

## SHORTEST PATH THROUGH RANDOM POINTS

BY SUNG JIN HWANG<sup>\*,1,3</sup>, STEVEN B. DAMELIN<sup>†,2</sup>  
AND ALFRED O. HERO III<sup>‡,1</sup>

*Google*<sup>\*</sup>, *the American Mathematical Society*<sup>†</sup> and *University of Michigan*<sup>‡</sup>

Let  $(M, g_1)$  be a complete  $d$ -dimensional Riemannian manifold for  $d > 1$ . Let  $\mathcal{X}_n$  be a set of  $n$  sample points in  $M$  drawn randomly from a smooth Lebesgue density  $f$  supported in  $M$ . Let  $x, y$  be two points in  $M$ . We prove that the normalized length of the power-weighted shortest path between  $x, y$  through  $\mathcal{X}_n$  converges almost surely to a constant multiple of the Riemannian distance between  $x, y$  under the metric tensor  $g_p = f^{2(1-p)/d} g_1$ , where  $p > 1$  is the power parameter.

**1. Introduction.** The shortest path problem [see, e.g., [Cormen et al. \(2009\)](#), [Dijkstra \(1959\)](#)] is of interest both in theory and in applications since it naturally arises in combinatorial optimization problems, such as optimal routing in communication networks, and efficient algorithms exist to solve the problem. In this paper, we are interested in the shortest paths over random sample points embedded in Euclidean and Riemannian spaces.

Many graph structures over Euclidean sample points have been studied in the context of the Beardwood–Halton–Hammersley (BHH) theorem and its extensions. The BHH theorem states that the law of large numbers (LLN) holds for certain spanning graphs over random samples. Such graph structures include the travelling salesman path (TSP), the minimal spanning tree (MST) and the nearest neighbor graphs ( $k$ -NNG). See [Steele \(1997\)](#) and [Yukich \(1998\)](#). The BHH theorem applies to graphs that span all of the points in the random sample. This paper establishes a BHH-type theorem for shortest paths between any two points.

In the last few years, the asymptotic theory for spanning graphs such as the MST, the  $k$ -NNG and the TSP has been extended to the Riemannian case, for example, [Costa and Hero \(2004\)](#) extended the MST asymptotics in the context of entropy and intrinsic dimensionality estimation. More general non-Euclidean ex-

---

Received February 2013; revised November 2015.

<sup>1</sup>Supported in part by ARO Grant W911NF-15-1-0479, ARO Grant W911NF-09-1-0310, and NSF Grant CCF-1217880.

<sup>2</sup>Supported in part by the American Mathematical Society and grant: “Computational Research Initiative in Imaging and Remote Sensing-7-2011,” Center for High Performance Computing, South Africa.

<sup>3</sup>Most of the work was done while the author was a student at University of Michigan.

*MSC2010 subject classifications.* Primary 60F15; secondary 60C05, 53B21.

*Key words and phrases.* Shortest path, power-weighted graph, Riemannian geometry, conformal metric.

tensions have been established by Penrose and Yukich (2013). This paper extends the BHH theorem in a different direction: the shortest path between random points in a Riemannian manifold.

The asymptotic properties of paths through random Euclidean sample points have been studied mainly in first-passage percolation (FPP) models [Hammersley (1966)]. Shortest paths have been studied in FPP models in the context of first passage time or travel time with lattice models [Kesten (1987)] or (homogeneous) continuum models [Howard and Newman (1997)]. Under the FPP lattice model, LaGatta and Wehr (2010) extended these results to the non-Euclidean case where interpoint distances are determined by a translation-invariant random Riemannian metric in  $\mathbb{R}^d$ . This paper makes a contribution in a different direction. We assume a nonhomogeneous continuum model and establish convergence of the shortest path lengths to density-dependent deformed Riemannian distances. The convergent limit reduces to the result of Howard and Newman (2001) when specialized to a homogeneous Euclidean continuum model.

**2. Main results.** In this paper, a smooth function is an infinitely differentiable function, that is,  $f \in C^\infty$ . A smooth manifold means its transition maps are smooth.

Let  $(M, g_1)$  be a smooth  $d$ -dimensional Riemannian manifold without boundary with Riemannian metric tensor  $g_1$  and  $d > 1$ . Recall that a Riemannian metric tensor, often simply called a Riemannian metric, on a manifold is a family of positive definite inner products on the tangent spaces of the manifold. When  $M = \mathbb{R}^d$ ,  $g_1$  is the standard Euclidean inner product. The use of the subscript on  $g_1$  will become clear shortly.

Consider a probability space  $(M, \mathbb{B}, P)$  where  $P$  is a probability distribution over Borel subsets  $\mathbb{B}$  of the sample space  $M$ . Assume that the distribution has a Lebesgue probability density function (p.d.f.)  $f$  with respect to  $g_1$ . Let  $X_1, X_2, \dots$  denote an i.i.d. sequence drawn from this density and let the first  $n$  samples from this sequence be denoted by  $\mathcal{X}_n = \{X_1, \dots, X_n\}$ . The sequence  $\mathcal{X}_n$  will be associated with the nodes in a undirected simple graph whose edges have weight equal to the power weighted Euclidean distance between pairs of nodes. We will use indexing by  $n$  of a generic nonrandom point  $x_n \in M$ . This point is not related in any way to the random variable  $X_n$ . For realizations, we will use the notation  $X_n(\omega)$  where  $\omega$  is an elementary outcome in the sample space.

For  $p > 1$ , called the power parameter, define a new metric tensor  $g_p = f^{2(1-p)/d} g_1$ . That is, if  $Z_x$  and  $W_x$  are two tangent vectors at a point  $x \in M$ , then  $g_p(Z_x, W_x) = f(x)^{2(1-p)/d} g_1(Z_x, W_x)$ . The deformed metric tensor  $g_p$  is well-defined for every  $x$  with  $f(x) > 0$ , and  $g_p$  is a Riemannian metric tensor when  $f \in C^\infty$ . In this paper, we assume  $p > 1$  except for a few places where we compare with the undeformed case  $p = 1$ .

The main result of this paper, stated as Theorem 1, establishes an asymptotic limit of the lengths of the shortest paths through locally finite point processes.

A subset  $A \subset M$  is locally finite if  $A \cap B$  is finite for every  $B \subset M$  of finite volume. For example, a homogeneous Poisson process in  $\mathbb{R}^d$  is locally finite with probability one. For  $x, y \in M$  and locally finite  $A \subset M$ , let  $L(x, y; A)$  denote the power-weighted shortest path length from  $x$  to  $y$  through  $A \cup \{x, y\}$ . Let the edge weight between two points  $u$  and  $v$  be defined as  $\text{dist}_1(u, v)^p$  where  $\text{dist}_1$  denotes the Riemannian distance under the metric tensor  $g_1$ . A path  $\pi$  through points  $x_0, \dots, x_k$  has power-weighted length  $\sum_{i=0}^{k-1} \text{dist}_1(x_i, x_{i+1})^p$ .

For  $x \in M$  and  $r > 0$ , we denote by  $B(x; r)$  the open ball in  $M$  of radius  $r$  centered at  $x$ , that is,  $B(x; r) = \{u \in M: \text{dist}_1(x, u) < r\}$ .

2.1. *Main result.* Let  $\text{dist}_p$  denote the deformed distance under  $g_p$ ,

$$(1) \quad \text{dist}_p(x, y) = \inf_{\gamma} \int_0^1 f(\gamma_t)^{(1-p)/d} \sqrt{g_1(\gamma'_t, \gamma'_t)} dt,$$

where the infimum is taken over all piecewise smooth curves  $\gamma: [0, 1] \rightarrow M$  such that  $\gamma_0 = x$  and  $\gamma_1 = y$ . When a curve achieves the infimum, we call the curve a  $g_p$ -geodesic.

The following is the main result of this paper.

**THEOREM 1.** *Assume that  $M$  is compact, and that  $f$  is continuous with  $\inf_M f > 0$ . There exists a constant  $C(d, p) > 0$ , which only depends on  $d$  and  $p$ , satisfying the following. Let  $b > 0$  and  $\varepsilon > 0$ . Then there exists  $\theta_0 > 0$  such that*

$$P\left(\sup_{x,y} \left| \frac{L(x, y; \mathcal{X}_n)}{n^{(1-p)/d} \text{dist}_p(x, y)} - C(d, p) \right| > \varepsilon \right) \leq \exp(-\theta_0 n^{1/(d+2p)}),$$

for all sufficiently large  $n$ , where the supremum is taken over  $x, y \in M$  with  $\text{dist}_1(x, y) \geq b$ .

The constant  $C(d, p)$  is fixed throughout this paper [this is the same constant that is denoted as  $\mu$  in Howard and Newman (1997, 2001)]. When  $p = 1$ , there is no power-weighting of the edges, and  $C(d, 1) = 1$ .

The requirement  $\text{dist}_1(x, y) \geq b > 0$  can be relaxed. The probability upper bound  $\exp(-\theta_0 n^{1/(d+2p)})$  can be written as  $\exp(-\theta'_0 (nr_n^d)^{1/(d+2p)} + O(\log n))$  where  $\theta'_0 = \theta_0 b^{-d/(d+2p)}$  and where  $x, y$  is constrained to satisfy  $\text{dist}_1(x, y) \geq r_n$  for some positive sequence  $r_n$ . Therefore, for the probability upper bound to be nontrivial,  $nr_n^d / \log n$  must go to infinity. The separation requirement  $\text{dist}_1(x, y) \geq b > 0$  is one sufficient condition that ensures this property.

A similar convergence result holds when  $M$  is complete, but not necessarily compact, giving the almost sure limit stated below.

**THEOREM 2.** *Assume that  $M$  is complete and that  $f$  is continuous with  $f(u) > 0$  for all  $u \in M$ . Fix  $x, y \in M$ . Then*

$$\lim_{n \rightarrow \infty} n^{(p-1)/d} L(x, y; \mathcal{X}_n) = C(d, p) \text{dist}_p(x, y) \quad \text{a.s.}$$

The constant  $C(d, p)$  is the same constant as in Theorem 1.

REMARK. In the case where the p.d.f.  $f \in C^\infty$ , then the deformed metric tensor  $g_p$  is a Riemannian metric tensor, and  $\text{dist}_p$  is a Riemannian distance. Theorems 1 and 2 connect an algorithmic quantity, power-weighted shortest path lengths, to a geometric quantity, Riemannian distances.

2.2. Discussion. We use shorthand notation  $\mathcal{L}_\lambda(x, y)$  to denote  $L(x, y; \mathcal{H}_\lambda)$  where  $\mathcal{H}_\lambda$  is a homogeneous Poisson point process of intensity  $\lambda > 0$  in  $\mathbb{R}^d$ .

Theorems 1 and 2 can be compared to analogous results in the continuum FPP model of Howard and Newman (2001). The main differences are the following: (i) The results of Howard and Newman (2001) are restricted to the case of uniformly distributed node locations  $\mathcal{H}_\lambda$  in Euclidean spaces while our results also hold for the case of nonuniformly distributed points in compact or complete manifolds; (ii) our convergence rates improve upon those of Howard and Newman (2001).

Specifically, Howard and Newman [(2001), Theorem 2.2] establish a bound on the shortest path lengths in a homogeneous Poisson point process. Recall that  $\mathcal{L}_\lambda(x, y)$  denotes the power-weighted shortest path length from  $x \in \mathbb{R}^d$  to  $y \in \mathbb{R}^d$  through random nodes in a homogeneous Poisson point process  $\mathcal{H}_\lambda$  of intensity  $\lambda > 0$ .

Howard and Newman [(2001), Theorem 2.2] state the following. Let  $\kappa_1 = \min(1, d/p)$ ,  $\kappa_2 = 1/(4p + 3)$ , and  $e_1 \in \mathbb{R}^d$  be a unit vector. For any  $0 < b < \kappa_2$ , there exists a constant  $C_0$  (depending on  $b$ ) such that for  $t > 0$  and  $t^b \leq s\sqrt{t} \leq t^{\kappa_2 - b}$ ,

$$(2) \quad P\left(\left|\frac{1}{t}\mathcal{L}_1(0, te_1) - C(d, p)\right| > s\right) \leq \exp(-C_0(s\sqrt{t})^{\kappa_1}).$$

Note that the bound in (2) decays to zero no faster than  $\exp(-C_0 t^{\kappa_1 \kappa_2})$  where  $\kappa_1 \kappa_2 = t^{\min(1, d/p)/(4p+3)}$ .

On the other hand, for arbitrary (uniform or nonuniform) density, our Theorem 7 implies, after simple Poissonization of the sequence  $\mathcal{X}_n$  that there exists some  $\theta > 0$  such that (see the Appendix)

$$(3) \quad P\left(\left|\frac{1}{t}\mathcal{L}_1(0, te_1) - C(d, p)\right| > s\right) \leq \exp(-\theta t^{d/(d+2p)})$$

for all sufficiently large  $t$ . Therefore, the decay is exponential in  $t^{d/(d+2p)}$ . Under the condition  $d \geq 1$  and  $p > 1$ , the decay rate (3) is faster than the rate (2).

It is interesting to compare Theorem 1 with BHH results. The convergence result established in this paper differs from previous BHH theorems in two ways. The first difference is that Theorem 1 specifies a limit of the shortest path through  $\mathcal{X}_n$  while BHH theory [Steele (1997), Yukich (1998)] specifies limits of the total length of a graph spanning  $\mathcal{X}_n$ , for example, the minimal spanning tree (MST)

or the solution to the traveling salesman problem (TSP). The second difference is that the shortest path has fixed anchor points, hence it is not translation-invariant. This is in contrast to BHH theory developed in Penrose and Yukich (2003, 2013) where Euclidean functionals are generalized to locally stable functionals while the translation-invariance requirement is maintained.

**3. Main proofs.** An obvious but important property of  $L(x, y; A)$  for  $x, y \in M$  and locally finite  $A \subset M$  is that if  $A' \subset A$  then  $L(x, y; A) \leq L(x, y; A')$ . This property is used in several places in the proofs.

We define the constant  $\alpha = 1/(d + 2p)$  that is used throughout the paper.

3.1. *Local convergence results.* Theorem 1 states a convergence result applying to random variables in Riemannian manifolds. Theorem 1 is obtained by an extension of a simpler theorem on Euclidean space.

We first prove an upper bound for shortest path edge lengths.

LEMMA 3. Let  $z \in \mathbb{R}^d$  and  $R > 0$ . Assume that  $\mathcal{X}_n$  is i.i.d. in  $\mathbb{R}^d$  with p.d.f.  $f$ , and that  $f_m = \inf\{f(u) : u \in B(z; R)\}$  is strictly positive. Fix  $b > 0$ .

Define the event  $H_n(i, j)$  for each pair  $1 \leq i \neq j \leq n$  as the intersection of the following events:

- (i) both  $X_i$  and  $X_j$  are in  $B(z; R)$ ,
- (ii)  $|X_i - X_j| > b^\alpha (nf_m)^{(\alpha-1)/d}$ , and
- (iii) the shortest path from  $X_i$  to  $X_j$  over  $\mathcal{X}_n$  contains no sample point  $X_k$  other than  $X_i$  and  $X_j$ .

Let  $F_n = \bigcap_{i,j} (H_n(i, j))^c$ , where the superscript  $c$  denotes set complement. Then there exists a constant  $\theta_1 > 0$  such that

$$1 - P(F_n) \leq \exp(-\theta_1 n^\alpha)$$

for all sufficiently large  $n$ .

PROOF. Define  $h(X_i, X_j; \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$(4) \quad h(X_i, X_j; u) = |X_i - u|^p + |X_j - u|^p - |X_i - X_j|^p,$$

and let  $\Theta(X_i, X_j) = \{u \in \mathbb{R}^d : h(X_i, X_j; u) < 0\}$ . Note that if  $X_k \in \Theta(X_i, X_j)$ , then  $X_i \rightarrow X_k \rightarrow X_j$  is shorter than  $X_i \rightarrow X_j$  as measured by the sum of power-weighted edge lengths. Note that the volume of  $\Theta(X_i, X_j)$  is a function of the distance  $|X_i - X_j|$  and that a portion of  $\Theta(X_i, X_j)$  intersects  $B(z; R)$ . Therefore, there exists a constant  $\theta'_1 = \theta'_1(d, p) > 0$  such that the intersection volume is at least  $\theta'_1 |X_i - X_j|^d$  for all sufficiently large  $n$ .

Suppose that event  $H_n(1, 2)$  occurs. Then the shortest path from  $X_1$  to  $X_2$  contains no sample point other than  $X_1$  and  $X_2$ , and the intersection of  $\Theta(X_1, X_2)$

and  $B(z; R)$  cannot contain any of  $X_3, X_4, \dots, X_n$ . Since it is assumed that  $|X_1 - X_2| > b^\alpha (nf_m)^{(\alpha-1)/d}$ ,

$$P(H_n(1, 2)) \leq (1 - \theta'_1(nf_m b^d)^\alpha / n)^{n-2}.$$

There are  $n(n - 1)/2 \leq n^2$  pairs of sample points, hence

$$1 - P(F_n) = P\left(\bigcup_{i < j} H_n(i, j)\right) \leq n^2(1 - \theta_1 n^\alpha / n)^{n-2},$$

where  $\theta_1 = \theta'_1(f_m b^d)^\alpha$ .  $\square$

Next, we provide results on the number of nodes in the shortest paths (Lemma 4), and the mean convergence of  $\mathbb{E}L_n$  (Proposition 5). We will establish these results using the theory of Poisson processes in Section 4. Proposition 5 involves the constant  $C(d, p)$  in Theorem 1. The definition of this constant will be given in (44) in the proof of Lemma 12.

LEMMA 4. *Let  $z \in \mathbb{R}^d$ ,  $R_2 > R_1 > 0$ . Assume that the p.d.f.  $f$  is uniform in  $B(z; R_2)$ , that is,  $f(u) = f(z) > 0$  for all  $u \in B(z; R_2)$ , but may have probability mass outside  $B(z; R_2)$ . Then there exists a constant  $C_* > 0$ , depending only on  $d$  and  $p$ , satisfying the following.*

*For  $x, y \in B(z; R_1)$  with  $x \neq y$ , let  $\#L(x, y; \mathcal{X}_n \cap B(z; R_2))$  denote the number of nodes in the shortest path, and let  $G_n(x, y)$  denote the event that*

$$\frac{\#L(x, y; \mathcal{X}_n \cap B(z; R_2))}{(nf(z))^{1/d}|x - y|} \leq C_*.$$

*For any given  $b \in (0, 2R_1)$ , there exists  $\theta_2 > 0$  and  $n_0 > 0$  such that if  $|x - y| \geq b$  and  $n \geq n_0$  then*

$$1 - P(G_n(x, y)) \leq \exp(-\theta_2 n^{1/(d+2p-1)}).$$

PROPOSITION 5. *Let  $z \in \mathbb{R}^d$ ,  $R_2 > R_1 > 0$ . Assume that the p.d.f.  $f$  is uniform in  $B(z; R_2)$  but may have probability mass outside  $B(z; R_2)$ . Fix  $\varepsilon > 0$  and  $b \in (0, 2R_1)$ . Then there exists  $n_0 > 0$  such that for all  $n \geq n_0$  and  $x, y \in B(z; R_1)$  with  $|x - y| \geq b$ ,*

$$\left| \frac{\mathbb{E}L(x, y; \mathcal{X}_n \cap B(z; R_2))}{(nf(z))^{(1-p)/d}|x - y|} - C(d, p) \right| < \varepsilon.$$

From Lemmas 3, 4 and Proposition 5, we obtain the following local convergence result.

PROPOSITION 6. *Let  $z \in \mathbb{R}^d$ ,  $R_2 > R_1 > 0$ . Assume that the p.d.f.  $f$  is uniform in  $B(z; R_2)$  but may have probability mass outside. Fix  $\varepsilon > 0$  and  $b \in$*

$(0, 2R_1)$ . Then there exists a constant  $\theta_3 > 0$  such that for all sufficiently large  $n$  and for all  $x, y \in B(z; R_1)$  with  $|x - y| \geq b$ ,

$$P\left(\left|\frac{L(x, y; \mathcal{X}_n \cap B(z; R_2))}{(nf(z))^{(1-p)/d}|x - y|} - C(d, p)\right| > \varepsilon\right) \leq \exp(-\theta_3 n^\alpha).$$

While it is possible to obtain a weakened form of Proposition 6 from Howard and Newman (2001), we provide an alternative proof with improved convergence rate.

In the proof of Proposition 6, Talagrand’s convex distance [see Talagrand (1995), Section 4.1] is used in the following form. Let  $\omega$  be an elementary outcome in the sample space, and let  $A$  be a measurable event with respect to  $n$  sample points  $X_1, \dots, X_n$ . Define the convex distance  $d_c(\omega; A)$  of  $\omega$  from  $A$ ,

$$(5) \quad d_c(\omega; A) = \sup_{s_1, \dots, s_n} \min_{\eta \in A} \sum_i s_i 1_{\{X_i(\omega) \neq X_i(\eta)\}},$$

where the supremum is taken over  $s_1, \dots, s_n \in \mathbb{R}$ ,  $\sum_i |s_i|^2 \leq 1$ . For  $t > 0$  define  $A_t$  as the enlargement of  $A$  by  $t$ ,

$$(6) \quad A_t = \{\omega: d_c(\omega; A) \leq t\}.$$

This notation will be used only in the proof of Proposition 6. Talagrand’s concentration inequality Talagrand [(1995), Theorem 4.1.1] is

$$(7) \quad P(A)(1 - P(A_t)) \leq \exp(-\frac{1}{4}t^2).$$

**PROOF OF PROPOSITION 6.** Our proof is structured similarly to that of Yuchich [(2000), Theorem 1.3] and Talagrand [(1995), Section 7.1]. For convenience, define  $\tau_n = b(nf(z))^{1/d}$  and  $\zeta_n = \tau_n^\alpha (nf(z))^{-1/d}$ . Let:

- $F_n$  be the event that all the edges of the shortest path have distances at most  $\zeta_n$  (see Lemma 3 for  $F_n$ ),
- $G_n$  be the event that  $\#L_n(x_n, y_n) \leq C_* \tau_n$  where the constant  $C_*$  is specified in Lemma 4,
- $H_n$  be the event that at every point  $u \in B(z; R_2)$ , at least one of the sample points is in  $B(u; \zeta_n)$ .

All these events occur with high probability. Both  $1 - P(F_n)$  and  $1 - P(G_n)$  are exponentially small in  $n^\alpha$  by Lemmas 3 and 4, respectively. The probability  $1 - P(H_n)$  may be shown to be exponentially small as well by an argument similar to the proof to Lemma 3, which we will outline here. Let  $\{B(w_i; 2^{-1}\zeta_n), 1 \leq i \leq m\}$  be an open cover of  $B(z; R_2)$  with  $m = O(n)$ . The probability that at least one of the open balls comprising the cover does not contain any sample point is bounded above by  $m(1 - 4^{-d} f(z) V_d \zeta_n^d)^n$ , where  $V_d$  denotes the volume of a unit ball. This upper bound is exponentially small in  $nf(z)\zeta_n^d = (nf(z)b^d)^\alpha$ , hence exponentially small in  $n^\alpha$  as  $n$  goes to infinity.

We use shorthand notation  $L_n^B$  for  $L(x, y; \mathcal{X}_n \cap B(z; R_2))$ . For  $a > 0$ , define  $W_n(a)$  to be the event that  $L_n^B \geq a$ . Let  $\omega \in F_n \cap G_n$  and  $\eta \in H_n \cap W_n(a)$  be two elementary outcomes in the sample space. If the shortest path  $L_n(\omega)$  from  $x$  to  $y$  through the realization  $\mathcal{X}_n(\omega) = \{X_1(\omega), \dots, X_n(\omega)\}$  is the sequence

$$x_n = \pi_0(\omega) \rightarrow \pi_1(\omega) \rightarrow \dots \rightarrow \pi_{k+1}(\omega) = y_n,$$

where  $k = \#L_n^B(\omega)$ , then we may build a path  $\pi(\eta)$  from  $x$  to  $y$  through another realization  $X_1(\eta), \dots, X_n(\eta)$  as follows. For each  $i \in \{1, \dots, k\}$ , let  $j$  denote the index where  $X_j(\omega) = \pi_i(\omega)$ . If  $X_j(\eta) = X_j(\omega)$ , then set  $\pi_i(\eta) = \pi_i(\omega)$ . Otherwise, since  $\eta \in H_n \cap W_n(a) \subset H_n$ , there exists some  $l$  such that  $X_l(\eta)$  is in  $B(z; R_2)$  and  $|X_l(\eta) - \pi_i(\omega)| < \zeta_n$ . Set  $\pi_i(\eta) = X_l(\eta)$ . Then it follows that  $|\pi_i(\eta) - \pi_i(\omega)| \leq \zeta_n$  for all  $i = 1, \dots, k$ . At the same time,  $|\pi_i(\omega) - \pi_{i+1}(\omega)| \leq \zeta_n$  for all  $i = 1, \dots, k$  since  $\omega \in F_n$ . It follows from the triangle inequality that

$$\begin{aligned} &|\pi_i(\eta) - \pi_{i+1}(\eta)| \\ &\leq |\pi_i(\eta) - \pi_i(\omega)| + |\pi_i(\omega) - \pi_{i+1}(\omega)| + |\pi_{i+1}(\omega) - \pi_{i+1}(\eta)| \\ &\leq 3\zeta_n. \end{aligned}$$

Let  $I$  be the set of indices  $i$  where  $\pi_i(\omega) \neq \pi_i(\eta)$ . Then the power-weighted length of the path  $\pi(\eta)$  is at most

$$\begin{aligned} \sum_{i=0}^k |\pi_i(\eta) - \pi_{i+1}(\eta)|^p &\leq \sum_{i=0}^k |\pi_i(\omega) - \pi_{i+1}(\omega)|^p + \sum_{\substack{i \in I \\ \text{or } i+1 \in I}} |\pi_i(\eta) - \pi_{i+1}(\eta)|^p \\ &\leq L_n^B(\omega) + 2|I|(3\zeta_n)^p. \end{aligned}$$

On the other hand,  $\eta \in W_n(a)$ , that is,  $L_n^B(\eta) \geq a$ . Hence,

$$a \leq L_n^B(\omega) + 2|I|(3\zeta_n)^p.$$

Let  $d_c(\omega; H_n \cap W_n(a))$  be the convex distance as defined in (5). By choosing  $s_i$  in (5) as

$$s_i = \begin{cases} 1/\sqrt{\#L_n^B(\omega)}, & \text{if } X_i \text{ is in the path } L_n^B(\omega), \\ 0, & \text{otherwise,} \end{cases}$$

there exists  $\eta \in H_n \cap W_n(a)$  such that  $|I| \leq d_c(\omega; H_n \cap W_n(a))\sqrt{\#L_n^B(\omega)}$ . Therefore,

$$L_n^B(\omega) \geq a - 2 \cdot (3\zeta_n)^p \cdot d_c(\omega; H_n \cap W_n(a))\sqrt{\#L_n^B(\omega)}.$$

In particular, if  $L_n^B(\omega) \leq a - u$  for  $u > 0$ , then

$$\begin{aligned} (8) \quad d_c(\omega; H_n \cap W_n(a)) &\geq \frac{u}{2 \cdot (3\zeta_n)^p \cdot \sqrt{\#L_n^B(\omega)}} \\ &\geq \frac{u}{2 \cdot (3\zeta_n)^p \cdot \sqrt{C_* \tau_n}} \end{aligned}$$

since  $\omega \in G_n$ . Let  $t > 0$  equal to the right-hand side of (8), and let  $(H_n \cap W_n(a))_t$  denote the enlargement of  $H_n \cap W_n(a)$  as defined in (6), that is, the collection of all elementary outcomes whose convex distance from  $H_n \cap W_n(a)$  is at most  $t$ . Then (8) implies that

$$(9) \quad P(\{L_n^B \leq a - u\} \cap F_n \cap G_n) \leq 1 - P((H_n \cap W_n(a))_t).$$

Let  $M_n$  be the median of  $L_n^B$ . Note that  $P(H_n \cap W_n(M_n))$  is arbitrarily close to  $1/2$  for sufficiently large  $n$  since  $P(W_n(M_n)) = P(L_n^B \geq M_n) = 1/2$ , and  $P(H_n)$  approaches one as  $n \rightarrow \infty$  by Lemma 3. In particular, for  $n$  sufficiently large,  $P(H_n \cap W_n(M_n)) \geq 1/3$ . Set  $a = M_n$  in (9), and apply Talagrand’s inequality (7) with  $A = H_n \cap W_n(M_n)$  to obtain

$$P(L_n^B \leq M_n - u) \leq 3 \exp\left(-\frac{C_2 u^2}{\zeta_n^{2p} \tau_n}\right) + (1 - P(F_n)) + (1 - P(G_n)),$$

for sufficiently large  $n$ , where  $C_2 = (2^4 3^{2p} C_*)^{-1}$ .

To obtain an upper bound, set  $a = M_n + u$ . From (9),  $1 - P((H_n \cap W_n(a))_t) \geq 1/3$  for sufficiently large  $n$  since both  $P(F_n)$  and  $P(G_n)$  converge to one as  $n \rightarrow \infty$ . Apply Talagrand’s inequality again for  $A = H_n \cap W_n(a) = H_n \cap \{L_n^B \geq M_n + u\}$ . This gives

$$P(L_n^B \geq M_n + u) \leq 3 \exp\left(-\frac{C_2 u^2}{\zeta_n^{2p} \tau_n}\right) + (1 - P(H_n))$$

for sufficiently large  $n$ . Combine the above two inequalities above and  $|x - y| \geq b$  to obtain

$$(10) \quad \begin{aligned} P\left(\frac{|L_n^B - M_n|}{(nf(z))^{(1-p)/d} |x - y|} > u\right) \\ \leq 6 \exp(-C_2 (nf(z)b^d)^\alpha u^2) + h_n \\ = 6 \exp(-C_3 n^\alpha u^2) + h_n, \end{aligned}$$

where  $h_n = (1 - P(F_n)) + (1 - P(G_n)) + (1 - P(H_n))$ , and  $C_3 = C_2 (f(z)b^d)^\alpha$ . The reader can verify the inequality by recalling the definitions  $\tau_n = b(nf(z))^{1/d}$ ,  $\zeta_n = \tau_n^\alpha (nf(z))^{-1/d}$ , and  $\alpha = 1/(d + 2p)$ .

Note that  $P(|L_n^B - M_n| > u) = 0$  when  $u \geq |x - y|^p$ . Integrate the right-hand side of (10) for  $u \geq 0$  to obtain the upper bound

$$\frac{|\mathbb{E}L_n^B - M_n|}{(nf(z))^{(1-p)/d} |x - y|} \leq 6 \sqrt{\frac{\pi}{C_3 n^\alpha}} + ((nf(z))^{1/d} |x - y|)^{p-1} h_n.$$

Since  $P(F_n)$ ,  $P(G_n)$ , and  $P(H_n)$  approach one exponentially fast in  $n^\alpha$ , so does  $h_n$ . Furthermore, the convergence rate is independent of the choice  $x, y$ . Therefore,

$$\lim_{n \rightarrow \infty} \frac{|\mathbb{E}L_n^B - M_n|}{(nf(z))^{(1-p)/d} |x - y|} = 0.$$

By Proposition 5, for all sufficiently large  $n$  and  $|x - y| \geq b$ ,

$$P\left(\left|\frac{L_n^B}{(nf(z))^{(1-p)/d}|x-y|} - C(d, p)\right| > \varepsilon\right) \leq P\left(\frac{|L_n^B - M_n|}{(nf(z))^{(1-p)/d}|x-y|} > \frac{\varepsilon}{2}\right).$$

Thus, the proposition follows from (10), Lemmas 3 and 4.  $\square$

**THEOREM 7.** *Let  $z \in \mathbb{R}^d$ ,  $R > 0$ , and  $b \in (0, R/2)$ . Assume that the p.d.f.  $f$  is uniform in  $B(z; R)$  but may have probability mass outside. Choose  $\varepsilon > 0$  sufficiently small so that  $(C(d, p) + \varepsilon)/2 < (C(d, p) - \varepsilon)5/8$ .*

*Denote by  $E_n(\varepsilon)$  the event that for all  $x \in B(z; R/4)$  and  $u \notin B(z; R)$ ,*

$$(11) \quad \frac{L(x, u; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}} > (C(d, p) - \varepsilon) \cdot \frac{5}{8}R.$$

*Denote by  $E'_n(\varepsilon)$  the event that for all  $x, y \in B(z; R/4)$  with  $|x - y| \geq b$ ,*

$$(12) \quad L(x, y; \mathcal{X}_n) = L(x, y; \mathcal{X}_n \cap B(z; R)),$$

*and*

$$(13) \quad \left|\frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|x-y|} - C(d, p)\right| \leq \varepsilon.$$

*Then there exists  $\theta_4 > 0$  such that*

$$1 - P(E_n(\varepsilon) \cap E'_n(\varepsilon)) \leq \exp(-\theta_4 n^\alpha)$$

*for all sufficiently large  $n$ .*

Theorem 7 asserts that with high probability, for sufficiently large  $n$  the shortest path between the points  $x, y$  in the open ball  $B(z; R/4)$  does not exit  $B(z; R)$ . We first prove a lemma that will be used to prove Theorem 7.

**LEMMA 8.** *Suppose the assumptions in Theorem 7 hold. Let  $u, v, x, y \in B(z; R)$ . If:*

- C.1: *the event  $F_n$  from Lemma 3 occurs,*
- C.2:  *$|x - u|$  and  $|y - v|$  are at most  $b^\alpha (nf(z))^{(\alpha-1)/d}$ ,*
- C.3:  *$|x - y| \geq b$ , and*
- C.4:  *$\frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|u-v|} \leq C(d, p) + \varepsilon$ ,*

*then there exists  $n_0 > 0$  independent of the choice  $u, v, x, y$  in  $B(z; R)$  such that*

$$(14) \quad \left|\frac{L(x, y; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} - \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|u-v|}\right| < \frac{\varepsilon}{2}$$

*for all  $n \geq n_0$ .*

PROOF. We bound the left-hand side of (14) from above by

$$(15) \quad \left| \frac{L(x, y; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} - \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} \right| + \left| \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} - \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|u-v|} \right|.$$

Since the event  $F_n$  occurred by C.1, every edge in the shortest path has length at most  $b^\alpha(nf(z))^{(\alpha-1)/d}$ . Combining with C.2, the difference between  $L(x, y; \mathcal{X}_n \cap B(z; R))$  and  $L(u, v; \mathcal{X}_n \cap B(z; R))$  is at most  $2 \cdot (2b^\alpha)^p(nf(z))^{(\alpha-1)p/d}$ . With C.3, the first term in (15) may be bounded from above as follows:

$$\left| \frac{L(x, y; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} - \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} \right| \leq \frac{2^{p+1}b^{\alpha p}(nf(z))^{(\alpha-1)p/d}}{b(nf(z))^{(1-p)/d}}.$$

Note that  $(\alpha - 1)p/d - (1 - p)/d = -(1 - \alpha p)/d < 0$ . Therefore, there exists  $n_1$  such that the first term in (15) is smaller than  $\varepsilon/4$  for all  $n \geq n_1$ , and  $n_1$  is independent of the choice  $u, v, x, y$ .

Since  $||u - v| - |x - y|| < 2b^\alpha(nf(z))^{(\alpha-1)/d}$  by C.2, the second term in (15) can be bounded from above as follows using C.4:

$$\left| \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x-y|} - \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|u-v|} \right| \leq \frac{L(u, v; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|u-v|} \left| \frac{|u-v| - |x-y|}{|x-y|} \right| \leq (C(d, p) + \varepsilon) \frac{2b^\alpha(nf(z))^{(\alpha-1)/d}}{b}.$$

Since  $(\alpha - 1)/d < 0$ , there exists  $n_2$  such that the second term in (15) is smaller than  $\varepsilon/4$  for all  $n \geq n_2$ , and again  $n_2$  is independent of the choice  $u, v, x, y$ . Lemma 8 follows by choosing  $n_0 = \max(n_1, n_2)$ .  $\square$

PROOF OF THEOREM 7. Let  $\zeta_n = b^\alpha(nf(z))^{(\alpha-1)/d}$ . For a set of points  $\{w_i\}_{i=1}^m$  in  $\mathbb{R}^d$ , let

$$\{B(w_i; \zeta_n) : w_i \in B(z; R/4), 1 \leq i \leq m\}$$

be a finite open cover of  $B(z; R/4)$  with  $m = O(n)$ . Likewise, for a set of points  $\{v_k\}_{k=1}^\ell$  in  $\mathbb{R}^d$ , let

$$\{B(v_k; \zeta_n) : v_k \in B(z; 7R/8), 1 \leq k \leq \ell\}$$

be a finite open cover of the boundary of  $B(z; 7R/8)$  with  $\ell = O(n)$ .

Suppose that:

- D.1: the event  $F_n$  from Lemma 3 occurs, and
- D.2:  $|\frac{L(w_i, v_k; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|w_i - v_k|} - C(d, p)| \leq \frac{\epsilon}{2}$  for all  $w_i, v_k$ .

We claim that (11) holds for sufficiently large  $n$ , under the assumptions D.1 and D.2.

Let  $L(x; r; \mathcal{X}_n)$ ,  $r > 0$ , denote the minimal power-weighted path length from  $x$  to the boundary of  $B(z; r)$ , that is,

$$L(x; r; \mathcal{X}_n) = \min_{|z-y|=r} L(x, y; \mathcal{X}_n).$$

Every path from  $x \in B(z; R/4)$  to  $u \notin B(z; R)$  crosses the boundary of  $B(z; 7R/8)$ , therefore  $L(x, u; \mathcal{X}_n) \geq L(x; 7R/8; \mathcal{X}_n)$ . It suffices to show that

$$(16) \quad \frac{L(x; 7R/8; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}} \geq (C(d, p) - \epsilon) \frac{5}{8} R$$

for all  $x \in B(z; R/4)$  to prove (11).

Note that  $L(x; 7R/8; \mathcal{X}_n) = L(x; 7R/8; \mathcal{X}_n \cap B(z; R))$ , that is,

$$L(x; 7R/8; \mathcal{X}_n) = \min_{|z-y|=7R/8} L(x, y; \mathcal{X}_n \cap B(z; R)).$$

If the shortest path to the boundary were to reach any point outside  $B(z; R) \cap B(z; 7R/8)$ , the path must have already passed through the boundary, which is a contradiction.

For every  $x \in B(z; R/4)$ , there exists  $w_i$  such that  $|x - w_i| < \zeta_n$ , and for every  $q$  on the boundary of  $B(z; 7R/8)$ , there exists  $v_k$  such that  $|q - v_k| < \zeta_n$ . Consequently, by Lemma 8 and assumptions D.1 and D.2, for sufficiently large  $n$ ,

$$\frac{L(x, q; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|x - q|} \geq C(d, p) - \epsilon,$$

for all  $x \in B(z; R/4)$  and for all  $q$  satisfying  $|z - q| = 7R/8$ . Use  $|x - q| \geq 5R/8$ , and we have proved (16), and in turn (11).

Now let  $x, y \in B(z; R/4)$  and  $|x - y| \geq b$ . Then there exist  $w_i, w_j$  such that  $|w_i - x| < \zeta_n$  and  $|w_j - y| < \zeta_n$ . Suppose the following condition holds in addition to D.1 and D.2:

$$D.3: \left| \frac{L(w_i, w_j; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|w_i - w_j|} - C(d, p) \right| \leq \frac{\epsilon}{2}.$$

We claim that (12) holds.

Assume to the contrary that the path  $L(x, y; \mathcal{X}_n)$  includes some point outside  $B(z; R)$ . Then the path has crossed the boundary of  $B(z; 7R/8)$ , hence

$L(x; 7R/8; \mathcal{X}_n) \leq L(x, y; \mathcal{X}_n)$ . We have already seen that (16) holds. Therefore,

$$(C(d, p) - \varepsilon) \frac{5}{8} R \leq \frac{L(x; 7R/8; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}} \leq \frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}}.$$

On the other hand, apply Lemma 8 with D.3 to have

$$\frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}} \leq \frac{L(x, y; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}} \leq (C(d, p) + \varepsilon) \frac{1}{2} R$$

since  $x, y \in B(z; R/4)$  and  $|x - y| \leq R/2$ . Recall that  $\varepsilon$  was assumed to be sufficiently small so that  $(C(d, p) + \varepsilon)/2 < (C(d, p) - \varepsilon)5/8$ . Therefore, we have a contradiction.

We now claim that (13) is true when assumptions D.1, D.2 and D.3 hold. Start with

$$\begin{aligned} (17) \quad & \left| \frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|x - y|} - C(d, p) \right| \\ & \leq \left| \frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|x - y|} - \frac{L(w_i, w_j; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|w_i - w_j|} \right| \\ & \quad + \left| \frac{L(w_i, w_j; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|w_i - w_j|} - C(d, p) \right|. \end{aligned}$$

From (11) and (12),  $L(x, y; \mathcal{X}_n)$  and  $L(w_i, w_j; \mathcal{X}_n)$  in the upper bound of (17) can be replaced by  $L(x, y; \mathcal{X}_n \cap B(z; R))$  and  $L(w_i, w_j; \mathcal{X}_n \cap B(z; R))$ , respectively. Therefore, for sufficiently large  $n$ , the first term in the upper bound is less than  $\varepsilon/2$  by Lemma 8, and the second term is less than  $\varepsilon/2$  by D.3. This establishes that (13) holds.

In summary, we have shown that (11), (12) and (13) hold when events D.1, D.2 and D.3 occur. If the event  $E_n(\varepsilon)$  or  $E'_n(\varepsilon)$  does not occur, either one of D.1, D.2 or D.3 does not occur:

$$\begin{aligned} 1 - P(E_n(\varepsilon) \cap E'_n(\varepsilon)) & \leq (1 - P(F_n)) \\ & \quad + \sum_{w_i, v_k} P\left(\left| \frac{L(w_i, v_k; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|w_i - v_k|} - C(d, p) \right| > \frac{\varepsilon}{2}\right) \\ & \quad + \sum_{w_i, w_j} P\left(\left| \frac{L(w_i, w_j; \mathcal{X}_n \cap B(z; R))}{(nf(z))^{(1-p)/d}|w_i - w_j|} - C(d, p) \right| > \frac{\varepsilon}{2}\right). \end{aligned}$$

The first sum is over all pairs  $w_i, v_k$ . The second sum is over all  $w_i, w_j$  with  $|w_i - w_j| \geq b/2$ . Theorem 7 now follows from Lemma 3 and Proposition 6 with  $R_1 = 7R/8$  and  $R_2 = R$ , since the number of summands are of polynomial order in  $n$ .  $\square$

COROLLARY 9. Assume that  $f$  is continuous at  $z \in \mathbb{R}^d$  and  $f(z) > 0$ . Fix  $\varepsilon > 0$  small enough so that  $(C(d, p) + \varepsilon)/2 < (C(d, p) - \varepsilon)5/8$ . For  $R > 0$  and  $b \in (0, R/2)$ , let  $H_n = H_n(z, R, \varepsilon, b)$  denote the event that

$$(18) \quad \left| \frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|x-y|} - C(d, p) \right| \leq \varepsilon,$$

for all  $x, y \in B(z; R/4)$  with  $|x - y| \geq b$ , and simultaneously

$$(19) \quad \frac{L(x, u; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}} \geq (C(d, p) - \varepsilon)\frac{5}{8}R,$$

for all  $x \in B(z; R/4)$  and  $u \notin B(z; R)$ . Then there exists  $R = R(z) > 0$  such that for all  $b \in (0, R/2)$  there exists  $\theta_5 > 0$  for which,

$$(20) \quad 1 - P(H_n(z, R, \varepsilon, b)) \leq \exp(-\theta_5 n^\alpha)$$

for all sufficiently large  $n$ .

PROOF. Let  $f_m = f_m(R)$  and  $f_M = f_M(R)$  denote the infimum and the supremum of  $f$  inside  $B(z; R)$ , respectively. Since  $f$  is continuous at  $z$ , we can choose sufficiently small  $R = R(z) > 0$  such that  $f_m$  and  $f_M$  are sufficiently close to  $f(z)$  to satisfy

$$(21) \quad (C(d, p) + \varepsilon/2) \left( \frac{f(z)}{f_m} \right)^{(p-1)/d} \leq C(d, p) + \varepsilon,$$

and

$$(22) \quad (C(d, p) - \varepsilon/2) \left( \frac{f(z)}{f_M} \right)^{(p-1)/d} \geq C(d, p) - \varepsilon.$$

In addition, shrink  $R$  if necessary to ensure  $f_M V_d R^d \leq 1$  where  $V_d$  denotes the volume of a unit ball. Now choose and fix an arbitrary  $b \in (0, R/2)$  for the entire proof.

For each sample point  $X_i \in \mathcal{X}_n$ , let  $Y_i$  be an arbitrary random point outside  $B(z; R)$ , and let  $Z_i$  be an independent Bernoulli random variable with  $P(Z_i = 1) = f_m/f(X_i)$ . Define a new random variable

$$X_i^m = \begin{cases} X_i, & \text{if } X_i \text{ is not in } B(z; R), \\ X_i Z_i + Y_i(1 - Z_i), & \text{if } X_i \text{ is in } B(z; R). \end{cases}$$

Let  $\mathcal{X}_n^m = \{X_1^m, \dots, X_n^m\}$ .  $\mathcal{X}_n^m$  is an i.i.d. sample and its p.d.f. restricted to  $B(z; R)$  is uniform with intensity  $f_m$ . Define:

E.1: both the events  $E_n(\varepsilon/2)$  and  $E'_n(\varepsilon/2)$  in Theorem 7 occur for  $\mathcal{X}_n^m$ .

Assume that the event E.1 occurs. Let  $x, y \in B(z; R/4)$  with  $|x - y| \geq b$ . Since  $(\mathcal{X}_n^m \cap B(z; R)) \subset (\mathcal{X}_n \cap B(z; R))$ , we have

$$L(x, y; \mathcal{X}_n) \leq L(x, y; \mathcal{X}_n \cap B(z; R)) \leq L(x, y; \mathcal{X}_n^m \cap B(z; R)) = L(x, y; \mathcal{X}_n^m).$$

The last equality comes from (12) for  $\mathcal{X}_n^m$ . From (13) and (21), we have

$$(23) \quad \frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|x-y|} \leq \frac{L(x, y; \mathcal{X}_n^m)}{(nf_m)^{(1-p)/d}|x-y|} \cdot \left(\frac{f(z)}{f_m}\right)^{(p-1)/d} \leq C(d, p) + \varepsilon.$$

This establishes the upper half of the inequality (18) under E.1.

It remains to establish the lower half of the inequality (18). This is established in two steps. First, we show that (19) holds assuming an event E.2 analogous to E.1. Then we show that E.1 and E.2 imply (18).

For each point  $X_i \in \mathcal{X}_n$ , define a new random variable  $X_i^M$  as follows. Let  $\sigma = \int (f_M - f(u)) du \geq 0$  where the integral is taken inside  $B(z; R)$ . By the assumption  $f_M V_d R^d \leq 1$ , we have  $0 \leq \sigma \leq 1$ . Let  $\tilde{Y}_i$  be a random point inside  $B(z; R)$  with p.d.f.  $\sigma^{-1}(f_M - f(u))$  for  $u \in B(z; R)$ , and let  $\tilde{Z}_i$  be a Bernoulli random variable with  $P(\tilde{Z}_i = 1) = 1 - \sigma$ . Define

$$X_i^M = \begin{cases} X_i, & \text{if } X_i \text{ is in } B(z; R), \\ X_i \tilde{Z}_i + \tilde{Y}_i(1 - \tilde{Z}_i), & \text{if } X_i \text{ is not in } B(z; R). \end{cases}$$

Let  $\mathcal{X}_n^M = \{X_1^M, \dots, X_n^M\}$ .  $\mathcal{X}_n^M$  is an i.i.d. sample and its p.d.f. restricted to  $B(z; R)$  is uniform with intensity  $f_M$ . Define:

E.2: both the events  $E_n(\varepsilon/2)$  and  $E'_n(\varepsilon/2)$  in Theorem 7 occur for  $\mathcal{X}_n^M$ .

Assume that the event E.2 occurs. Let  $x \in B(z; R/4)$  and  $v = \arg \min_y L(x, y; \mathcal{X}_n)$  over all  $|z - y| = R$ . Then  $L(x, v; \mathcal{X}_n) = L(x, v; \mathcal{X}_n \cap B(z; R))$ , otherwise the shortest path from  $x$  to  $v$  has passed through another point on the boundary of  $B(z; R)$ , and this contradicts the choice of  $v$ . Since  $(\mathcal{X}_n \cap B(z; R)) \subset (\mathcal{X}_n^M \cap B(z; R))$ ,

$$L(x, v; \mathcal{X}_n) = L(x, v; \mathcal{X}_n \cap B(z; R)) \geq L(x, v; \mathcal{X}_n^M \cap B(z; R)) \geq L(x, v; \mathcal{X}_n^M).$$

As  $v \notin B(z; R)$ , it follows from (22) and (11) for  $\mathcal{X}_n^M$  that

$$\frac{L(x, v; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}} \geq \frac{L(x, v; \mathcal{X}_n^M)}{(nf_M)^{(1-p)/d}} \cdot \left(\frac{f(z)}{f_M}\right)^{(p-1)/d} \geq (C(d, p) - \varepsilon) \frac{5}{8} R.$$

If  $u \notin B(z; R)$ , the path  $L(x, u; \mathcal{X}_n)$  crosses the boundary of  $B(z; R)$  at some point  $u'$ , hence  $L(x, u; \mathcal{X}_n) \geq L(x, u'; \mathcal{X}_n) \geq L(x, v; \mathcal{X}_n)$  by the minimality of  $v$ . This establishes (19) under E.2.

Now we show that the upper bound of (18) holds under the conditions E.1 and E.2. Let  $x, y \in B(z; R/4)$ . Then  $L(x, y; \mathcal{X}_n) = L(x, y; \mathcal{X}_n \cap B(z; R))$ . Otherwise, that is, if  $L(x, y; \mathcal{X}_n)$  reaches some point  $u \notin B(z; R)$ , then  $L(x, u; \mathcal{X}_n) \leq L(x, y; \mathcal{X}_n)$  but this contradicts (19) and (23) since  $|x - y| \leq R/2$  and  $\varepsilon$  was assumed to satisfy  $(C(d, p) + \varepsilon)/2 < (C(d, p) - \varepsilon)5/8$ . We can now repeat the

same argument used to establish that the upper half of the inequality (18) follows from E.1 to show that

$$\begin{aligned} \frac{L(x, y; \mathcal{X}_n)}{(nf(z))^{(1-p)/d}|x-y|} &\geq \frac{L(x, y; \mathcal{X}_n^M)}{(nf_M)^{(1-p)/d}|x-y|} \cdot \left(\frac{f(z)}{f_M}\right)^{(p-1)/d} \\ &\geq C(d, p) - \varepsilon \end{aligned}$$

by (12) and (13) for  $\mathcal{X}_n^M$ , and (22).

Applying Theorem 7 once to  $\mathcal{X}_n^m$  and once to  $\mathcal{X}_n^M$  asserts that E.1 and E.2 occur with exponentially small probability, which establishes Corollary 9.  $\square$

3.2. *Convergence in Riemannian manifolds.* We adapt Corollary 9 to the case when the probability distribution is supported on a Riemannian manifold  $M$  instead of on a Euclidean space. For  $z \in M$  and  $R > 0$ ,  $B(z; R)$  denotes the set  $\{u \in M: \text{dist}_1(u, z) < R\}$ . Recall that  $\alpha = 1/(d + 2p)$ .

LEMMA 10. *Let  $(M, g_1)$  be a Riemannian manifold equipped with metric tensor  $g_1$ . Let  $z \in M$  be a point, and let  $\varepsilon > 0$  be a fixed constant. For  $R > 0$  and  $b \in (0, 2R)$ , we denote by  $E_n(B(z; R), \varepsilon, b)$  the event that:*

(i) *if a shortest path passes through  $B(z; R)$  then its edges in  $B(z; R)$  have  $\text{dist}_1$ -lengths at most  $b^\alpha (nf(z))^{\alpha-1/d}$ , and*

(ii) *for every pair  $x, y \in B(z; R)$  with  $\text{dist}_1(x, y) \geq b$ ,*

$$(24) \quad \left| \frac{L(x, y; \mathcal{X}_n)}{n^{(1-p)/d} \text{dist}_p(x, y)} - C(d, p) \right| \leq \varepsilon.$$

*Assume that  $f(z) > 0$  and that  $f$  is continuous at  $z$ . Then there exists  $R = R(z) > 0$  such that for all  $b \in (0, 2R)$  there exists  $\theta_6 > 0$  for which,*

$$(25) \quad 1 - P(E_n(B(z; R), \varepsilon, b)) \leq \exp(-\theta_6 n^\alpha)$$

*for all sufficiently large  $n$ .*

In (i), a shortest path edge between two sample points is contained in  $B(z; R)$  when a  $g_1$ -geodesic between these points is contained in  $B(z; R)$ .

PROOF OF LEMMA 10. We prove this lemma by showing that both the events (i) and (ii) satisfy the probability tail bound as in (25). That the statement (25) holds for the event (i) follows from similar arguments as used in Lemma 3. We focus on the event (ii).

Choose  $\delta > 0$  small enough to ensure that

$$(26) \quad 2(1 + \delta)^p \leq \frac{5}{2}(1 - \delta)^p,$$

$$(27) \quad ((1 + \delta)/(1 - \delta))^p (C(d, p) + \varepsilon/2) \leq C(d, p) + \varepsilon, \quad \text{and}$$

$$(1 - \delta)/(1 + \delta))^p (C(d, p) - \varepsilon/2) \geq C(d, p) - \varepsilon.$$

Define  $U = B(z; 4R) = \{u \in M: \text{dist}_1(u, z) < 4R\}$  for  $R > 0$ . Since  $f$  is continuous, we may choose  $R > 0$  small enough so that there exists a normal chart map  $\varphi: U \subset M \rightarrow V \subset \mathbb{R}^d$  such that the event  $H_n = H_n(\varphi(z), 4R, 2^{-1}\varepsilon, (1 + \delta)^{-1}b)$  from Corollary 9 satisfies (20), and that

$$(28) \quad (1 - \delta)^d \sup_U f \leq f(z) = f(\varphi(z)) \leq (1 + \delta)^d \inf_U f,$$

$$(29) \quad 1 - \delta \leq \frac{\text{dist}_1(u, v)}{|\varphi(u) - \varphi(v)|} \leq 1 + \delta,$$

for all  $u, v \in U$  with  $u \neq v$ . Recall that  $f(z) = f(\varphi(z))$  follows from the basic properties of normal coordinates at  $z \in M$ . See, for example, O’Neill (1983), page 73. The denominator in (29) is a Euclidean distance.

We claim that (ii) is true when the event  $H_n(\varphi(z), 4R, 2^{-1}\varepsilon, (1 + \delta)^{-1}b)$  occurs. Then (25) would follow from Corollary 9. In the remainder of this proof, we prove this claim.

Let  $x, y \in B(z; R) \subset U$ . Then a  $g_1$ -geodesic curve from  $x$  to  $y$  is contained in  $U$  by the triangle inequality. It follows from the definition of  $\text{dist}_p$  in (1) that

$$(30) \quad \text{dist}_p(x, y) \leq \text{dist}_1(x, y) \left( \inf_U f \right)^{(1-p)/d}.$$

Furthermore, if a  $g_p$ -geodesic curve from  $x$  to  $y$  were contained in  $U$ , then

$$(31) \quad \text{dist}_p(x, y) \geq \text{dist}_1(x, y) \left( \sup_U f \right)^{(1-p)/d}.$$

If a  $g_p$ -geodesic curve from  $x$  to  $y$  exits  $U$ , then  $\text{dist}_p(x, y)$  must be at least  $(3R)(\sup_U f)^{(1-p)/d}$  by the assumptions  $\text{dist}_1(x, z) < R$  and  $U = B(z; 4R)$ . On the other hand, it follows from (26), (28) and (30) that

$$\begin{aligned} \text{dist}_p(x, y) &\leq \text{dist}_1(x, y) \left( \inf_U f \right)^{(1-p)/d} \\ &\leq (2R) \left( \sup_U f \right)^{(1-p)/d} \left( \frac{1 + \delta}{1 - \delta} \right)^{p-1} \\ &< \frac{5}{2} R \left( \sup_U f \right)^{(1-p)/d}, \end{aligned}$$

and this is a contradiction. Therefore, a  $g_p$ -geodesic curve from  $x$  to  $y$  does not exit  $U$ , hence (31) holds.

Next, we show that  $L(x, y; \mathcal{X}_n) = L(x, y; \mathcal{X}_n \cap U)$ , that is, the shortest path between  $x, y \in B(z; R)$  is contained in  $U$ , when  $H_n(\varphi(z), 4R, 2^{-1}\varepsilon, (1 + \delta)^{-1}b)$  occurs. Assume to the contrary that the path  $L(x, y; \mathcal{X}_n)$  from  $x$  exits  $U$ . Then the corresponding path in  $V$  starts from  $\varphi(x)$  and exits  $V$ , and its power-weighted length is at least  $(C(d, p) - \varepsilon)5R/2$ . Note that in this proof Corollary 9 was applied with  $R$  replaced by  $4R$ . By (29) and (19), it implies that

$L(x, y; \mathcal{X}_n)/(nf(z))^{(1-p)/d}$  is at least  $(C(d, p) - \varepsilon)(1 - \delta)^p 5R/2$ . On the other hand, by (29) and (18),  $L(x, y; \mathcal{X}_n \cap U)/(nf(z))^{(1-p)/d}$  is at most  $(C(d, p) + \varepsilon)(1 + \delta)^p 2R$ . This is a contradiction by (26), so we conclude  $L(x, y; \mathcal{X}_n) = L(x, y; \mathcal{X}_n \cap U)$ .

Let  $L(\varphi(x), \varphi(y); \varphi(\mathcal{X}_n \cap U))$  denote the shortest path length between  $\varphi(x), \varphi(y) \in V$  in Euclidean space  $\mathbb{R}^d$ . Then

$$\begin{aligned} \frac{L(x, y; \mathcal{X}_n)}{n^{(1-p)/d} \text{dist}_p(x, y)} &\leq \frac{L(x, y; \mathcal{X}_n)}{(n \sup_U f)^{(1-p)/d} \text{dist}_1(x, y)} \\ &\leq \left(\frac{1 + \delta}{1 - \delta}\right)^p \frac{L(\varphi(x), \varphi(y); \varphi(\mathcal{X}_n \cap U))}{(nf(\varphi(z)))^{(1-p)/d} |\varphi(x) - \varphi(y)|} \\ &\leq C(d, p) + \varepsilon. \end{aligned}$$

The first inequality follows from (31), the second one follows from (28) and (29) and the third one follows from (27) and the assumption that  $H_n(\varphi(z), 4R, 2^{-1}\varepsilon, (1 + \delta)^{-1}b)$  occurred. Repeat the same argument for the lower bound to obtain

$$\frac{L(x, y; \mathcal{X}_n)}{n^{(1-p)/d} \text{dist}_p(x, y)} \geq C(d, p) - \varepsilon.$$

The last two inequalities imply (24).  $\square$

Our main result Theorem 1 can now be obtained by applying Lemma 10 to a finite open cover of the compact manifold  $M$ .

**PROOF OF THEOREM 1.** The crux of the proof is that the shortest path length has near sub- and super-additivity with high probability. We will show that if Lemma 10 holds locally at every point of  $M$ , then the local convergences may be assembled together to yield global convergence of the curve length.

For convenience, define  $L_n(x, y) = L(x, y; \mathcal{X}_n)$  in this proof.

For each  $w \in M$ , we may associate positive  $R(w) > 0$  such that Lemma 10 holds within open ball  $V(w) = B(w; 3R(w))$  with error  $2^{-1}\varepsilon$ , that is, the event  $E_n(V(w), 2^{-1}\varepsilon, b')$ , defined in Lemma 10, satisfies (25) with any  $b' \in (0, 6R(w))$ . Let  $U(w) = B(w; R(w))$ . By compactness, there exists a finite  $m > 0$ ,  $\{w_i \in M\}_{i=1}^m$  such that the collection  $\{U(w_i)\}_{i=1}^m$  is a finite open cover of  $M$ . Define  $R_i = R(w_i)$ ,  $U_i = U(w_i)$ , and  $V_i = V(w_i)$  for  $i = 1, \dots, m$ .

Reorder the indices if necessary so that  $x \in U_1$ . Define  $z_1 = x$ . If  $L_n(x, y)$  ever exits  $V_1$ , then a point  $z_2 \in V_1$  on the shortest path may be chosen such that  $z_2 \notin U_1$  and  $\text{dist}_1(z_1, z_2) \geq R_1$ . Note that shortest paths are piece-wise  $g_1$ -geodesics, and  $z_2 \in M$  need not be in  $\mathcal{X}_n$ . Reorder the indices of the open cover again if necessary so that  $z_2$  is in  $U_2$ . Repeat the procedure until  $L_n(x, y)$  ends at  $y$  in an open ball, say,  $V_k$ . Set  $z_{k+1} = y$ . Then points  $x = z_1, z_2, \dots, z_k, z_{k+1} = y$  satisfy the conditions  $z_i, z_{i+1} \in V_i$  for  $i = 1, 2, \dots, k$ , and  $\text{dist}_1(z_i, z_{i+1}) \geq R_i \geq R$  for  $i = 1, 2, \dots, k - 1$ , where  $R = \min_i R_i$ . The last edge length  $\text{dist}_1(z_k, z_{k+1})$  may

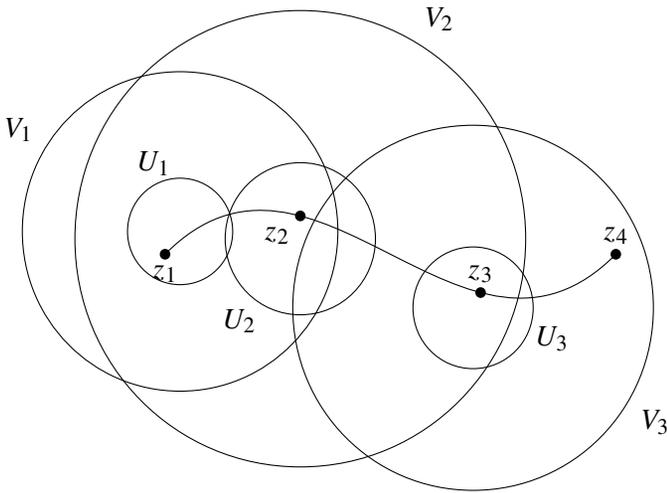


FIG. 1. Path division procedure described in the proof of Theorem 1. Here  $k = 4$ . Note that  $z_i \in U_i$  and  $z_{i+1} \in V_i$  for  $i = 1, 2, 3$ . Shortest path is depicted as a smooth curve for illustration purpose only and it is actually piecewise smooth.

be less than  $R$ . However, note that  $z_{k-1} \in U_{k-1}$  and  $y = z_{k+1} \notin V_{k-1}$  by definition, hence  $\text{dist}_1(z_{k-1}, z_{k+1}) > 2R_{k-1} \geq 2R$ . Therefore,  $z_k$  may be adjusted so that  $\text{dist}_1(z_k, z_{k+1}) \geq R$  as well and  $z_k \in V_k$ . See Figure 1 for illustration.

Suppose that

$$(32) \quad (C(d, p) - \varepsilon) \text{dist}_p(z_i, z_{i+1}) \leq n^{(p-1)/d} L_n(z_i, z_{i+1}),$$

holds for all  $i = 1, 2, \dots, k$ . Then by the triangle inequality and the property  $v^p + \omega^p \leq (v + \omega)^p$  for  $v, \omega \geq 0$  and  $p \geq 1$ ,

$$\begin{aligned} (C(d, p) - \varepsilon) \text{dist}_p(x, y) &\leq (C(d, p) - \varepsilon) \sum_{i=1}^k \text{dist}_p(z_i, z_{i+1}) \\ &\leq \sum_{i=1}^k n^{(p-1)/d} L_n(z_i, z_{i+1}) \\ &\leq n^{(p-1)/d} L_n(x, y). \end{aligned}$$

Since  $m$  is finite and the event  $E_n(V_i, 2^{-1}\varepsilon, \min\{R, b\})$  from Lemma 10 satisfies (25), for  $i = 1, \dots, m$ , we have

$$P\left(\inf_{x,y} \frac{L_n(x, y)}{n^{(1-p)/d} \text{dist}_p(x, y)} < C(d, p) - \varepsilon\right) \leq m \exp(-\theta_6 n^\alpha)$$

for all sufficiently large  $n$ .

For the upper tail, we follow a similar strategy to Bernstein et al. (2000). Recall that  $\alpha = 1/(d + 2p)$ . If  $z_1 = x, z_{k+1} = y$ , and  $z_i$  are points on a  $g_p$ -geodesic curve

from  $x$  to  $y$ , then  $\text{dist}_p(x, y) = \sum_{i=1}^k \text{dist}_p(z_i, z_{i+1})$ . We showed above that the points may be chosen and indices of the open cover may be rearranged such that  $z_i, z_{i+1} \in V_i$  and  $\text{dist}_1(z_i, z_{i+1}) \geq R$  for all  $i = 1, 2, \dots, k$ . The shortest path from  $z_{i-1}$  to  $z_i$  and another shortest path from  $z_i$  to  $z_{i+1}$  may be pasted together to create a path from  $z_{i-1}$  to  $z_{i+1}$  by removing  $z_i$  and connecting two nodes that were incident to  $z_i$ . This pasting procedure can be repeated to create a path from  $x = z_1$  to  $y = z_{k+1}$ . Since Lemma 10 applies in  $V_1, \dots, V_m$ , every edge length of the shortest path from  $z_i$  to  $z_{i+1}$  is at most  $b^\alpha (n \inf f)^{(\alpha-1)/d}$  for  $i = 1, \dots, k$ . Therefore, each pasting procedure may incur an additional cost of at most  $(2b^\alpha)^p (n \inf f)^{(\alpha-1)p/d}$  so that

$$(33) \quad L_n(x, y) \leq \sum_{i=1}^k L_n(z_i, z_{i+1}) + k(2b^\alpha)^p (n \inf f)^{(\alpha-1)p/d}.$$

Therefore, if event  $E_n(V_i, 2^{-1}\varepsilon, \min\{R, b\})$  in Lemma 10 holds for  $V_1, V_2, \dots, V_m$ , then

$$\begin{aligned} & n^{(p-1)/d} L_n(x, y) \\ & \leq \text{dist}_p(x, y) \left( C(d, p) + \frac{\varepsilon}{2} \right) + k(2b^\alpha)^p n^{(p-1)/d} (n \inf f)^{(\alpha-1)p/d} \end{aligned}$$

since  $\text{dist}_p(x, y) \leq \text{dist}_1(x, y)(\inf f)^{(1-p)/d}$ . For sufficiently large  $n$ , we have  $n^{(p-1)/d} L_n(x, y) \leq (C(d, p) + \varepsilon) \text{dist}_p(x, y)$  since  $n^{\alpha p-1}$  shrinks to zero as  $n \rightarrow \infty$ . Therefore, Theorem 1 is established by applications of Lemma 10 to  $V_1, V_2, \dots, V_m$ .  $\square$

We turn to the proof of Theorem 2. Note that when  $M$  is complete, for every  $x, y \in M$  there exists a geodesic path  $\gamma$  between  $x, y$  in  $M$  by the Hopf–Rinow theorem [O’Neill (1983), Theorem 5.21, page 138].

**PROOF OF THEOREM 2.** Define  $L_n(x, y) = L(x, y; \mathcal{X}_n)$ . Let  $0 < \varepsilon < C(d, p)$ . Define

$$A = \{u \in M: (C(d, p) - \varepsilon) \text{dist}_p(x, u) \leq (C(d, p) + \varepsilon) \text{dist}_p(x, y)\}.$$

$A$  is compact by the Hopf–Rinow theorem.

The proof of Theorem 2 is similar to the proof of Theorem 1, with  $M$  replaced by  $A$ . Let  $V_1, \dots, V_m$  be an open cover of  $A$  chosen as in the previous proof for compact  $M$ .

Suppose that the event  $E_n(V_i, 2^{-1}\varepsilon, \min\{R_1, \dots, R_m, b\})$  from Lemma 10 holds for  $i = 1, \dots, m$ . By the construction of  $A$ ,  $g_p$ -geodesics from  $x$  to  $y$  is contained in  $A$ . Repeat the same argument used in the proof of Theorem 1 to obtain

$$(34) \quad n^{(p-1)/d} L_n(x, y) \leq (C(d, p) + \varepsilon) \text{dist}_p(x, y).$$

Similarly, if the shortest path  $L_n(x, y)$  does not exit  $A$ , then

$$(35) \quad n^{(p-1)/d} L_n(x, y) \geq (C(d, p) - \varepsilon) \text{dist}_p(x, y).$$

We claim that the shortest path  $L_n(x, y)$  does not exit  $A$ , so that (35) is true. Assume to the contrary that the path  $L_n(x, y)$  exits  $A$ . Let  $u = \arg \min_{u'} L_n(x, u')$  where  $u'$  is over all boundary points of  $A$ . Since the path  $L_n(x, y)$  exits  $A$ , we have  $L_n(x, y) > L_n(x, u)$ . Since the path  $L_n(x, u)$  is contained in  $A$ , (35) holds with  $u$  in place of  $y$ . Since  $u$  is a point on the boundary of  $A$ , we have  $(C(d, p) - \varepsilon) \text{dist}_p(x, u) = (C(d, p) + \varepsilon) \text{dist}_p(x, y)$ . Combine these with (34) to obtain

$$\begin{aligned} (C(d, p) + \varepsilon) \text{dist}_p(x, y) &\geq n^{(p-1)/d} L_n(x, y) > n^{(p-1)/d} L_n(x, u) \\ &\geq (C(d, p) - \varepsilon) \text{dist}_p(x, u) \\ &= (C(d, p) + \varepsilon) \text{dist}_p(x, y), \end{aligned}$$

and we have a contradiction. We have shown that (35) holds.

Combine Lemma 10 with (34) and (35) to obtain that

$$P\left(\left|\frac{L_n(x, y)}{n^{(1-p)/d} \text{dist}_p(x, y)} - C(d, p)\right| > \varepsilon\right)$$

has exponential decay in  $n^\alpha = n^{1/(d+2p)}$ . Almost-sure convergence, and the limit stated in Theorem 2, follow by the Borel–Cantelli lemma.  $\square$

**4. Mean convergence and node cardinality.** In this section, we prove Lemma 4 and Proposition 5. Since they were stated for sequences  $\mathcal{X}_n$  in a Euclidean space, we return to the Euclidean case  $M = \mathbb{R}^d$ . We introduce a few additional notations used in this section.

The proofs in this section use Poissonization arguments. We denote by  $\mathcal{H}_\lambda$  a homogeneous Poisson point process in  $\mathbb{R}^d$  of constant intensity  $\lambda > 0$ . Specifically, for any Borel set  $B$  of Lebesgue measure  $\nu(B)$  the cardinality  $N_B$  of  $\mathcal{H}_\lambda \cap B$  is a Poisson random variable with mean  $\lambda\nu(B)$  and, conditioned on  $N_B$ , the points  $\mathcal{H}_\lambda \cap B$  are i.i.d. uniform over  $B$ . We use a shorthand notation  $\mathcal{L}_\lambda(x, y) = L(x, y; \mathcal{H}_\lambda)$ .

Let  $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^d$  denote the unit vector. By the translation and rotation invariance of  $\mathcal{H}_\lambda$ , the distribution of  $\mathcal{L}_\lambda(x, y)$  for  $x, y \in \mathbb{R}^d$  is the same as the distribution of  $\mathcal{L}_\lambda(0, te_1)$  where  $t = |x - y|$ . This observation is used frequently in this section.

Let  $T(u, v; b)$  for  $u, v \in \mathbb{R}^d, b > 0$ , denote the set

$$(36) \quad T(u, v; b) = \bigcup_{0 \leq s \leq 1} B(su + (1 - s)v; b).$$

Note that  $\bigcup_{b>0} T(u, v; b) = \mathbb{R}^d$ . For convenience, define

$$(37) \quad \mathcal{L}_\lambda(u, v; b) = L(u, v; \mathcal{H}_\lambda \cap T(u, v; b)).$$

4.1. *Percolation lemma.* The following lemma on percolation will be used in the proof of Lemma 4.

LEMMA 11. *Let  $\pi$  be a graph path in  $\mathcal{H}_\lambda$  starting at  $0 \in \mathbb{R}^d$ . Suppose that  $\pi$  has power-weighted path length at most  $c_0\lambda^{(1-p)/d}$  and has at least  $c_1\lambda^{1/d}$  nodes for some  $c_0, c_1 > 0$ . Then there exists a constant  $\rho_0 > 0$ , dependent on  $d$  and  $p$ , such that if  $c_1 > \rho_0 c_0$  then the probability that such path  $\pi$  exists is exponentially small in  $c_1\lambda^{1/d}$ .*

PROOF. The structure of the proof is similar to that of Meester and Roy (1996), Theorem 6.1. We first define a Galton–Watson process  $\mathbb{X}_n$ . Let  $\mathbb{X}_0 = \{x_0 = 0 \in \mathbb{R}^d\}$  be the ancestor of the family, and associate with it the parameter  $r_0 > 0$ . Then define the offspring  $\mathbb{X}_1(r_0)$  to be  $\mathcal{H}_\lambda \cap B(x_0; r_0^{1/p})$ .  $\mathbb{X}_1(r_0)$  is the set of points in  $\mathcal{H}_\lambda$  that may be reached from  $x_0$  with a single edge with path length at most  $r_0$  in power-weighted sense. Note that  $\mathbb{E}|\mathbb{X}_1(r_0)| = \lambda V_d r_0^{d/p}$  where  $|\mathbb{X}_1(r_0)|$  denotes the cardinality of  $\mathbb{X}_1(r_0)$ , and  $V_d$  denotes the volume of  $B(0; 1)$ .

For each offspring  $x_{1,k} \in \mathbb{X}_1(r_0)$ , we associate the parameter  $r_{1,k} = r_0 - |x_{1,k} - x_0|^p$ . Then  $\mathcal{H}_\lambda$  in the union of  $B(x_{1,k}; r_{1,k}^{1/p}) - \{x_{1,k}\}$  over  $k$  is the set of points that may be reached from  $x_0$  with exactly two edges, while the power-weighted path length is at most  $r_0$ . Define  $\mathbb{X}_2(r_0)$  to be the collection of all the second generation offspring, and define recursively the  $n$ th generation offspring  $\mathbb{X}_n(r_0)$ . Then  $\mathbb{X}_n(r_0)$  is the set of all the points that may be reached in  $n$  hops from the ancestor  $x_0$  within path length  $r_0$ . See Figure 2. We prove by induction that

$$(38) \quad \mathbb{E}|\mathbb{X}_n(r_0)| \leq (\lambda V_d r_0^{d/p})^n \frac{\Gamma(1 + d/p)^n}{\Gamma(1 + nd/p)}.$$

We mentioned above that  $\mathbb{E}|\mathbb{X}_1(r_0)| = \lambda V_d r_0^{d/p}$ , and (38) is true for  $n = 1$ . For general  $n$ , apply the Campbell–Mecke formula [Baddeley (2007), Theorem 3.2, page 48] to see that

$$(39) \quad \begin{aligned} \mathbb{E}|\mathbb{X}_n(r_0)| &\leq \lambda \int_{B(x_0; r_0^{1/p})} \mathbb{E}|\mathbb{X}_{n-1}(r_0 - |x - x_0|^p)| dx \\ &\leq \lambda^n V_d^{n-1} \frac{\Gamma(1 + d/p)^{n-1}}{\Gamma(1 + (n-1)d/p)} \\ &\quad \times \int_{B(x_0; r_0^{1/p})} (r_0 - |x - x_0|^p)^{(n-1)d/p} dx. \end{aligned}$$

The last integral evaluates to

$$\begin{aligned} &\int_{B(x_0; r_0^{1/p})} (r_0 - |x - x_0|^p)^{(n-1)d/p} dx \\ &= V_d r_0^{(n-1)d/p} d \int_0^{r_0^{1/p}} \left(1 - \frac{u^p}{r_0}\right)^{(n-1)d/p} u^{d-1} du \end{aligned}$$

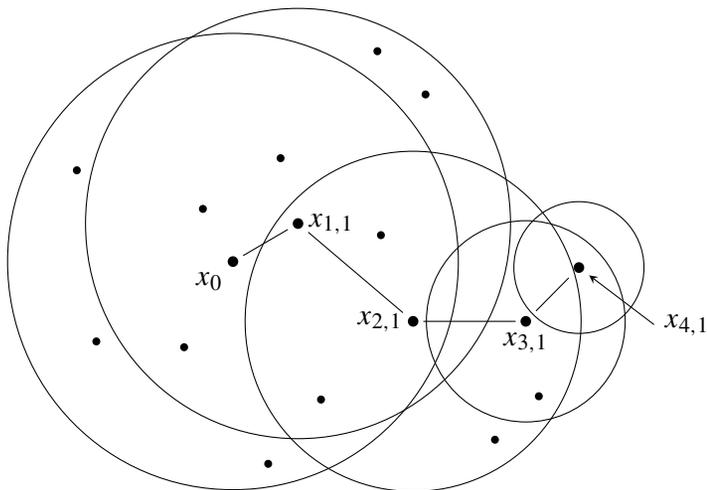


FIG. 2. A run through the family tree generated by  $\mathbb{X}_n$  with  $p = 2$ . The point  $x_0$  is the ancestor with parameter  $r_0 = 9$ . This means that all the runs through the family tree are paths with power-weighted length less than  $r_0^{1/p} = 3$ . Here,  $x_{1,1} \in \mathbb{X}_1$  is among the first generations since it is within  $B(x_0; r_0^{1/p})$ , and  $x_{2,1} \in \mathbb{X}_2$  is among the second generations since it is within the balls centered at the first generation offsprings, for example,  $x_{1,1}$ . This particular run ends at  $x_{4,1}$  as there is no point in the vicinity. In this example, the power-weighted path length is  $\sqrt{1^2 + 2^2 + 1.5^2 + 1^2} = \sqrt{8.25} < 3$ . Note that  $x_{2,1}$  is also in the ball centered at  $x_0$ , so it is also a first generation offspring. Some other runs through the family tree will have the point  $x_{2,1}$  as a first generation offspring.

$$\begin{aligned}
 &= V_d r_0^{nd/p} \frac{d}{p} \int_0^1 (1-v)^{(n-1)d/p} v^{d/p-1} dv \\
 &= V_d r_0^{nd/p} \frac{\Gamma(1+d/p)\Gamma(1+(n-1)d/p)}{\Gamma(1+nd/p)}.
 \end{aligned}$$

Note that a spherical coordinate transformation was used in the first equality, a transformation  $v = u^p/r_0$  was used in the second equality, and the third equality was obtained by properties of the beta function. Substituting the expression in the last line into (39) establishes (38).

Using the Markov inequality and Stirling’s approximation, we have

$$\log P(\mathbb{X}_n(r_0) \neq \emptyset) \leq n \log \left( V_d \Gamma \left( 1 + \frac{d}{p} \right) \left( \frac{c_0}{c_1} \cdot \frac{pe}{d} \right)^{d/p} \right) + O(\log n)$$

as  $n \rightarrow \infty$ . Note that if a path starting at  $x$  passes through more than  $n \geq c_1 \lambda^{1/d}$  nodes and has path length less than  $r_0 \leq c_0 \lambda^{(1-p)/d}$ , then the  $n$ th generation set  $\mathbb{X}_n(r_0)$  will not be empty. Lemma 11 follows since, if the ratio  $c_1/c_0$  is sufficiently large, the logarithm term above is negative.  $\square$

4.2. Mean convergence for Poisson point processes.

LEMMA 12. Consider the shortest path length  $\mathcal{L}_1(0, te_1)$  from  $0 \in \mathbb{R}^d$  to  $te_1 \in \mathbb{R}^d$  in  $\mathcal{H}_1$  for  $t > 0$ . Then the limit

$$(40) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \mathcal{L}_1(0, te_1) = C(d, p)$$

exists.

In addition, if  $b = b_t$  is a function of  $t$  satisfying  $\liminf_t b_t = \infty$ , then

$$(41) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \mathcal{L}_1(0, te_1; b_t) = C(d, p).$$

Recall that  $\mathcal{L}_1(0, te_1; b_t)$  denotes  $L(0, te_1; \mathcal{H}_1 \cap T(0, te_1; b_t))$  from (36) and (37).

When  $T(0, te_1; b) = \mathbb{R}^d$ , that is,  $b = +\infty$ , (40) is a consequence of, for example, Howard and Newman (2001), Section 4. The main difference is the case when  $b < +\infty$ . Howard and Newman [(2001), Theorem 2.4] states that the probability that  $\mathcal{L}_1(0, te_1) \neq \mathcal{L}_1(0, te_1; b_t)$  is exponentially small of order at least  $t^{3p\varepsilon/4}$  when  $b_t \geq t^{3/4+\varepsilon}$  for some  $\varepsilon > 0$ . Lemma 12 is weaker in the sense that it only asserts closeness in the mean. On the other hand, Lemma 12 is stronger in the sense that the assumption on  $b_t$  is relaxed so that  $b_t$  need only diverge to infinity, and the rate of growth may even be sub-polynomial.

PROOF OF LEMMA 12. Initially, we let  $b > 0$  be a constant instead of a function of  $t$ . This assumption is removed later in the proof. Recall the definition of function  $h$  in (4),

$$h(x, y; u) = |x - u|^p + |u - y|^p - |x - y|^p.$$

Let

$$T(b) = \bigcup_{s>0} T(-se_1, +se_1; b),$$

and let

$$\xi_t(\lambda, b) = \sup\{|u - te_1| : u \in T(b), h(u, te_1; v) \geq 0 \text{ for all } v \in \mathcal{H}_\lambda \cap T(b)\}.$$

In other words,  $\xi_t(\lambda, b)$  denotes an upper bound distance of  $u \in T(b)$  from  $te_1$  such that the shortest path from  $te_1$  to  $u$  is the direct path  $te_1 \rightarrow u$ . From the continuity of function  $h$ , it is not difficult to show that there exist constants  $A, \delta > 0$  and constant integers  $k, m > 0$ , all independent of  $b$  and  $\lambda$ , such that for all  $t \in \mathbb{R}$ ,

$$(42) \quad \mathbb{E} \xi_t(\lambda, b)^p \leq \frac{k\Gamma(1 + p/d)}{(\lambda A)^{p/d}} + \frac{m2^p \Gamma(1 + p)}{\lambda^p (\delta b)^{p(d-1)}}.$$

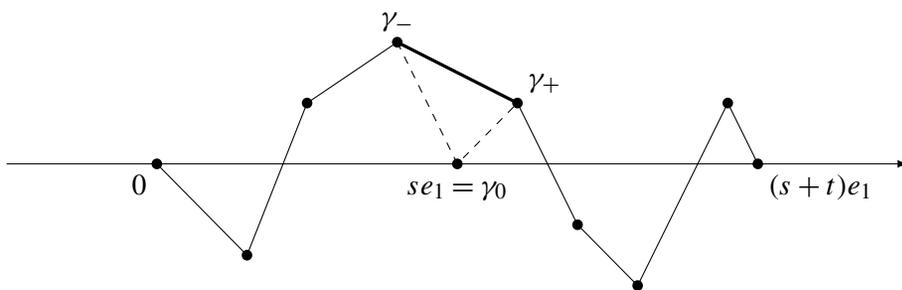


FIG. 3. An illustration of the path pasting procedure. A new path from 0 to  $(s + t)e_1$  is created by removing  $se_1 = \gamma_0$  and joining  $\gamma_-$  and  $\gamma_+$ . Only the end points are fixed points in the new path.

It is not surprising that the upper bound does not depend on  $t$  since  $\mathcal{H}_\lambda$  is homogeneous. For a simple proof of this, see Hwang (2012), Lemma 2.5, equation (2.14).

Let  $s, t > 0$ . Consider the shortest path  $\mathcal{L}_1(0, se_1; b)$  between 0 and  $se_1$ , and let  $\gamma_-$  denote the node that directly connects to  $se_1$ . Similarly, consider the shortest path for  $\mathcal{L}_1(se_1, (s + t)e_1; b)$  and let  $\gamma_+$  denote the node that directly connects to  $se_1$ . Therefore,  $\gamma_-$  and  $\gamma_+$  are Poisson sample points incident to  $se_1$ . For convenience let  $\gamma_0 = se_1$ . Remove  $\gamma_0 = se_1$  in the two paths, and join the nodes  $\gamma_-$  and  $\gamma_+$  so that we have a new path connecting 0 and  $(s + t)e_1$ , as indicated in Figure 3. This new path has length that is an upper bound on  $\mathcal{L}_1(0, (s + t)e_1; b)$ ,

$$\begin{aligned} \mathcal{L}_1(0, (s + t)e_1; b) &\leq \mathcal{L}_1(0, se_1; b) + \mathcal{L}_1(se_1, (s + t)e_1; b) \\ &\quad + (|\gamma_0 - \gamma_-| + |\gamma_+ - \gamma_0|)^p. \end{aligned}$$

Note that both  $|\gamma_0 - \gamma_-|$  and  $|\gamma_+ - \gamma_0|$  are bounded above by  $\xi_s(1, b)$ , and  $\mathbb{E}\xi_s(1, b)^p$  is finite by (42). Therefore,  $\mathbb{E}\mathcal{L}_1(0, (s + t)e_1; b)$  is bounded above by

$$\begin{aligned} &\mathbb{E}\mathcal{L}_1(0, se_1; b) + \mathbb{E}\mathcal{L}_1(se_1, (s + t)e_1; b) + \mathbb{E}(2\xi_s(1, b))^p \\ &= \mathbb{E}\mathcal{L}_1(0, se_1; b) + \mathbb{E}\mathcal{L}_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p. \end{aligned}$$

The equality holds by the translation invariant property of the distribution of  $\mathcal{H}_1$ . Therefore,  $\mathbb{E}\mathcal{L}_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p$  is a sub-additive function of  $t$ . Note that  $\mathbb{E}\mathcal{L}_1(0, te_1; b) \leq t^p$ . A standard proof of Fekete’s lemma [see, e.g., Steele (1997), Lemma 1.2.1] may be easily adapted to sub-additive functions that are bounded in bounded intervals. Apply Fekete’s lemma to the sub-additive function  $\mathbb{E}\mathcal{L}_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p$ , then

$$\begin{aligned} (43) \quad &\lim_{t \rightarrow \infty} \frac{\mathbb{E}\mathcal{L}_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p}{t} \\ &= \inf_{t > 0} \frac{\mathbb{E}\mathcal{L}_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p}{t}, \end{aligned}$$

and we denote the limit by  $\kappa(d, p; b)$ . Note that  $\mathbb{E}\xi_0(1, b)^p$  does not depend on  $t$ , hence (43) implies that  $\lim_t t^{-1}\mathbb{E}\mathcal{L}_1(0, te_1; b) = \kappa(d, p; b)$ .

Define

$$(44) \quad C(d, p) = \lim_{t \rightarrow \infty} \frac{\mathbb{E}\mathcal{L}_1(0, te_1)}{t}.$$

We now show that  $\kappa(d, p; b)$  converges to  $C(d, p)$  when  $b \rightarrow \infty$ . Choose an arbitrary  $\varepsilon > 0$ . By (42) and by the fact that  $C(d, p)$  is the limit of  $t^{-1}\mathbb{E}\mathcal{L}_1(0, te_1)$ , there exists  $T > 0$  such that

$$\frac{1}{T}\mathbb{E}\mathcal{L}_1(0, Te_1) < C(d, p) + \frac{\varepsilon}{3},$$

and

$$\frac{1}{T}\mathbb{E}(2\xi_0(1, b))^p < \frac{\varepsilon}{3},$$

for all  $b > 1$ . For this fixed  $T$ , note that  $\lim_{b \rightarrow \infty} \mathcal{L}_1(0, Te_1; b) = \mathcal{L}_1(0, Te_1)$  monotonically from above almost surely, and by the monotone convergence theorem, there exists  $B > 1$  such that for all  $b > B$  and fixed  $T$ ,

$$\frac{1}{T}\mathbb{E}\mathcal{L}_1(0, Te_1; b) \leq \frac{1}{T}\mathbb{E}\mathcal{L}_1(0, Te_1) + \frac{\varepsilon}{3}.$$

Combining the three inequalities above with (43), we obtain

$$\kappa(d, p; b) \leq \frac{1}{T}(\mathbb{E}\mathcal{L}_1(0, Te_1; b) + \mathbb{E}(2\xi_0(1, b))^p) \leq C(d, p) + \varepsilon,$$

for all  $b > B$ . Therefore,  $\kappa(d, p; b)$  converges to  $C(d, p)$  as  $b \rightarrow \infty$ .

Finally, suppose  $b = b_t$  is a function of  $t$  rather than a constant. If  $\liminf_t b_t = \infty$ , then

$$\begin{aligned} C(d, p) &\leq \lim_{t \rightarrow \infty} \frac{1}{t}\mathbb{E}\mathcal{L}_1(0, te_1; b_t) \\ &\leq \lim_{t \rightarrow \infty} \frac{1}{t}\mathbb{E}\mathcal{L}_1(0, te_1; B) = \kappa(d, p; B), \end{aligned}$$

for any fixed  $B > 0$ . (41) follows as  $B \rightarrow \infty$  on the right-hand side.  $\square$

For the readers' benefit, we establish two use cases of Lemma 12.

**COROLLARY 13.** *The following two cases follow from Lemma 12:*

(i) *For every  $\varepsilon > 0$  there exists a constant  $t_0 > 0$  such that for all  $\lambda > 0$  and  $r > 0$  satisfying  $\lambda^{1/d}r > t_0$ ,*

$$\left| \frac{\mathbb{E}\mathcal{L}_\lambda(0, re_1; r)}{r\lambda^{(1-p)/d}} - C(d, p) \right| < \varepsilon.$$

(ii) Let  $z \in \mathbb{R}^d$  and  $R_2 > R_1 > 0$ . Let  $b > 0$  and  $\varepsilon > 0$ . Then there exists  $\lambda_0 > 0$  such that for all  $\lambda \geq \lambda_0$  and  $x, y \in B(z; R_1)$  with  $|x - y| \geq b$ , we have

$$\left| \frac{\mathbb{E}L(x, y; \mathcal{H}_\lambda \cap B(z; R_2))}{\lambda^{(1-p)/d}|x - y|} - C(d, p) \right| < \varepsilon.$$

PROOF.  $\mathcal{H}_1$  scaled by factor of  $\lambda^{-1/d}$  has identical distribution to  $\mathcal{H}_\lambda$ . At the same time, power-weighted shortest path lengths are scaled by factor of  $\lambda^{-p/d}$ . From (41), we have

$$(45) \quad \frac{\mathbb{E}\mathcal{L}_\lambda(0, \lambda^{-1/d}te_1; \lambda^{-1/d}b_t)}{\lambda^{(1-p)/d}\lambda^{-1/d}t} = \frac{1}{t}\mathbb{E}\mathcal{L}_1(0, te_1; b_t).$$

Choose  $b_t = t$  and  $t = \lambda^{1/d}r$  to obtain (i) from Lemma 12.

For (ii), note that

$$\mathcal{L}_\lambda(x, y) \leq L(x, y; \mathcal{H}_\lambda \cap B(z; R_2)) \leq \mathcal{L}_\lambda(x, y; R_2 - R_1)$$

since  $x, y \in B(z; R_1)$ . By translation- and rotation-invariance of  $\mathcal{H}_\lambda$ ,

$$\mathbb{E}\mathcal{L}_\lambda(0, |x - y|e_1) \leq \mathbb{E}L(x, y; \mathcal{H}_\lambda \cap B(z; R_2)) \leq \mathbb{E}\mathcal{L}_\lambda(0, |x - y|e_1; R_2 - R_1).$$

Choose  $t = \lambda^{1/d}|x - y|$  and  $b_t = \lambda^{1/d}(R_2 - R_1)$  for (45). Then (ii) follows from Lemma 12.  $\square$

4.3. *Shortest path size.* In order to prove Lemma 4, we need an upper bound for shortest path lengths in  $\mathcal{H}_\lambda$ .

LEMMA 14. Let  $z \in \mathbb{R}^d$ ,  $R_2 > R_1 > 0$ . Let  $b > 0$  and  $\varepsilon > 0$ . For every  $x, y \in B(z; R_1)$ , let  $E_\lambda(x, y, \varepsilon)$  denote the event that

$$(46) \quad \frac{L(x, y; \mathcal{H}_\lambda \cap B(z; R_2))}{\lambda^{(1-p)/d}|x - y|} \leq C(d, p) + \varepsilon.$$

Then there exist  $\lambda_0 > 0$  and  $\theta_7 > 0$  such that for all  $\lambda \geq \lambda_0$  and  $|x - y| \geq b$ ,

$$(47) \quad 1 - P(E_\lambda(x, y, \varepsilon)) \leq \exp(-\theta_7\lambda^{1/(d+2p-1)}).$$

PROOF. Let  $0 < r < R_2 - R_1$ . Recall the notation  $T(x, y; r)$  and  $\mathcal{L}_\lambda(x, y; r)$  from (36) and (37). Since  $x, y \in B(z; R_1)$  and  $r < R_2 - R_1$ ,

$$\mathcal{H}_\lambda \cap T(x, y; r) \subset \mathcal{H}_\lambda \cap B(z; R_2),$$

and hence

$$L(x, y; \mathcal{H}_\lambda \cap B(z; R_2)) \leq \mathcal{L}_\lambda(x, y; r).$$

Let  $E'_\lambda(x, y, \varepsilon)$  denote the event that (46) holds with  $\mathcal{L}_\lambda(x, y; r)$  in place of  $L(x, y; \mathcal{H}_\lambda \cap B(z; R_2))$ . By the inequality above, if  $E'_\lambda(x, y, \varepsilon)$  occurs then

$E_\lambda(x, y, \varepsilon)$  occurs, hence  $P(E'_\lambda(x, y, \varepsilon)) \leq P(E_\lambda(x, y, \varepsilon))$ . Therefore, it is sufficient to show that (47) holds with  $1 - P(E'_\lambda(x, y, \varepsilon))$  in place of  $1 - P(E_\lambda(x, y, \varepsilon))$ .

As in Lemma 12, by the convex property of the power functions,  $\mathcal{L}_\lambda(0, 2re_1; r)$  may be bounded above by  $\mathcal{L}_\lambda(0, re_1; r) + \mathcal{L}_\lambda(re_1, 2re_1; r) + (2^{p-1} - 1)(Z_1^p + Y_0^p)$ , where  $Z_k$  and  $Y_k$  are the first and the last edge lengths in  $\mathcal{L}_\lambda(kre_1, (k + 1)re_1; r)$ , respectively. In Figure 3, when  $s = r$  and  $s + t = 2r$ ,  $Z_1$  and  $Y_0$  correspond to  $|\gamma_+ - \gamma_0|$  and  $|\gamma_0 - \gamma_-|$ , respectively.

Note that the shortest path for  $\mathcal{L}_\lambda(kre_1, (k + 1)re_1; r)$  is not likely to be the direct path  $kre_1 \rightarrow (k + 1)re_1$ . That is, if it were the direct path, then as in the proof of Lemma 3, there exists  $\delta > 0$  such that  $\mathcal{H}_\lambda$  is empty in the open ball of radius  $\delta r$  centered at the middle of  $kre_1$  and  $(k + 1)re_1$ . Such event happens with probability at most  $\exp(-\lambda\theta' r^d)$ , where  $\theta'$  denotes the volume of an open ball of radius  $\delta$ . If none of the shortest paths for  $\mathcal{L}_\lambda(kre_1, (k + 1)re_1; r)$  is a direct path, then the previous pasting procedure used in Lemma 12 may be repeated so that

$$(48) \quad \mathcal{L}_\lambda(0, mre_1; r) \leq \sum_{k=0}^{m-1} (\mathcal{L}_\lambda(kre_1, (k + 1)re_1; r) + (2^{p-1} - 1)(Z_k^p + Y_k^p)),$$

with probability at least  $1 - m \exp(-\lambda\theta' r^d)$ .

If  $k, l$  are integers and  $l - k \geq 3$ , then  $T(kre_1, (k + 1)re_1; r)$  and  $T(lre_1, (l + 1)re_1; r)$  are disjoint, hence  $\mathcal{L}_\lambda(kre_1, (k + 1)re_1; r)$  and  $\mathcal{L}_\lambda(lre_1, (l + 1)re_1; r)$  are mutually independent, and so are  $Z_k$  and  $Z_l$ , as well as  $Y_k$  and  $Y_l$ . Then the sum in (48) may split into  $K \geq 3$  sums of independent variables, and each sum has at least  $\lfloor m/K \rfloor$  summands. Note that each summand is almost surely bounded since  $Z_k^p + Y_k^p \leq \mathcal{L}_\lambda(kre_1, (k + 1)re_1; r) \leq r^p$ . Apply Azuma–Hoeffding’s inequality [Azuma (1967)] for  $K = 4$  separate sequences to obtain

$$(49) \quad P\left(\frac{\mathcal{L}_\lambda(0, mre_1; r)}{\lambda^{(1-p)/d}mr} \geq \mu_r + \varepsilon\right) \leq m \exp(-\lambda\theta' r^d) + 4 \exp\left(-\frac{(m - 3)\varepsilon^2}{2^{1+2p}(\lambda^{1/d}r)^{2(p-1)}}\right),$$

where

$$\mu_r = \frac{\mathbb{E}\mathcal{L}_\lambda(0, re_1; r) + (2^{p-1} - 1)(\mathbb{E}Z_0^p + \mathbb{E}Y_0^p)}{r\lambda^{(1-p)/d}}.$$

Let  $\beta = 1/(d + 2p - 1)$ . Set  $mr = |x - y|$  and  $m = \lfloor (\lambda^{1/d}|x - y|)^{1-\beta} \rfloor$ . Note that  $r = |x - y|/m \leq 2R_1/m$  is less than  $R_2 - R_1$  when  $\lambda^{1/d}|x - y|$  is sufficiently large. By the definition, both  $\mathbb{E}Z_k^p$  and  $\mathbb{E}Y_{k-1}^p$  are bounded above by  $\mathbb{E}\xi_{kb}(\lambda, r)^p = \mathbb{E}\xi_0(\lambda, r)^p$  in (42), and a direct computation with (42) shows that  $\mathbb{E}\xi_0(\lambda, r)^p$  divided by  $\lambda^{(1-p)/d}r$  shrinks to zero when  $\lambda^{1/d}r \geq (\lambda^{1/d}|x - y|)^\beta \rightarrow \infty$ . See Hwang

[(2012), Lemma 2.5] for more details. Apply Corollary 13 to see that  $\mu_r$  converges to  $C(d, p)$  as  $\lambda^{1/d}r \rightarrow \infty$ . Then (49) becomes

$$\begin{aligned} P\left(\frac{\mathcal{L}_\lambda(x, y; r)}{\lambda^{(1-p)/d}|x-y|} \geq C(d, p) + 2\varepsilon\right) &\leq (\lambda|x-y|^d)^{(1-\beta)/d} \exp(-\theta'(\lambda|x-y|^d)^\beta) + 4 \exp\left(-\frac{(\lambda|x-y|^d)^\beta}{2^{2(p+1)}}\varepsilon^2\right) \\ &\leq (\lambda(2R_1)^d)^{(1-\beta)/d} \exp(-\theta'(\lambda b^d)^\beta) + 4 \exp\left(-\frac{(\lambda b^d)^\beta}{2^{2(p+1)}}\varepsilon^2\right), \end{aligned}$$

for all sufficiently large  $\lambda$ .  $\square$

PROOF OF LEMMA 4. Fix constants  $A > 1$  and  $0 < A' < 1$ . Let  $N$  and  $N'$  be independent Poisson variables with mean  $nA$  and  $nA'$ , respectively. Let  $a = Af(z)$  and  $a' = A'f(z)$ . Let  $H_n$  denote the event that  $N \geq n$  and  $N' \leq n$ . Let  $K_n$  denote the event that

$$(50) \quad \frac{L(x, y; \mathcal{H}_{na'} \cap B(z; R_2))}{(na')^{(1-p)/d}|x-y|} \leq C(d, p) + \frac{\varepsilon}{2}.$$

We first show that if both  $H_n$  and  $K_n$  occur, then the following conditions are satisfied:

- (i)  $L(x, y; \mathcal{X}_n \cap B(z; R_2))$  is a path in  $\mathcal{H}_{na}$ .
- (ii)  $L(x, y; \mathcal{X}_n \cap B(z; R_2)) \leq (C(d, p) + \varepsilon/2)(na')^{(1-p)/d}|x-y|$ .

Note that restriction of  $\mathcal{H}_{na}$  to  $B(z; R_2)$  may be realized as  $\mathcal{X}_N \cap B(z; R_2)$  since  $\mathbb{E}N = nA$ . Since  $H_n$  is assumed to occur, it follows that  $N \geq n$ , and  $\mathcal{X}_n \cap B(z; R_2) \subset \mathcal{X}_N \cap B(z; R_2) = \mathcal{H}_{na} \cap B(z; R_2) \subset \mathcal{H}_{na}$ . Therefore, (i) holds.

For condition (ii),  $H_n$  is assumed to occur, so we have  $N' \leq n$ . Then similar to the previous argument,  $\mathcal{H}_{na'} \cap B(z; R_2) = \mathcal{X}_{N'} \cap B(z; R_2) \subset \mathcal{X}_n \cap B(z; R_2)$  and it follows that  $L(x, y; \mathcal{X}_n \cap B(z; R_2)) \leq L(x, y; \mathcal{H}_{na'} \cap B(z; R_2))$ . Condition (ii) follows by (50).

Recall that  $G_n = G_n(x, y)$  denotes the event  $\#L(x, y; \mathcal{X}_n \cap B(z; R_2))$  is less than or equal to  $C_*(nf(z))^{1/d}|x-y|$ . We have shown that when  $H_n$  and  $K_n$  occur, (i) and (ii) hold, and an application of Lemma 11 shows that

$$1 - P(G_n | H_n \cap K_n) \leq \exp(-C(nf(z))^{1/d}|x-y|) \leq \exp(-C'n^{1/d})$$

for some  $C, C' > 0$  when  $C_* > (C(d, p) + \varepsilon/2)A^{p/d}A'^{(1-p)/d}\rho_0$ . See Lemma 11 for the constant  $\rho_0$ .

By Lemma 14,  $1 - P(K_n)$  is bounded above by  $\exp(-\theta_7(na')^{1/(d+2p-1)})$  for sufficiently large  $n$ , since  $a' > 0$  is a fixed constant. By the Chernoff bound [Billingsley (1995), Theorem 9.3],  $1 - P(H_n)$ , that is, the probability that either  $N < n$  or  $N' > n$ , is exponentially small in  $n$ .

Note that  $1 - P(G_n)$  is bounded above by the sum of  $(1 - P(G_n | H_n \cap K_n))$ ,  $(1 - P(H_n))$ , and  $(1 - P(K_n))$ . The lemma follows from the observation that the overall decay is determined by the summand with slowest decay rate, and it is  $(1 - P(K_n))$ , which is exponentially small in  $n^{1/(d+2p-1)}$ .  $\square$

4.4. Mean convergence in i.i.d. cases: de-Poissonization.

PROOF OF PROPOSITION 5. For convenience, let  $L_k$  denote  $L(x, y; \mathcal{X}_k \cap B(z; R_2))$  for all  $k \geq 0$ . Recall that  $\alpha = 1/(d + 2p)$ . Let  $\tau_k = (kf(z))^{1/d}|x - y|$ , and  $\zeta_k = (kf(z))^{(\alpha-1)/d}|x - y|^\alpha$ .

Let  $C_* > 0$  as in Lemma 4 and suppose that the number of nodes  $\#L_k$  in the shortest path  $L_k$  is less than  $C_*\tau_k$ . Suppose that the event  $F_k(|x - y|)$  from Lemma 3 occurred so that all the shortest path edge lengths are at most  $\zeta_k$ . When a sample point from  $\mathcal{X}_k$  is discarded,  $L_{k-1}$  remains the same as  $L_k$  if the discarded sample point were not a node in  $L_k$ . Furthermore, since edge lengths are at most  $\zeta_k$ ,  $L_{k-1}$  and  $L_k$  may differ at most by  $(2\zeta_k)^p$ . Therefore,

$$(51) \quad \mathbb{E}L_{k-1} - \mathbb{E}L_k \leq \frac{C_*\tau_k}{k}(2\zeta_k)^p + h_k\mathbb{E}L_0,$$

where  $h_k$  denotes the probability that either  $\#L_k > C_*\tau_k$ , or the event  $F_k(|x - y|)$  does not occur.  $\mathbb{E}L_0$  in the last term is chosen because  $\mathbb{E}L_k \leq \mathbb{E}L_0$  for all  $k > 0$ .

Let  $N$  be a Poisson variable with mean  $n$ . Write

$$\mathbb{E}L_N = \sum_{k \geq 0} \mathbb{E}L_k P(N = k).$$

The difference  $|\mathbb{E}L_n - \mathbb{E}L_N|$  is bounded above by

$$\begin{aligned} & \sum_{k \geq 0} |\mathbb{E}L_n - \mathbb{E}L_k| P(N = k) \\ & \leq \mathbb{E}L_0 P\left(N < \frac{n}{2}\right) + \sum_{k \geq 2^{-1}n} |\mathbb{E}L_n - \mathbb{E}L_k| P(N = k). \end{aligned}$$

Note that the first term on the right of (51) is monotonically decreasing in  $k$ , since both  $\tau_k/k$  and  $\zeta_k$  monotonically decrease in  $k$  for fixed  $n$ . Therefore, for  $k \geq 2^{-1}n$ ,

$$|\mathbb{E}L_n - \mathbb{E}L_k| \leq \frac{C_*\tau_k}{2^{-1}n} \cdot (2\zeta_k)^p |n - k| + \mathbb{E}L_0 \sum_{l > 2^{-1}n} h_l.$$

Since  $\mathbb{E}|N - n| \leq \sqrt{n}$  and  $\mathbb{E}L_0 = |x - y|^p$ , after expanding  $\tau_k$  and  $\zeta_k$  we have

$$\begin{aligned} & \frac{|\mathbb{E}L_n - \mathbb{E}L_N|}{(nf(z))^{(1-p)/d}|x - y|} \\ & \leq O\left(\frac{((nf(z))^{1/d}|x - y|)^{\alpha p}}{\sqrt{n}}\right) + \frac{P(N < 2^{-1}n) + \sum h_l}{((nf(z))^{1/d}|x_n - y_n|)^{1-p}}, \end{aligned}$$

where the summation  $\sum h_l$  is still for  $l > 2^{-1}n$ . The first term on the right decays to zero since  $n^{\alpha p/d} < \sqrt{n}$  and  $|x - y| < 2R_1$ . The second term also decays to zero since, while the denominator has at most polynomial decay in  $n$ ,  $\sum h_l$  in the numerator has exponential decay in  $n^\alpha$  by Lemmas 3 and 4, and  $P(N < 2^{-1}n)$  in the numerator has exponential decay in  $n$  by the Chernoff bound [Billingsley (1995), Theorem 9.3]. Note that  $\mathcal{X}_N \cap B(z; R_2)$  is identically distributed as  $\mathcal{H}_{nf(z)} \cap B(z; R_2)$ , and the proposition follows since the difference of  $\mathbb{E}L_N = \mathbb{E}L(x, y; \mathcal{H}_{nf(z)} \cap B(z; R_2))$  from  $C(d, p)$  is less than  $\varepsilon$  for sufficiently large  $n$  by Corollary 13.  $\square$

APPENDIX

Here, we show how (3) can be derived from Theorem 7.

As we did in the proof of Corollary 13, scale the space by factor of  $\lambda^{-1/d}$  with choice  $\lambda = t^d$  to obtain

$$(52) \quad P\left(\left|\frac{1}{t}\mathcal{L}_1(0, te_1) - C(d, p)\right| > s\right) = P\left(\left|\frac{\mathcal{L}_\lambda(0, e_1)}{\lambda^{(1-p)/d}} - C(d, p)\right| > s\right).$$

Let  $N$  be a Poisson random variable with mean  $\lambda V_d R^d$ , where  $V_d$  denotes the volume of a unit ball. Fix  $0 < \delta < 1$  and  $0 < \varepsilon < C(d, p)$ . Suppose that:

- (i)  $(1 - \delta)\mathbb{E}N \leq N \leq (1 + \delta)\mathbb{E}N$ , and
- (ii) the events  $E_n(\varepsilon)$  from Theorem 7 occur for all  $n$  in the range  $(1 - \delta)\mathbb{E}N \leq n \leq (1 + \delta)\mathbb{E}N$ .

For Theorem 7, choose  $z = 0 \in \mathbb{R}^d$ ,  $b = 0$ , and pick  $R$  so that  $R > 8(C(d, p) + \varepsilon)/5(C(d, p) - \varepsilon)$  and  $R > 4$  so that  $e_1 \in B(z; R/4)$ .

Note that  $\mathcal{H}_\lambda \cap B(z; R)$  may be realized as  $\mathcal{X}_N$  where  $X_1, X_2, \dots$  are uniform i.i.d. random variables in  $B(z; R)$ , and thus  $L(0, e_1; \mathcal{H}_\lambda \cap B(z; R)) = L(0, e_1; \mathcal{X}_N)$ . Therefore, under the assumptions (i) and (ii), we have

$$\left|\frac{L(0, e_1; \mathcal{H}_\lambda \cap B(z; R))}{(Nf(z))^{(1-p)/d}} - C(d, p)\right| \leq \varepsilon,$$

and

$$\frac{L(0, u; \mathcal{H}_\lambda \cap B(z; R))}{(Nf(z))^{(1-p)/d}} \geq \frac{5}{8}R(C(d, p) - \varepsilon) > C(d, p) + \varepsilon$$

for all  $u \notin B(0; R)$  by the choice of  $R$ . Therefore, the path  $\mathcal{L}_\lambda(0, e_1) = L(0, e_1; \mathcal{H}_\lambda)$  is contained in  $B(z; R)$  and

$$\left|\frac{\mathcal{L}_\lambda(0, e_1)}{(Nf(z))^{(1-p)/d}} - C(d, p)\right| \leq \varepsilon.$$

Note that  $f(z) = (V_d R^d)^{-1}$ . Using the condition (i) and  $\mathbb{E}N = \lambda V_d R^d$ , we have

$$\frac{C(d, p) - \varepsilon}{(1 + \delta)^{(p-1)/d}} \leq \frac{\mathcal{L}_\lambda(0, e_1)}{\lambda^{(1-p)/d}} \leq \frac{C(d, p) + \varepsilon}{(1 - \delta)^{(p-1)/d}}.$$

We can choose  $\varepsilon$  and  $\delta$  small enough so that

$$\left| \frac{\mathcal{L}_\lambda(0, e_1)}{\lambda^{(1-p)/d}} - C(d, p) \right| \leq s.$$

In summary, the probability in (52) is bounded above by the probability that either (i) or (ii) does not occur.

From Theorem 7, there exists a constant  $\theta_4 > 0$  such that  $1 - P(E_n(\varepsilon)) \leq \exp(-\theta_4(nf(z))^\alpha)$  for all sufficiently large  $n$ . By (ii),  $nf(z) \geq (1 - \delta)\lambda$ . Denote by  $H_n$  the event of (i). Then for sufficiently large  $\lambda = t^d$ , (52) is bounded above by

$$\begin{aligned} & (1 - P(H_n)) + 2\delta\lambda V_d R^d \exp(-\theta_4((1 - \delta)\lambda)^\alpha) \\ &= (1 - P(H_n)) + \exp(-\theta_4(1 - \delta)^\alpha t^{d/(d+2p)} + O(\log t)). \end{aligned}$$

Note that  $1 - P(H_n)$  is exponentially small in  $n$  by the Chernoff bound [Billingsley (1995), Theorem 9.3], so that the tail is dominated by the second term.

## REFERENCES

- AZUMA, K. (1967). Weighted sums of certain dependent random variables. *Tôhoku Math. J.* (2) **19** 357–367. [MR0221571](#)
- BADDELEY, A. (2007). Spatial point processes and their applications. In *Stochastic Geometry* (W. Weil, ed.). *Lecture Notes in Math.* **1892** 1–75. Springer, Berlin. [MR2327290](#)
- BERNSTEIN, M., DE SILVA, V., LANGFORD, J. C. and TENENBAUM, J. B. (2000). Graph approximations to geodesics on embedded manifolds. Available at <http://isomap.stanford.edu/BdSLT.pdf>.
- BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. Wiley, New York. [MR1324786](#)
- CORMEN, T. H., LEISERSON, C. E., RIVEST, R. L. and STEIN, C. (2009). *Introduction to Algorithms*, 3rd ed. MIT Press, Cambridge, MA. [MR2572804](#)
- COSTA, J. A. and HERO, A. O. (2004). Geodesic entropic graphs for dimension and entropy estimation in manifold learning. *IEEE Trans. Signal Process.* **52** 2210–2221. [MR2085582](#)
- DIJKSTRA, E. W. (1959). A note on two problems in connexion with graphs. *Numer. Math.* **1** 269–271. [MR0107609](#)
- HAMMERSLEY, J. M. (1966). First-passage percolation. *J. Roy. Statist. Soc. Ser. B* **28** 491–496. [MR0214163](#)
- HOWARD, C. D. and NEWMAN, C. M. (1997). Euclidean models of first-passage percolation. *Probab. Theory Related Fields* **108** 153–170. [MR1452554](#)
- HOWARD, C. D. and NEWMAN, C. M. (2001). Geodesics and spanning trees for Euclidean first-passage percolation. *Ann. Probab.* **29** 577–623. [MR1849171](#)
- HWANG, S. J. (2012). Geometric representations of high dimensional random data. Ph.D. thesis, Univ. Michigan. Available at <http://hdl.handle.net/2027.42/96097>.
- KESTEN, H. (1987). Percolation theory and first-passage percolation. *Ann. Probab.* **15** 1231–1271. [MR0905330](#)
- LAGATTA, T. and WEHR, J. (2010). A shape theorem for Riemannian first-passage percolation. *J. Math. Phys.* **51** 053502, 14. [MR2666980](#)
- MEESTER, R. and ROY, R. (1996). *Continuum Percolation*. *Cambridge Tracts in Mathematics* **119**. Cambridge Univ. Press, Cambridge. [MR1409145](#)
- O’NEILL, B. (1983). *Semi-Riemannian Geometry*. *Pure and Applied Mathematics* **103**. Academic Press, New York. [MR0719023](#)
- PENROSE, M. D. and YUKICH, J. E. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Probab.* **13** 277–303. [MR1952000](#)

- PENROSE, M. D. and YUKICH, J. E. (2013). Limit theory for point processes in manifolds. *Ann. Appl. Probab.* **23** 2161–2211. [MR3127932](#)
- STEELE, J. M. (1997). *Probability Theory and Combinatorial Optimization. CBMS-NSF Regional Conference Series in Applied Mathematics* **69**. SIAM, Philadelphia, PA. [MR1422018](#)
- TALAGRAND, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* **81** 73–205. [MR1361756](#)
- YUKICH, J. E. (1998). *Probability Theory of Classical Euclidean Optimization Problems. Lecture Notes in Math.* **1675**. Springer, Berlin. [MR1632875](#)
- YUKICH, J. E. (2000). Asymptotics for weighted minimal spanning trees on random points. *Stochastic Process. Appl.* **85** 123–138. [MR1730615](#)

S.-J. HWANG  
GOOGLE INC  
1600 AMPHITHEATRE PKWY  
MOUNTAIN VIEW, CALIFORNIA 94043  
USA  
E-MAIL: [sjhwang@google.com](mailto:sjhwang@google.com)

S. B. DAMELIN  
MATHEMATICAL REVIEWS  
THE AMERICAN MATHEMATICAL SOCIETY  
416 FOURTH ST.  
ANN ARBOR, MICHIGAN 48103  
USA  
E-MAIL: [damelin@umich.edu](mailto:damelin@umich.edu)

A. O. HERO III  
UNIVERSITY OF MICHIGAN  
1301 BEAL AVE. 4417 EECS  
ANN ARBOR, MICHIGAN 48109-2122  
USA  
E-MAIL: [hero@umich.edu](mailto:hero@umich.edu)