,
(2) the endpoints v_q and v_{q+1} of Γ_q satisfy (10), for each $q = 0, 1, \dots, Q-1$.

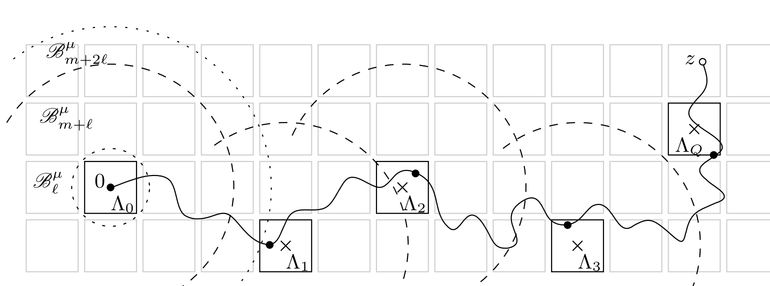

FIG. 1. *The decomposition of a path into segments, where dots represent the v_k's.*

Since $T(\Gamma) \geq \sum_{q=0}^{Q-1} T(\Gamma_q)$, together with (9), we obtain that

$$\{T(0,z) < x\} \subseteq \bigcup_{Q+1 \geq n/(m+b\ell)} \bigcup_{w \in W_Q} A(x,w), \tag{12}$$

where $b = 2 + \mu(\mathbf{e}_1)$. Note that given w_q, the passage time of any path between two vertices v and v' such that $\mu(v - w_q) \leq \ell$ and $\mu(v' - w_q) > \ell + m$ is stochastically larger than $T(\mathscr{B}^\mu_\ell, \neg\mathscr{B}^\mu_{\ell+m})$. Hence, via a BK-like inequality (e.g., Theorem 4.8, or (4.13), in [16]), it is for each $w \in W_Q$ possible to bound the probability of the event $A(x,w)$ from above by

$$\mathbf{P}\big(X^{(1)}_{\ell,\ell+m} + X^{(2)}_{\ell,\ell+m} + \cdots + X^{(Q)}_{\ell,\ell+m} < x\big). \tag{13}$$

It remains to count the number of elements $(w_0, w_1, \dots, w_{Q-1})$ in W_Q. Assuming that w_q has already been chosen, the number of choices for w_{q+1} is restricted by (11). In particular, w_{q+1} has to be contained in a cube centered at w_q and whose side length is a multiple of $(5 + \mu(\mathbf{e}_1))m$. This cube is intersected by at most $(Cm/\ell)^d$ boxes of the form $(-r\ell, r\ell]^d$ in the tiling of $\mathbf{Z}^d$, for some $C < \infty$. Since w_{q+1} is the center of one of these boxes, this is also an upper bound for its number of choices. Consequently, the total number of choices for $w_1, w_2, \dots, w_{Q-1}$ is at most $(Cm/\ell)^{d(Q-1)}$. Together with (12) and (13), we conclude that

$$\mathbf{P}\big(T(0,z) < x\big) \leq \sum_{Q+1 \geq n/(m+C\ell)} \left(C\frac{m}{\ell}\right)^{d(Q-1)} \mathbf{P}\left(\sum_{q=1}^{Q} X^{(q)}_{\ell,\ell+m} < x\right),$$

for some $C < \infty$. The lemma follows observing that the number of $z \in \mathbf{Z}^d$ that satisfies $\mu(z) > n$ and has a neighbor within $\mathscr{B}^\mu_n$ is of order n^{d-1}. $\square$

LEMMA 9. *Assume that $\mu \not\equiv 0$. For every $\varepsilon > 0$,*

$$\lim_{m\to\infty} \max_{\ell \leq m} \mathbf{P}\big(T\big(\mathscr{B}^\mu_\ell, \neg\mathscr{B}^\mu_{\ell+m}\big) < m(1-\varepsilon)\big) = 0.$$

PROOF. Let Γ be a path with endpoints z and y satisfying $\mu(z) \leq \ell$ and $\mu(y) > \ell + m$. Then

$$T\big(0, \neg\mathscr{B}^\mu_{\ell+m}\big) \leq T(0, \Delta_z) + \bar{t}|\Delta_z| + T(\Gamma),$$

and we may choose Γ so that $T(\Gamma) = T(\mathscr{B}^\mu_\ell, \neg\mathscr{B}^\mu_{\ell+m})$. It follows that

$$\begin{aligned} T\big(0, \neg\mathscr{B}^\mu_{\ell+m}\big) &\leq \max\{T(0,\Delta_z) : \mu(z) \leq \ell\} \\ &\quad + \bar{t} \cdot \max\{|\Delta_z| : \mu(z) \leq \ell\} + T\big(\mathscr{B}^\mu_\ell, \neg\mathscr{B}^\mu_{\ell+m}\big). \end{aligned}$$

For every $\varepsilon > 0$, we have $\mathbf{P}(T(0, \neg\mathscr{B}^\mu_{\ell+m}) < (\ell+m)(1-\varepsilon/4)) < \varepsilon$ for all large enough m, by Proposition 7. Thus, it suffices to show that for all large m and $\ell \leq m$ also

$$\mathbf{P}\left(\max\{T(0,\Delta_z) : \mu(z) \leq \ell\} > \ell + \frac{\varepsilon m}{4}\right) + \mathbf{P}\left(\max\{|\Delta_z| : \mu(z) \leq m\} > \frac{\varepsilon m}{4\bar{t}}\right)$$

is at most ε. The latter of the two probabilities vanishes as $m \to \infty$ as a consequence of the exponential decay of the diameter of a shell in (5). To show that the former probability is small, we will use (7).

Let N be an integer and note that if $\max\{T(0, \Delta_z) : \mu(z) \le \ell\} > \ell + \varepsilon m/4$, then either $\max\{T(0, \Delta_z) : \mu(z) \le N\} > \varepsilon m/4$ or $T(0, \Delta_z) - \mu(z) > \ell - \mu(z) + \varepsilon m/4$ for some z satisfying $\mu(z) \in [N, \ell]$. Since $m \ge \ell \ge \mu(z)$, we have $\ell - \mu(z) + \varepsilon m/4 \ge \varepsilon\mu(z)/4$. We thus obtain the inequality

$$\begin{aligned}
&\mathbf{P}\Big(\max\{T(0, \Delta_z) : \mu(z) \le \ell\} > \ell + \frac{\varepsilon m}{4}\Big) \\
&\quad \le \mathbf{P}\Big(\max\{T(0, \Delta_z) : \mu(z) \le N\} > \frac{\varepsilon m}{4}\Big) \\
&\qquad + \mathbf{P}\big(T(0, \Delta_z) > (1 + \varepsilon/4)\mu(z) \text{ for some } \mu(z) \ge N\big).
\end{aligned}$$

From (7), we know that the right-hand side can be made arbitrarily small by choosing N large and sending m to infinity. □

PROOF OF THEOREM 3. We may assume that $\mu \not\equiv 0$. We will prove that for every $\varepsilon > 0$ there exist $M = M(\varepsilon)$ and $\gamma = \gamma(\varepsilon)$ such that for every $x \ge n \ge 1$

$$\mathbf{P}\big(T\big(0, \neg\mathscr{B}_n^{\mu}\big) < n - \varepsilon x\big) \le Me^{-\gamma x},$$

from which Theorem 3 is an easy consequence.

Let $X_{\ell,\ell+m}^{(1)}, X_{\ell,\ell+m}^{(2)}, \dots, X_{\ell,\ell+m}^{(Q)}$ and $C < \infty$ be as in Lemma 8, and fix $\varepsilon \in (0, 4C)$. For some integer m, let $\ell = \ell(m)$ be the largest integer such that $\ell \le \frac{m\varepsilon}{4C}$. Markov's inequality and independence give that for any $\xi > 0$

$$\mathbf{P}\Bigg(\sum_{q=1}^{Q} X_{\ell,\ell+m}^{(q)} < n - \varepsilon x\Bigg) \le e^{\xi(n-\varepsilon x)}\mathbf{E}\big[e^{-\xi X_{\ell,\ell+m}^{(1)}}\big]^Q.$$

Writing $n - \varepsilon x = n(1 - \varepsilon) - \varepsilon(x - n)$, we obtain for $(Q + 1)(m + C\ell) \ge n$ the upper bound

$$e^{-\varepsilon\xi(x-n)}e^{\xi(m+C\ell)}\big[e^{\xi(m+C\ell)(1-\varepsilon)}\big(e^{-\xi m(1-\varepsilon/2)} + \mathbf{P}\big(X_{\ell,\ell+m}^{(1)} < m(1 - \varepsilon/2)\big)\big)\big]^Q.$$

Since $C\ell - m\varepsilon/2 \le -m\varepsilon/4$ and $m + C\ell \le (1 + \varepsilon/4)m$, the expression within square brackets is at most

$$e^{-\xi m\varepsilon/4} + e^{(1+\varepsilon/4)\xi m}\mathbf{P}\big(X_{\ell,\ell+m}^{(1)} < m(1 - \varepsilon/2)\big). \tag{14}$$

According to Lemma 9, we can make (14) arbitrarily small by choosing ξ and m such that ξm is large and m is as large as necessary. Fix ξ and m such that $\ell \ge 1$ and (14) is not larger than

$$(2C)^{-d}\bigg(\frac{8C}{\varepsilon}\bigg)^{-d} \le \bigg(2C\frac{m}{\ell}\bigg)^{-d}.$$

Finally, apply Lemma 8 with these ξ, m and ℓ to obtain

$$\begin{aligned}
&\mathbf{P}\big(T\big(0, \neg\mathscr{B}_n^{\mu}\big) < n - \varepsilon x\big) \\
&\quad \le e^{-\varepsilon\xi(x-n)} e^{\xi(m+C\ell)} \sum_{(Q+1)\ge n/(m+C\ell)} n^{d-1}\left(C\frac{m}{\ell}\right)^{d(Q-1)}\left(2C\frac{m}{\ell}\right)^{-dQ} \\
&\quad \le e^{-\varepsilon\xi(x-n)} \cdot e^{\xi(m+C\ell)} \cdot n^{d-1} \cdot 2^{-d(n/(m+C\ell)-1)+1},
\end{aligned}$$

which is of the required form. $\square$

4. A regenerative approach. We will in this section explore a regenerative approach that can be used to study the asymptotics of travel times along cylinders. This approach was previously studied in more detail in [1]. It will for the sake of this paper be sufficient to obtain a sequence which is approximately regenerative, which in turn avoids some additional technicalities. Some additional notation will be required however.

Given $z \in \mathbf{Z}^d$ and $r \ge 0$, let $\mathcal{C}(z, r) := \bigcup_{a\in\mathbf{R}} B(az, r)$ denote the cylinder in direction z of radius r, where $B(x, r) := \{y \in \mathbf{R}^d : |y - x| \le r\}$ denotes the closed Euclidean ball. The travel time between two points x and y over paths restricted to the cylinder $\mathcal{C}(z, r)$ will be denoted by $T_{\mathcal{C}(z,r)}(x, y)$. The regenerative approach referred to will consist of a comparison between $T_{\mathcal{C}(z,r)}(0, nz)$ and the sum of travel times between randomly chosen "cross-sections" of $\mathcal{C}(z, r)$.

Due to symmetry it means no restriction assuming that $z \in \mathbf{Z}^d$ lies in the first orthant, that is, that the coordinate $z_i \ge 0$ for each $i = 1, 2, \dots, d$. Let $\mathcal{H}_n := \{z \in \mathbf{Z}^d : z_1 + z_2 + \cdots + z_d = n\}$, $r \ge 0$, and pick $\bar{t} \in \mathbf{R}_+$ such that $\mathbf{P}(\tau_e \le \bar{t}) > 0$. The following notation will be used, and is illustrated in Figure 2 below:

$$\begin{aligned}
&V_n(z, r) := \mathcal{C}(z, r) \cap \mathcal{H}_{n\|z\|}, \\
&E_n(z, r) := \big\{\text{edges connecting } \mathcal{C}(z, r) \cap \mathcal{H}_{n\|z\|-1} \text{ with } \mathcal{C}(z, r) \cap \mathcal{H}_{n\|z\|}\big\}, \\
&A_n(z, r) := \big\{\tau_e \le \bar{t} \text{ for all } e \in E_n(z, r)\big\}, \\
&\rho_j(z, r) := \min\big\{n > \rho_{j-1}(z, r) : A_n(z, r) \text{ occurs}\big\} \qquad \text{for } j \ge 1, \rho_0 = 0.
\end{aligned}$$

When understood from the context, the reference to z and r will be dropped.

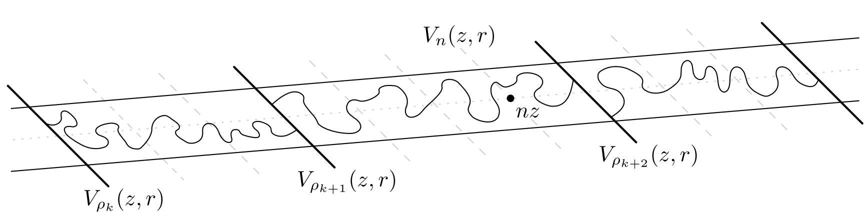

FIG. 2. *A piece of $\mathcal{C}(z, r)$. The thick diagonal lines indicate $\{V_{\rho_j}(z, r)\}_{j\ge 0}$ and the curly lines $\{T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j})\}_{j\ge 1}$. In this illustration, we have $\nu(n) = k + 2$.*

Note that $\{A_n(z,r)\}_{n\geq 1}$ are i.i.d., so the increments $\{\rho_j - \rho_{j-1}\}_{j\geq 1}$ are independent geometrically distributed with success probability $\mathbf{P}(A_0(z,r))$. Consequently, $\{T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j})\}_{j\geq 1}$ are i.i.d. Introduce the following notation for their means:

$$\mu_\tau(z,r) := \mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_0}, V_{\rho_1})\big],$$
$$\mu_\rho(z,r) := \mathbf{E}[\rho_1 - \rho_0],$$

and, for the time constant for travel times restricted to cylinders, let

$$\mu_{\mathcal{C}(z,r)} := \lim_{n\to\infty} \frac{\mathbf{E}[T_{\mathcal{C}(z,r)}(V_0, V_n)]}{n}.$$

The existence of the above limit is given by Fekete's lemma, since that $(-\mathbf{E}[T_{\mathcal{C}(z,r)}(V_0, V_n)])_{n\geq 1}$ is a subadditive sequence. A sufficient condition for the limit $\mu_{\mathcal{C}(z,r)}$ to be finite will be achieved with Proposition 13 below.

Finally, travel times on $\mathcal{C}(z,r)$ and the sequence $\{T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j})\}_{j\geq 1}$ will be compared via optimal stopping. We define for that purpose the stopping time

$$\nu(m) = \nu(m,z,r) := \min\{j \geq 1 : \rho_j(z,r) > m\}.$$

Note that $\nu(m) - 1$ equals the number of $n \in \{1,2,\dots,m\}$ for which $A_n(z,r)$ occurs, which is binomially distributed with success probability $\mathbf{P}(A_0(z,r)) = \mu_\rho(z,r)^{-1}$.

REMARK. A geometrical constraint should be noted. For some $z \in \mathbf{Z}^d$, there may not be any paths at all between x and y that only passes through points in $\mathcal{C}(z,r)$, when r is small. [In this case set $T_{\mathcal{C}(z,r)}(x,y) = \infty$.] However, it is not hard to realize that for every $k \geq 1$, there is $R = R(d,k)$ such that for every $z \in \mathbf{Z}^d$ and $r \geq R$ there are k edge-disjoint paths from $V_0(z,r)$ to $V_1(z,r)$ of length $\|z\|$, which are all contained in $\mathcal{C}(z,r)$. ($R = k\sqrt{d}$ is sufficient.)

4.1. *Tail and moment comparisons*. The next task will be to relate tail probabilities of travel times and moments of $T_{\mathcal{C}(z,r)}(V_{\rho_0}, V_{\rho_1})$ with the corresponding quantities for Y. The latter will provide a sufficient condition for $\mu_{\mathcal{C}(z,r)}$ to be finite and converge to $\mu(z)$ as $r \to \infty$. We begin with a well-known tail comparison.

LEMMA 10. *For every $z \in \mathbf{Z}^d$, $x \geq 0$ and large enough r,*

$$\mathbf{P}\big(T_{\mathcal{C}(z,r)}(0,z) > 9\|z\|x\big) \leq 9^{2d}\|z\|\mathbf{P}(Y > x).$$

PROOF. Note that there are $2d$ edge disjoint paths between the origin and $\mathbf{e}_1$ of length at most 9. Denote these paths by $\Gamma_1, \Gamma_2, \dots, \Gamma_{2d}$, and assume that Γ_1 is the longest among them. Clearly,

$$\mathbf{P}\Big(\min_{i=1,2,\dots,2d} T(\Gamma_i) > 9x\Big) \leq \mathbf{P}\big(T(\Gamma_1) > 9x\big)^{2d} \leq 9^{2d}\mathbf{P}(\tau_e > x)^{2d}.$$

For r large, $T_{\mathcal{C}(z,r)}(0,z)$ is dominated by $\|z\|$ random variables distributed as $\min_{i=1,2,\dots,2d} T(\Gamma_i)$. Consequently,

$$\begin{aligned}\mathbf{P}\big(T_{\mathcal{C}(z,r)}(0,z) > 9\|z\|x\big) &\le \|z\|\mathbf{P}\Big(\min_{i=1,2,\dots,2d} T(\Gamma_i) > 9x\Big)\\ &\le 9^{2d}\|z\|\mathbf{P}(Y > x),\end{aligned}$$

as required. $\square$

In preparation for the second aim, we have a couple of lemmata of general character.

LEMMA 11. *Let $\{\tau_i\}_{i\ge 1}$ be a collection of nonnegative i.i.d. random variables. For any $\alpha, \beta > 0$ and integers $L \ge K \ge 1$ such that $\beta K \le \alpha L$, then*

$$\mathbf{E}\Big[\Big(\min_{i\le L}\tau_i\Big)^{\beta}\Big] \le 1 + \frac{\beta}{\alpha}\mathbf{E}\Big[\Big(\min_{i\le K}\tau_i\Big)^{\alpha}\Big]^{L/K}.$$

PROOF. Recall the formula $\mathbf{E}[X^\alpha] = \alpha\int_0^\infty x^{\alpha-1}\mathbf{P}(X > x)\,dx$, valid for nonnegative random variables and $\alpha > 0$. Note that for any $x \ge 1$ Markov's inequality gives that

$$\mathbf{P}(\tau_i > x) = \mathbf{P}\Big(\min_{i\le K}\tau_i > x\Big)^{1/K} \le \frac{\mathbf{E}[(\min_{i\le K}\tau_i)^\alpha]^{1/K}}{x^{\alpha/K}},$$

from which one, under the imposed conditions, easily obtains

$$x^{\beta-1}\mathbf{P}(\tau_i > x)^L \le x^{\alpha-1}\mathbf{P}\Big(\min_{i\le K}\tau_i > x\Big)\cdot\mathbf{E}\Big[\Big(\min_{i\le K}\tau_i\Big)^{\alpha}\Big]^{(L-K)/K}.$$

Finally, integrating over the intervals $[0,1)$ and $[1,\infty)$ separately yields

$$\begin{aligned}\mathbf{E}\Big[\Big(\min_{i\le L}\tau_i\Big)^{\beta}\Big] &\le 1 + \beta\mathbf{E}\Big[\Big(\min_{i\le K}\tau_i\Big)^{\alpha}\Big]^{(L-K)/K}\int_{x\ge 1} x^{\alpha-1}\mathbf{P}\Big(\min_{i\le K}\tau_i > x\Big)\,dx\\ &= 1 + \frac{\beta}{\alpha}\mathbf{E}\Big[\Big(\min_{i\le K}\tau_i\Big)^{\alpha}\Big]^{1+(L-K)/K},\end{aligned}$$

as required. $\square$

LEMMA 12. *Let $\{\tau_{i,j}\}_{i,j\ge 1}$ be a collection of nonnegative i.i.d. random variables. For any $\alpha, \beta > 0$ and integers $K, N \ge 1$ and $L \ge K$ satisfying $\beta K \le \alpha L$,*

$$\mathbf{E}\Big[\Big(\min_{i\le L}\sum_{j\le N}\tau_{i,j}\Big)^{\beta}\Big] \le N^{L+\beta}\Big(1 + \frac{\beta}{\alpha}\mathbf{E}\Big[\Big(\min_{i\le K}\tau_{i,j}\Big)^{\alpha}\Big]^{L/K}\Big).$$

PROOF. First, since if a sum of N nonnegative numbers is greater than x, then at least one of the terms has to be greater than x/N, it follows that

$$\mathbf{P}\Big(\min_{i\le L}\sum_{j\le N}\tau_{i,j}>x\Big)=\mathbf{P}\Big(\sum_{j\le N}\tau_{i,j}>x\Big)^L\le N^L\mathbf{P}(\tau_{i,j}>x/N)^L.$$

Thus, via the substitution $x=Ny$, we conclude that

$$\begin{aligned}\mathbf{E}\Big[\Big(\min_{i\le L}\sum_{j\le N}\tau_{i,j}\Big)^\beta\Big]&\le\beta N^L\int_0^\infty x^{\beta-1}\mathbf{P}(\tau_{i,j}>x/N)^L\,dx\\&=N^{L+\beta}\mathbf{E}\Big[\Big(\min_{i\le L}\tau_{i,j}\Big)^\beta\Big],\end{aligned}$$

from which the statement follows via Lemma 11. □

PROPOSITION 13. *For every $\beta\ge\alpha>0$ and $z\in\mathbf{Z}^d$, there is a finite constant $R_1=R_1(\alpha,\beta,d)$ such that for $r\ge R_1$ and some finite constant $M_1=M_1(\alpha,\beta,d,z,r)$,*

$$\mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_0},V_{\rho_1})^\beta\big]\le M_1\big(1+\mathbf{E}\big[Y^\alpha\big]\big)^{\beta/\alpha+1}.$$

PROOF. If $\mathbf{P}(\tau_e>\bar{t})=0$, then $\rho_1=1$ and the statement is a consequence of Lemma 12. Assume instead the contrary, in which case a bit more care is needed before appealing to Lemma 12.

Let $\eta=\{\eta_e\}_{e\in\mathcal{E}}$ denote the family of indicator functions $\eta_e=\mathbf{1}_{\{\tau_e>\bar{t}\}}$. Independently of $\{\tau_e\}_{e\in\mathcal{E}}$, let $\{\tilde{\tau}_e\}_{e\in\mathcal{E}}$ be a collection of independent random variables distributed as $\mathbf{P}(\tilde{\tau}_e\in\cdot)=\mathbf{P}(\tau_e\in\cdot|\tau_e>\bar{t})$, and define $\{\sigma_e\}_{e\in\mathcal{E}}$ as

$$\sigma_e:=\begin{cases}\tau_e, & \text{if } \eta_e=1,\\ \tilde{\tau}_e, & \text{if } \eta_e=0.\end{cases}$$

Note that $\{\sigma_e\}_{e\in\mathcal{E}}$ is an i.i.d. family independent of η, but that η determines $\{A_n(z,r)\}_{n\ge1}$, and hence $\{\rho_j-\rho_{j-1}\}_{j\ge1}$, for every z and r. In particular, $\{\sigma_e\}_{e\in\mathcal{E}}$ and $\{\rho_j-\rho_{j-1}\}_{j\ge1}$ are independent. Let $T'_{\mathcal{C}}(x,y)$ denote the passage time between x and y with respect to $\{\sigma_e\}_{e\in\mathcal{E}}$. By construction $\tau_e\le\sigma_e$ for every $e\in\mathcal{E}$, so $T_{\mathcal{C}}(x,y)\le T'_{\mathcal{C}}(x,y)$.

Fix $\beta\ge\alpha>0$ and $z\in\mathbf{Z}^d$. Choose $r=r(\alpha,\beta,d)$ large so that there are at least $2d\beta/\alpha$ disjoint paths between $V_0(z,r)$ and $V_1(z,r)$ of length $\|z\|$, contained in $\mathcal{C}(z,r)$. Similarly, there are equally many paths between $V_{\rho_0}(z,r)$ and $V_{\rho_1}(z,r)$ of length $(\rho_1-\rho_0)\|z\|$. Hence, by Lemma 12

$$\begin{aligned}&\mathbf{E}\big[T'_{\mathcal{C}}(V_{\rho_0},V_{\rho_1})^\beta|\eta\big]\\&\quad\le\big((\rho_1-\rho_0)\|z\|\big)^{2d\beta/\alpha+\beta+1}\Big(1+\frac{\beta}{\alpha}\mathbf{E}\Big[\Big(\min_{i\le2d}\sigma_i\Big)^\alpha\Big]^{\beta/\alpha+1/(2d)}\Big),\end{aligned}$$

where $\sigma_1, \sigma_2, \dots, \sigma_{2d}$ denote independent variables distributed as σ_e. In addition,

$$\begin{aligned}\mathbf{E}\Big[\Big(\min_{i\le 2d}\sigma_i\Big)^\alpha\Big] &= \alpha\int_0^\infty x^{\alpha-1}\mathbf{P}(\tau_e > x|\tau_e > \bar{t})^{2d}\,dx\\ &\le \mathbf{E}\Big[\Big(\min_{i\le 2d}\tau_i\Big)^\alpha\Big]\mathbf{P}(\tau_e > \bar{t})^{-2d}.\end{aligned}$$

Since $T_{\mathcal{C}}(V_{\rho_0}, V_{\rho_1}) \le T'_{\mathcal{C}}(V_{\rho_0}, V_{\rho_1})$, and $\rho_1 - \rho_0$ is geometrically distributed, the bound follows easily. $\square$

4.2. *Time constant comparison*. Proposition 13 gives, in particular, a criterion for $\mu_\tau(z,r)$ and $\mu_{\mathcal{C}(z,r)}$ to be finite.

LEMMA 14. *Assume that* $\mathbf{E}[Y^\alpha] < \infty$ *for some* $\alpha > 0$. *Then* $\mu_\tau(z,r)$ *is finite for all* $z \in \mathbf{Z}^d$ *and* $r \ge R_1$, *where* R_1 *is given by Proposition* 13 (*with* $\beta = 1$). *Moreover*,

$$\frac{\mu_\tau(z,r)}{\mu_\rho(z,r)} \le \mu_{\mathcal{C}(z,r)} \le \frac{\mu_\tau(z,r)}{\mu_\rho(z,r)} + \frac{\bar{t}|E_n|}{\mu_\rho(z,r)}.$$

PROOF. The former assertion is the content of Proposition 13. For the latter, note that

$$\begin{aligned}\sum_{j=1}^{\nu(n)-1} T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j}) &\le T_{\mathcal{C}(z,r)}(V_0, V_n)\\ &\le \sum_{j=1}^{\nu(n)} T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j}) + \bar{t}|E_n|\big(\nu(n)-1\big).\end{aligned}$$

Taking expectations, the right-hand side is via Wald's lemma turned into

$$\mathbf{E}\big[\nu(n)\big]\mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_0}, V_{\rho_1})\big] + \bar{t}|E_n|\mathbf{E}\big[\nu(n)-1\big],$$

which after division by n gives

$$\Big(\frac{1}{\mu_\rho(z,r)} + \frac{1}{n}\Big)\mu_\tau(z,r) + \frac{\bar{t}|E_n|}{\mu_\rho(z,r)}.$$

Sending n to infinity thus gives the upper bound.

For the lower bound, it suffices to show that $\mathbf{E}[T_{\mathcal{C}(z,r)}(V_{\rho_{\nu(n)-1}}, V_{\rho_{\nu(n)}})]$ is bounded. This follows from $T_{\mathcal{C}(z,r)}(V_{\rho_{\nu(n)-1}}, V_{\rho_{\nu(n)}})$ being stochastically dominated by $T_{\mathcal{C}(z,r)}(V_{\rho_{-1}}, V_{\rho_1})$, where $\rho_{-1} = \max\{n \le 0 : A_n(z,r) \text{ occurs}\}$. Now, we have not proven that $T_{\mathcal{C}(z,r)}(V_{\rho_{-1}}, V_{\rho_1})$ has finite mean, but that follows readily from the proof of Proposition 13. $\square$

Since $\mu_{\mathcal{C}(z,r)}$ is decreasing in r, it seems reasonable that it should converge to its lower bound $\mu(z)$, as r tends to infinity. An indication of this was given already in [6] and [16], but proofs of this fact may have appeared only more recently, see, for example, [1, 3]. These proofs assume finite expectation of Y, and to extend them to a minimal moment condition requires some care.

PROPOSITION 15. *Assume that* $\mathbf{E}[Y^\alpha] < \infty$ *for some* $\alpha > 0$. *Then, for every* $z \in \mathbf{Z}^d$,

$$\lim_{r\to\infty} \mu_{\mathcal{C}(z,r)} = \mu(z).$$

PROOF. Fix $z \in \mathbf{Z}^d$ and, based on Proposition 13, pick $s > 0$ sufficiently large for $\mathbf{E}[T_{\mathcal{C}(z,s)}(V_{\rho_0(z,s)}, V_{\rho_1(z,s)})^2]$ to be finite. To the end of this proof, let $V_n = V_n(z,s)$, $E_n = E_n(z,s)$, $\rho_j = \rho_j(z,s)$, $\nu(n) = \nu(n,z,s)$, and $\epsilon(s) = \bar{t}|E_0(z,s)|$. For $r \geq s$, let

$$a_n(r,s) := \mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_1(z,s)}, V_{\rho_{\nu(n)}(z,s)})\big] + \epsilon(s).$$

The sequence $(a_n(r,s))_{n\geq 1}$ is subadditive, that is, $a_{n+m}(r,s) \leq a_n(r,s) + a_m(r,s)$ for all $n, m \geq 1$, as a consequence of the inequality (note that $\rho_1 = \rho_{\nu(0)}$)

$$\begin{aligned}\mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_1}, V_{\rho_{\nu(n+m)}})\big] + \epsilon(s) &\leq \mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_{\nu(0)}}, V_{\rho_{\nu(n)}})\big] \\ &\quad + \mathbf{E}\big[T_{\mathcal{C}(z,r)}(V_{\rho_{\nu(n)}}, V_{\rho_{\nu(n+m)}})\big] + 2\epsilon(s).\end{aligned}$$

Recall Fekete's lemma, which says that for any subadditive sequence the limit $\lim_{n\to\infty} \frac{1}{n} a_n(r,s)$ exists and equals $\inf_{n\geq 1} \frac{1}{n} a_n(r,s)$. This holds for every $r \geq s$, including the case $r = \infty$ in which the cylinder equals the whole lattice.

Next note the inequality

$$\begin{aligned}T_{\mathcal{C}(z,r)}\big(V_0(z,r), V_n(z,r)\big) &\leq T_{\mathcal{C}(z,r)}\big(V_0(z,s), V_n(z,s)\big) \\ &\leq T_{\mathcal{C}(z,r)}(V_0, V_{\rho_1}) + T_{\mathcal{C}(z,r)}(V_{\rho_1}, V_{\rho_{\nu(n)}}) \\ &\quad + T_{\mathcal{C}(z,r)}(V_n, V_{\rho_{\nu(n)}}) + 2\epsilon(s),\end{aligned}$$

which shows that $\mu_{\mathcal{C}(z,r)} \leq \lim_{n\to\infty} \frac{1}{n} a_n(r,s)$ for every $r \geq s$. Consequently,

$$\begin{aligned}\lim_{r\to\infty} \mu_{\mathcal{C}(z,r)} = \inf_{r\geq 0} \mu_{\mathcal{C}(z,r)} &\leq \inf_{r\geq s}\inf_{n\geq 1} \frac{a_n(r,s)}{n} = \inf_{n\geq 1}\inf_{r\geq s} \frac{a_n(r,s)}{n} \\ &= \inf_{n\geq 1}\lim_{r\to\infty} \frac{a_n(r,s)}{n} = \lim_{n\to\infty} \frac{\mathbf{E}[T(V_{\rho_1}, V_{\rho_{\nu(n)}})]}{n},\end{aligned}$$

where we in the second to last step have used that $T_{\mathcal{C}(z,r)}(\cdot,\cdot)$ is decreasing in r, and in the finial step have appealed to the monotone convergence theorem and used that $(a_n(\infty,s))_{n\geq 1}$ is subadditive. It remains to prove that the limit equals $\mu(z)$.

We will proceed by showing that $\frac{1}{n} T(V_{\rho_1}, V_{\rho_{\nu(n)}}) \to \mu(z)$ in probability as $n \to \infty$, and then argue that the limit carries over in the mean due to uniform integrability of the family $\{\frac{1}{n}T(V_{\rho_1}, V_{\rho_{\nu(n)}})\}_{n\geq 1}$. We begin proving convergence in probability. By subadditivity,

$$\left|T(V_{\rho_1}, V_{\rho_{\nu(n)}}) - T(0, nz)\right| \leq T(0, V_{\rho_{\nu(0)}}) + T(nz, V_{\rho_{\nu(n)}}).$$

Since the distributions of the dominating two terms are independent of n, the convergence of $\frac{1}{n}T(V_{\rho_1}, V_{\rho_{\nu(n)}})$ in probability to $\mu(z)$ follows from the convergence of $\frac{1}{n}T(0, nz)$ in (1).

The condition $\sup_{n\geq 1}\mathbf{E}[(\frac{1}{n}T(V_{\rho_1}, V_{\rho_{\nu(n)}}))^\alpha] < \infty$, for some $\alpha > 1$, is sufficient for uniform integrability. To see that this holds, observe that

$$T(V_{\rho_1}, V_{\rho_{\nu(n)}}) \leq T_{\mathcal{C}(z,s)}(V_{\rho_1}, V_{\rho_{n+1}}) \leq \sum_{j=2}^{n+1} T_{\mathcal{C}(z,s)}(V_{\rho_{j-1}}, V_{\rho_j}) + n\epsilon(s),$$

where we have used that $\nu(n) - 1 \leq n$. Thus, by convexity of the function x^2, we find that

$$\left(\frac{1}{n}T(V_{\rho_1}, V_{\rho_{\nu(n)}})\right)^2 \leq 2\left(\frac{1}{n}\sum_{j=2}^{n+1} T_{\mathcal{C}(z,s)}(V_{\rho_{j-1}}, V_{\rho_j})^2 + \epsilon(s)^2\right).$$

Since the terms in the sum are i.i.d., we obtain

$$\mathbf{E}\left[\left(\frac{1}{n}T(V_{\rho_1}, V_{\rho_{\nu(n)}})\right)^2\right] \leq 2\mathbf{E}\left[T_{\mathcal{C}(z,s)}(V_{\rho_0}, V_{\rho_1})^2\right] + 2\epsilon(s)^2,$$

which is finite and independent of n. Thus, $\{\frac{1}{n}T(V_{\rho_1}, V_{\rho_{\nu(n)}})\}_{n\geq 1}$ is uniformly integrable, and

$$\lim_{n\to\infty} \frac{\mathbf{E}[T(V_{\rho_1}, V_{\rho_{\nu(n)}})]}{n} = \mu(z),$$

as required. $\square$

5. Large deviations above the time constant. In this section, we estimate the probability of large deviations above the time constant and prove Theorem 4. Recall that it suffices to consider z in the first orthant, due to symmetry. The regenerative approach set up for in the previous section will serve to obtain a first modest estimate on the tail decay. This first step of the proof is as follows.

LEMMA 16. *Assume that $\mathbf{E}[Y^\alpha] < \infty$ for some $\alpha > 0$. There exists $R_2 = R_2(\alpha, d)$ such that for every $\varepsilon > 0$, $z \in \mathbf{N}^d$ and $r \geq R_2$, there is a finite constant $M_2 = M_2(\alpha, \varepsilon, d, z, r)$ such that for every $n \in \mathbf{N}$ and $x \geq n$*

$$\mathbf{P}\big(T_{\mathcal{C}(z,r)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|}) - n\mu_{\mathcal{C}(z,r)} > \varepsilon x\|z\|\big) \leq \frac{M_2}{x}.$$

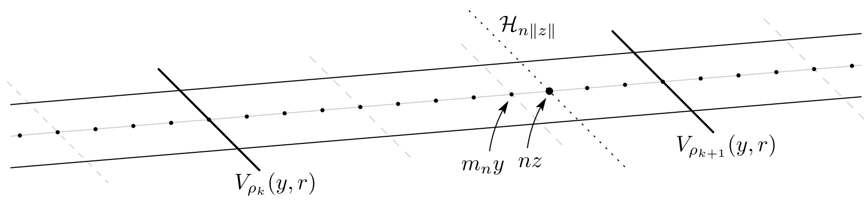

FIG. 3. *Dots indicate* $\{jz\}_{j\geq 0}$, *dashed diagonal lines* $\{V_j(y,r)\}_{j\geq 0}$, *and thick diagonal lines* $\{V_{\rho_j}(y,z)\}_{j\geq 0}$. *In the illustration we have* $y=4z$ *and* $\nu(m_n)=\nu(m_n,y,r)=k+1$.

PROOF. We may without loss of generality assume that $\alpha\in(0,2]$. Let $\beta=2$ and let $R_1=R_1(\alpha,d)$ be given as in Proposition 13. In particular, $\mu_{\mathcal{C}(z,r)}$ is finite for $r\geq R_1$. Fix $r\geq R_1$, $\varepsilon>0$ and choose $N\in\mathbf{N}$ large enough for $2\bar{t}|E_0(z,r)|\leq \varepsilon N\|z\|$ to hold. Set $y=Nz$ and let $m_n=\max\{m\geq 0: mN\leq n\}$. To the end of this proof, let $V_n=V_n(y,r)$, $E_n=E_n(y,r)$, $\rho_j=\rho_j(y,r)$, $\nu(n)=\nu(n,y,r)$, $\mu_\tau=\mu_\tau(y,r)$, and $\mu_\rho=\mu_\rho(y,r)$. (For a relation between nz, $m_n y$ and $\rho_{\nu(m_n)}y$, see Figure 3.) Recall that $\rho_0=0$, so $V_{\rho_0}(y,r)=V_0(z,r)$. By subadditivity,

$$\begin{aligned}T_{\mathcal{C}(z,r)}(\mathcal{H}_0,\mathcal{H}_{n\|z\|})-n\mu_{\mathcal{C}(z,r)} &\leq \sum_{j=1}^{\nu(m_n)}\bigl(T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}},V_{\rho_j})-\mu_\tau\bigr)\\ &\quad+T_{\mathcal{C}(z,r)}(V_{\rho_{\nu(m_n)}},\mathcal{H}_{n\|z\|})\\ &\quad+\bigl(\nu(m_n)\mu_\tau-n\mu_{\mathcal{C}(z,r)}\bigr)+\sum_{j=1}^{\nu(m_n)}\bar{t}|E_0|.\end{aligned}$$

Label the four terms on the right-hand side as X_1,X_2,X_3,X_4. Since that $\sum_{i\leq 4}X_i>4\varepsilon x\|z\|$ would imply that $X_i>\varepsilon x\|z\|$ for some $i=1,2,3,4$, and since $\varepsilon>0$ was arbitrary, it suffices to obtain a bound on $\mathbf{P}(X_i>\varepsilon x\|z\|)$ of the desired form, for each $i=1,2,3,4$ separately.

Starting from behind, since $\nu(m_n)\leq m_n+1\leq n/N+1$, it follows that for $n\geq N$, $\mathbf{P}(\nu(m_n)\bar{t}|E_0|>\varepsilon n\|z\|)=0$ by the choice of N. So, the last term satisfies a bound on the desired form. Since $\mu_{\mathcal{C}(y,r)}=N\mu_{\mathcal{C}(z,r)}$, the third term is via Lemma 14 bounded above by

$$\nu(m_n)\mu_\rho\mu_{\mathcal{C}(y,r)}-m_n\mu_{\mathcal{C}(y,r)}.$$

Recall that $\nu(m)-1$ counts the number of $k\in\{1,2,\dots,m\}$ for which $A_k(y,r)$ occurs, and is therefore binomially distributed with success probability $\mathbf{P}(A_n(y,r))=1/\mu_\rho$. For large n, we will have $\varepsilon x\|z\|/2>\mu_\rho\mu_{\mathcal{C}(y,r)}$. Thus, for large n Chebyshev's inequality may be applied to give

$$\mathbf{P}\biggl(\nu(m_n)\mu_\rho-m_n>\varepsilon\frac{x\|z\|}{\mu_{\mathcal{C}(y,r)}}\biggr)\leq 4\mu^2_{\mathcal{C}(y,r)}\frac{\mu_\rho-1}{\varepsilon^2 N\|z\|^2 x},$$

which also meets the requirement.

For $\beta = 2$ and $y = Nz$, let $M_1 = M_1(\alpha, d, y, r)$ be given as in Proposition 13. Since $m_n\|y\| \leq n\|z\| < \rho_{\nu(m_n)}\|y\|$, then $T_{\mathcal{C}(z,r)}(\mathcal{H}_{n\|z\|}, V_{\rho_{\nu(m_n)}}) \leq T_{\mathcal{C}(z,r)}(V_{m_n}, V_{\rho_{\nu(m_n)}})$, which is distributed as $T_{\mathcal{C}(z,r)}(V_{\rho_0}, V_{\rho_1})$. Consequently, Markov's inequality and Proposition 13 give that

$$\mathbf{P}\big(T_{\mathcal{C}(z,r)}(V_{\rho_{\nu(m_n)}}, \mathcal{H}_{n\|z\|}) > \varepsilon x\|z\|\big) \leq M_1 \frac{(1+\mathbf{E}[Y^\alpha])^{1+1/\alpha}}{\varepsilon\|z\|x},$$

for $r \geq R_1$. Finally, by Wald's lemma $\sum_{j=1}^{\nu(m_n)}(T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j}) - \mu_\tau)$ has mean zero and second moment

$$\mathbf{E}\Bigg[\Bigg(\sum_{j=1}^{\nu(m_n)}(T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j}) - \mu_\tau)\Bigg)^2\Bigg] = \mathrm{Var}\big(T_{\mathcal{C}(z,r)}(V_{\rho_0}, V_{\rho_1})\big)\mathbf{E}\big[\nu(m_n)\big].$$

Using Chebyshev's inequality, Proposition 13 and the identity $\mathbf{E}[\nu(m_n)] = 1 + m_n\mu_\rho^{-1}$ shows that $\mathbf{P}(\sum_{j=1}^{\nu(m_n)}(T_{\mathcal{C}(z,r)}(V_{\rho_{j-1}}, V_{\rho_j}) - \mu_\tau) > \varepsilon x\|z\|)$ is bounded above by

$$M_1 \frac{(1+\mathbf{E}[Y^\alpha])^{1+2/\alpha}(\mu_\rho^{-1}N^{-1} + x^{-1})}{\varepsilon^2\|z\|^2 x},$$

for all $r \geq R_1$, as required. $\square$

In the second step, we improve upon the above decay by aligning disjoint cylinders.

PROPOSITION 17. *Assume that* $\mathbf{E}[Y^\alpha] < \infty$ *for some* $\alpha > 0$. *For every* $\varepsilon > 0$, $q \geq 1$ *and* $z \in \mathbf{Z}^d$ *there exists* $M_3 = M_3(\varepsilon, \alpha, q, d, z)$ *such that for all* $n \in \mathbf{N}$ *and* $x \geq n$

$$\mathbf{P}\big(T(0, nz) - n\mu(z) > \varepsilon x\|z\|\big) \leq M_3\mathbf{P}(Y > x/M_3) + \frac{M_3}{x^q}.$$

PROOF. We may assume that z lies in the first orthant due to symmetry. Fix $\varepsilon > 0$, $q \in \mathbf{Z}_+$ and choose $r \geq R_2$ large enough for $\mu_{\mathcal{C}(z,r)} - \mu(z) \leq \varepsilon\|z\|$ to hold, where $R_2 = R_2(\alpha, d)$ is as in Lemma 16. Pick $v_1, v_2, \ldots, v_q \in \mathcal{H}_0$ such that the transposed cylinders $v_i + \mathcal{C}(z, r)$ are pairwise disjoint, and choose $s > r$ so that $v_i + \mathcal{C}(z, r) \subseteq \mathcal{C}(z, s)$ for all $i = 1, 2, \ldots, q$. For such s, the travel time $T_{\mathcal{C}(z,s)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|})$ is clearly dominated by the minimum of q independent random variables distributed as $T_{\mathcal{C}(z,r)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|})$. Thus,

$$\begin{aligned}&\mathbf{P}\big(T_{\mathcal{C}(z,s)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|}) - n\mu_{\mathcal{C}(z,r)} > \varepsilon x\|z\|\big)\\ &\quad\leq \mathbf{P}\big(T_{\mathcal{C}(z,r)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|}) - n\mu_{\mathcal{C}(z,r)} > \varepsilon x\|z\|\big)^q,\end{aligned}$$

which by Lemma 16 is at most M_2^q/x^q. Since $\mu_{\mathcal{C}(z,r)} - \mu(z) \le \varepsilon\|z\|$ by the choice of r, we obtain

$$\mathbf{P}\big(T_{\mathcal{C}(z,s)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|}) - n\mu(z) > 2\varepsilon x\|z\|\big) \le \frac{M_2^q}{x^q}. \tag{15}$$

It remains to connect to the starting and ending points 0 and nz. Subadditivity gives that

$$T(0, nz) \le \sum_{v \in V_0(z,s)} \big(T(0, v) + T(v + nz, nz)\big) + T_{\mathcal{C}(z,s)}(\mathcal{H}_0, \mathcal{H}_{n\|z\|}),$$

so we may via Lemma 10 find an M_3', depending on ε, d and s, such that

$$\mathbf{P}\big(T(0, nz) - n\mu(z) > 3\varepsilon x\|z\|\big) \le M_3'\mathbf{P}\big(Y > x/M_3'\big) + \frac{M_2^q}{x^q}.$$

Since $\varepsilon > 0$ was arbitrary, the proof is complete. $\square$

Before completing the proof of Theorem 4, we will show that travel times cannot be too large. This result will help us to loose the dependence on z still present in Proposition 17.

PROPOSITION 18. *Assume that* $\mathbf{E}[Y^\alpha] < \infty$ *for some* $\alpha > 0$. *For every* $q \ge 1$ *there is a constant* $M_4 = M_4(\alpha, q, d)$ *such that for all* $z \in \mathbf{Z}^d$ *and* $x \ge \|z\|$

$$\mathbf{P}\big(T(0, z) > M_4 x\big) \le M_4\mathbf{P}(Y > x) + \frac{1}{x^q}.$$

PROOF. Again, assume that z lies in the first orthant. Let M_3 denote the constant figuring in Proposition 17 (for given α and q, and with $\varepsilon = 1$ and $z = \mathbf{e}_1$). The point z can be reached from 0 in d steps by in each step taking z_i steps in direction $\mathbf{e}_i$, for $i = 1, 2, \dots, d$. Thus, due to subadditivity and Proposition 17,

$$\mathbf{P}\big(T(0, z) > dM_4'x\big) \le \sum_{i=1}^{d} \mathbf{P}\big(T(0, z_i\mathbf{e}_i) > M_4'x\big) \le dM_3\mathbf{P}(Y > x/M_3) + \frac{dM_3}{x^q}$$

for any $M_4' \ge \mu(\mathbf{e}_1) + 1$. Hence, $M_4 = d^2M_3M_4'$ is sufficient. $\square$

PROOF OF THEOREM 4. Fix $\varepsilon > 0$. To start, we will choose a finite set of directions so that each $z \in \mathbf{Z}^d$ will be within distance $\varepsilon\|z\|$ of some straight line intersecting the origin, and continues in one of the chosen directions. The set of directions can be chosen as $\{y \in \mathbf{Z}^d : \|y\| = N\}$, given that N is large enough. More precisely, pick $N \ge d/\varepsilon$ and note that any $z \in \mathbf{Z}^d$ satisfying $mN \le \|z\| < (m+1)N$ will be within ℓ^1-distance $N + dm$ of some point in $\{my : \|y\| = N\}$. In particular, for $\|z\| \ge N/\varepsilon$ we have

$$N + dm \le \varepsilon\|z\| + d\|z\|/N \le 2\varepsilon\|z\|.$$

Given $z \in \mathbf{Z}^d$, let $m_z := \max\{m \geq 0 : mN \leq \|z\|\}$. First, note that if $\|z\| \leq N/\varepsilon$, then

$$\mathbf{P}\big(T(0,z) - \mu(z) > \varepsilon x\big) \leq \mathbf{P}\big(T(0,z) > \varepsilon x\big) \leq 9^{2d}(N/\varepsilon)\mathbf{P}\big(Y > \varepsilon^2 x/(9N)\big),$$

by Lemma 10. So, we may proceed assuming that $\|z\| \geq N/\varepsilon$. For any y with $\|y\| = N$, we have

$$T(0,z) - \mu(z) \leq T(0, m_z y) - m_z\mu(y) + T(m_z y, z) + \big(m_z\mu(y) - \mu(z)\big),$$

and, for at least one of these y (the one closest to z), since μ satisfies the properties of a norm,

$$m_z\mu(y) - \mu(z) \leq \mu(m_z y - z) \leq \mu(\mathbf{e}_1)\|m_z y - z\| \leq 2\mu(\mathbf{e}_1)\varepsilon\|z\|. \tag{16}$$

Moreover, with M_4 as in Proposition 18, and $x \geq \|z\| \geq \|m_z y - z\|/(2\varepsilon)$,

$$\mathbf{P}\big(T(m_z y, z) > 2M_4\varepsilon x\big) \leq M_4\mathbf{P}(Y > 2\varepsilon x) + \frac{1}{(2\varepsilon x)^q}. \tag{17}$$

Since N depends on nothing but ε, there is a constant $M_3 = M_3(\varepsilon, \alpha, q, d)$, given by Proposition 17, such that for every y satisfying $\|y\| = N$, and $x \geq \|z\| \geq m_z N$, then

$$\mathbf{P}\big(T(0, m_z y) - m_z\mu(y) > \varepsilon x\big) \leq M_3\mathbf{P}\big(Y > x/(M_3 N)\big) + \frac{M_3 N^q}{x^q}. \tag{18}$$

Combining (16), (17) and (18), we conclude that also for $x \geq \|z\| \geq N/\varepsilon$,

$$\mathbf{P}\big(T(0,z) - \mu(z) > \big(1 + 2\mu(\mathbf{e}_1) + 2M_4\big)\varepsilon x\big) \leq M\mathbf{P}(Y > x/M) + \frac{M}{x^q},$$

where M can be taken as the maximum of $M_4 + M_3$, $1/(2\varepsilon)^q$, and $M_3 N^q$. Since $\varepsilon > 0$ was arbitrary, this completes the proof. $\square$

6. Proof of the Hsu–Robbins–Erdős strong law. Both Theorem 1 and Corollary 5 may be thought of as strong laws of the kind introduced by Hsu, Robbins and Erdős. They are easily derived in a similar fashion from the large deviation estimates presented in Theorems 3 and 4. For that reason, we only present a proof of the former.

PROOF OF THEOREM 1. Deviations below and above the time constant are easily handled separately via the identity

$$\begin{aligned}\mathbf{P}\big(\big|T(0,z) - \mu(z)\big| > \varepsilon\|z\|\big) &= \mathbf{P}\big(T(0,z) - \mu(z) < -\varepsilon\|z\|\big)\\ &\quad + \mathbf{P}\big(T(0,z) - \mu(z) > \varepsilon\|z\|\big).\end{aligned}$$

Summability of the probabilities of deviations below the time constant is immediate from Theorem 3, since $\mathbf{P}(T(0,z) - \mu(z) < -\varepsilon\|z\|)$ decays exponentially in $\|z\|$, while the number of sites satisfying $\|z\| = n$ grows polynomially in n.

Consider instead deviations above the time constant. Assume first that $\mathbf{E}[Y^\alpha] < \infty$ for some $\alpha > 0$. According to Theorem 4, with $q = \alpha + 1$, there is a constant $M = M(\alpha, \varepsilon, d)$ such that

$$\sum_{z\in\mathbf{Z}^d} \|z\|^{\alpha-d}\mathbf{P}\big(T(0,z) - \mu(z) > \varepsilon\|z\|\big)$$
$$\leq M \sum_{z\in\mathbf{Z}^d} \left(\|z\|^{\alpha-d}\mathbf{P}(Y > \|z\|/M) + \frac{1}{\|z\|^{d+1}}\right).$$

Observe that the number of $z \in \mathbf{Z}^d$ for which $\|z\| = n$ is of order n^{d-1}. The above summation is therefore finite since $\mathbf{E}[Y^\alpha] < \infty$ implies that

$$\sum_{n=1}^{\infty} n^{\alpha-1}\mathbf{P}(Y > n/M) + \sum_{n=1}^{\infty} n^{-2} < \infty.$$

For the necessity of $\mathbf{E}[Y^\alpha]$ being finite, note that $T(0,z)$ is at least as large as the minimum value among the $2d$ edges adjacent to z. For all large enough M, we therefore have

$$\sum_{z\in\mathbf{Z}^d} \|z\|^{\alpha-d}\mathbf{P}\big(T(0,z) - \mu(z) > \varepsilon\|z\|\big) \geq \sum_{z\in\mathbf{Z}^d} \|z\|^{\alpha-d}\mathbf{P}\big(Y > M\|z\|\big)$$
$$\geq \sum_{n=1}^{\infty} n^{\alpha-1}\mathbf{P}(Y > Mn),$$

which is finite only if $\mathbf{E}[Y^\alpha] < \infty$. $\square$

7. The sets of points in space and time of linear order deviations. In this section, we study the set of times t for which the random set of sites reachable within time t from the origin deviates by as much as a constant factor from the asymptotic shape. In particular, we will see how Theorem 1 and the estimates on large deviation above and below the time constants can be used to estimate moments of $|\mathscr{T}_\varepsilon|$ and prove Theorem 2. This will be done by estimating the contribution of each site $z \in \mathbf{Z}^d$ to $\mathscr{T}_\varepsilon$ separately.

Let $Y(z)$ denote the minimum of the $2d$ weights associated with the edges incident to z. Note that $Y(y)$ and $Y(z)$ are independent as soon as y and z are at ℓ^1-distance at least 2. Since $T(0,z)$ is at least as large as $Y(z)$, it is possible to obtain a sufficient condition for z to be contained in $\mathscr{Z}_\varepsilon$ in terms of $Y(z)$. Recall that either $\mu \equiv 0$ or μ is bounded away from 0 and infinity on compact sets not containing the origin. Consequently, $\bar{\mu} := \sup_{\|x\|=1} \mu(x)$ is finite. Thus, $Y(z) > (\bar{\mu} + \varepsilon)\|z\|$ implies that $z \in \mathscr{Z}_\varepsilon$, and

$$|\mathscr{Z}_\varepsilon| = \sum_{z\in\mathbf{Z}^d} \mathbf{1}_{\{|T(0,z)-\mu(z)|>\varepsilon\|z\|\}} \geq \sum_{z\in\mathbf{Z}^d} \mathbf{1}_{\{Y(z)>\beta\|z\|\}}, \tag{19}$$

for large enough $\beta = \beta(\varepsilon)$.

A similar estimate can be obtained for the Lebesgue measure of $\mathscr{T}_\varepsilon$ as well. Assume until this end that $\mu \not\equiv 0$, in which case $\underline{\mu} := \inf_{\|x\|=1} \mu(x)$ is strictly positive. For $t \geq 0$, let

$$A_t := \{z \in \mathbf{Z}^d : T(0, z) > t \text{ and } \mu(z) \leq t(1-\varepsilon)\},$$
$$B_t := \{z \in \mathbf{Z}^d : T(0, z) \leq t \text{ and } \mu(z) > t(1+\varepsilon)\}.$$

Note that $A_t \neq \varnothing$ is equivalent to $\mathscr{B}^\mu_{(1-\varepsilon)t} \not\subset \mathscr{B}_t$. Similarly, $B_t \neq \varnothing$ if and only if $\mathscr{B}_t \not\subset \mathscr{B}^\mu_{(1+\varepsilon)t}$. Thus, $\mathscr{T}_\varepsilon = \{t \geq 0 : A_t \cup B_t \neq \varnothing\}$, and the contribution of a site z is given by the interval of time for which z is contained in either A_t or B_t. Denote these intervals by $I_A(z)$ and $I_B(z)$, respectively, and note that

$$\mathscr{T}_\varepsilon = \bigcup_{z \in \mathbf{Z}^d} I_A(z) \cup I_B(z).$$

Crude but useful upper bounds on the length of $I_A(z)$ and $I_B(z)$ are given by $T(0, z)$ and $\mu(z)/(1+\varepsilon)$, respectively. More precisely, we have

$$\begin{aligned} |I_A(z)| &= \big(T(0, z) - \mu(z)/(1-\varepsilon)\big)\mathbf{1}_{\{I_A(z)\neq\varnothing\}} \\ &\leq \big(T(0, z) - \mu(z)\big)\mathbf{1}_{\{T(0,z)-\mu(z)>\beta\|z\|\}} \end{aligned} \tag{20}$$

and

$$\begin{aligned} |I_B(z)| &= \big(\mu(z)/(1+\varepsilon) - T(0, z)\big)\mathbf{1}_{\{I_B(z)\neq\varnothing\}} \\ &\leq \mu(z)\mathbf{1}_{\{T(0,z)-\mu(z)<-\beta\|z\|\}}, \end{aligned} \tag{21}$$

for any $\beta > 0$ not larger than $\varepsilon\underline{\mu}/(1+\varepsilon)$. Moreover, for $\beta > \bar{\mu}/(1-\varepsilon)$

$$|I_A(z)| \geq \big(Y(z) - \beta\|z\|\big)\mathbf{1}_{\{Y(z)>\beta\|z\|\}}. \tag{22}$$

A first consequence of these representations for $|\mathscr{Z}_\varepsilon|$ and $|\mathscr{T}_\varepsilon|$ is the following simple observation.

PROPOSITION 19. *If $\mathbf{E}[Y^d] = \infty$, then $|\mathscr{Z}_\varepsilon|$ and $|\mathscr{T}_\varepsilon|$ are infinite for all $\varepsilon > 0$, almost surely.*

PROOF. Recall that the $Y(z)$'s are independent for points at ℓ^1-distance at least 2. $\mathbf{E}[Y^d] = \infty$ implies that $\sum_{\|z\|\in 2\mathbf{N}} \mathbf{P}(Y(z) > \beta\|z\|) = \infty$ for every $\beta > 0$. Consequently, $Y(z) > \beta\|z\|$ for infinitely many $z \in \mathbf{Z}^d$ via the Borel–Cantelli lemma, almost surely, so $|\mathscr{Z}_\varepsilon|$ is almost surely infinite. The same argument shows that also $Y(z) > \beta\|z\| + 1$ for infinitely many $z \in \mathbf{Z}^d$, and hence (assuming $\mu \not\equiv 0$) $|\mathscr{T}_\varepsilon|$ is infinite almost surely. $\square$

On the other hand, that $\mathbf{E}[Y^d] < \infty$ is sufficient for the expected cardinality of $\mathscr{Z}_\varepsilon$ to be finite is immediate from Theorem 1. The first hint to why $\mathbf{E}[Y^{d+1}] < \infty$

is required in order for the expected Lebesgue measure of $\mathscr{T}_\varepsilon$ to be finite is that although the cardinality of $\mathscr{Z}_\varepsilon$ is finite, the furthest point may lie very far from the origin. For the furthest point to be expected within finite distance, it is necessary that $\mathbf{E}[Y^{d+1}] < \infty$. With a slight abuse of notation, we let $\sup \mathscr{Z}_\varepsilon$ denote the ℓ^1-distance to the furthest point in $\mathscr{Z}_\varepsilon$.

PROPOSITION 20. *For every $\alpha > 0$ and $\varepsilon > 0$,*

$$\mathbf{E}\big[Y^{d+\alpha}\big] < \infty \quad \iff \quad \mathbf{E}\big[(\sup \mathscr{Z}_\varepsilon)^\alpha\big] < \infty.$$

PROOF. Fix $\alpha > 0$ and $\varepsilon > 0$. The sufficiency of $\mathbf{E}[Y^{d+\alpha}] < \infty$ is immediate from Theorem 1, since

$$\mathbf{E}\big[(\sup \mathscr{Z}_\varepsilon)^\alpha\big] \le \sum_{z\in\mathbf{Z}^d} \|z\|^\alpha \mathbf{P}\big(\big|T(0,z) - \mu(z)\big| > \varepsilon\|z\|\big).$$

It remains to show that $\mathbf{E}[Y^{d+\alpha}] < \infty$ is also necessary. That $\mathbf{E}[Y^d] < \infty$ is necessary is a consequence of Proposition 19, so there is no restriction assuming that $\mathbf{E}[Y^d]$ is finite. In order to obtain a lower bound on $\sup \mathscr{Z}_\varepsilon$, we are, in contrast to the lower bound in (19), looking for the largest integer n such that there exists $z \in \mathbf{Z}^d$ for which $\|z\| = n$ and $Y(z) > (\bar{\mu} + \varepsilon)\|z\|$. Since $Y(z)$'s are independent for points at ℓ^1-distance at least 2, we restrict focus further to even values of n.

For $\beta > 0$, let

$$\eta_\beta := \big\{n \in 2\mathbf{N} : \exists z \in \mathbf{Z}^d \text{ for which } \|z\| = n \text{ and } Y(z) > \beta n\big\}.$$

For $\beta = \beta(\varepsilon)$ sufficiently large ($\beta \ge \bar{\mu} + \varepsilon$ will do), we have the lower bound

$$\begin{aligned}\mathbf{E}\big[(\sup \mathscr{Z}_\varepsilon)^\alpha\big] &\ge \sum_{n\in 2\mathbf{N}} n^\alpha \mathbf{P}(\sup \eta_\beta = n)\\ &= \sum_{n\in 2\mathbf{N}} n^\alpha \mathbf{P}\Big(\max_{\|z\|=n} Y(z) > \beta n\Big)\mathbf{P}(\sup \eta_\beta \le n),\end{aligned} \tag{23}$$

where the equality follows by independence.

It is well known that the probability of a binomially distributed random variable being strictly positive is comparable to its mean. Let X be binomially distributed with parameters n and p. The union bound shows that its mean np is an upper bound on $\mathbf{P}(X > 0)$, but an application of Cauchy–Schwarz's inequality gives as well the lower bound $\mathbf{E}[X]^2/\mathbf{E}[X^2] \ge np/(1+np)$, which if np is small compared to 1, is at least $np/2$.

Fix $\beta = \beta(\varepsilon)$ such that (23) holds. Let X_n denote the number of z for which $\|z\| = n$ and $Y(z) > \beta n$. Since $Y(z)$'s are independent for different points at the same ℓ^1-distance, X_n is binomial. The number of points at distance n from the origin are of order n^{d-1}, and (since $\mathbf{E}[Y^d] < \infty$ is assumed) $\mathbf{P}(Y > \beta n)$ decays at least as n^{-d} via Markov's inequality. Consequently, for n large $\mathbf{E}[X_n] \leq 1$, and there is $\delta > 0$ such that

$$\mathbf{P}\Big(\max_{\|z\|=n} Y(z) > \beta n\Big) \geq \mathbf{E}[X_n]/2 \geq \delta n^{d-1}\mathbf{P}(Y > \beta n). \tag{24}$$

We next claim that $\mathbf{P}(\sup \eta_\beta \leq n) \geq 1/2$ for large n. Via the lower bound in (23), we conclude that for some $\delta > 0$ and $N < \infty$

$$\mathbf{E}\big[(\sup \mathscr{X}_\varepsilon)^\alpha\big] \geq \frac{\delta}{2} \sum_{n\in 2\mathbf{N}:\, n\geq N} n^{d+\alpha-1}\mathbf{P}(Y > \beta n),$$

for which $\mathbf{E}[Y^{d+\alpha}] < \infty$ is necessary in order to be finite.

It remains to show that $\mathbf{P}(\sup \eta_\beta \leq n) \geq \frac{1}{2}$ for large enough n. Note that

$$\mathbf{P}(\sup \eta_\beta > n) \leq \sum_{k>n} \mathbf{P}\Big(\max_{\|z\|=k} Y(z) > \beta k\Big) \leq C \sum_{k>n} k^{d-1}\mathbf{P}(Y > \beta k),$$

for some $C < \infty$. Since $\mathbf{E}[Y^d]$ is assumed finite, the right-hand side is summable, and becomes arbitrarily small as n increases. This proves the claim. $\square$

What remains is to prove Theorem 2. The proof will be similar to that of Proposition 20, but will require a couple of additional estimates. The difference is a consequence of the difference in the upper and lower bounds between (19) and (20)–(22). As such, we will encounter moments of products on the form $X \cdot \mathbf{1}_{\{X>a\}}$. For every $\alpha > 0$, $a \geq 0$ and random variable X, we have the following formula:

$$\mathbf{E}\big[X^\alpha \cdot \mathbf{1}_{\{X>a\}}\big] = a^\alpha \mathbf{P}(X > a) + \alpha \int_a^\infty x^{\alpha-1}\mathbf{P}(X > x)\,dx. \tag{25}$$

Since it will be used more than once, we separate the following bound on a double summation as a lemma.

LEMMA 21. *For every $\alpha \geq 0$ and $\beta \geq 0$, there are $c = c(\alpha,\beta)$ and $C = C(\alpha,\beta)$ such that*

$$\begin{aligned} c\sum_{n=1}^\infty (n-1)^{\alpha+\beta+1}\mathbf{P}(X > n) &\leq \sum_{n=1}^\infty n^\alpha \int_n^\infty x^\beta \mathbf{P}(X > x)\,dx \\ &\leq C\sum_{n=1}^\infty (n+1)^{\alpha+\beta+1}\mathbf{P}(X > n). \end{aligned}$$

PROOF. Split the integration domain into unit intervals and bound the integrand from below and above. The lower and upper bonds follow via the estimate

$$(m/2)^{\alpha+1} \leq \sum_{n=1}^m n^\alpha \leq m^{\alpha+1}.$$

$\square$

PROOF OF THEOREM 2. Fix $\alpha > 0$ and $\varepsilon > 0$. We will prove the implications, one by one, between the three expressions: (a) $\mathbf{E}[Y^{d+\alpha}] < \infty$; (b) $\mathbf{E}[(\sup \mathscr{T}_\varepsilon)^\alpha] < \infty$; and (c) $\mathbf{E}[|\mathscr{T}_\varepsilon|^\alpha] < \infty$, starting with the following:

(a) $\Rightarrow$ (b) Since $\sup \mathscr{T}_\varepsilon = \sup\{\max I_A(z) \cup I_B(z) : z \in \mathbf{Z}^d\}$, we obtain in similarity to (20) and (21) the upper bound

$$\begin{aligned}\mathbf{E}\big[(\sup \mathscr{T}_\varepsilon)^\alpha\big] &\le \sum_{z\in\mathbf{Z}^d} \mathbf{E}\big[\big(\max I_A(z) \cup I_B(z)\big)^\alpha\big] \\ &\le \sum_{z\in\mathbf{Z}^d} \mathbf{E}\big[T(0,z)^\alpha \mathbf{1}_{\{T(0,z)-\mu(z)>\beta\|z\|\}}\big] \\ &\quad + \sum_{z\in\mathbf{Z}^d} \mu(z)^\alpha \mathbf{P}\big(T(0,z)-\mu(z) < -\beta\|z\|\big),\end{aligned}$$

valid for all sufficiently small $\beta = \beta(\varepsilon) > 0$. The latter sum in the right-hand side is finite as of Theorem 1. The former sum takes via (25) the form

$$\begin{aligned}&\sum_{z\in\mathbf{Z}^d} \big(\mu(z)+\beta\|z\|\big)^\alpha\, \mathbf{P}\big(T(0,z)-\mu(z) > \beta\|z\|\big) \\ &\quad + \sum_{z\in\mathbf{Z}^d} \alpha \int_{\mu(z)+\beta\|z\|}^\infty x^{\alpha-1}\, \mathbf{P}\big(T(0,z) > x\big)\, dx.\end{aligned}$$

The former of these two sums is again finite according to Theorem 1. Once the latter sum is broken up into two, one over $n \in \mathbf{N}$ and the other over $\|z\| = n$, and the integral has gone through a chance of variables $x \mapsto x - \mu(z)$, then Theorem 4 can be used to relate the probability tail of $T(0,z) - \mu(z)$ with that of Y. Since the number of points at distance n from the origin is of order n^{d-1}, an upper bound is given by

$$\alpha C M \sum_{n=1}^\infty n^{d-1} \int_{\beta n}^\infty x^{\alpha-1}\Big(\mathbf{P}(Y > x/M) + \frac{1}{x^{d+\alpha+1}}\Big)\, dx,$$

for some finite constants C and $M = M(\alpha, \beta, d)$. Integrating over the two terms separately breaks the sum in two, of which the latter is easily seen to be finite. The former can instead be estimated via Lemma 21. An upper bound on this part is obtained as

$$M' \sum_{n=1}^\infty (n+1)^{d+\alpha-1} \mathbf{P}(Y > \beta n/M),$$

for some constant $M' = M'(\alpha, \beta, d)$, which is finite when $\mathbf{E}[Y^{d+\alpha}] < \infty$.

(b) $\Rightarrow$ (c) This step is trivial.

(c) $\Rightarrow$ (a) A lower bound on $|\mathscr{T}_\varepsilon|$ is given by the contribution of a single site z. However, looking at a particular site is not going to give a bound of the right order.

Instead we will pick a site randomly, and more precisely, the site furthest from the origin among those contributing to $\mathscr{T}_\varepsilon$. In case this site is not unique, then we pick the one contributing more. The contribution of each site z was in (22) seen to be at least $(Y(z) - \beta\|z\|)\mathbf{1}_{\{Y(z)>\beta\|z\|\}}$, for every sufficiently large $\beta = \beta(\varepsilon)$. Similar to that of (23), we have

$$\mathbf{E}\big[|\mathscr{T}_\varepsilon|^\alpha\big] \ge \sum_{n\in 2\mathbf{N}} \mathbf{E}\Big[\Big(\max_{\|z\|=n} Y(z) - \beta n\Big)^\alpha \mathbf{1}_{\{\max_{\|z\|=n} Y(z)>\beta n\}}\Big]\mathbf{P}(\sup \eta_\beta \le n).$$

Here, like in the proof of Proposition 20, we sum over even integers for the sake of independence between the events $\{\sup \eta_\beta = n\}$ and $\{\sup \eta_\beta \le n\}$. Combining the identity (25) and the bound (24) with a change of variables, we arrive at

$$\mathbf{E}\big[|\mathscr{T}_\varepsilon|^\alpha\big] \ge \alpha\delta \sum_{n\in 2\mathbf{N}} n^{d-1}\mathbf{P}(\sup \eta_\beta \le n) \int_{\beta n}^\infty (x - \beta n)^{\alpha-1}\mathbf{P}\big(Y(z) > x\big)\,dx.$$

Again, since $\mathbf{P}(\sup \eta_\beta \le n) \ge 1/2$ for n large enough, we may via the double summation estimate in Lemma 21 conclude that $\mathbf{E}[Y^{d+\alpha}] < \infty$ is necessary for $\mathbf{E}[|\mathscr{T}_\varepsilon|^\alpha]$ to be finite. □

APPENDIX A: CONVERGENCE TOWARD THE TIME CONSTANT

The time constant was in (1) defined for $z \in \mathbf{Z}^d$ as the limit in probability of $\frac{1}{n}T(0, nz)$ as $n \to \infty$. Existence of the limit (almost surely and in L^1) under the assumption $\mathbf{E}[Y] < \infty$ follows from a straightforward application of the subadditive ergodic theorem [18]. Existence of the limit (in probability) without a moment condition was later derived in [8, 16]. There is a unique extension of μ to all of $\mathbf{R}^d$ that retains the properties of a semi-norm. For example, we may define $\mu(x)$ for $x \in \mathbf{R}^d$ via the limit

$$\mu(x) := \lim_{n\to\infty} \frac{\mu(v_n)}{n},$$

where $v_1, v_2, \dots$ is any sequence in $\mathbf{Z}^d$ such that $v_n/n \to x$ as $n \to \infty$.

Existence of this limit is well known and follows from the properties of μ as a semi-norm on $\mathbf{Z}^d$. That these properties are preserved in the limit is similarly verified. We would here like to emphasize, perhaps, a less-known fact, which is easily seen to follow the results reported in this paper. Namely, the necessary and sufficient condition under which $\mu(x)$, for $x \in \mathbf{R}^d$, appears as the almost sure limit for some sequence of travel times. We are not aware of such a condition previously appearing in the literature.

PROPOSITION 22. *Fix $x \in \mathbf{R}^d$ and let $v_1, v_2, \dots$ be any sequence of points in $\mathbf{Z}^d$ such that $v_n/n \to x$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} \frac{T(0, v_n)}{n} = \mu(x) \qquad \textit{in probability}.$$

Moreover, the limit holds almost surely, in L^1 and completely if and only if $\mathbf{E}[Y] < \infty$.

PROOF. A first observation due to the triangle inequality is that

$$\left|T(0, v_n) - n\mu(x)\right| \leq \left|T(0, v_n) - \mu(v_n)\right| + \left|\mu(v_n) - n\mu(x)\right|.$$

Let $\varepsilon > 0$. By the properties of μ as a norm, the latter term in the right-hand side is bounded above by $\mu(\mathbf{e}_1)\|v_n - nx\|$, which is at most εn when n is large. It follows that

$$\limsup_{n\to\infty} \mathbf{P}\big(\left|T(0, v_n) - n\mu(x)\right| > 2\varepsilon n\big) \leq \limsup_{n\to\infty} \mathbf{P}\big(\left|T(0, v_n) - \mu(v_n)\right| > \varepsilon n\big),$$

which by (2) has to equal zero. This proves convergence in probability.

Necessity of $\mathbf{E}[Y] < \infty$ for almost sure and L^1-convergence follows as before, due to the fact that a lower bound on the travel time from the origin to any other point is bounded from below by the minimum of the $2d$ weights associated with the edges adjacent to the origin. To conclude that $\mathbf{E}[Y] < \infty$ is sufficient for the convergence to hold almost surely it suffices to note that the sequence $(nz)_{n\geq 1}$ in Corollary 5 can be exchanged for any sequence $(v_n)_{n\geq 1}$ for which $\|v_n\|/n$ is bounded away from 0 and ∞. In particular, by Theorems 3 and 4 (with $\alpha = 1$ and $q = 2$) there exists M (depending on $\varepsilon > 0$ and the upper bound on $\|v_n\|/n$) such that

$$\sum_{n=1}^{\infty} \mathbf{P}\big(\left|T(0, v_n) - \mu(v_n)\right| > \varepsilon n\big) \leq M \sum_{n=1}^{\infty} \left(\mathbf{P}(Y > n/M) + \frac{1}{n^2}\right),$$

which is finite since $\mathbf{E}[Y] < \infty$. This proves almost sure and complete convergence. Finally, L^1-convergence is due to Corollary 6, again since $\|v_n\|/n$ is assumed bounded. $\square$

APPENDIX B: KESTEN'S CONSTRUCTION OF SHELLS

For completeness, let us present a precise construction of the shells in Section 2.2. The construction will follow closely that of Kesten [16], Section 2.

Given $\delta > 0$, pick $\bar{t} = \bar{t}(\delta)$ such that $\mathbf{P}(\tau_e \leq \bar{t}) \geq 1 - \delta$. As before, color each vertex in $\mathbf{Z}^d$ either black or white; black if at least one of the edges adjacent to it has weight larger than $\bar{t}$, and white otherwise. We shall below introduce a notion of black and white clusters, for which we will need to distinguish paths from $\star$-paths. A *path* will refer to a sequence of vertices $v_0, v_1, \dots, v_n$ of the $\mathbf{Z}^d$ lattice such that two consecutive points are at ℓ^1-distance one. A $\star$*-path* will similarly refer to a nearest-neighbor sequence of vertices with respect to ℓ^∞-distance on $\mathbf{Z}^d$. A path or a $\star$-path will be called *black* or *white* if all its points are black or white, respectively.

Given $A \subset \mathbf{Z}^d$, define the black and white clusters of A as

$$C(A, b) := A \cup \{z \in \mathbf{Z}^d : z \leftrightarrow y \text{ by a black } \star\text{-path}, \\ \text{for some } y \text{ at } \ell^\infty\text{-distance 1 from } A\},$$

$$C(A, w) := A \cup \{z \in \mathbf{Z}^d : z \leftrightarrow y \text{ by a white path}, \\ \text{for some } y \text{ at } \ell^1\text{-distance 1 from } A\}.$$

The exterior boundary $\partial_{\text{ext}} C$ of a set $C \subset \mathbf{Z}^d$ is defined as the set of points $z \in \mathbf{Z}^d \setminus C$ for which there is a point $y \in C$ at ℓ^∞-distance 1 from z, and for which there is a path connecting z to infinity without intersecting C. Next, let $D_n(z)$ denote the box of side-length $2n + 1$ centered at z, and let

$$n(z) := \min\{n \geq 0 : |C(y, w)| = \infty \text{ for some } y \in D_n(z)\}.$$

Let $S_z := \partial_{\text{ext}} C(D_{n(z)}(z), b)$. By construction, all vertices in S_z are white, and S_z must contain a white vertex belonging to an infinite white cluster. Kesten continues with two lemmas.

LEMMA 23 ([16], Lemma 2.23). *The exterior boundary of any finite $\star$-connected set is connected. In particular, if $C(D_{n(z)}(z), b)$ is finite, then S_z is connected. Moreover, in that case S_z separates z from infinity in the sense that every path from z to infinity has to intersect S_z.*

LEMMA 24 ([16], Lemma 2.24). *If $\delta > 0$ is sufficiently small, then the set $C(D_n(z), b)$ is almost surely finite for every $n \geq 0$. Moreover, there are constants $M < \infty$ and $\gamma > 0$ such that for every $k \geq 0$*

$$\mathbf{P}(n(z) > k) \leq M e^{-\gamma k} \quad \textit{and} \quad \mathbf{P}(\operatorname{diam}(S_z) > k) \leq M e^{-\gamma k}.$$

The first lemma distinguishes an "inside" of S_z, consisting of points separated from infinity by S_z. Finally, let Δ_z denote the union of S_z and each point on the inside of S_z which is white and connected to S_z by a white path.

Since each point in S_z is white, it follows that Δ_z is all white. That Δ_z almost surely satisfies the first three properties stated in Section 2.2 is a consequence of the above two lemmas. The final property follows from the following lemma. The lemma is a slight variant of (2.30) in [16], and is proved similarly. For completeness, we present an argument also here.

LEMMA 25. *Either every path between y and z in $\mathbf{Z}^d$ intersects both Δ_y and Δ_z, or $\Delta_y \cap \Delta_z \neq \varnothing$.*

PROOF. Assume that there is a path γ connecting y and z which does not intersect Δ_y. We will show that this implies that $\Delta_y \cap \Delta_z \neq \varnothing$. Let Γ be a path from z to infinity. Γ must intersect both S_y and S_z, since there would otherwise be a path from y to infinity (the concatenation of γ and Γ) not intersecting S_y or a path from z to infinity (Γ) not intersecting S_z, thus contradicting Lemma 23. Let v denote the first point in $S_y \cup S_z$ visited by Γ. We claim that either $v \in S_y \cap S_z$, or v is contained in just one of them, but connected to the other by a white path. The claim, if true, would imply that $v \in \Delta_y \cap \Delta_z$, since v would either be contained in S_y, or contained in its interior but connected to S_y by a white path.

It remains to prove the claim. Assume that $v \in S_z$. We will next construct a white path from v to S_y. Since $D_{n(z)}(z)$ contains a vertex in an infinite white cluster, there has to be a vertex $u \in S_z$ that is connected to infinity by a white path. A white path connecting v to infinity may now be obtained by first connecting v to u within S_z, and thereafter connect u to infinity. This path will necessarily intersect S_y, since we otherwise would have found a path from y to infinity avoiding S_y. The remaining case when $v \in S_y$ is analogous. $\square$

Acknowledgement. The author is grateful to Simon Griffiths for a helpful discussion.

REFERENCES

[1] AHLBERG, D. (2015). Asymptotics of first-passage percolation on one-dimensional graphs. *Adv. in Appl. Probab.* **47** 182–209. MR3327321

[2] AHLBERG, D. (2015). Convergence towards an asymptotic shape in first-passage percolation on cone-like subgraphs of the integer lattice. *J. Theoret. Probab.* **28** 198–222. MR3320965

[3] AHLBERG, D. (2008). Asymptotics of first-passage percolation on 1-dimensional graphs. Licentiate thesis, Univ. Gothenburg.

[4] ANTAL, P. and PISZTORA, A. (1996). On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24** 1036–1048. MR1404543

[5] BENJAMINI, I., KALAI, G. and SCHRAMM, O. (2003). First passage percolation has sublinear distance variance. *Ann. Probab.* **31** 1970–1978. MR2016607

[6] CHAYES, J. T. and CHAYES, L. (1986). Percolation and random media. In *Phénomènes Critiques, Systèmes Aléatoires, Théories de Jauge, Part I, II* (*Les Houches*, 1984) 1001–1142. North-Holland, Amsterdam. MR0880545

[7] CHOW, Y. and ZHANG, Y. (2003). Large deviations in first-passage percolation. *Ann. Appl. Probab.* **13** 1601–1614. MR2023891

[8] COX, J. T. and DURRETT, R. (1981). Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.* **9** 583–603. MR0624685

[9] CRANSTON, M., GAUTHIER, D. and MOUNTFORD, T. S. (2009). On large deviation regimes for random media models. *Ann. Appl. Probab.* **19** 826–862. MR2521889

[10] ERDÖS, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Stat.* **20** 286–291. MR0030714

[11] ERDÖS, P. (1950). Remark on my paper "On a theorem of Hsu and Robbins." *Ann. Math. Stat.* **21** 138. MR0032970

[12] GARET, O. and MARCHAND, R. (2007). Large deviations for the chemical distance in supercritical Bernoulli percolation. *Ann. Probab.* **35** 833–866. MR2319709

[13] GRIMMETT, G. and KESTEN, H. (1984). First-passage percolation, network flows and electrical resistances. *Z. Wahrsch. Verw. Gebiete* **66** 335–366. MR0751574

[14] HAMMERSLEY, J. M. and WELSH, D. J. A. (1965). First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif.* 61–110. Springer, New York. MR0198576

[15] HSU, P. L. and ROBBINS, H. (1947). Complete convergence and the law of large numbers. *Proc. Natl. Acad. Sci. USA* **33** 25–31. MR0019852

[16] KESTEN, H. (1986). Aspects of first passage percolation. In *École d'Été de Probabilités de Saint-Flour, XIV—1984. Lecture Notes in Math.* **1180** 125–264. Springer, Berlin. MR0876084

[17] KESTEN, H. (1993). On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** 296–338. MR1221154

[18] KINGMAN, J. F. C. (1968). The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B* **30** 499–510. MR0254907

[19] RICHARDSON, D. (1973). Random growth in a tessellation. *Proc. Cambridge Philos. Soc.* **74** 515–528. MR0329079

[20] TALAGRAND, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* **81** 73–205. MR1361756

[21] ZHANG, Y. (2010). On the concentration and the convergence rate with a moment condition in first passage percolation. *Stochastic Process. Appl.* **120** 1317–1341. MR2639748

INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA
ESTRADA DONA CASTORINA 110
22460-320 RIO DE JANEIRO
BRASIL
E-MAIL: ahlberg@impa.br