TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 377, Number 3, March 2024, Pages 1641–1670
https://doi.org/10.1090/tran/9035
Article electronically published on December 22, 2023

ON THE NUMBER AND SIZE OF HOLES IN THE GROWING BALL OF FIRST-PASSAGE PERCOLATION

MICHAEL DAMRON, JULIAN GOLD, WAI-KIT LAM, AND XIAO SHEN

Abstract. First-passage percolation is a random growth model defined on $\mathbb{Z}^d$ using i.i.d. nonnegative weights (τ_e) on the edges. Letting $T(x,y)$ be the distance between vertices x and y induced by the weights, we study the random ball of radius t centered at the origin, $\mathbf{B}(t) = \{x \in \mathbb{Z}^d : T(0,x) \leq t\}$. It is known that for all such τ_e, the number of vertices (volume) of $\mathbf{B}(t)$ is at least order t^d, and under mild conditions on τ_e, this volume grows like a deterministic constant times t^d. Defining a hole in $\mathbf{B}(t)$ to be a bounded component of the complement $\mathbf{B}(t)^c$, we prove that if τ_e is not deterministic, then a.s., for all large t, $\mathbf{B}(t)$ has at least ct^{d-1} many holes, and the maximal volume of any hole is at least $c\log t$. Conditionally on the (unproved) uniform curvature assumption, we prove that a.s., for all large t, the number of holes is at most $(\log t)^C t^{d-1}$, and for $d = 2$, no hole in $\mathbf{B}(t)$ has volume larger than $(\log t)^C$. Without curvature, we show that no hole has volume larger than $Ct \log t$.

1. Introduction

1.1. **Backgound and definitions.** In the '60s, Hammersley-Welsh introduced first-passage percolation (FPP) on the cubic lattice $\mathbb{Z}^d$ as model for fluid flow in a porous medium. FPP is now often viewed in other ways: as a random growth model, a particle system, or a random metric space; see [1, 9] for recent surveys. In addition to the usual questions, like passage time asymptotics, the geometry of geodesics, and concentration bounds, attention has recently been paid to the boundary of the growing set $\mathbf{B}(t)$ [3, 5, 11] and its topological properties [12]. The purpose of the current paper is to continue some of these newer questions, addressing the number and size of holes in $\mathbf{B}(t)$.

Consider $\mathbb{Z}^d$, the d-dimensional integer lattice with nearest-neighbor edges $\mathbb{E}^d$. Let $(\tau_e)_{e\in\mathbb{E}^d}$ be an i.i.d. family of nonnegative random variables (the edge-weights) assigned to the edges. A path from a vertex x to a vertex y is an alternating sequence $x = x_0, e_0, x_1, e_1, \ldots, x_{n-1}, e_{n-1}, x_n = y$ of vertices and edges such that

Received by the editors May 20, 2022, and, in revised form, January 24, 2023.
2020 *Mathematics Subject Classification.* Primary 60K35.
The research of the first author was supported by an NSF grant DMS-2054559 and an NSF CAREER award.
The research of the second author was supported by NSF Postdoctoral Research Fellowship DMS-1803622 while at Northwestern. The second author gratefully acknowledges financial support from the Schmidt DataX Fund at Princeton University made possible through a major gift from the Schmidt Futures Foundation.
The research of the third author was supported by the National Science and Technology Council in Taiwan Grant 110-2115-M002-012-MY3 and NTU New Faculty Founding Research Grant NTU-111L7452.

$e_i = \{x_i, x_{i+1}\} \in \mathbb{E}^d$ for all $i = 0, \ldots, n-1$. The passage time of a path γ is

$$T(\gamma) = \sum_{i=0}^{n-1} \tau_{e_i},$$

and the first-passage time from x to y is

$$T(x, y) = \inf_{\gamma : x \to y} T(\gamma),$$

where the infimum is over all paths γ from x to y.

We study the random "ball"

$$\mathbf{B}(t) = \{x \in \mathbb{Z}^d : T(0, x) \le t\} \text{ for } t \ge 0.$$

The shape theorem of FPP gives a type of law of large numbers for $\mathbf{B}(t)$, and states that the rescaled set $\mathbf{B}(t)/t$ converges to a deterministic limiting shape as $t \to \infty$. The usual assumptions are that

$$\mathbb{P}(\tau_e = 0) < p_c(d), \tag{1.1}$$

where $p_c(d)$ is the critical value for d-dimensional Bernoulli bond percolation (a constant known to be in the open interval $(0,1)$ for $d \ge 2$ and to be equal to $1/2$ for $d = 2$), and $\mathbb{E} \min\{\tau_1^d, \ldots, \tau_{2d}^d\} < \infty$, where the τ_i are i.i.d. copies of τ_e. Under these conditions [10, Thm. 1.7], there exists a nonrandom convex set $\mathcal{B}$ which is invariant under permutations of the coordinates and under reflections in the coordinate hyperplanes, has nonempty interior, and which is compact, such that for all $\epsilon > 0$,

$$\mathbb{P}\left((1-\epsilon)\mathcal{B} \subset \frac{1}{t}\widetilde{\mathbf{B}}(t) \subset (1+\epsilon)\mathcal{B} \text{ for all large } t \right) = 1. \tag{1.2}$$

Here, $\widetilde{\mathbf{B}}(t)$ is the sum set $\{x + y : x \in \mathbf{B}(t), y \in [0,1)^d\}$. This $\mathcal{B}$ is the unit ball of a norm g on $\mathbb{R}^d$:

$$\mathcal{B} = \{z \in \mathbb{R}^d : g(z) \le 1\}.$$

Hence, in the limit, the set $\mathbf{B}(t)$ has no holes, but holes may be present for finite t.

Only assuming (1.1), Kesten's lemma [10, Lem. 5.8] implies that $\mathbf{B}(t)$ is a.s. finite for each $t \ge 0$ and so the complement $\mathbf{B}(t)^c$ is a union of finitely many connected components. All but one of these components is finite. We then define the number of "holes" in $\mathbf{B}(t)$ as

$$N(t) = \text{number of finite connected components of } \mathbf{B}(t)^c$$

and the volume of the largest hole as

$$M(t) = \max\left\{\#\mathbf{S} : \mathbf{S} \text{ is a finite connected component of } \mathbf{B}(t)^c\right\}.$$

If τ_e is deterministic, then $N(t) = 0$ for all t, so we will assume

$$\text{the distribution of } \tau_e \text{ is non-trivial.} \tag{1.3}$$

In other words, the support of the distribution of τ_e contains at least two points.

1.2. **Main results.** Our results give upper and lower bounds on $N(t)$ and $M(t)$ under some conditions on the weights (τ_e). First are the lower bounds.

Theorem 1.1. *Suppose* (1.1) *and* (1.3) *hold.*

(1) *There exists* $c > 0$ *such that*

$$\mathbb{P}\left(M(t) \geq c \log t \text{ for all large } t\right) = 1.$$

(2) *There exists* $c > 0$ *such that*

$$\mathbb{P}\left(N(t) \geq c t^{d-1} \text{ for all large } t\right) = 1.$$

The proof of Theorem 1.1 appears in Section 2. Close inspection of the proof reveals that a.s., for all large t, the number of holes of $\mathbf{B}(t)$ of volume at least $c \log t$ is at least $t^{d-1-\alpha}$ for some α which satisfies $\alpha(c) \to 0$ as $c \to 0$.

The authors of [12] study the Betti numbers associated with the growing set in the Eden model, a simple model for cell growth. Their results give asymptotics for these numbers and, in particular, show that with high probability, the number of holes at time t is the same order as the perimeter, which is at least t^{d-1}. The same bound therefore holds for a site-FPP model with exponential weights, because it is equivalent, through the memoryless property of exponentials, to the Eden model. Our proof of item 2 of Theorem 1.1 has a similar structure to theirs. For large t, we condition on $\mathbf{B}(t)$ and find order t^{d-1} many disjoint sets in $\mathbf{B}(t)^c$ near the boundary of $\mathbf{B}(t)$. Each such set has a positive probability to contain a special configuration that will develop into a hole in $\mathbf{B}(t + C)$ for a constant $C > 0$. Because we cannot use the memoryless property, finding and constructing these holes is more complicated. First, if the weights are bounded, we cannot just create a hole by increasing the weights of the $2d$ edges incident to a particular vertex in $\mathbf{B}(t)^c$. Instead, in step 1 of the proof, we must define a more detailed high-weight event that ensures the existence of holes. Second, if the weights are unbounded, high-weight boundary edges may prevent $\mathbf{B}(t+C)$ from enveloping our high-weight configurations outside $\mathbf{B}(t)$ in constant time. We must therefore show in step 2 that for large t, the boundary of $\mathbf{B}(t)$ contains many sections of low-weight edges that are near large areas in $\mathbf{B}(t)^c$.

Remark 1.2. Holes in $\mathbf{B}(t)$ were also previously studied in the proof of lower bounds on the size of the edge boundary of $\mathbf{B}(t)$ in [5, Thm. 1.3]. Their argument involves constructing order $t^d(1 - F_Y(t))$ many unit-size holes in $\mathbf{B}(t)$, where F_Y is the distribution function of $\min\{\tau_1, \ldots, \tau_{2d}\}$ and the τ_i are i.i.d. copies of τ_e. These holes arise from isolated vertices all of whose incident edges have high weight. When τ_e has a heavy tail, this number can be made arbitrarily close to t^d. The strategy from [5] does not obviously extend to lighter-tailed distributions, and the holes built in the proof of Theorem 1.1 above arise instead from large regions of slightly large edge-weights.

Remark 1.3. As mentioned, if we remove assumption (1.3), we obtain $N(t) = 0$ for all t. Regarding assumption (1.1), if $F(0) > p_c$ but (1.3) holds (that is, τ_e is not identically zero), then there is an infinite component of zero-weight edges a.s., and $\#\mathbf{B}(t) = \infty$ for all large t. In addition, we have $N(t) = M(t) = \infty$ for all such t a.s. On the other hand, the situation when $F(0) = p_c$ is more complicated because the growth rate of $\mathbf{B}(t)$ can depend on the distribution of τ_e [8]. For some τ_e, we still have $\#\mathbf{B}(t) = \infty$ for all large t (and so $N(t) = M(t) = \infty$), but for others, $\#\mathbf{B}(t) < \infty$ for all t a.s. Our proof of Theorem 1.1 can be used for $d = 2$ to give a (probably nonoptimal) lower bound for $N(t)$ in terms of the growth rate of $\mathbf{B}(t)$. For $d \geq 3$, there is not currently a simple condition on τ_e to determine if $\#\mathbf{B}(t) = \infty$ for finite t. For these reasons, we leave this case for a future study.

Turning to upper bounds, each bounded component of $\mathbf{B}(t)^c$ contributes at least one edge to the edge boundary of $\mathbf{B}(t)$

$$\partial_e \mathbf{B}(t) = \{\{x, y\} : x \in \mathbf{B}(t), y \notin \mathbf{B}(t)\}.$$

Therefore $N(t) \leq \#\partial_e \mathbf{B}(t)$, and any upper bound for the size of the edge boundary holds also for $N(t)$. In [5], Damron-Hanson-Lam gave some such inequalities, proving in particular that if Y is the minimum of $2d$ many i.i.d. edge-weights, then $\#\partial_e \mathbf{B}(t)$ is at most order $t^{d-1}\mathbb{E}\min\{Y, t\}$ for "most" times (see [5, Thm. 1.2]). This gives a weak complement to the inequality in item 2 of Theorem 1.1 when $\mathbb{E}Y < \infty$. We focus instead on a different result of [5] which involves the "uniform curvature condition" of Newman. This condition is unproved, but believed to be true for distributions of τ_e that are, say, continuous; see [1, Sec. 2.8] for more details.

Definition 1.4. We say that the limit shape $\mathcal{B}$ satisfies the uniform curvature condition if there exist constants $c > 0, \eta > 0$ such that for all $z_1, z_2 \in \partial\mathcal{B}$ and $z = (1 - \lambda)z_1 + \lambda z_2$ with $\lambda \in [0, 1]$,

$$1 - g(z) \geq c \min\{g(z - z_1), g(z - z_2)\}^\eta,$$

where g is the norm associated to $\mathcal{B}$.

This condition is typically used in concert with an exponential moment condition:

$$\mathbb{E}e^{\alpha\tau_e} < \infty \text{ for some } \alpha > 0, \tag{1.4}$$

but it is possible to define $\mathcal{B}$ and therefore uniform curvature without any moment condition on τ_e.

As a consequence of the bound on $\#\partial_e \mathbf{B}(t)$ from [5, Thm. 1.5], we immediately obtain the following.

Proposition 1.5. *Suppose* (1.1) *and* (1.4) *hold, and assume the uniform curvature condition for $\mathcal{B}$. There exists $C > 0$ such that*

$$\mathbb{P}\left(N(t) \leq (\log t)^C t^{d-1} \textit{ for all large } t\right) = 1.$$

This result does not directly imply a good upper bound on the maximal hole size $M(t)$. For that, we give the following result in two dimensions.

Theorem 1.6. *Let $d = 2$. Suppose* (1.1) *and* (1.4) *hold, and assume the uniform curvature condition for $\mathcal{B}$. There exists $C > 0$ such that*

$$\mathbb{P}\left(M(t) \leq (\log t)^C \textit{ for all large } t\right) = 1.$$

The proof of Theorem 1.6 is in Section 3. The argument bounds the diameter of a hole in both the radial direction and the lateral direction by $(\log t)^C$. The radial estimate (see "The first case ..." above (3.9)) is valid in general dimensions. To bound the diameter in the lateral direction (below (3.15)), we must use planarity to trap a hole between two geodesics. This second part of the proof only works for $d = 2$. It would be interesting to study the geometry of holes in more detail. Do the largest holes have larger diameter in the radial direction than in the lateral one? Is there an asymptotic shape for these holes?

Remark 1.7. In fact, one can modify the proof of Theorem 1.6 slightly to obtain a weaker statement for $d \geq 3$. For any finite set of vertices $\mathbf{A}$, we define the in-volume of $\mathbf{A}$ to be

$$\mathrm{InVol}(\mathbf{A}) = \max\{\#\boldsymbol{\Xi}(n) : \exists x \in \mathbf{A}, n \geq 0 \text{ such that } x + \boldsymbol{\Xi}(n) \subset \mathbf{A}\},$$

where $\boldsymbol{\Xi}(n) = [-n, n]^d \cap \mathbb{Z}^d$. One can show that for $d \geq 3$, under (1.1), (1.4) and the uniform curvature assumption, there exists a constant $C > 0$ such that a.s. for all large t, we have $\max\{\mathrm{InVol}(\mathbf{S}) : \mathbf{S} \text{ is a finite connected component of } \mathbf{B}(t)^c\} \leq (\log t)^C$. In a certain sense, this means that the size of the largest "d-dimensional" hole (which is more or less symmetric) is no larger than $(\log t)^C$, but it does not rule out the possibility that there is a larger asymmetric hole (possibly of a "lower dimension").

Without the curvature assumption, the method of proof of Theorem 1.6 still works in some form, and produces the following weaker result. It gives a bound on the diameter of a hole in both the radial and lateral direction of order $\sqrt{t \log t}$. Its proof is in Section 4.

Theorem 1.8. *Let $d = 2$. Suppose* (1.1) *and* (1.4) *hold. There exists $C > 0$ such that*

$$\mathbb{P}\left(M(t) \leq Ct \log t \textit{ for all large } t\right) = 1.$$

1.3. **Outline of the paper.** The rest of the paper consists of proofs of the main results. First, in Section 2, we prove Theorem 1.1. The proof contains three steps. In step 1, we construct a high-weight event contained in an ℓ^1-ball $\boldsymbol{\Lambda}(n)$ that is used to create holes in $\mathbf{B}(t)$. In step 2, we show how to find translates of $\boldsymbol{\Lambda}(n)$ that are directly outside $\mathbf{B}(t)$. In step 3, we put these tools together to prove that a.s., for all large t, many of the translates of $\boldsymbol{\Lambda}(n)$ outside of $\mathbf{B}(t)$ have high-weight configurations that turn into holes in $\mathbf{B}(t)$ after a short time. In Section 3, we move to the proof of Theorem 3. The argument shows that a.s., for all large t, the largest hole in $\mathbf{B}(t)$ must be contained in a sector of an annulus with volume of order $(\log t)^C$ (see Fig. 6). Last, in Section 4, we show how to modify the proof from Section 3 without the curvature assumption to prove Theorem 1.8.

2. Proof of Theorem 1.1

Throughout this section, we suppose that (1.1) and (1.3) hold. Therefore we can pick a, b with $0 < a < b$ such that for every $\delta > 0$,

$$\mathbb{P}(\tau_e \in [a - \delta, a]) > 0 \text{ and } \mathbb{P}(\tau_e \in [b, 2b]) > 0. \tag{2.1}$$

Step 1 (Contruction of a high-weight event). We first construct a high-weight event that ensures the existence of holes. For $n \geq 1$, let

$$\boldsymbol{\Lambda}(n) = \{x \in \mathbb{Z}^d : \|x\|_1 \leq n\}$$

and write $\mathbf{e}_i$ for the i-th coordinate vector. For $m_1, m_2, m_3 \geq 1$, define the region

$$\mathbf{R} = \mathbf{R}(m_1, m_2, m_3) = \left\{x \in \mathbb{Z}^d : -m_1 \leq x \cdot \mathbf{e}_1 \leq m_2, \sum_{i=2}^{d} |x \cdot \mathbf{e}_i| \leq m_3\right\},$$

with interior boundary

$$\hat{\mathbf{R}} = \hat{\mathbf{R}}(m_1, m_2, m_3) = \{x \in \mathbf{R} : \exists y \in \mathbb{Z}^d \setminus \mathbf{R} \text{ with } \|x - y\|_1 = 1\}.$$

Also define the discrete line segment

$$\mathbf{L} = \{k\mathbf{e}_1 : k = -n, \ldots, -m_1\}.$$

(See the left side of Fig. 1.)

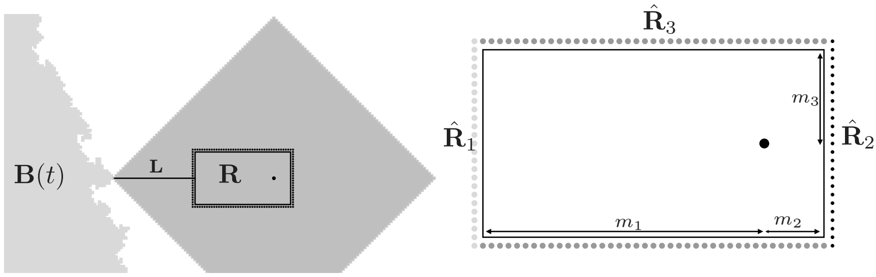

FIGURE 1. On the left: the rectangle $\mathbf{R}$ inscribed in the set $\mathbf{\Lambda}(n)$, translated to touch the growing ball $\mathbf{B}(t)$ at a corner. On the right: the interior vertex boundary of $\mathbf{R}$, defined as $\hat{\mathbf{R}}$, is partitioned into $\hat{\mathbf{R}}_1$, $\hat{\mathbf{R}}_2$, and $\hat{\mathbf{R}}_3$, indicated by the left (light grey), right (black), and top (dark grey) vertices. The origin is represented by the solid ball inside $\mathbf{R}$.

Given these geometric definitions, we now define our high-weight event. It is E_n, the event that

(1) $\tau_e \in [a-\delta, a]$ for all $e = \{x, y\}$ with $x, y \in \hat{\mathbf{R}} \cup \mathbf{L}$, and
(2) $\tau_e \in [b, 2b]$ for all other $e = \{x, y\}$ with $x, y \in \mathbf{\Lambda}(n)$.

In step 3, we will use this event to create a hole in $\mathbf{B}(t)$. The edges in item 1 allow one to enter $\mathbf{\Lambda}(n)$ at $-n\mathbf{e}_1$, travel along $\mathbf{L}$, and quickly encircle the high-weight region in $\mathbf{R}$, where a hole can appear.

Lemma 2.1. *Let a, b be as in (2.1) and $\epsilon < (b-a)/(2b+3a)$. If*

$$1 \le m_2 \le \epsilon m_3 \le \epsilon^2 m_1 \le \epsilon^3 n,$$

then, for any $\delta > 0$, on E_n,

$$T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, y) \le a(n + 2m_3) + am_2 \text{ for all } y \in \hat{\mathbf{R}}, \tag{2.2}$$

and

$$T_{\mathbf{\Lambda}(n)}(x, 0) \ge (a-\delta)(n+2m_3) + bm_2 \text{ for all } x \in \mathbb{Z}^d \text{ with } \|x\|_1 = n, \tag{2.3}$$

where $T_{\mathbf{\Lambda}(n)}$ is the minimal passage time over paths whose vertices are in $\mathbf{\Lambda}(n)$.

Proof. Throughout the proof we will use the sides of $\hat{\mathbf{R}}$:

$$\hat{\mathbf{R}}_1 = \{w \in \hat{\mathbf{R}} : w \cdot \mathbf{e}_1 = -m_1\},\ \hat{\mathbf{R}}_2 = \{w \in \hat{\mathbf{R}} : w \cdot \mathbf{e}_1 = m_2\},$$

and

$$\hat{\mathbf{R}}_3 = \{w \in \hat{\mathbf{R}} : -m_1 < w \cdot \mathbf{e}_1 < m_2\}.$$

(See the right side of Fig. 1.)

To show (2.2), let $y \in \hat{\mathbf{R}}$; we will construct a path γ from $-n\mathbf{e}_1$ to y and estimate its passage time. By symmetry, we may assume that $y \cdot \mathbf{e}_i \ge 0$ for $i = 2, \ldots, d$. If $y \in \hat{\mathbf{R}}_1 \cup \hat{\mathbf{R}}_3$ then there is a γ from $-n\mathbf{e}_1$ to y with $\|-n\mathbf{e}_1 - y\|_1$ many edges all of which have both endpoints in $\hat{\mathbf{R}} \cup \mathbf{L}$. To build γ, start at $-n\mathbf{e}_1$ and move to

$-m_1\mathbf{e}_1$ along $\mathbf{L}$. If $y \in \hat{\mathbf{R}}_1$, move to y by increasing each i-th coordinate for $i = 2$, $\ldots$, d in sequence. If $y \in \hat{\mathbf{R}}_3$, move to $-m_1\mathbf{e}_1 + \sum_{i=2}^d (y \cdot \mathbf{e}_i)\mathbf{e}_i$ by increasing each i-th coordinate for $i = 2, \ldots, d$ in sequence, and then move to y by increasing the first coordinate. The path γ as constructed has the desired properties, and

$$T(\gamma) \le a\| - n\mathbf{e}_1 - y\|_1 \le a(n + m_3 + m_2) \le \text{RHS of (2.2)}.$$

If, instead, $y \in \hat{\mathbf{R}}_2$, then we again move from $-n\mathbf{e}_1$ to $-m_1\mathbf{e}_1$ along $\mathbf{L}$, and then to the vertex

$$q = -m_1\mathbf{e}_1 + \left(m_3 - \sum_{i=2}^{d} (y \cdot \mathbf{e}_i)\right)\mathbf{e}_2 + \sum_{i=2}^{d} (y \cdot \mathbf{e}_i)\mathbf{e}_i$$

by increasing each i-th coordinate for $i = 2, \ldots, d$ in sequence. Then we move to $q + (m_1 + m_2)\mathbf{e}_1$ by increasing the first coordinate, and finally decrease the second coordinate to reach y. This γ as constructed has

$$(n - m_1) + m_3 + (m_1 + m_2) + (m_3 - \sum_{i=2}^{d} |y \cdot \mathbf{e}_i|) \le n + 2m_3 + m_2$$

many edges with weight $\le a$, so we obtain (2.2).

For (2.3), we first show that if $x \in \mathbb{Z}^d$ has $\|x\|_1 = n$, then

$$T_{\mathbf{\Lambda}(n)}(x, 0) \ge T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, 0). \tag{2.4}$$

To do this, let u be the first intersection of any $T_{\mathbf{\Lambda}(n)}$-optimal path from x to 0 with the set $\hat{\mathbf{R}} \cup \mathbf{L}$. Write γ_1 for the segment from x to u and γ_2 for the remaining segment. Then

$$\begin{aligned} T_{\mathbf{\Lambda}(n)}(x, 0) - T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, 0) &\ge (T(\gamma_1) + T(\gamma_2)) - (T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, u) + T(\gamma_2)) \\ &= T(\gamma_1) - T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, u). \end{aligned} \tag{2.5}$$

Because γ_1 uses only edges with weight $\ge b$, $T(\gamma_1) \ge b\|x - u\|_1 \ge b(n - \|u\|_1)$. If $u \in \mathbf{L}$ equals $-k\mathbf{e}_1$, then this is $b(n-k)$, but $T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, u) \le a(n-k)$, so (2.5) is nonnegative. If, on the other hand, $u \in \hat{\mathbf{R}}$, then $n - \|u\|_1 \ge n - (m_1 + m_3)$ so by (2.2),

$$\begin{aligned} T(\gamma_1) - T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, u) &\ge b(n - m_1 - m_3) - a(n + 2m_3 + m_2) \\ &\ge b(n - 2m_1) - a(n + 3m_1) \\ &\ge (b - a - 2b\epsilon - 3a\epsilon)n, \end{aligned}$$

which is > 0. This shows (2.4).

To prove (2.3), it now suffices by (2.4) to give the same lower bound for the variable $T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, 0)$. Consider any $T_{\mathbf{\Lambda}(n)}$-optimal path ρ from $-n\mathbf{e}_1$ to 0 and let u_1 be the last vertex of ρ with $u_1 \cdot \mathbf{e}_1 = -m_1$. First, if ρ contains no point in $\hat{\mathbf{R}}_3$ after u_1, then let u_1' be its first point after u_1 with $u_1' \cdot \mathbf{e}_1 = 0$. Then all edges on ρ between u_1 and u_1' have weight $\ge b$, so we obtain

$$T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, 0) = T(\rho) \ge (a-\delta)\|-n\mathbf{e}_1 - u_1\|_1 + b\|u_1 - u_1'\|_1 \ge (a-\delta)(n-m_1) + bm_1. \tag{2.6}$$

Otherwise, ρ contains a point in $\hat{\mathbf{R}}_3$ after u_1. Let u_3 be the last such point. If ρ does not contain a point of $\hat{\mathbf{R}}_2$ after u_3, then all edges on ρ after u_3 have weight

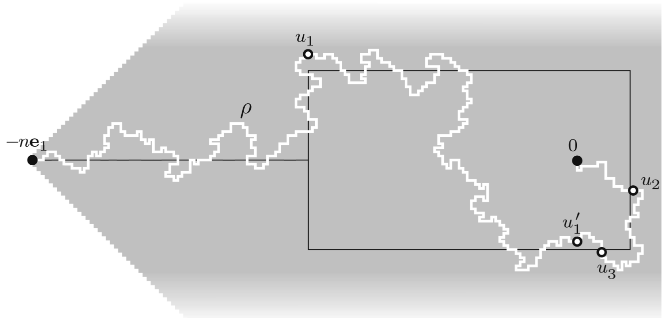

FIGURE 2. Illustration of the last part of the proof of Lemma 2.1. The path depicted in white, ρ, is a $T_{\mathbf{\Lambda}(n)}$-optimal path from $-n\mathbf{e}_1$ to 0.

$\geq b$, and we obtain

$$\begin{aligned} T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, 0) = T(\rho) \geq (a-\delta)\|-n\mathbf{e}_1 - u_3\|_1 + b\|u_3\|_1 \\ = (a-\delta)(n + u_3 \cdot \mathbf{e}_1) + (a-\delta)\sum_{i=2}^d |u_3 \cdot \mathbf{e}_i| + b\sum_{i=1}^d |u_3 \cdot \mathbf{e}_i| \\ \geq (a-\delta)(n+m_3) + bm_3. \end{aligned} \tag{2.7}$$

Here we have used that $\sum_{i=2}^d |u_3 \cdot \mathbf{e}_i| = m_3$.

The last possibility, shown in Fig. 2, is that ρ contains a point of $\hat{\mathbf{R}}_2$ after u_3; let u_2 be the last such one. Again, all edges on ρ after u_2 must have weight $\geq b$, so $T_{\mathbf{\Lambda}(n)}(-n\mathbf{e}_1, 0)$ is at least

$$\begin{aligned} &(a-\delta)\|-n\mathbf{e}_1 - u_3\|_1 + (a-\delta)\|u_3 - u_2\|_1 + b\|u_2\|_1 \\ &= (a-\delta)(n + u_3 \cdot \mathbf{e}_1 + m_3) + (a-\delta)\left((u_2 - u_3)\cdot \mathbf{e}_1 + \sum_{i=2}^d |(u_3 - u_2)\cdot \mathbf{e}_i|\right) \\ &+ b\sum_{i=1}^d |u_2 \cdot \mathbf{e}_i| \\ &\geq (a-\delta)(n+m_3) + (a-\delta)\sum_{i=2}^d (|(u_3-u_2)\cdot \mathbf{e}_i| + |u_2 \cdot \mathbf{e}_i|) + (b + a - \delta)(u_2 \cdot \mathbf{e}_1) \\ &\geq (a-\delta)(n + 2m_3) + bm_2. \end{aligned} \tag{2.8}$$

We claim that among (2.6)–(2.8),

$$\text{the term in (2.8) is minimal.} \tag{2.9}$$

Combining this fact with (2.4) will complete the proof of (2.3). To see why (2.9) holds, we write the difference between the terms in (2.6) and (2.8) as

$$\begin{aligned} (a-\delta)(n-m_1) + bm_1 - (a-\delta)(n+2m_3) - bm_2 \geq b(m_1 - m_2) - a(m_1 + 2m_3) \\ \geq (b - a - b\epsilon^2 - 2a\epsilon)m_1, \end{aligned}$$

which is > 0. Also, the difference between the terms in (2.7) and (2.8) is

$$(a-\delta)(n+m_3)+bm_3-(a-\delta)(n+2m_3)-bm_2 = -(a-\delta)m_3+b(m_3-m_2) \geq (b-\epsilon b-a)m_3,$$

which is also > 0. This completes the proof of (2.9). □

Step 2 (Placing the high-weight events outside $\mathbf{B}(t)$). Now that we have our high-weight event E_n which takes place in $\mathbf{\Lambda}(n)$, we describe a procedure to find translates of $\mathbf{\Lambda}(n)$ that are directly outside the growing ball $\mathbf{B}(t)$. These will house images of the event E_n, and will force holes in the ball at a time soon after t.

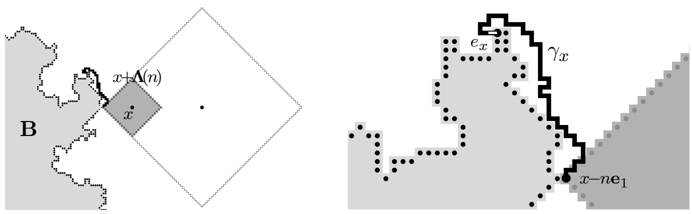

FIGURE 3. On the left: a (b', n)-good vertex x, the center of the small diamond $x+\mathbf{\Lambda}(n)$. The larger diamond and its center (corresponding to some $4ny$ for $y \in \partial^\infty \mathbf{B}_n$, in notation introduced later) are not labeled, but are part of the statement in display (2.13). On the right: a close-up of the path γ_x for the good vertex x. The initial vertex of γ_x is $x - n\mathbf{e}_1$, and its terminal edge is e_x.

Let $\mathbf{B}$ be a finite connected set of vertices (like $\mathbf{B}(t)$) and let $n \geq 1$ and $b' \geq 0$. We say that a vertex $x \in \mathbf{B}^c$ is (b', n)-good for $\mathbf{B}$ if

(1) $x + \mathbf{\Lambda}(n) \subset \mathbf{B}^c$ but $x + \mathbf{\Lambda}(n+1)$ intersects $\mathbf{B}$, and
(2) there exists a path γ_x starting at a vertex of the form $x \pm n\mathbf{e}_j$ such that
  (a) γ_x uses no vertices of either $x + \mathbf{\Lambda}(n-1)$ or $\mathbf{B}$,
  (b) some edge e_x connects the final point of γ_x to a vertex in $\mathbf{B}$ and has $\tau_{e_x} \leq b'$, and
  (c) γ_x has at most $\sqrt{n} + d$ many edges.

Roughly speaking, if x is (b', n)-good for $\mathbf{B}$, then we are able to find a path from $\mathbf{B}$ to x with relatively small passage-time. Moreover, from 2(a), the passage time of this path is independent of all the edge-weights in a neighborhood of x, $\mathbf{B}$ and its boundary. Once we fix the shape of $\mathbf{B}(t)$ to be $\mathbf{B}$, the existence of a (b', n)-good vertex for $\mathbf{B}$ will allow us to create a hole at a later time $t' > t$ in $\mathbf{B}(t')$ with high probability, by using the high-weight events defined in the previous step. In fact, we will show that there are many (b', n)-good vertices for $\mathbf{B}$.

Fix a constant $c_1 > 0$. We say that $\mathbf{B}$ is (b', n)-good if there is a set $\mathbf{S}(\mathbf{B})$ of vertices x that are (b', n)-good for $\mathbf{B}$ such that

$$\text{any distinct } x, x' \in \mathbf{S}(\mathbf{B}) \text{ have } \|x - x'\|_1 \geq 4dn \tag{2.10}$$

and

$$\#\mathbf{S}(\mathbf{B}) \geq \frac{c_1}{n^{d-1}} \#\mathbf{B}^{\frac{d-1}{d}}. \tag{2.11}$$

Fig. 3 illustrates (b', n)-good vertices and Fig. 4 illustrates a (b', n)-good set $\mathbf{B}$.

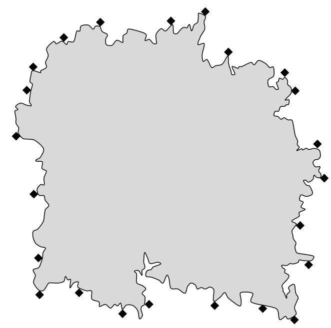

FIGURE 4. Illustration of a (b', n)-good connected set $\mathbf{B}$ of vertices. The translates $x + \mathbf{\Lambda}(n)$, as x ranges over the set $\mathbf{S}(\mathbf{B})$ of vertices that are (b', n)-good for $\mathbf{B}$, are shown in black.

Proposition 2.2. *There exist $b', c_1, C_2, c_3 > 0$ such that*

$$\mathbb{P}\left(\exists \text{ connected } \mathbf{B} \text{ with } 0 \in \mathbf{B}, \#\mathbf{B} = N \text{ and } \mathbf{B} \text{ is not } (b', n)\text{-good}\right)$$
$$\leq C_2 \left(\frac{N}{n}\right)^d \exp\left(-\frac{c_3}{n^{d-1}} N^{\frac{d-1}{d}}\right)$$

for all large n and for all $N \geq 1$.

Proof. Let $\mathbf{B}$ be a connected set with $\#\mathbf{B} = N$ and $0 \in \mathbf{B}$. By taking C_2 large, we may assume that $N \geq (4n)^d$, so that

$$0 \in \mathbf{B} \text{ and } \mathbf{B} \text{ is not contained in } [-2n, 2n-1]^d. \tag{2.12}$$

To verify that $\mathbf{B}$ is (b', n)-good with high probability, we first consider vertices of the form $4ny$ which are directly outside of $\mathbf{B}$. So, we cover $\mathbf{B}$ with boxes to get

$$\mathbf{B}_n = \{z \in \mathbb{Z}^d : (4nz + [-2n, 2n-1]^d) \cap \mathbf{B} \neq \emptyset\},$$

which is also a finite connected set. Using the notation

$$\partial^\infty \mathbf{V} = \{y \in \mathbf{V}^c : y \text{ is in the infinite component of } \mathbf{V}^c, \exists z \in \mathbf{V} \text{ with } \|y - z\|_\infty = 1\}$$

for the exterior $*$-boundary of a finite connected $\mathbf{V} \subset \mathbb{Z}^d$, we remark that $\partial^\infty \mathbf{B}_n$ is connected [14, Thm. 3]. For $v \in \mathbb{Z}^d$, define $\sigma_v = \max_e \tau_e$, where the maximum is taken over all edges e within distance d of v. For $y \in \mathbb{Z}^d$, let F_y be the event that there exists a vertex self-avoiding path in $4ny + [-7dn, 7dn-1]^d$ with $\lfloor \sqrt{n} \rfloor$ many edges and whose vertices v satisfy $\sigma_v > b'$. In this first part of the proof, we show that

$$\begin{array}{c}\text{if } y \in \partial^\infty \mathbf{B}_n \text{ and } F_y^c \text{ occurs, then some vertex in} \\ 4ny + [-6dn, 6dn]^d \text{ is } (b', n)\text{-good for } \mathbf{B}.\end{array} \tag{2.13}$$

(See Fig. 5.)

To prove (2.13), suppose that $y \in \partial^\infty \mathbf{B}_n$ and F_y^c occurs. Because $y \notin \mathbf{B}_n$, we have $4ny + [-2n, 2n-1]^d \subset \mathbf{B}^c$. Choose $y_0 \in \mathbf{B}$ such that $\|y_0 - 4ny\|_1 = \min_{y' \in \mathbf{B}} \|y' - 4ny\|_1$. Because there is a $z \in \mathbf{B}_n$ with $\|z - y\|_\infty = 1$, we know $4ny + [-6n, 6n-1]^d$ intersects $\mathbf{B}$, and so

$$2n \leq \|y_0 - 4ny\|_\infty \leq \|y_0 - 4ny\|_1 \leq 6dn.$$

To select our point x which will be (b', n)-good for $\mathbf{B}$, we assume without loss in generality that $(y_0 - 4ny) \cdot \mathbf{e}_i \geq 0$ for all $i = 1, \ldots, d$, and that the first coordinate of $y_0 - 4ny$ is maximal. Because $\|y_0 - 4ny\|_\infty \geq 2n$, we find $(y_0 - 4ny) \cdot \mathbf{e}_1 \geq n + 1$, and we define

$$x = y_0 - (n+1)\mathbf{e}_1.$$

Then $\|x - y_0\|_1 = n+1$ and so $x + \mathbf{\Lambda}(n+1)$ intersects $\mathbf{B}$ at the point $y_0 = x + (n+1)\mathbf{e}_1$. However, if $w \in x + \mathbf{\Lambda}(n)$, then $\|w - 4ny\|_1 \leq \|w - x\|_1 + \|x - 4ny\|_1 \leq n + \|x - 4ny\|_1$ and

$$\|x - 4ny\|_1 = \|y_0 - 4ny - (n+1)\mathbf{e}_1\|_1 = \|y_0 - 4ny\|_1 - (n+1),$$

so $\|w - 4ny\|_1 \leq \|y_0 - 4ny\|_1 - 1$, giving by minimality of y_0 that $w \notin \mathbf{B}$. Therefore $x + \mathbf{\Lambda}(n) \subset \mathbf{B}^c$. Furthermore, because $(y_0 - 4ny) \cdot \mathbf{e}_1 = \|y_0 - 4ny\|_\infty \geq n + 1$, we have $\|x - 4ny\|_\infty = \|y_0 - 4ny - (n+1)\mathbf{e}_1\|_\infty \leq \|y_0 - 4ny\|_\infty \leq 6dn$, so we conclude that $x \in 4ny + [-6dn, 6dn]^d$. This shows that

$$x \in 4ny + [-6dn, 6dn]^d,\ x + \mathbf{\Lambda}(n) \subset \mathbf{B}^c, \text{ and } x + \mathbf{\Lambda}(n+1) \text{ intersects } \mathbf{B} \text{ at a point } y_0,$$

where $y_0 = x + (n+1)\mathbf{e}_1$. Even without our assumptions on $(y_0 - 4ny) \cdot \mathbf{e}_i$, we obtain the same statement, but y_0 is then of the form $x \pm (n+1)\mathbf{e}_j$. This shows item 1 of the definition of (b', n)-good for the vertex x.

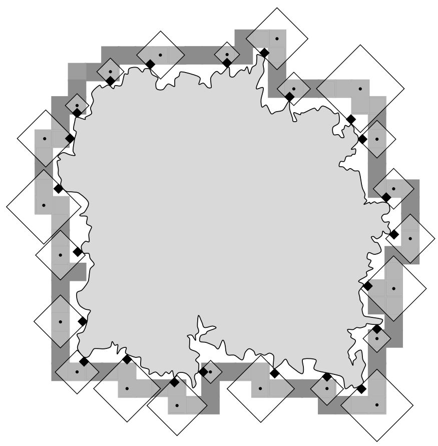

FIGURE 5. Extracting a collection of (b', n)-good vertices (the centers of the black diamonds) surrounding a connected set $\mathbf{B}$ of vertices, in light grey. The smaller diamonds are centered at vertices $4ny$ for y in the set $\partial^\infty \mathbf{B}_n$ (shown in grey) such that F_y^c occurs.

If the edge connecting y_0 to the unique vertex w_0 in $x + \mathbf{\Lambda}(n)$ has weight $\leq b'$, then we can simply set γ_x to be the path with no edges and a single vertex w_0. In general, though, we must find a nearby edge satisfying this weight constraint. We observe that w_0 is in the exterior $*$-boundary $\partial^\infty \mathbf{B}$ of $\mathbf{B}$. This is because it is adjacent to y_0, which is in $\mathbf{B}$, but also can be connected to $4ny$ without touching $\mathbf{B}$, and $y \in \partial^\infty \mathbf{B}_n$. Assumption (2.12) ensures that $\mathbf{B}$ is not contained in $w_0 + [-\sqrt{n}, \sqrt{n}]^d$. Neither is $\partial^\infty \mathbf{B}$ if n is large, and so we can select $w_0' \in \partial^\infty \mathbf{B}$ which is not in $w_0 + [-\sqrt{n}, \sqrt{n}]^d$. Because $\partial^\infty \mathbf{B}$ is connected, there is a vertex self-avoiding path $\mathfrak{p}$ from w_0 to w_0' in $\partial^\infty \mathbf{B}$ and since $x + \mathbf{\Lambda}(n) \subset \mathbf{B}^c$, the path $\mathfrak{p}$ cannot use any vertices of $x + \mathbf{\Lambda}(n-1)$. Let $\mathfrak{p}_x$ be the initial segment of $\mathfrak{p}$ consisting of the first $\lfloor\sqrt{n}\rfloor$ many edges and list the vertices of $\mathfrak{p}_x$ as $w_0, w_1, \ldots, w_{\lfloor\sqrt{n}\rfloor}$. Each w_i is 1 unit away from $\mathbf{B}$ in ℓ^∞-distance. If there exist an i and a path γ_i from w_i to $\mathbf{B}$ with ℓ^1-length equal to $\min\{\|w_i - \bar{u}\|_1 : \bar{u} \in \mathbf{B}\}$ (which is at most d) such that the final edge f_i of γ_i has $\tau_{f_i} \le b'$ for some i, we let i_0 be the first such i and define γ_x to be the initial segment of $\mathfrak{p}_x$ from w_0 to w_{i_0} concatenated with γ_{i_0} (with the point in $\mathbf{B}$ and the edge f_i removed). This γ_x satisfies conditions (a)–(c) of the definition of (b', n)-good. If i_0 does not exist, then the entire path $\mathfrak{p}_x$ must have vertices with $\sigma_v > b'$, meaning that F_y occurs. This shows (2.13).

Given (2.13), we can now return to the main proof. Let $\mathbf{B}$ be connected with $\#\mathbf{B} = N$ and such that (2.12) holds. The set $\mathbf{B}_n$ satisfies $(4n)^d \#\mathbf{B}_n \ge \#\mathbf{B} = N$, so the isoperimetric inequality implies

$$\#\partial^\infty \mathbf{B}_n \ge \frac{c_4}{n^{d-1}} N^{\frac{d-1}{d}}. \tag{2.14}$$

Suppose that $\mathbf{B}$ is not (b', n)-good. We claim that for some constant $C_5 > 0$, and c_1 from the definition of (b', n)-good,

$$\sum_{y \in \partial^\infty \mathbf{B}_n} \mathbf{1}_{F_y^c} \le \frac{C_5 c_1}{n^{d-1}} N^{\frac{d-1}{d}}. \tag{2.15}$$

To see why, partition $\partial^\infty \mathbf{B}_n$ into $C_5 = C_5(d)$ many subsets $\mathbf{S}_1, \ldots, \mathbf{S}_{C_5}$ such that if for a fixed i, we select distinct $y, y' \in \mathbf{S}_i$, then $\|y - y'\|_\infty \ge 4d$. If for some such i, F_y^c and $F_{y'}^c$ both occur, then let x, x' be the corresponding (b', n)-good points from (2.13). We have

$$\begin{aligned}\|x - x'\|_1 \ge \|x - x'\|_\infty &\ge \|4ny - 4ny'\|_\infty - \|4ny - x\|_\infty - \|4ny' - x'\|_\infty \\ &\ge 16dn - 6dn - 6dn = 4dn.\end{aligned}$$

The definition of (b', n)-good then implies $\sum_{y \in \mathbf{S}_i} \mathbf{1}_{F_y^c} < (c_1/n^{d-1}) N^{(d-1)/d}$ for $i = 1, \ldots, C_5$, and this gives (2.15).

Since $\mathbf{B} \subset [-N, N]^d$, we have $\mathbf{B}_n \subset [-(N/n) - 1, (N/n) + 1]^d$ and so $\partial^\infty \mathbf{B}_n \subset [-(N/n) - 2, (N/n) + 2]^d$. Taking c_4 from (2.14), if $A_{N,n}$ is the event that there exists a finite connected set $\mathbf{S} \subset \mathbb{Z}^d$ such that

$$\#\mathbf{S} \ge \frac{c_4}{n^{d-1}} N^{\frac{d-1}{d}} \text{ and } \mathbf{S} \subset \left[-\frac{N}{n} - 2, \frac{N}{n} + 2\right]^d,$$

but $\sum_{y \in \mathbf{S}} \mathbf{1}_{F_y^c} \le (c_1 C_5 / c_4) \#\mathbf{S}$, then

$$\mathbb{P}(\exists \text{ connected } \mathbf{B} \text{ with } 0 \in \mathbf{B}, \#\mathbf{B} = N \text{ and } \mathbf{B} \text{ is not } (b', n)\text{-good}) \le \mathbb{P}(A_{N,n}). \tag{2.16}$$

Let $\mathcal{S}_k$ be the collection of connected $\mathbf{S} \subset \mathbb{Z}^d$ such that $\#\mathbf{S} = k$ and $\mathbf{S}$ contains the origin. Using the bound $\#\mathcal{S}_k \le (2de)^k \le e^{C_6 k}$ [2] for some $C_6 > 0$, we obtain for $\ell \ge 0$

$$\begin{aligned}&\# \text{ connected } \mathbf{S} \subset \mathbb{Z}^d \cap [-\ell, \ell]^d \text{ with } \#\mathbf{S} = k \\ &\le \sum_{v \in [-\ell, \ell]^d} \# \text{ connected } \mathbf{S} \subset \mathbb{Z}^d \text{ containing } v \text{ with } \#\mathbf{S} = k \\ &\le (2\ell + 1)^d e^{C_6 k}.\end{aligned}$$

Applying this with $\ell = N/n + 2$, we see that

$$\mathbb{P}(A_{N,n}) \le \left(2\frac{N}{n} + 5\right)^d \sum_{k \ge \frac{c_4}{n^{d-1}} N^{\frac{d-1}{d}}} e^{C_6 k} \max_{\mathbf{S} \in \mathcal{S}_k} \mathbb{P}\left(\sum_{y \in \mathbf{S}} \mathbf{1}_{F_y^c} \le \frac{c_1 C_5}{c_4} k\right). \tag{2.17}$$

For a given $\mathbf{S} \in \mathcal{S}_k$, the events $\mathbf{1}_{F_y^c}$ are not independent as y ranges over $\mathbf{S}$, but they are only finitely dependent. Therefore we can extract a subset of size at least $c_7 k$ such that as y ranges over the subset, the events $\mathbf{1}_{F_y^c}$ are independent. This implies that

$$\max_{\mathbf{S} \in \mathcal{S}_k} \mathbb{P}\left(\sum_{y \in \mathbf{S}} \mathbf{1}_{F_y^c} \le \frac{c_1 C_5}{c_4} k\right) \le \mathbb{P}\left(\sum_{i=1}^{\lfloor c_7 k \rfloor} Z_i \le \frac{c_1 C_5}{c_4} k\right),$$

where Z_i are i.i.d. and have the same distribution as $\mathbf{1}_{F_0^c}$. The right side is bounded by

$$\mathbb{P}\left(\sum_{i=1}^{\lfloor c_7 k \rfloor} (1 - Z_i) \ge \lfloor c_7 k \rfloor - \frac{c_1 C_5}{c_4} k\right) \le 2^{c_7 k} \mathbb{P}(F_0)^{\lfloor c_7 k \rfloor - \frac{c_1 C_5}{c_4} k},$$

so we can return to (2.17) to state, for some $C_8 > 0$,

$$\mathbb{P}(A_{N,n}) \le \left(2\frac{N}{n} + 5\right)^d \sum_{k \ge \frac{c_4}{n^{d-1}} N^{\frac{d-1}{d}}} e^{C_8 k} \mathbb{P}(F_0)^{\lfloor c_7 k \rfloor - \frac{c_1 C_5}{c_4} k}. \tag{2.18}$$

Last, we must estimate $\mathbb{P}(F_0)$. For a given vertex self-avoiding path γ in the box $[-7dn, 7dn - 1]^d$ with $\lfloor \sqrt{n} \rfloor$ many edges, the events $\{\sigma_v > b\}$ as v ranges over the vertices of γ are not independent, but they are finitely dependent. Again, we can find a subset of the vertices of size at least $c_9 \sqrt{n}$ such that as v ranges over the subset, the events are independent. This gives

$$\mathbb{P}(\text{for all } v \in \gamma, \sigma_v > b') \le \mathbb{P}(\sigma_0 > b')^{c_9 \sqrt{n} - 1} \le ((2d)^d \mathbb{P}(\tau_e > b'))^{c_9 \sqrt{n} - 1}.$$

The number of such paths γ is at most $(14dn)^d (2d)^{\sqrt{n}}$, so

$$\mathbb{P}(F_0) \le (14dn)^d (2d)^{\sqrt{n}} ((2d)^d \mathbb{P}(\tau_e > b'))^{c_9 \sqrt{n} - 1}.$$

Putting this in (2.18), we find

$$\begin{aligned} \mathbb{P}(A_{N,n}) &\le \left(2\frac{N}{n} + 5\right)^d \\ &\quad \times \sum_{k \ge \frac{c_4}{n^{d-1}} N^{\frac{d-1}{d}}} e^{C_8 k} ((14dn)^d (2d)^{\sqrt{n}} ((2d)^d \mathbb{P}(\tau_e > b'))^{c_9 \sqrt{n} - 1})^{\lfloor c_7 k \rfloor - \frac{c_1 C_5}{c_4} k}. \end{aligned}$$

First choose c_1 so small that $\lfloor c_7 k \rfloor - c_1 C_5 k / c_4$ is at least $c_7 k/2$. After this, we may choose b' so large that the entire summand is at most 2^{-k}. This produces the bound

$$\mathbb{P}(A_{N,n}) \le 2 \cdot \left(2\frac{N}{n} + 5\right)^d 2^{-\frac{c_4}{n^{d-1}} N^{\frac{d-1}{d}}}.$$

Combined with (2.16), this implies the statement of Proposition 2.2. $\square$

Step 3 (Construction of holes). In this step we use the tools from the previous steps to construct holes in $\mathbf{B}(t)$. First, by [4, Eq. (3)], our assumption (1.1) gives a $c_{10} > 0$ such that

$$\mathbb{P}(c_{10}t^d \leq \#\mathbf{B}(t) < \infty \text{ for all large } t) = 1. \tag{2.19}$$

To place the translates of $\mathbf{\Lambda}(n)$ from step 2 around the set $\mathbf{B}(t)$, we choose a size of one of the two forms

$$n = n_t = C_{11} \in \mathbb{N} \text{ or } \lfloor c_{12}(\log t)^{\frac{1}{d}} \rfloor. \tag{2.20}$$

We fix the rest of our parameters as follows:

(1) a, b are as in (2.1) and b' is from Proposition 2.2,
(2) let $\delta = \epsilon^4$, where $\epsilon < (b-a)/(2b+3a)$ (compare to Lemma 2.1) will be taken small in the proof of (2.25),
(3) $m_1 = \lfloor \epsilon n \rfloor, m_3 = \lfloor \epsilon m_1 \rfloor, m_2 = \lfloor \epsilon m_3 \rfloor$ and set $H = \lfloor \epsilon m_2 \rfloor$.

If C_{11} and t are large with c_{12} fixed, the m_i's satisfy the constraints in Lemma 2.1. The parameter H will be a lower bound on the radius of a hole. Now we apply Proposition 2.2 for

$$\begin{aligned}
&\mathbb{P}(\exists \text{ connected } \mathbf{B} \text{ with } 0 \in \mathbf{B}, c_{10}t^d \leq \#\mathbf{B} < \infty \text{ and } \mathbf{B} \text{ is not } (b', n)\text{-good}) \\
&\leq \sum_{N \geq c_{10}t^d} \mathbb{P}(\exists \text{ connected } \mathbf{B} \text{ with } 0 \in \mathbf{B}, \#\mathbf{B} = N \text{ and } \mathbf{B} \text{ is not } (b', n)\text{-good}) \\
&\leq C_2 \sum_{N \geq c_{10}t^d} \left(\frac{N}{n}\right)^d \exp\left(-\frac{c_3}{n^{d-1}} N^{\frac{d-1}{d}}\right).
\end{aligned} \tag{2.21}$$

The application of Proposition 2.2 requires that n is large, and this holds for large C_{11} and t for fixed c_{12}. For either choice of n from (2.20), the expression in (2.21) is summable in t, so for any large C_{11} and any fixed c_{12},

$$\sum_{t \in \mathbb{N}} \mathbb{P}\left(c_{10}t^d \leq \#\mathbf{B}(t) < \infty \text{ but } \mathbf{B}(t) \text{ is not } (b', n_t)\text{-good}\right) < \infty. \tag{2.22}$$

From the definition of (b', n)-good, we get boxes of the form $x + \mathbf{\Lambda}(n)$ situated around our set $\mathbf{B}(t)$, so now we must populate them with versions of the event E_n from step 1. To do this properly, we need to decouple the variables inside $\mathbf{B}(t)$ from those outside. For a given finite, connected $\mathbf{B}$ containing the origin that is (b', n) good, we may choose at least $(c_1/n^{d-1})\#\mathbf{B}^{(d-1)/d}$ many vertices x that are (b', n)-good for $\mathbf{B}$ and distinct x, x' satisfy inequality (2.10). These vertices come with edges e_x and paths γ_x as in the definition. The edges and paths are contained in the boxes $[-n - \sqrt{n} - d - 1, n + \sqrt{n} + d + 1]^d + x$ because of item 2(c) in the definition, and by (2.10), these boxes are disjoint for distinct x, x'. Enumerate the first

$$r = \left\lceil \frac{c_1}{n^{d-1}} \#\mathbf{B}^{\frac{d-1}{d}} \right\rceil$$

many of these points in some deterministic way as $x_1, \ldots, x_r$. All of the x_i, γ_{x_i}, and e_{x_i} are random, so we must fix their values for a large $t \in \mathbb{N}$ as

$$\begin{aligned}
&\mathbb{P}\left(c_{10}t^d \leq \#\mathbf{B}(t) < \infty \text{ and } \mathbf{B}(t) \text{ is } (b', n)\text{-good}\right) \\
&= \sum_{\mathbf{B}: c_{10}t^d \leq \#\mathbf{B} < \infty} \sum_{(z_i, \pi_i, e_i)_{i=1}^r} \mathbb{P}\left(\begin{array}{c} \mathbf{B}(t) = \mathbf{B} \text{ is } (b', n)\text{-good}, \\ x_i = z_i, \gamma_{x_i} = \pi_i, e_{x_i} = e_i \ \forall i \end{array}\right).
\end{aligned} \tag{2.23}$$

We observe that the event in the probability depends only on edges with at least one endpoint in $\mathbf{B}$, so it is independent of the weights of edges with both endpoints outside of $\mathbf{B}$.

For a given choice of $(z_i, \pi_i, e_i)_{i=1}^r$, and $\mathbf{B}$, we define events $(A_i)_{i=1}^r$ by the following conditions. A_i is the event that:

(1) all edges e of π_i have $\tau_e \leq a$, and
(2) the event $T_i E_n$ occurs.

In item 2, $T_i E_n$ is a certain translation and rotation of the high-weight event E_n from step 1. Precisely, the initial point of π_i is one of the points of the form $z_i \pm n\mathbf{e}_j$, and we define T_i to be an isometry of $\mathbb{R}^d$ that maps $\mathbf{\Lambda}(n)$ to $z_i + \mathbf{\Lambda}(n)$ and $-n\mathbf{e}_1$ to the initial point of π_i. Then $T_i E_n$ is the event that the image configuration $(\tau_{T_i^{-1}(e)})$ is in E_n. Not only does the definition of E_n depend on n from (2.20) and a, b, it also depends on $\delta = \epsilon^4$ from (2.1) and the numbers m_1, m_2, m_3. Regardless of the values of the m_i, since they are $\leq n$, there exists $C_{13} > 0$ depending only on a, b, ϵ such that $\mathbb{P}(E_n) \geq e^{-C_{13} n^d}$. Using this in the definition of A_i, there exists $C_{14} > 0$ also depending only on a, b, ϵ such that

$$\mathbb{P}(A_i) \geq \exp\left(-C_{14} n^d\right) \text{ for all } i = 1, \ldots, r.$$

Because the A_i's are independent, we may bound the family $(\mathbf{1}_{A_i})_{i=1}^r$ stochastically from below by a family $(W_i)_{i=1}^r$ of i.i.d. Bernoulli variables with parameter $p = e^{-C_{14} n^d}$. By Hoeffding's bound for Bernoulli random variables, we have $\mathbb{P}\left(W_1 + \cdots + W_r \leq r\frac{p}{2}\right) \leq \exp\left(-\frac{r}{2}p^2\right)$, and therefore

$$\mathbb{P}\left(\sum_{i=1}^r \mathbf{1}_{A_i} \leq \frac{r}{2}\exp\left(-C_{14} n^d\right)\right) \leq \exp\left(-\frac{r}{2}\exp\left(-2C_{14} n^d\right)\right).$$

For any large C_{11} and small c_{12}, we have $r \geq c_{15}(t/n)^{d-1}$ for all large t, so

$$\frac{r}{2}\exp\left(-2C_{14} n^d\right) \geq t^{d-1}\exp\left(-C_{16} n^d\right) \geq c_{17} t^{d-\frac{3}{2}} \text{ for all large } t.$$

This implies for any large C_{11} and small c_{12},

$$\mathbb{P}\left(\sum_{i=1}^r \mathbf{1}_{A_i} \leq \exp\left(-C_{16} n^d\right) t^{d-1}\right) \leq \exp\left(-c_{17} t^{d-\frac{3}{2}}\right) \text{ for all large } t.$$

Returning to the right side of (2.23), independence gives for any large C_{11} and small c_{12},

$$\begin{aligned}&\left(1 - \exp\left(-c_{17} t^{d-\frac{3}{2}}\right)\right)\mathbb{P}\left(c_{10} t^d \leq \#\mathbf{B}(t) < \infty \text{ and } \mathbf{B}(t) \text{ is } (b', n)\text{-good}\right) \\ &\leq \sum_{\mathbf{B}: c_{10} t^d \leq \#\mathbf{B} < \infty} \sum_{(z_i, \pi_i, e_i)_{i=1}^r} \mathbb{P}\left(\begin{array}{c}\mathbf{B}(t) = \mathbf{B} \text{ is } (b', n)\text{-good}, x_i = z_i, \gamma_{x_i} = \pi_i \\ e_{x_i} = e_i \ \forall i, \ \sum_{i=1}^r \mathbf{1}_{A_i} \geq \exp\left(-C_{16} n^d\right) t^{d-1}\end{array}\right)\end{aligned} \tag{2.24}$$

for all large t.

We will now argue that there exists $\epsilon > 0$ such that on the event on the right of (2.24), if C_{11} is any large number and c_{12} is any fixed number, then for all large t, and all i such that A_i occurs,

$$x_i + \mathbf{\Lambda}(H) \text{ is in a bounded component of } \mathbf{B}(s)^c \text{ for all } s \in [t+\kappa, t+\kappa+\epsilon^4 n], \tag{2.25}$$

where

$$\kappa = \kappa_t = \epsilon^4 n + a(n + 2\epsilon^2 n) + a\epsilon^3 n, \tag{2.26}$$

and these components are distinct for distinct values of i. In this statement, as before, $n = n_t$, so that H (defined below (2.20)) and κ are also functions of t (not s). To prove this, pick an outcome in this event with i such that A_i occurs, and let u_i be the endpoint of e_i in $\mathbf{B}$. Let v_i be the endpoint of π_i that is in $x_i+\mathbf{\Lambda}(n)$. Let $y \in T_i\hat{\mathbf{R}}$ (this is the corresponding image of the set $\hat{\mathbf{R}}$ from Lemma 2.1 inside $x + \mathbf{\Lambda}(n)$). Because $u_i \in \mathbf{B}(t)$ and π_i has at most $\sqrt{n}$ many edges, we have

$$\begin{aligned} T(0, y) \le T(0, u_i) + \tau_{e_i} + T(\pi_i) + T(v_i, y) &\le t + b' + a\sqrt{n} + a(n + 2m_3) + am_2 \\ &\le t + \kappa. \end{aligned} \tag{2.27}$$

We have used (2.2) to estimate $T(v_i, y)$ and used $b' + a\sqrt{n} \le \epsilon^4 n$, which is valid for any ϵ and c_{12} so long as C_{11} and t are large. On the other hand, if $z \in \mathbb{Z}^d$ has $\|z - x_i\|_1 \le H \le \min\{m_i\}$, condition 2 of the definition of E_n implies

$$T(0, z) \ge T(0, x_i) - T(x_i, z) \ge T(0, x_i) - 2Hb.$$

Let σ be any path from 0 to x_i, let σ_1 be the initial segment until its first vertex outside $\mathbf{B}$, and let σ_2 be its terminal segment starting at the point at which it enters $x_i + \mathbf{\Lambda}(n)$ for the last time. Then because σ_1 connects 0 to $\mathbf{B}(t)^c$,

$$\begin{aligned} T(\sigma) \ge T(\sigma_1) + T(\sigma_2) &\ge t + \min_{x:\|x - x_i\|_1 = n} T_{T_i\mathbf{\Lambda}(n)}(x, x_i) \\ &\ge t + (a - \delta)(n + 2m_3) + bm_2. \end{aligned}$$

The last inequality follows from (2.3). Take the infimum over σ to obtain

$$\begin{aligned} T(0, z) \ge T(0, x_i) - 2Hb &\ge t + (a - \delta)(n + 2m_3) + bm_2 - 2Hb \\ &\ge t + (a - \epsilon^4)(n + 2\epsilon^2 n) + b\epsilon^3 n - (2b + 1)\epsilon^4 n \\ &= t + \kappa + (b - a)\epsilon^3 n - \epsilon^4(n + 2\epsilon^2 n) - (2b + 2)\epsilon^4 n. \end{aligned}$$

Again we have assumed that ϵ is fixed, c_{12} is fixed, and C_{11} and t are large to remove the floor function in the definition of the m_i's. From the above, we can choose ϵ so small such that for any c_{12}, and for any large C_{11},

$$T(0, z) \ge t + \kappa + \epsilon^4 n \text{ for all large } t.$$

This inequality and (2.27) show that for any s in the interval described in (2.25), the set $x_i + \mathbf{\Lambda}(H)$ is in $\mathbf{B}(s)^c$, but $T_i\hat{\mathbf{R}}$ is in $\mathbf{B}(s)$. This implies (2.25). Furthermore, the sets $x_i + \mathbf{\Lambda}(n)$ are disjoint, so since the components described in (2.25) are contained in these sets, they are distinct for distinct values of i.

Given (2.25), we can finish the proof. Any component listed in (2.25) contains $x_i + [0, H/d]^d$, so it has at least $(H/d)^d$ many vertices. If we define

$$Y_t = \min_{s \in [t + \kappa_t, t + \kappa_t + \epsilon^4 n_t]} \begin{array}{c} \#\text{ bounded components of } \mathbf{B}(s)^c \text{ with at} \\ \text{least } \left(\frac{H_t}{d}\right)^d \text{ many vertices,} \end{array}$$

then we can continue from (2.24) with our ϵ from (2.25), any large C_{11} and any small c_{12} to obtain

$$\begin{aligned}
&\left(1-\exp\left(-c_{17}t^{d-\frac{3}{2}}\right)\right)\mathbb{P}\left(c_{10}t^d \leq \#\mathbf{B}(t)<\infty \text{ and } \mathbf{B}(t) \text{ is } (b',n)\text{-good}\right)\\
&\leq \sum_{\mathbf{B}:c_{10}t^d\leq\#\mathbf{B}<\infty}\ \sum_{(z_i,\pi_i,e_i)_{i=1}^r}\mathbb{P}\left(\begin{array}{c}\mathbf{B}(t)=\mathbf{B} \text{ is } (b',n)\text{-good}, x_i=z_i, \gamma_{x_i}=\pi_i,\\ e_{x_i}=e_i\ \forall i,\ Y_t\geq \exp\left(-C_{16}n^d\right)t^{d-1}\end{array}\right)
\end{aligned}$$

(2.28)
$$= \mathbb{P}\left(\mathbf{B}(t) \text{ is } (b',n)\text{-good}, c_{10}t^d\leq\#\mathbf{B}(t)<\infty,\ Y_t\geq\exp\left(-C_{16}n^d\right)t^{d-1}\right)$$

for all large t. This implies for any large C_{11} and any small c_{12}

$$\begin{aligned}
&\sum_{t\in\mathbb{N}}\mathbb{P}\left(c_{10}t^d\leq\#\mathbf{B}(t)<\infty \text{ and } Y_t<\exp\left(-C_{16}n_t^d\right)t^{d-1}\right)\\
&\leq \sum_{t\in\mathbb{N}}\mathbb{P}\left(c_{10}t^d\leq\#\mathbf{B}(t)<\infty \text{ but } \mathbf{B}(t) \text{ is not } (b',n_t)\text{-good}\right)\\
&+\sum_{t\in\mathbb{N}}\mathbb{P}\left(Y_t<\exp\left(-C_{16}n_t^d\right)t^{d-1}\ \middle|\ \mathbf{B}(t) \text{ is } (b',n_t)\text{-good}, c_{10}t^d\leq\#\mathbf{B}(t)<\infty\right).
\end{aligned}$$

The sum in the second line is finite by (2.22). By (2.28), the summands of the third are bounded for large t by the summands of $\sum_{t\in\mathbb{N}}\exp\left(-c_{17}t^{d-\frac{3}{2}}\right)<\infty$. The Borel-Cantelli lemma combined with (2.19) therefore implies that for any large C_{11} and any small c_{12}, a.s.,

$$Y_t\geq\exp\left(-C_{16}n_t^d\right)t^{d-1} \text{ for all large } t\in\mathbb{N}. \tag{2.29}$$

Last, we use (2.29) to prove Theorem 1.1. First take $n_t=C_{11}$. Then $\kappa_t=(\epsilon^4+a+2\epsilon^2+a\epsilon^3)C_{11}$, and so the interval $I_t=[t+\kappa_t,t+\kappa_t+\epsilon^4 n_t]$ satisfies $I_t\cap I_{t+1}\neq\emptyset$ for all $t\geq 1$ so long as C_{11} is large. Therefore (2.29) gives that a.s. $\mathbf{B}(t)^c$ has at least $\exp\left(-C_{16}C_{11}^d\right)t^{d-1}$ many bounded components for all large t. This proves item 2 of Theorem 1.1. If we take $n_t=\lfloor c_{12}(\log t)^{1/d}\rfloor$, then the intervals I_t and I_{t+1} also intersect for large t, if c_{12} is fixed. For small c_{12}, we have $Y_t\geq t^{d-3/2}$ for all large t so, in particular, $Y_t>0$. This gives that a.s., for all large s, the maximum hole size $M(s)$ is at least equal to $(H_t/d)^d$, where t is any number such that $s\in I_t$. If t is large and c_{12} is fixed, then this t satisfies $t\geq s/2$, so we obtain

$$\text{a.s., } M(s)\geq\left(\frac{H_{\frac{s}{2}}}{d}\right)^d \text{ for all large } s.$$

This implies item 1 of Theorem 1.1 and completes the proof.

3. Proof of Theorem 1.6

In this section, we assume (1.1), (1.4), and the uniform curvature condition. We first describe the idea of the proof. Let t be large and let x_0, if it exists, be any vertex in the largest bounded component $\mathbf{C}$ of $\mathbf{B}(t)^c$ with maximal Euclidean norm $\|x_0\|_2$. Let $\angle(v,w)$ be the angle (in $(-\pi,\pi]$) between two vectors $v,w\in\mathbb{R}^2$ and define the sector portion

$$\mathbf{S}_{x_0}=\left\{v\in\mathbb{R}^2:|\angle(v,x_0)|\leq J_{x_0},\ 1-K_{x_0}\leq\frac{\|v\|_2}{\|x_0\|_2}\leq 1\right\}, \tag{3.1}$$

where

$$J_{x_0} = \frac{(\log \|x_0\|_2)^{C_{18}-3}}{\|x_0\|_2}, \; K_{x_0} = \frac{(\log \|x_0\|_2)^{C_{18}}}{\|x_0\|_2}, \tag{3.2}$$

and $C_{18} > 3$ is a large constant to be chosen later; see Fig. 6. The component $\mathbf{C}$ containing x_0 is connected, and by extremality of x_0, it cannot cross the far side of $\mathbf{S}_{x_0}$. Once we show that it cannot cross the left, right, and near sides, then we can deduce that $\mathbf{C} \subset \mathbf{S}_{x_0}$. Because

$$\mathbf{S}_{x_0} \text{ contains at most } C_{19}(\log \|x_0\|_2)^{2C_{18}} \text{ many vertices,} \tag{3.3}$$

and $\|x_0\|_2$ must be of order t to be in a bounded component of $\mathbf{B}(t)^c$, we conclude the result.

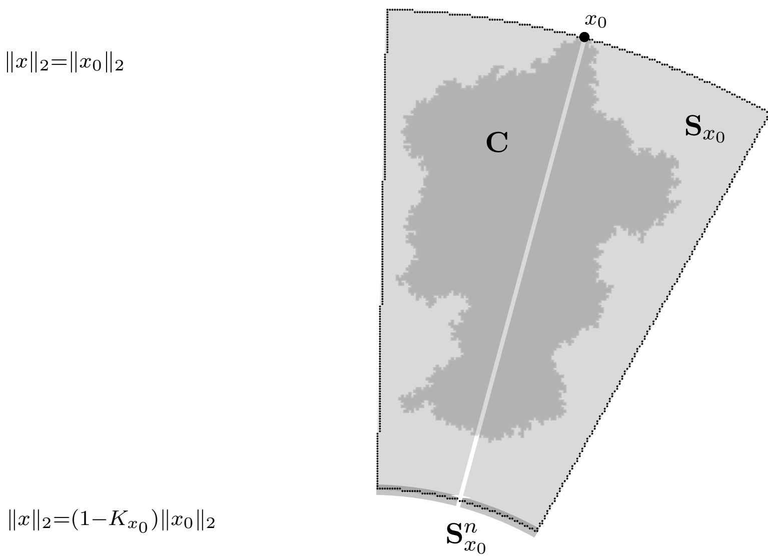

FIGURE 6. The set $\mathbf{C}$, depicted above as the darker shaded region, is the largest hole in the ball $\mathbf{B}(t_0)$. The set $\mathbf{S}_{x_0}$ is a sector (lighter shaded region) centered on the vertex x_0 with maximal Euclidean norm among all those in $\mathbf{C}$. The boundary segment of $\mathbf{S}_{x_0}$ nearest to the origin, $\mathbf{S}^n_{x_0}$, is also shaded.

Step 1 (Setup of the proof). To start the proof, we let $s > 0$ and recall the notation $\widetilde{\mathbf{B}}(t) = \mathbf{B}(t) + [0,1)^d$ from the introduction. Define the events

$$E_1(s) = \left\{ \frac{1}{2}\mathcal{B} \subset \frac{1}{t}\widetilde{\mathbf{B}}(t) \subset 2\mathcal{B} \text{ for all } t \geq s \right\}$$

and

$$E_2(s) = \{\tau_e \leq C_{20} \log t \text{ for all } e \text{ with an endpoint in } 3t\mathcal{B} \text{ and all } t \geq s\}.$$

We have

$$\begin{aligned}&\mathbb{P}(M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s)\\ &\leq \mathbb{P}(E_1(s)^c) + \mathbb{P}(E_2(s)^c)\\ &+ \mathbb{P}\left(E_1(s) \cap E_2(s) \cap \{M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s\}\right).\end{aligned}$$

By the shape theorem in (1.2), $\mathbb{P}(E_1(s)^c) \to 0$ as $s \to \infty$. To estimate $\mathbb{P}(E_2(s)^c)$, we write $\mathbb{P}(\tau_e > C_{20} \log n) \leq \mathbb{E}e^{\alpha\tau_e}/e^{\alpha C_{20} \log n}$ for the α in (1.4), so

$$\mathbb{P}(\tau_e > C_{20} \log n \text{ for some } e \text{ with an endpoint in } 3n\mathcal{B}) \leq C_{21} n^2 n^{-C_{20}\alpha}.$$

By a union bound,

$$\mathbb{P}(\tau_e > C_{20} \log n \text{ for some } e \text{ with an endpoint in } 3n\mathcal{B} \text{ and some } n \geq N) \to 0$$

as $N \to \infty$ if we choose $C_{20} > 4\alpha$. By increasing C_{20} further, this implies that $\mathbb{P}(E_2(s)^c) \to 0$ as $s \to \infty$.

From the above arguments, we obtain

$$\begin{aligned}&\lim_{s\to\infty} \mathbb{P}(M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s)\\ &= \lim_{s\to\infty} \mathbb{P}\left(E_1(s) \cap E_2(s) \cap \{M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s\}\right).\end{aligned} \tag{3.4}$$

To show the limit in (3.4) is zero, we use the sector construction from the proof idea above. Fix an outcome in the event in the probability in (3.4) and let $t_0 \geq s$ be any value of t for which $M(t) \geq (\log t)^{3C_{18}}$. Choose x_0 as any vertex with maximal Euclidean norm in a bounded component $\mathbf{C}$ of $\mathbf{B}(t_0)^c$ with the largest number of vertices, and let $\mathbf{S}_{x_0}, J_{x_0}, K_{x_0}$ be as in (3.1) and (3.2). We first argue that for large s,

$$\mathbf{C} \text{ contains a vertex in } \mathbf{S}_{x_0}^c. \tag{3.5}$$

To do this, we note that there exists $c_{22} > 0$ such that

$$[0, c_{22}] \subset \{\|w\|_2 : w \in \mathcal{B}\} \subset \left[0, c_{22}^{-1}\right]. \tag{3.6}$$

Because x_0 is in $\mathbf{C}$ and $E_1(s)$ occurs, we have

$$\|x_0\|_2 \leq \max_{x\in\mathbf{B}(t_0)} \|x\|_2 \leq 2t_0 \max_{x\in\mathcal{B}} \|x\|_2 \leq 2c_{22}^{-1} t_0.$$

As $x_0 \in \mathbf{B}(t_0)^c$, we have $\|x_0\|_2 \geq (t_0/2)\max_{x\in\mathcal{B}} \|x\|_2 \geq c_{22}t_0/2$. In summary,

$$\frac{c_{22}}{2} t_0 \leq \|x_0\|_2 \leq \frac{2}{c_{22}} t_0. \tag{3.7}$$

Now for a contradiction, assume that $\mathbf{C} \subset \mathbf{S}_{x_0}$. Then from (3.3), we get

$$M(t_0) \leq C_{19}(\log \|x_0\|_2)^{2C_{18}}. \tag{3.8}$$

Combining this with (3.7), we obtain

$$M(t_0) \leq C_{19}(\log(2c_{22}^{-1} t_0))^{2C_{18}}.$$

This contradicts $M(t_0) \geq (\log t_0)^{3C_{18}}$ for large s because $t_0 \geq s$, and shows (3.5).

We have now shown that for our outcome in the probability in (3.4), (3.5) holds. Let γ be a path contained in $\mathbf{C}$ starting at x_0 that ends at a vertex outside of $\mathbf{S}_{x_0}$; we may assume only its final vertex, say p_0, is outside of $\mathbf{S}_{x_0}$. Let γ' be the continuous plane curve produced by following γ from x_0 to its last point p_0' on the boundary of $\mathbf{S}_{x_0}$ (directly before γ touches p_0). We examine the possibility that p_0' is on the left or right sides of $\mathbf{S}_{x_0}$, or on the near side.

Step 2 (The hole cannot touch the near side). The first case is that p_0' is in the near side

$$\mathbf{S}_{x_0}^n = \{v \in \mathbf{S}_{x_0} : \|v\|_2 = (1 - K_{x_0})\|x_0\|_2\}.$$

If this holds, let $x_0' = (1 - K_{x_0})x_0$, which is in $\mathbf{S}_{x_0}^n$; we will show that $T(0, x_0')$ is abnormally large. (Here we use the definition $T(y, z) = T([y], [z])$, where $[y]$ is the point of $\mathbb{Z}^d$ with $y \in [y] + [0, 1)^d$, and similarly for z.)

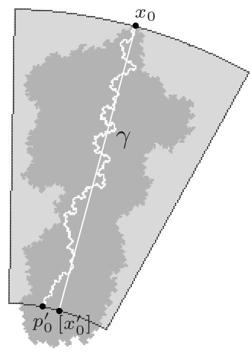

FIGURE 7. The first case of the argument supposes that $\mathbf{C}$ exits the sector $\mathbf{S}_{x_0}$ through its near side $\mathbf{S}_{x_0}^n$. The path γ in $\mathbf{C}$ starts at x_0, intersects the boundary of $\mathbf{S}_{x_0}^n$ first at $p_0' \in \mathbf{S}_{x_0}^n$, and it ends immediately after p_0' at a vertex $p_0 \notin \mathbf{S}_{x_0}$ (not pictured). Above, $x_0' = (1 - K_{x_0})x_0$, and $[x_0']$ is the closest lattice point to x_0'.

Because $p_0 \in \mathbf{C} \subset \mathbf{B}(t_0)^c$,

$$T(0, x_0') = T(0, p_0) + (T(0, x_0') - T(0, p_0)) > t_0 + (T(0, x_0') - T(0, p_0)). \tag{3.9}$$

For large s, the points x_0' and p_0 are in $3t_0\mathcal{B}$, and by occurrence of $E_2(s)$, there exists a path from $[x_0']$ to p_0 with $\|[x_0'] - p_0\|_1$ many edges whose weights are at most $C_{20} \log t_0 \leq C_{20} \log(2c_{22}^{-1}\|x_0\|_2)$ (see (3.7)). This gives $T(0, x_0') - T(0, p_0) \geq -(C_{20} \log(2c_{22}^{-1}\|x_0\|_2))\|[x_0'] - p_0\|_1$. However $\|[x_0'] - p_0\|_1 \leq \|x_0' - p_0'\|_1 + 3 \leq \sqrt{2}\|x_0' - p_0'\|_2 + 3$, and x_0', p_0' are in $\mathbf{S}_{x_0}^n$, so if s is large, then $\|x_0' - p_0'\|_2 \leq J_{x_0}\|x_0\|_2 = (\log \|x_0\|_2)^{C_{18}-3}$. Together, for large s,

$$\begin{aligned} T(0, x_0') - T(0, p_0) &\geq -(C_{20} \log(2c_{22}^{-1}\|x_0\|_2))(3 + \sqrt{2}(\log \|x_0\|_2)^{C_{18}-3}) \\ &\geq -(\log \|x_0\|_2)^{C_{18}-1}. \end{aligned}$$

Putting this in (3.9) gives

$$T(0, x_0') > t_0 - (\log \|x_0\|_2)^{C_{18}-1}. \tag{3.10}$$

To use (3.10), we relate the left side to $T(0, x_0)$. Although x_0 is not in $\mathbf{B}(t_0)$, it is the endpoint of an edge that has an endpoint in $3t_0\mathcal{B}$, so since $E_2(s)$ occurs, $T(0, x_0) \leq t_0 + C_{20} \log t_0 \leq t_0 + C_{20} \log(2c_{22}^{-1}\|x_0\|_2)$. With (3.10), we obtain for large s

$$T(0, x_0) - T(0, x_0') \leq C_{20} \log(2c_{22}^{-1}\|x_0\|_2) + (\log \|x_0\|_2)^{C_{18}-1} < 2(\log \|x_0\|_2)^{C_{18}-1}. \tag{3.11}$$

We now use a bound on passage time differences established in [5, Prop. 3.7] under the uniform curvature assumption. The result is that for some $c_{23}, C_{24}, c_{25} > 0$, any $z \in \mathbb{R}^d$ with $\|z\|_2 = 1$, and any $k, \ell \geq 0$ with $k \geq \ell$,

$$\mathbb{P}(T(0, kz) - T(0, \ell z) \geq c_{23}(k - \ell)) \geq 1 - C_{24} e^{-(k-\ell)^{c_{25}}}. \tag{3.12}$$

We put $z = x_0/\|x_0\|_2$, $k = \|x_0\|_2$, and $\ell = \|x_0'\|_2 = (1 - K_{x_0})\|x_0\|_2$ to produce the bound

$$\mathbb{P}\left(T(0, x_0) - T(0, x_0') < c_{23}(\log \|x_0\|_2)^{C_{18}}\right) \leq C_{24} \exp\left(-(\log \|x_0\|_2)^{C_{18} c_{25}}\right). \tag{3.13}$$

If we define the event $G(s)$ to be

$$G(s) = \left\{ \begin{array}{c} T(0, x_0) - T(0, x_0') \geq 2(\log \|x_0\|_2)^{C_{18}-1} \\ \text{for all } x_0 \in \mathbb{Z}^d \text{ with } \|x_0\|_2 \geq \frac{c_{22}}{2} s \end{array} \right\},$$

then, by (3.7) and (3.11), if p_0' is in the near side $\mathbf{S}_{x_0}^n$, then $G(s)^c$ must occur, and by (3.13), we get

$$\begin{aligned} \mathbb{P}(G(s)^c) &\leq \sum_{\|x_0\|_2 \geq \frac{c_{22}}{2} s} \mathbb{P}\left(T(0, x_0) - T(0, x_0') < 2(\log \|x_0\|_2)^{C_{18}-1}\right) \\ &\leq C_{24} \sum_{\|x_0\|_2 \geq \frac{c_{22}}{2} s} \exp\left(-(\log \|x_0\|_2)^{C_{18} c_{25}}\right). \end{aligned}$$

Here we have used that for large s, $2(\log \|x_0\|_2)^{C_{18}-1} < c_{23}(\log \|x_0\|_2)^{C_{18}}$. Assuming C_{18} is chosen larger than c_{25}^{-1}, we get $\mathbb{P}(G(s)^c) \to 0$ as $s \to \infty$. In summary, we can return to (3.4) and write

$$\begin{aligned} &\lim_{s \to \infty} \mathbb{P}(M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s) \\ &= \lim_{s \to \infty} \mathbb{P}\left(E_1(s) \cap E_2(s) \cap G(s) \cap \{M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s\}\right), \end{aligned} \tag{3.14}$$

observing now that any outcome in the event in the probability in (3.14) must have the property that p_0' is on the union of the left and right sides of $\mathbf{S}_{x_0}$:

$$|\angle(x_0, p_0')| = J_{x_0}. \tag{3.15}$$

Step 3 (The hole cannot touch the left and right sides). This brings us to deal with the second case, that (3.15) holds in our outcome. Here, the idea is that geodesics (optimal paths in the definition of $T(x, y)$—these exist a.s. from [1, Thm. 4.2]) between some point nearby x_0 and the origin must avoid ("go around") the component $\mathbf{C}$, and therefore deviate significantly from the straight line connecting the point to the origin. This is unlikely due to geodesic wandering estimates from [13].

Our two possible "nearby" points are $y_0, z_0 \in \mathbb{R}^2$, defined to have $\|y_0\|_2 = \|z_0\|_2 = (1 + K_{x_0})\|x_0\|_2$, $\angle(y_0, x_0) = J_{x_0}/2$, and $\angle(z_0, x_0) = -J_{x_0}/2$. Let $A_{x_0}^{(i)}, i = 1, 2$ be defined as follows.

(1) $A_{x_0}^{(1)}$ is the event that some geodesic from $[y_0]$ to 0 has a point $x \in \mathbb{R}^2$ with $\|x\|_2 \geq (1 - K_{x_0})\|x_0\|_2$ and $\angle(x, x_0) = 0$ or J_{x_0}.

(2) $A_{x_0}^{(2)}$ is the event that some geodesic from $[z_0]$ to 0 has a point $x \in \mathbb{R}^2$ with $\|x\|_2 \geq (1 - K_{x_0})\|x_0\|_2$ and $\angle(x, x_0) = -J_{x_0}$ or 0.

We claim that because (3.15) holds,

$$\text{at least one of } A_{x_0}^{(1)} \text{ or } A_{x_0}^{(2)} \text{ occurs.} \tag{3.16}$$

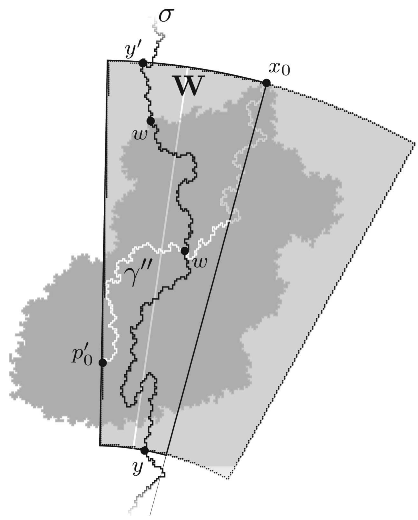

FIGURE 8. In the figure, the region $\mathbf{W}$ is the left half of the sector $\mathbf{S}_{x_0}$. In the second case considered for the argument, $\mathbf{C}$ exits $\mathbf{S}_{x_0}$ through one of its sides, the side of $\mathbf{W}$ above. The path γ plays an analogous role to the first case of the argument, excepting that p_0' is no longer on the near boundary of $\mathbf{S}_{x_0}$, and contains a subpath γ'' spanning opposite sides of $\mathbf{W}$. Under the event $A_{x_0}^{(1)}$, planarity forces a geodesic σ joining $[y_0]$ (not pictured) to the origin to cross γ'' at a vertex w. The path σ is further decomposed at the last point z' on σ with $\|z'\|_2 = \|x_0\|_2$ and the first point z on σ with $\|z\|_2 = (1 - K_{x_0})\|x_0\|_2$, denoted y' and y respectively.

To see why, let us assume first that $\angle(p_0', x_0) = J_{x_0}$. Then γ', which we defined in the paragraph following (3.8), contains a segment γ'' which crosses the region

$$\mathbf{W} = \left\{ v \in \mathbb{R}^2 : \frac{\|v\|_2}{\|x_0\|_2} \in [(1 - K_{x_0}), 1],\ \angle(v, x_0) \in [0, J_{x_0}] \right\}$$

between its two side boundaries; see Fig. 8. This is because γ' cannot exit $\mathbf{S}_{x_0}$ through the far or near boundaries. Assume for a contradiction that $A_{x_0}^{(1)}$ does not occur, and let σ be any geodesic from $[y_0]$ to 0. Observe that for large s, we have $\angle([y_0], x_0) \in (0, J_{x_0})$. The segment of σ from $[y_0]$ to its first point y with $\|y\|_2 = (1 - K_{x_0})\|x_0\|_2$ cannot contain any x with $\angle(x, x_0) = 0$ or J_{x_0}, so it must contain a segment σ' of σ (starting at its last point y' with $\|y'\|_2 = \|x_0\|_2$ before y and ending at y) that crosses $\mathbf{W}$ from its far boundary to its near boundary. By planarity, σ' must intersect γ', and they must intersect at a vertex w. We know $w \in \mathbf{C}$, so $T(0, w) > t_0$. Furthermore, $T(0, y_0) \geq T(0, w)$, so $[y_0] \notin \mathbf{B}(t_0)$. But $[y_0] \notin \mathbf{C}$ by maximality of x_0, so $[y_0]$ is in a different component of $\mathbf{B}(t_0)^c$. Starting from $[y_0]$, the geodesic σ must therefore touch some $\hat{w} \in \mathbf{B}(t_0)$ before it reaches w. This gives a contradiction because then $t_0 \geq T(0, \hat{w}) \geq T(0, w) > t_0$. We conclude that $A_{x_0}^{(1)}$ occurs. If we suppose that $\angle(p_0', x_0) = -J_{x_0}$ instead, a similar argument shows that $A_{x_0}^{(2)}$ occurs.

Returning to (3.14), the last paragraph plus a union bound gives

$$\lim_{s\to\infty} \mathbb{P}(M(t) \geq (\log t)^{3C_{18}} \text{ for some } t \geq s) \leq \lim_{s\to\infty} \sum_{x_0\in\mathbb{Z}^2:\|x_0\|_2\geq \frac{c_{22}}{2}s} \mathbb{P}(A_{x_0}^{(1)} \cup A_{x_0}^{(2)}).$$

To complete the proof of Theorem 1.6, we will show that this limit is zero, and to do this, we will prove that

$$\sum_{x_0\in\mathbb{Z}^2} \mathbb{P}(A_{x_0}^{(1)}) < \infty. \tag{3.17}$$

A symmetric argument will establish the same bound for the sum of $\mathbb{P}(A_{x_0}^{(2)})$, and this will finish the proof.

Assertion (3.17) will follow from a lemma that summarizes some estimates from [13]. If $x, y \in \mathbb{Z}^d$, we write

$$\mathrm{out}(y,x) = \{z \in \mathbb{Z}^d : T(y,z) = T(y,x) + T(x,z)\}$$

for the set of vertices z such that a geodesic from y to z goes through x. The lemma states that with high probability, vertices in $\mathrm{out}(0,x)$ have small angle from x. In [13], this is used to show that the origin has a "$r^{-1/4}$-straight geodesic tree." The argument from [13] assumes that the distribution of τ_e is continuous, but this is not needed. Only the uniform curvature assumption is required. Recall Definition 1.4, which introduces the number η.

Lemma 3.1. *Let $p \in (0, 1/(2\eta))$. There exist $C_{26}, c_{27} > 0$ such that for any $r \geq 1$,*

$$\begin{aligned}\mathbb{P}(|\angle(x,z)| \leq C_{26}\|x\|_2^{-p} \text{ for all } z \in \mathrm{out}(0,x) \text{ and } x \text{ with } \|x\|_2 \geq r)\\ \geq 1 - C_{26}\exp\left(-r^{c_{27}}\right).\end{aligned}$$

Proof. The proof is nearly the same as that of [13, Prop. 3.2], so we omit some details. For a vertex $x \neq 0$, let $\mathbf{C}_x$ be the sector portion

$$\mathbf{C}_x = \{z \in \mathbb{Z}^d : g(z) \in [g(x) - g(x)^{1-\eta p}, 2g(x)], |\angle(z,x)| \leq g(x)^{-p}\}.$$

The vertices in the boundary set $\{y \in \mathbf{C}_x^c : \exists z \in \mathbf{C}_x \text{ such that } \|z-y\|_1 = 1\}$ split into three sets: $\partial_i\mathbf{C}_x$ is those y with $g(y) < g(x) - g(x)^{1-\eta p}$, $\partial_o\mathbf{C}_x$ is those y with $g(y) > 2g(x)$, and $\partial_s\mathbf{C}_x$ is those y with $\angle(x,y) > g(x)^{-p}$. Let G_x be the event $\{\mathrm{out}(0,x) \cap (\partial_i\mathbf{C}_x \cup \partial_s\mathbf{C}_x) \neq \emptyset\}$. (The function g was defined below (1.2).) Then the argument leading to [13, Eq. (3.3)] gives that for some C_{28}, c_{29}, we have $\mathbb{P}(G_x) \leq C_{28}\|x\|_2^d \exp\left(-c_{29}\|x\|_2^{1/2-\eta p}\right)$. (The only difference is that [13] takes $\eta = 2$ but we have general η.) By a union bound, if $c_{30} < 1/2 - \eta p$,

$$\mathbb{P}(G_x \text{ occurs for some } x \in \mathbb{Z}^d \text{ with } \|x\|_2 \geq r) \leq C_{31}\exp\left(-r^{c_{30}}\right). \tag{3.18}$$

Fix an outcome in the event $\cap_{\|x\|_2\geq r} G_x^c$ and let $x \in \mathbb{Z}^d$ with $\|x\|_2 \geq r$. If $z \in \mathrm{out}(0,x)$, consider a geodesic from 0 to z that contains x, and define a sequence of points inductively by $x_0 = x$, and for $i \geq 1$, x_i is the first point of the geodesic after x_{i-1} that lies in $\partial_o\mathbf{C}_{x_{i-1}}$. (It must touch this set and not the set $\partial_i\mathbf{C}_{x_{i-1}} \cup \partial_s\mathbf{C}_{x_{i-1}}$ if it leaves $\mathbf{C}_{x_{i-1}}$ because $G_{x_{i-1}}^c$ occurs.) If such a point does not exist for a particular $i = I$, because the geodesic does not leave $\mathbf{C}_{x_{i-1}}$ before touching z, we set $x_I = x_{I+1} = \cdots = z$. Because x_i is adjacent to $\mathbf{C}_{x_{i-1}}$, we have $|\angle(x_i, x_{i-1})| \le C_{32}/\|x_{i-1}\|_2^p$, and so

$$|\angle(z,x)| \le \sum_{i=1}^{\infty} |\angle(x_i, x_{i-1})| \le C_{32} \sum_{i=1}^{I} \|x_{i-1}\|_2^{-p}.$$

However, for $i = 1, \ldots, I-1$, we have $\|x_{i-1}\|_2 \ge C_{33}^{i-1}\|x\|_2$ for some $C_{33} > 1$, since $x_{i-1} \in \partial_o \mathbf{C}_{x_{i-2}}$. Therefore $|\angle(z,x)| \le C_{32}\|x\|_2^{-p} \sum_{i=1}^{\infty} C_{33}^{-p(i-1)}$. In other words, for $C_{26} = C_{32} \sum_{i=1}^{\infty} C_{33}^{-p(i-1)}$, any outcome in $\cap_{\|x\|_2 \ge r} G_x^c$ has $|\angle(z,x)| \le C_{26}\|x\|_2^{-p}$ so long as $\|x\|_2 \ge r$ and $z \in \mathrm{out}(0,x)$. The estimate (3.18) finishes the proof. $\square$

Using Lemma 3.1, we can show (3.17), and therefore finish the proof of Theorem 1.6. Suppose that $A_{x_0}^{(1)}$ occurs. Choose a point $x \in \mathbb{R}^2$ such that $\|x\|_2 \ge (1-K_{x_0})\|x_0\|_2$ and $\angle(x,x_0) = 0$ or J_{x_0}, but that x is on a geodesic from $[y_0]$ to 0. Let x' be a vertex on this geodesic such that $\|x-x'\|_1 \le 1$. The law of sines from trigonometry implies that if $\angle_{[y_0]}(0,x')$ is the angle between 0 and x' as measured from $[y_0]$, then

$$\|x'\|_2 \sin|\angle([y_0],x')| = \|[y_0]-x'\|_2 \sin|\angle_{[y_0]}(0,x')|. \tag{3.19}$$

To estimate these quantities, we observe first that for large $\|x_0\|_2$, we have

$$\|x'\|_2 \ge \|x_0\|_2/2. \tag{3.20}$$

Next, because $|\angle(x,y_0)| = J_{x_0}/2$, we have

$$|\angle([y_0],x')| \in \left(\frac{J_{x_0}}{3}, 2\frac{J_{x_0}}{3}\right) \tag{3.21}$$

so long as $\|x_0\|_2$ is large enough. In particular, if $\|x_0\|_2$ is large, then $|\angle([y_0],x')|$ is small, and so

$$\sin|\angle([y_0],x)| \ge \frac{|\angle([y_0],x')|}{2} \ge \frac{J_{x_0}}{6}. \tag{3.22}$$

The term $\sin|\angle_{[y_0]}(0,x')|$ can be bounded using Lemma 3.1. For $u \in \mathbb{Z}^d$ and $p \in (0, 1/(2\eta))$ fixed, write $F_u(r)$ for the event described in Lemma 3.1, translated in the natural way so that the origin is mapped to u. Precisely, if T_u is the translation of $\mathbb{R}^d$ such that $T_u(0) = u$, then $F_u(r)$ is the event that the image configuration $(\tau_{T_u^{-1}(e)})$ is in the event described in Lemma 3.1. We observe that $\|[y_0]-x'\|_2 \ge \|y_0 - x\|_2 - \|[y_0]-y_0\|_2 - \|x-x'\|_2 \ge J_{x_0}\|x_0\|_2/2 - 3$, so if $\|x_0\|_2$ is large and $F_{[y_0]}(r)$ occurs for $r = J_{x_0}\|x_0\|_2/3$, then we must have

$$|\angle_{[y_0]}(0,x')| \le C_{26}\|[y_0]-x'\|_2^{-p}. \tag{3.23}$$

Putting this, (3.20), and (3.22) into (3.19) produces

$$\frac{1}{12}(\log\|x_0\|_2)^{C_{18}-3} = \frac{\|x_0\|_2 J_{x_0}}{12} \le \|[y_0]-x'\|_2 \sin|\angle_{[y_0]}(0,x')| \le C_{26}\|[y_0]-x'\|_2^{1-p}. \tag{3.24}$$

For large $\|x_0\|_2$, we conclude

$$\|y_0 - x\|_2 \ge c_{34}(\log\|x_0\|_2)^{\frac{C_{18}-3}{1-p}}. \tag{3.25}$$

To proceed from (3.25), we assume for a contradiction that $F_{[y_0]}(r)$ occurs (so that (3.25) holds) and consider two cases. If $\|x\|_2 \le \|y_0\|_2$, then $\|x\|_2/\|x_0\|_2 \in [1-K_{x_0}, 1+K_{x_0}]$. If $\angle(x,x_0)=0$, then

$$\|y_0-x\|_2 \le \|y_0-x_0\|_2+\|x-x_0\|_2 \le \left(\frac{J_{x_0}}{2}+K_{x_0}\right)\|x_0\|_2+K_{x_0}\|x_0\|_2 \le 3K_{x_0}\|x_0\|_2.$$

By symmetry, the inequality $\|y_0-x\|_2 \le 3K_{x_0}\|x_0\|_2$ also holds if $\angle(x,x_0)=J_{x_0}$. Putting it in (3.25), we find

$$3(\log\|x_0\|_2)^{C_{18}} \ge c_{34}(\log\|x_0\|_2)^{\frac{C_{18}-3}{1-p}},$$

which is false if $C_{18} > 3/p$ and $\|x_0\|_2$ is large. Otherwise, if $\|x\|_2 \ge \|y_0\|_2$, then

$$\|x'\|_2 \ge \|[y_0]\|_2 - 1 - \sqrt{2} \ge \|[y_0]\|_2 - 3. \tag{3.26}$$

Because $|\angle_{[y_0]}(0,x')| + |\angle(x',[y_0])| + |\angle_{x'}(0,[y_0])| = \pi$, we see for large $\|x_0\|_2$ from (3.21) and (3.23) that $|\angle_{x'}(0,[y_0])| \ge 3\pi/4$, and so $\cos|\angle_{x'}(0,[y_0])| \le -1/\sqrt{2}$. The law of cosines along with (3.24) and (3.26) then gives for large $\|x_0\|_2$

$$\begin{aligned}\|[y_0]\|_2^2 &= \|x'\|_2^2 + \|[y_0]-x'\|_2^2 - 2\|x'\|_2\|[y_0]-x'\|_2\cos\angle_{x'}(0,[y_0])\\ &\ge \|x'\|_2^2 + \sqrt{2}\|x'\|_2\|[y_0]-x'\|_2\\ &\ge \|x'\|_2^2 + 7\|x'\|_2\\ &\ge (\|[y_0]\|_2-3)^2 + 7\|[y_0]\|_2 - 21.\end{aligned}$$

This is a contradiction if $\|x_0\|_2$ is large.

We conclude that if $\|x_0\|_2$ is sufficiently large, then $A_{x_0}^{(1)} \subset F_{[y_0]}(r)^c$ with $r = J_{x_0}\|x_0\|_2/3$. Lemma 3.1 gives the bound

$$\mathbb{P}(A_{x_0}^{(1)}) \le \mathbb{P}(F_{[y_0]}(r)^c) \le C_{26}\exp\left(-(J_{x_0}\|x_0\|_2/3)^{c_{27}}\right).$$

This is summable over $x_0 \in \mathbb{Z}^2$ so long as $C_{18} > 3 + c_{27}^{-1}$. This completes the proof of (3.17).

4. Proof of Theorem 1.8

The proof of Theorem 1.8 is like that of Theorem 1.6, and will use similar constructions, so we give fewer details and focus on the modifications needed to apply the argument. There are two main differences. First, instead of using the bound (3.12) on passage time differences (which requires the uniform curvature assumption), we will use a general concentration inequality. Second, instead of using Lemma 3.1 on the straightness of geodesics (also requiring curvature), we will use Kesten's lemma.

The concentration inequality states the there exists $C_{35} > 0$ such that for all large $x \in \mathbb{Z}^d$,

$$\mathbb{P}\left(|T(0,x)-g(x)| \ge C_{35}\sqrt{g(x)\log g(x)}\right) \le \|x\|_1^{-100}. \tag{4.1}$$

This inequality follows from standard results. First, it suffices to prove it with $\sqrt{g(x)\log g(x)}$ replaced by $\sqrt{\|x\|_1\log\|x\|_1}$. In this form, it follows from the result [7, Prop. 1.1], which says that for some $C_{36} > 0$, we have $0 \le \mathbb{E}T(0,x) - g(x) \le C_{36}\sqrt{\|x\|_1\log\|x\|_1}$, and [6, Thm. 1.1], which says that $\mathbb{P}(|T(0,x)-\mathbb{E}T(0,x)| \ge \lambda\sqrt{\|x\|_1/\log\|x\|_1}) \le e^{-c_{37}\lambda}$ for some constant $c_{37} > 0$ and all $\lambda \ge 0$ and nonzero

$x \in \mathbb{Z}^d$. From these two, we just have to choose $\lambda = 2C_{35} \log \|x\|_1$ for large enough C_{35}.

The second tool, Kesten's lemma [10, Prop. 5.8], states that there exist $a, c_{38} > 0$ such that

(4.2)

$$\mathbb{P}(\exists \text{ edge-self-avoiding path } \gamma \text{ containing } 0 \text{ with } \#\gamma \geq n \text{ but } T(\gamma) < an) \leq e^{-c_{38}n}.$$

Here, $\#\gamma$ is the number of edges in γ. This result will allow us to show in (4.10) that if a geodesic deviates too far from a straight line, it must have a long segment with high passage time.

Before giving the proof, we first outline what changes are necessary. In this section, since we are not assuming uniform curvature, (3.12) and Lemma 3.1 are no longer available. However, we are still able to show that if a hole is too large, then it will take a large passage-time to cross one of the sides of a large sector, which will violate (4.1) and (4.2).

Step 1 (Setup of the proof). As in the proof of Theorem 1.6, we define events $E_i(s)$ for $s \geq 0$ and a constant $C_{39} > 0$ as

$$E_1(s) = \left\{ \frac{1}{2}\mathcal{B} \subset \frac{1}{t}\widetilde{\mathbf{B}}(t) \subset 2\mathcal{B} \text{ for all } t \geq s \right\}$$

$$E_2(s) = \{\tau_e \leq C_{39} \log t \text{ for all } e \text{ with an endpoint in } 3t\mathcal{B} \text{ and all } t \geq s\}$$

$$E_3(s) = \left\{ \begin{array}{c} |T(0,x) - g(x)| \leq C_{35}\sqrt{g(x)\log g(x)} \text{ for all} \\ \text{integer } x \in 3t\mathcal{B} \setminus ((t/3)\mathcal{B}) \text{ and all } t \geq s \end{array} \right\}.$$

As in (3.4), for some C_{39} large enough, and any $C_{40} > 0$,

$$\begin{aligned} &\lim_{s\to\infty} \mathbb{P}\left(M(t) \geq C_{40} t \log t \text{ for some } t \geq s\right) \\ (4.3) \qquad &= \lim_{s\to\infty} \mathbb{P}\left(E_1(s) \cap E_2(s) \cap \{M(t) \geq C_{40} t \log t \text{ for some } t \geq s\}\right). \end{aligned}$$

Using (4.1) with a union bound, we obtain

$$\mathbb{P}(E_3(s)^c) \leq \sum_{x \in ((s/3)\mathcal{B})^c} \|x\|_1^{-100} \to 0 \text{ as } s \to \infty. \tag{4.4}$$

Last, we let $E_4(s)$ be the event that, for all $t \geq s$, and all vertices $x \in 3t\mathcal{B} \setminus ((t/3)\mathcal{B})$, any edge-self-avoiding path Γ containing x with at least $(12C_{35}/a)\sqrt{g(x)\log g(x)}$ many edges satisfies $T(\Gamma) \geq 12C_{35}\sqrt{g(x)\log g(x)}$. To prove that

$$\mathbb{P}(E_4(s)^c) \to 0 \text{ as } s \to \infty, \tag{4.5}$$

we use (4.2) with a union bound. We obtain

$$\mathbb{P}(E_4(s)^c) \leq \sum_{x \in \left(\frac{s}{3}\mathcal{B}\right)^c} e^{-c_{38} 12 C_{35} a^{-1} \sqrt{g(x)\log g(x)}} \to 0 \text{ as } s \to \infty.$$

This shows (4.5). Putting (4.4) and (4.5) into (4.3), we find

$$\begin{aligned} &\lim_{s\to\infty} \mathbb{P}\left(M(t) \geq C_{40} t \log t \text{ for some } t \geq s\right) \\ (4.6) \qquad &= \lim_{s\to\infty} \mathbb{P}\left(\left(\cap_{i=1}^4 E_i(s)\right) \cap \{M(t) \geq C_{40} t \log t \text{ for some } t \geq s\}\right). \end{aligned}$$

The rest of the proof serves to show that if C_{40} is large, then (4.6) is zero. To do this, we choose an outcome in the event in (4.6), and let $t_0 \geq s$. Pick x_0 as any vertex with maximal value of $g(x_0)$ in a bounded component $\mathbf{C}$ of $\mathbf{B}(t_0)^c$ with the largest number of vertices. Analogously to (3.1), let

$$\mathbf{S}_{x_0} = \left\{ v \in \mathbb{R}^2 : |\angle(v, x_0)| \le J_{x_0} \text{ and } 1 - K_{x_0} \le \frac{g(v)}{g(x_0)} \le 1 \right\},$$

where

$$K_{x_0} = \frac{3C_{35}\sqrt{g(x_0)\log g(x_0)}}{g(x_0)}$$

and

$$J_{x_0} = \frac{64}{ac_{22}} K_{x_0}.$$

By a similar argument to that which gave (3.5), if s is large, because $M(t_0) \ge C_{40} t_0 \log t_0$,

$$\mathbf{C} \text{ contains a vertex in } \mathbf{S}_{x_0}^c,$$

so long as C_{40} is fixed to be large enough. Because of this, we can find a path γ contained in $\mathbf{C}$ starting at x_0 that ends at a vertex outside of $\mathbf{S}_{x_0}$; we may assume only its final vertex, say p_0, is outside of $\mathbf{S}_{x_0}$. We also let γ' be the continuous plane curve produced by following γ from x_0 to its last point p_0' on the boundary of $\mathbf{S}_{x_0}$ (directly before γ touches p_0). As we have done in the last section, we must exclude the possibility that p_0' is on the left or right sides of $\mathbf{S}_{x_0}$, or on the near side. The point p_0' cannot be on the far side only because $g(x_0)$ is maximal among vertices in $\mathbf{C}$.

Step 2 (The hole cannot touch the near side). The first case is that p_0' is in the near side

$$\mathbf{S}_{x_0}^n = \{v \in \mathbf{S}_{x_0} : g(v) = (1 - K_{x_0}) g(x_0)\}.$$

If s is large, then $p_0 \in 3t_0\mathcal{B} \setminus ((t_0/3)\mathcal{B})$, so since $E_3(s)$ occurs, we have for some $C_{41} > 0$

$$\begin{aligned} T(0, p_0) &\le g(p_0) + C_{35}\sqrt{g(p_0)\log g(p_0)} \\ &\le C_{41} + g(p_0') + C_{35}\sqrt{g(p_0')\log g(p_0')} \\ &= C_{41} + g(x_0) - 3C_{35}\sqrt{g(x_0)\log g(x_0)} + C_{35}\sqrt{g(p_0')\log g(p_0')} \\ &\le g(x_0) - 2C_{35}\sqrt{g(x_0)\log g(x_0)}. \end{aligned}$$

Because $T(0, x_0) \ge g(x_0) - C_{35}\sqrt{g(x_0)\log g(x_0)}$, we obtain

$$T(0, x_0) - T(0, p_0) \ge C_{35}\sqrt{g(x_0)\log g(x_0)} \tag{4.7}$$

as long as s is large. On the other hand, $p_0 \notin \mathbf{B}(t_0)$, so $T(0, p_0) > t_0$. Furthermore, x_0 is an endpoint of an edge with an endpoint in $\mathbf{B}(t_0)$, and this edge must have weight at most $C_{39} \log t_0$ because $E_2(s)$ occurs. Therefore

$$T(0, x_0) - T(0, p_0) \le C_{39} \log t_0 + t_0 - t_0 = C_{39} \log t_0.$$

Because $t_0 \le (2/c_{22})\|x_0\|_2$ from (3.7), this contradicts (4.7).

Step 3 (The hole cannot touch the left and right sides). The second case is that p_0' satisfies $|\angle(x_0, p_0')| = J_{x_0}$. We will suppose that $\angle(x_0, p_0') = J_{x_0}$, as the other possibility is dealt with using a similar argument. Let $y_0 \in \mathbb{R}^2$ satisfy $\angle(y_0, x_0) = J_{x_0}/2$ and $g(y_0) = g(x_0)$, and choose a vertex $\bar{y}_0$ with $g(\bar{y}_0) > g(y_0)$ but $\|y_0 - \bar{y}_0\|_1 = 1$. Let σ be any geodesic from $\bar{y}_0$ to 0. As in the proof of (3.16), as σ proceeds from $\bar{y}_0$ to 0, planarity implies it must touch one of the rays

$$\mathbf{M} = \{v \in \mathbb{R}^2 : \angle(v, x_0) = 0\} \text{ or } \mathbf{M}' = \{v \in \mathbb{R}^2 : \angle(v, x_0) = J_{x_0}\}$$

before touching the set $\mathbf{B}' = \{v : g(v) = (1 - K_{x_0})g(x_0)\}$. Indeed, if this were false, then because the curve γ' connecting x_0 to p_0' must contain a segment crossing the region

$$\mathbf{W} = \left\{ v \in \mathbb{R}^2 : \frac{g(v)}{g(x_0)} \in [(1 - K_{x_0}), 1], \angle(v, x_0) \in [0, J_{x_0}] \right\}$$

between its two side boundaries, σ would have to intersect γ' at a vertex w. As in the last section, this gives a contradiction because $w \in \mathbf{C}$, so $T(0, w) > t_0$, but because σ originates outside of $\mathbf{C}$, it must touch some $\hat{w} \in \mathbf{B}(t_0)$ before reaching w, and so $t_0 \geq T(0, \hat{w}) \geq T(0, w) > t_0$.

Without loss of generality, we suppose that σ touches some $p_1 \in \mathbf{M}'$ before some $p_2 \in \mathbf{B}'$. Let $\bar{p}_1$ be the vertex we encounter on σ directly before p_1 as we proceed from $\bar{y}_0$ to 0, and let $\bar{p}_2$ be the vertex we encounter on σ directly after p_2. Because $p_2 \in \mathbf{B}'$, we have $g(p_2) = g(x_0) - 3C_{35}\sqrt{g(x_0) \log g(x_0)}$. The event $E_2(s) \cap E_3(s)$ occurs, so for large s,

$$\begin{aligned}
T(0, \bar{p}_2) &\geq g(\bar{p}_2) - C_{35}\sqrt{g(\bar{p}_2) \log g(\bar{p}_2)} \\
&\geq g(p_2) - C_{35}\sqrt{g(p_2) \log g(p_2)} - C_{42} \\
&= g(x_0) - 3C_{35}\sqrt{g(x_0) \log g(x_0)} - C_{35}\sqrt{g(p_2) \log g(p_2)} - C_{42} \\
&\geq g(x_0) - 4C_{35}\sqrt{g(x_0) \log g(x_0)}.
\end{aligned}$$

Here, $C_{42} > 0$ is a constant. Because $\bar{p}_1$ appears first on σ, we have $T(0, \bar{p}_1) \geq T(0, \bar{p}_2)$, so

$$T(0, \bar{p}_1) \geq g(x_0) - 4C_{35}\sqrt{g(x_0) \log g(x_0)}. \tag{4.8}$$

To obtain an upper bound on $T(0, \bar{p}_1)$, we use the occurrence of $E_2(s) \cap E_3(s)$ to estimate

$$\begin{aligned}
T(0, \bar{p}_1) &= T(0, \bar{y}_0) - T(\bar{y}_0, \bar{p}_1) \\
&\leq T(0, y_0) + C_{39} \log t_0 - T(\bar{y}_0, \bar{p}_1) \\
&\leq g(x_0) + C_{35}\sqrt{g(x_0) \log g(x_0)} + C_{39} \log t_0 - T(\bar{y}_0, \bar{p}_1). \qquad (4.9)
\end{aligned}$$

Any path from $\bar{y}_0$ to $\bar{p}_1$ must have at least $\|\bar{y}_0 - \bar{p}_1\|_1$ many edges, and by (3.6), if s is large,

$$\begin{aligned}
\|\bar{y}_0 - \bar{p}_1\|_1 \geq \|y_0 - p_1\|_2 - 2 \geq \sin\left(\frac{J_{x_0}}{2}\right) \|y_0\|_2 - 2 \\
&\geq \frac{J_{x_0}}{4}\|y_0\|_2 - 2 \\
&= \frac{3}{4} \cdot \frac{64}{ac_{22}} C_{35}\sqrt{g(x_0) \log g(x_0)} \frac{\|y_0\|_2}{g(y_0)} - 2 \\
&\geq \frac{3}{8} c_{22} \cdot \frac{64}{ac_{22}} C_{35}\sqrt{g(x_0) \log g(x_0)}.
\end{aligned}$$

If s is large, this is bigger than $(12C_{35}/a)\sqrt{g(\bar{y}_0)\log g(\bar{y}_0)}$, so since $E_4(s)$ occurs,

$$T(\bar{y}_0, \bar{p}_1) \geq 12C_{35}\sqrt{g(\bar{y}_0)\log g(\bar{y}_0)} \geq 6C_{35}\sqrt{g(x_0)\log g(x_0)}. \tag{4.10}$$

Returning to (4.9), for large s, we get

$$T(0, \bar{p}_1) \leq g(x_0) + (C_{35} - 6C_{35})\sqrt{g(x_0)\log g(x_0)} + C_{39}\log t_0.$$

This contradicts (4.8) for large s, since $t_0 \leq (2/c_{22})\|x_0\|_2$ from (3.7).

Acknowledgments

The authors thank anonymous reviewers for suggestions that improved the paper. They also thank Tzu-Han Chou for very careful reading and pointing out some mistakes in a previous version.

References

[1] Antonio Auffinger, Michael Damron, and Jack Hanson, *50 years of first-passage percolation*, University Lecture Series, vol. 68, American Mathematical Society, Providence, RI, 2017, DOI 10.1090/ulect/068. MR3729447

[2] Ronnie Barequet, Gill Barequet, and Günter Rote, *Formulae and growth rates of high-dimensional polycubes*, Combinatorica **30** (2010), no. 3, 257–275, DOI 10.1007/s00493-010-2448-8. MR2728490

[3] G. Bouch, *The expected perimeter in Eden and related growth processes*, J. Math. Phys. **56**.

[4] Raphaël Cerf and Marie Théret, *Weak shape theorem in first passage percolation with infinite passage times* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 3, 1351–1381, DOI 10.1214/15-AIHP686. MR3531712

[5] Michael Damron, Jack Hanson, and Wai-Kit Lam, *The size of the boundary in first-passage percolation*, Ann. Appl. Probab. **28** (2018), no. 5, 3184–3214, DOI 10.1214/18-AAP1388. MR3847985

[6] Michael Damron, Jack Hanson, and Philippe Sosoe, *Subdiffusive concentration in first-passage percolation*, Electron. J. Probab. **19** (2014), no. 109, 27, DOI 10.1214/EJP.v19-3680. MR3286463

[7] Michael Damron and Naoki Kubota, *Rate of convergence in first-passage percolation under low moments*, Stochastic Process. Appl. **126** (2016), no. 10, 3065–3076, DOI 10.1016/j.spa.2016.04.001. MR3542626

[8] Michael Damron, Wai-Kit Lam, and Xuan Wang, *Asymptotics for 2D critical first passage percolation*, Ann. Probab. **45** (2017), no. 5, 2941–2970, DOI 10.1214/16-AOP1129. MR3706736

[9] Michael Damron, Firas Rassoul-Agha, and Timo Seppäläinen, *Random growth models*, Notices Amer. Math. Soc. **63** (2016), no. 9, 1004–1008, DOI 10.1090/noti1400. MR3525713

[10] Harry Kesten, *Aspects of first passage percolation*, École d'été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 125–264, DOI 10.1007/BFb0074919. MR876084

[11] F. Leyvraz, *The "active perimeter" in cluster growth models: a rigorous bound*, J. Phys. A. **18**, L941–L945.

[12] Fedor Manin, Érika Roldán, and Benjamin Schweinhart, *Topology and local geometry of the Eden model*, Discrete Comput. Geom. **69** (2023), no. 3, 771–799, DOI 10.1007/s00454-022-00474-w. MR4555869

[13] C. M. Newman, *A surface view of first-passage percolation*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, pp. 1017–1023.

[14] Ádám Timár, *Boundary-connectivity via graph theory*, Proc. Amer. Math. Soc. **141** (2013), no. 2, 475–480, DOI 10.1090/S0002-9939-2012-11333-4. MR2996951

School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, Georgia 30332
Email address: mdamron6@protonmail.com

Center for Statistics and Machine Learning, Princeton University, 26 Prospect Ave., Princeton, New Jersey 08544
Email address: julian.thomas.gold@gmail.com

Department of Mathematics, National Taiwan University, Astronomy Mathematics Building 5F, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan (R.O.C.)
Email address: waikitlam@ntu.edu.tw

Department of Mathematics, University of Utah, 155 South 1400 East, JWB 233, Salt Lake City, Utah 84112
Email address: xiao.shen@utah.edu