Shortest Path through Random Points

Sung Jin Hwang\(^1\), Steven B. Damelin\(^2\), and Alfred O. Hero III\(^1\)

\(^1\)University of Michigan
\(^2\)Wayne Country Day School

Abstract

Let \((M,g_1)\) be a compact \(d\)-dimensional Riemannian manifold for \(d > 1\). Let \(X_n\) be a set of \(n\) sample points in \(M\) drawn randomly from a smooth Lebesgue density \(f\), bounded away from zero. Let \(x, y\) be two points in \(M\). We prove that the normalized length of the power-weighted shortest path between \(x, y\) through \(X_n\) converges to the Riemannian distance between \(x, y\) under the metric \(g_p = f^{2(1-p)/d}g_1\), 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 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 Riemannian case, e.g., Costa.

\textit{AMS 2000 subject classifications:} Primary 60F15; secondary 60C05, 53B21

\textit{Keywords and phrases:} shortest path, power-weighted graph, Riemannian geometry, conformal metric
and Hero (2004) extended the MST asymptotics in the context of entropy and intrinsic dimensionality estimation. More general non-Euclidean extensions have been established by Penrose and Yukich (2011). 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; Howard and Newman 2001). 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 non-homogeneous continuum model and establish convergence of the shortest path lengths to density-dependent conformally 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 differentiable function is an infinitely differentiable function. 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 \( g_1 \) and \( d > 1 \). When \( M = \mathbb{R}^d \), we fix \( g_1 \) as the standard Euclidean metric. The use of the subscript for \( g_1 \) will become clear shortly.

Consider a probability space \((M, \mathcal{B}, P)\) where \( P \) is a probability distribution over Borel subsets \( \mathcal{B} \) of \( M \). Assume that the distribution has a Lebesgue probability density function \( f \) with respect to \( g_1 \). Let \( X_1, X_2, \ldots \) denote an i.i.d. sequence drawn from this density. For convenience we denote this sequence by \( X_n = \{X_1, \ldots, X_n\} \).

The sequence \( X_n \) 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. Observe that we will use indexing by \( n \) of generic non-random 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 conformal Riemannian metric \( g_p = f^{2(1-p)/d} g_1 \). 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 Riemannian metric \( g_p \) is well-defined for every \( x \) with \( f(x) > 0 \). In this paper, we assume \( p > 1 \) except for a few places where we compare with the un-deformed 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 when $M = \mathbb{R}^d$. 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 $A \cup \{x, y\}$. Here the edge weight between two points $u$ and $v$ is $\text{dist}_1(u, v)^p$ where $\text{dist}_1$ denotes the Riemannian distance under the metric $g_1$. For convenience, we use the shorthand notation $L_n(x, y)$ for $L(x, y; X_n)$.

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

### 2.1 Main result

The following is the main result of this paper.

**Theorem 1.** Assume that $M$ is compact, and that $f$ is smooth with $\inf_M f > 0$. Let $0 \leq b < 1$ and $c > 0$ be constants. Then for every fixed $\varepsilon > 0$,

$$\limsup_{n \to \infty} (n \inf f)^{\frac{b-1}{p+d}} \mathbb{P}\left\{ \sup_{x, y} \left| \frac{L_n(x, y)}{n^{(1-p)/d} \text{dist}_p(x, y)} - C(d, p) \right| > \varepsilon \right\} < 0,$$

where the supremum is taken over $x, y \in M$ such that

$$\text{dist}_1(x, y) \geq c(n \inf f)^{-b/(d+2p)},$$

and $\text{dist}_p$ denotes the Riemannian distance under $g_p$. $C(d, p)$ is a positive constant that only depends on $d$ and $p$.

The constant $C(d, p)$ is fixed throughout this paper. This constant is also denoted as $\mu$ in Howard and Newman (1997); Howard and Newman (2001).

An immediate consequence of the Borel-Cantelli lemma is the almost-sure convergence of $L_n(x, y)$.

**Corollary 2.** Suppose that the conditions assumed in Theorem 1 hold. Let $x, y \in M$. Then

$$\lim_{n \to \infty} n^{(p-1)/d} L_n(x, y) = C(d, p) \text{dist}_p(x, y) \quad \text{a.s.}$$

As in Theorem 1, $\text{dist}_p$ denotes the Riemannian distance under $g_p$,

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

where the infimum is taken over all piece-wise smooth curves $\gamma : [0, 1] \to M$ such that $\gamma(0) = x$ and $\gamma(1) = y$.

When $p = 1$, there is no power-weighting of the edges, $C(d, 1) = 1$ and the shortest path length is the Riemannian distance $\text{dist}_1(x, y)$ from $x$ to $y$. 


2.2 Discussion

Theorem 1 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 $X_n$ while our results hold for the important case of non-uniformly distributed nodes; (ii) our convergence rates improve upon those of Howard and Newman (2001).

Specifically, Howard and Newman (2001, Theorem 2.2) show the following bound on the shortest path lengths in a homogeneous Poisson point process. Let $L_\lambda(x, y)$ denote 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 $H_\lambda$ 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 exist constants $C_0$ and $C_1$ (depending on $b$) such that for $t > 0$ and $t^b \leq s \sqrt{t} \leq t^{\kappa_2 - b}$,

$$(1) \quad \Pr \left\{ \left| \frac{1}{t} L_1(0, te_1) - C(d, p) \right| > s \right\} \leq C_1 \exp \left( -C_0(s \sqrt{t})^{\kappa_1} \right).$$

Note that the fastest decay rate achievable by (1) is bounded above by an exponential decay of $t^{\min(1, d/p)/(4p+3)}$.

On the other hand, our Corollary 8 implies, after simple Poissonization of the sequence $X_n$,

$$(2) \quad \limsup_{\lambda \to \infty} t^{-d/(4p+2)} \Pr \left\{ \left| \frac{1}{t} L_1(0, te_1) - C(d, p) \right| > s \right\} < 0,$$

so that the decay is exponential in $t^{d/(d+2p)}$. Under the condition $d \geq 1$, $p > 1$, the decay rate (2) is faster than the rate (1).

It is useful 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 $X_n$ while BHH theory (Steele 1997; Yukich 1998) specifies limits of the total length of a graph spanning $X_n$, e.g., 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) and Penrose and Yukich (2011) 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.
3.1 Local convergence results

Theorem 1 states a convergence result of random variables in manifolds. Theorem 1 is obtained by 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 $X_1, \ldots, X_n$ is i.i.d. in $\mathbb{R}^d$ with probability density function (pdf) $f$ and assume that there exists a constant $f_m > 0$ such that the pdf $f(u) \geq f_m$ for all $u \in B(z; R)$. Fix a number $\alpha \in (0, 1)$.

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) $|X_i - X_j| > (nf_m)^{(\alpha-1)/d}$, and

(iii) the shortest path from $X_i$ to $X_j$ 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

$$\limsup_{n \to \infty} \frac{1}{(nf_m)^\alpha} \log(1 - P(F_n)) < 0.$$

**Proof.** Suppose that the direct path $X_i \to X_j$ is not the shortest path between $X_i$ and $X_j$, i.e., there exists $k \neq i, j$ such that the path $X_i \to X_k \to X_j$ is shorter than $X_i \to X_j$ as measured by the sum of power-weighted edge lengths: $|X_i - X_k|^p + |X_k - X_j|^p < |X_i - X_j|^p$.

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

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

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 $X_i \to X_k \to X_j$ is shorter than $X_i \to X_j$. Note that the volume of $\Theta(X_i, X_j)$ is a function of the distance $|X_i - X_j|$ and that a certain proportion of $\Theta(X_i, X_j)$ intersects with $B(z; R)$. 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 $X_3, X_4, \ldots, X_n$. Since it is assumed that $|X_1 - X_2| > (nf_m)^{(\alpha-1)/d},$

$$P(H_n(1, 2)) \leq (1 - \theta_1 f_m^\alpha n^{\alpha-1})^{n-2}.$$

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

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

and the claim follows. \qed
Next we provide following two propositions on the cardinality of the shortest paths, and mean convergence of $\mathbb{E}L_n$. Their proofs require some results from the theory of Poisson processes, and we defer them to section 4.

**Lemma 4.** Let $z \in \mathbb{R}^d$, $R_2 > R_1 > 0$, and $\alpha = (d + 2p)^{-1}$. Assume that the pdf $f$ is uniform and supported in $B(z; R_2)$. Let $x_n, y_n$ be sequences in $B(z; R_1)$ satisfying $\lim \inf_n (nf(z))^\alpha |x_n - y_n| = +\infty$.

Let $#L_n(x_n, y_n)$ denote the number of nodes in the shortest path. Then there exists $C_\ast > 0$ such that if $G_n$ denotes the event

\[
\#
\left\{ x_n, y_n \right\} \in G_n \Rightarrow \frac{#L_n(x_n, y_n)}{(nf(z))^{1/d} |x_n - y_n|} \leq C_\ast,
\]

then

\[
\limsup_{n \to \infty} \frac{1}{(nf(z))^\alpha |x_n - y_n|} \log(1 - P(G_n)) < 0.
\]

**Proposition 5.** Let $z \in \mathbb{R}^d$, $R_2 > R_1 > 0$, and $\alpha = (d + 2p)^{-1}$. Assume that the pdf $f$ is uniform and supported in $B(z; R_2)$. Let $x_n, y_n$ be sequences in $B(z; R_1)$ satisfying $\lim \inf_n (nf(z))^{\alpha} |x_n - y_n| = +\infty$. Then

\[
\lim_{n \to \infty} \frac{\mathbb{E}L_n(x_n, y_n)}{(nf(z))^{(1-p)/d} |x_n - y_n|} = C(d, p).
\]

From the previous three results, we obtain the following preliminary local convergence result.

**Proposition 6.** Let $z \in \mathbb{R}^d$, $R_2 > R_1 > 0$, and $\alpha = (d + 2p)^{-1}$. Assume that the pdf $f$ is uniform and supported in $B(z; R_2)$. Let $x_n, y_n$ be sequences in $B(z; R_1)$ satisfying $\lim \inf_n (nf(z))^{\alpha} |x_n - y_n| = +\infty$. Fix $\varepsilon > 0$ and let $E_n$ denote the event that

\[
\frac{L_n(x_n, y_n)}{(nf(z))^{(1-p)/d} |x_n - y_n|} - C(d, p) \leq \varepsilon,
\]

then

\[
\limsup_{n \to \infty} \frac{1}{(nf(z))^{\alpha} |x_n - y_n|} \log(1 - P(E_n)) < 0.
\]

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.

**Proof of Proposition 6.** Our proof is structured similarly to that of Yukich (2000, Theorem 4.1) and Talagrand (1995, Section 7.1). Let
• $F_n$ be the event that all the shortest path link distances are at most $(nf(z))^{(\alpha-1)/d}$ (See Lemma 3 for $F_n$),

• $G_n$ be the event that $\#L_n(x_n, y_n) \leq C_*(nf(z))^{1/d}|x_n - y_n|$ for the constant $C_*$ 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 $B(u; (nf(z))^{(\alpha-1)/d})$.

All these events occur with high probability. Both $1 - P(F_n)$ and $1 - P(G_n)$ are exponentially small in $(nf(z))^\alpha|x_n - y_n|$ by Lemma 3 and Lemma 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.

We use shorthand notation $L_n$ for $L_n(x_n, y_n)$. For every $a > 0$, define $W_n(a)$ to be the event that $L_n \geq a$. 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. Let $\pi^*(\omega)$ be the shortest path $L_n(\omega)$ from $x_n$ to $y_n$ through the realization $X_n(\omega) = \{X_1(\omega), \ldots, X_n(\omega)\}$. If $\pi^*(\omega)$ is the sequence

$$x_n = \pi_0(\omega) \rightarrow \pi_1(\omega) \rightarrow \cdots \rightarrow \pi_k(\omega) = y_n,$$

where $k = \#L_n(\omega)$, then we may build a path $\pi(\eta)$ from $x_n$ to $y_n$ through another realization $X_1(\eta), \ldots, X_n(\eta)$ as follows. For each $i \in \{1, \ldots, k-1\}$, let $j$ denote the index where $X_j(\omega) = \pi_i(\omega)$. If $X_j(\omega) = X_j(\eta)$, then set $\pi_i(\eta) = \pi_i(\omega)$. Otherwise, since $\eta$ was assumed to be in $H_n$, there exists some $l$ such that $X_1(\eta) \in B(\pi_l(\omega); (nf(z))^{(\alpha-1)/d})$. Set $\pi_i(\eta) = X_i(\eta)$. Then it follows that $|\pi_i(\eta) - \pi_i(\omega)| \leq (nf(z))^{(\alpha-1)/d}$ for all $i = 1, \ldots, k$.

Let $I$ be the set of indices $i$ where $\pi_i(\omega) \neq \pi_i(\eta)$. If $L(\pi)$ denotes the power-weighted length of the path $\pi$, then $L(\pi(\eta)) \leq L(\pi^*(\omega)) + 2|I|3^p(nf(z))^{(\alpha-1)p/d}$ since $\omega \in F_n$. On the other hand, $\eta \in W_n(a)$ and hence

$$L(\pi^*(\omega)) = L_n(\omega) \geq a - 2|I|3^p(nf(z))^{(\alpha-1)p/d}.$$

Let $d_c(\omega; H_n \cap W_n(a))$ be the convex distance of $\omega$ to $H_n \cap W_n(a)$ (See Talagrand 1995, Section 4.1). By Lemma 4.1.2 of Talagrand (1995), there exists $\eta \in H_n \cap W_n(a)$ such that $|I| \leq d_c(\omega; H_n \cap W_n(a)) \sqrt{\#L_n(\omega)}$, and hence

$$L_n(\omega) \geq a - 2 \cdot 3^p \cdot d_c(\omega; H_n \cap W_n(a)) \sqrt{\#L_n(\omega)} (nf(z))^{(\alpha-1)p/d}.$$

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

$$d_c(\omega; H_n \cap W_n(a)) \geq \frac{u(nf(z))^{(1-\alpha)p/d}}{2 \cdot 3^p \cdot \sqrt{\#L_n(\omega)}} \geq \frac{u(nf(z))^{(1-\alpha)p/d}}{2 \cdot 3^p \cdot \sqrt{C_1(nf(z))^{1/d}|x_n - y_n|}},$$

(8)
since $\omega \in G_n$. 

Let $M_n$ be the median of $L_n$. Set $a = M_n$ in (8) and apply Theorem 4.1.1 of Talagrand (1995) to obtain, for $u > 0$,

$$
P\{L_n \leq M_n - u\} \leq 3 \exp\left(-\frac{C_2 u^2}{|x_n - y_n|} (nf(z))^{2(1-\alpha)-1}\right) + (1 - P(F_n)) + (1 - P(G_n)),
$$

where $C_2 = (2^4 3^2 C_1)^{-1}$, since $P(H_n)$ approaches one as $n \to \infty$. For an upper bound, set $a = M_n + u$ in (8). Then, similarly,

$$
P\{L_n \geq M_n + u\} \leq 3 \exp\left(-\frac{C_2 u^2}{|x_n - y_n|} (nf(z))^{2(1-\alpha)-1}\right) + (1 - P(H_n)),
$$

for sufficiently large $n$ since both $P(F_n)$ and $P(G_n)$ converge to one as $n \to \infty$. Therefore

$$
\text{(9)} \quad P\left\{\frac{|L_n - M_n|}{(nf(z))^{(1-p)/d}|x_n - y_n|} > u\right\} \leq 6 \exp\left(-C_2|x_n - y_n|(nf(z))^\alpha u^2\right) + h_n,
$$

where $h_n = (1 - P(F_n)) + (1 - P(G_n)) + (1 - P(H_n))$.

Now we show that the median $M_n$ and the mean $\mathbb{E}L_n$ are close. By Jensen’s inequality, $|\mathbb{E}L_n - M_n| \leq \mathbb{E}|L_n - M_n| = \int_0^\infty P\{|L_n - M_n| > u\} du$. Note that $P\{|L_n - M_n| > u\} = 0$ when $u \geq |x_n - y_n|^p$ since $L_n = L_n(x_n, y_n)$ is bounded above by $|x_n - y_n|^p$. Integrate the first term on the right side of (9) for $u \geq 0$, and integrate the second term from $u = 0$ to $(nf(z))^{(p-1)/d}|x_n - y_n|^{p-1}$, we obtain an upper bound

$$
\frac{|\mathbb{E}L_n - M_n|}{(nf(z))^{(1-p)/d}|x_n - y_n|} \leq 6 \sqrt{\frac{\pi}{C_2|x_n - y_n|(nf(z))^\alpha}} + \left((nf(z))^{1/d}|x_n - y_n|\right)^{p-1} h_n.
$$

The probability of the events $F_n$, $G_n$, and $H_n$ approach one exponentially fast in $(nf(z))^\alpha|x_n - y_n|$, therefore

$$
\lim_{n \to \infty} \frac{|\mathbb{E}L_n - M_n|}{(nf(z))^{(p-1)/d}|x_n - y_n|} = 0.
$$

By Proposition 5, for sufficiently large $n$,

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

Thus the claim follows from (9). \qed

Next we show that Proposition 6 can be extended to non-uniform probability density $f$. 

8
**Theorem 7.** Let $z \in \mathbb{R}^d$ and $R > 0$. Assume that the pdf $f$ is uniform in $B(z; R)$ but may have probability mass outside of $B(z; R)$.

Let $0 < b < 1$ and $c > 0$ be constants and let $\alpha = (d + 2p)^{-1}$. Fix $\varepsilon > 0$ and denote by $E_n$ the event that

\[
\sup_{x, y} \left| \frac{L_n(x, y)}{(nf(z))^{(1-p)/d}} - C(d, p) \right| \leq \varepsilon,
\]

where the supremum is taken over $x, y \in B(z; R/4)$ and $(nf(z))^{b\alpha}|x - y| \geq c$. Then

\[
\limsup_{n \to \infty} \frac{1}{nf(z)^{b\alpha}} \log \left( 1 - P(E_n \cap F_n) \right) < 0,
\]

where $F_n$ is the event defined in Lemma 3.

The condition $(nf(z))^{b\alpha}|x - y| \geq c > 0$ is introduced to guarantee that $n^\alpha \inf|x - y|$ has polynomial order and to prevent sub-polynomial, e.g., logarithmic, growth.

**Proof.** Let $L(x, r; \mathcal{X}_n)$, $r > 0$, denote the minimal power-weighted path length over all the shortest paths from $x$ to all boundary points of $B(x; r)$, i.e., $L(x, r; \mathcal{X}_n) = \inf L_n(x, u)$ over all $u \in \mathbb{R}^d$ and $|x - u| = r$. For $x, r$ satisfying $B(x; r) \subset B(z; R)$, we claim that

\[
\limsup_{n \to \infty} \frac{1}{r(nf(z))^{\alpha}} \log \left\{ \left| \frac{L(x, r; \mathcal{X}_n)}{r(nf(z))^{(1-p)/d}} - C(d, p) \right| > \varepsilon \right\} < 0.
\]

To establish (12), first note that the boundary of $B(x; r)$ may be covered with open balls of radii $(nf(z))^{(\alpha-1)/d}$, and the number of cover elements may be chosen less than $(2nf(z)r)^d$. Let $v_1, v_2, \ldots, v_m$ be the centers of the cover elements. If the event $F_n$ in Lemma 3 occurs and

\[
\left| \frac{L_n(x, v_k)}{r(nf(z))^{(1-p)/d}} - C(d, p) \right| \leq \varepsilon,
\]

for all $k = 1, \ldots, m$, then $(nf(z))^{(p-1)/d}r L_n(x, u) - C(d, p) \leq \varepsilon$ for all $u$ on the boundary of $B(x; r)$ for sufficiently large $n$. Note that $L(x, r; \mathcal{X}_n) = L(x, r; \mathcal{X}_n \cap B(x; r))$, since the boundary of $B(x; r)$ disconnects the interior and the exterior. If the shortest path to the boundary were to reach any point outside $B(x; r)$, the path must have passed through the boundary, which is a contradiction. An application of Proposition 6 to $L(x, v_k; \mathcal{X}_n \cap B(x; r))$ and an application of the Chernoff bound (Billingsley 1995, Theorem 9.3) proves the inequality (12).

Let $x_n, y_n \in B(z; R/4)$ be two points satisfying $(nf(z))^{b\alpha}|x_n - y_n| \geq c$. Let $H_n$ denote the events that

\[
\frac{L_n(x_n, y_n)}{(nf(z))^{(1-p)/d}|x_n - y_n|} \leq C(d, p) + \varepsilon,
\]
and let $K_n$ denote the event that
\[
\frac{L_n(x_n, y_n)}{(n f(z))^{(1-p)/d} |x_n - y_n|} \geq C(d, p) - \varepsilon.
\]

We claim that
\[
\limsup_{n \to \infty} \frac{1}{(n f(z))^{(1-b)\alpha}} \log(1 - P(H_n \cap K_n)) < 0.\tag{13}
\]

We break the proof of claim (13) into two parts dealing with $H_n$ and $K_n$ separately. The part for $H_n$ is simple. The inequality $L_n(x_n, y_n) = L(x_n, y_n; X_n) \leq L(x_n, y_n; X_n \cap B(z; R))$ and an application of Proposition 6 with the Chernoff bound show that $1 - P(H_n)$ has an exponential decay in $(n f(z))^\alpha |x_n - y_n| \geq c(n f(z))^{(1-b)\alpha}$.

For the $K_n$ part, let $A_n$ denote the event that $L(x_n, y_n; X_n) = L(x_n, y_n; X_n \cap B(z; R))$. Suppose that $A_n$ did not occur, i.e., $L(x_n, y_n; X_n)$ contains a node outside $B(z; R)$ hence outside $B(x_n; 5R/8)$. The reader may find Figure 1 helpful. It follows that $L(x_n, 5R/8; X_n) \leq L(x, y; X_n)$, and
\[
1 - P(K_n) = P(K_n^c \cap A_n) + P(K_n^c \cap A_n^c)
\leq P \left\{ \frac{L(x_n, y_n; X_n \cap B(z; R))}{(n f(z))^{(1-p)/d} |x_n - y_n|} < C(d, p) - \varepsilon \right\} + P \left\{ \frac{L(x_n, 5R/8; X_n)}{(n f(z))^{(1-p)/d}(5R/8)} < C(d, p) - \varepsilon \right\}.
\]

Apply Proposition 6 with the Chernoff bound to the first term on the right side of the above inequality, and apply (12) to the second term. Therefore $1 - P(K_n)$ decays exponentially in $(n f(z))^{\alpha} |x_n - y_n| \geq c(n f(z))^{(1-b)\alpha}$. We have shown that claim (13) holds.

Now we are ready to prove the theorem. Suppose that the event $F_n$ occurs, and let $\{B(w_i; (n f(z))^{(\alpha-1)/d}) : w_i \in B(z; R/4), 1 \leq i \leq m\}$ be an open cover of $B(z; R/4)$ with $m \leq (n f(z)R)^d$. Then for every $x, y \in B(z; R/4)$ there exists $w_i, w_j$ such that $|x - w_i| < (n f(z))^{(\alpha-1)/d}$ and $|y - w_j| < (n f(z))^{(\alpha-1)/d}$, hence
\[
|x - y| - |w_i - w_j| < 2(n f(z))^{(\alpha-1)/d}
\]
and
\[
|L_n(x, y) - L_n(w_i, w_j)| \leq 2^{p+1}(n f(z))^{(\alpha-1)p/d}.
\]

Since $|x - y| \geq c(n f(z))^{-\alpha b}$ and $F_n$ is assumed to occur, an algebraic manipulation shows that if $n$ is sufficiently large,
\[
\frac{L_n(x, y)}{|x - y|(n f(z))^{(1-p)/d}} - \frac{L_n(w_i, w_j)}{|w_i - w_j|(n f(z))^{(1-p)/d}} < \varepsilon/2.
\]

10
Figure 1: When $x_n \in B(z; R/4)$, then $B(z; R/4) \subset B(x_n; 5R/8) \subset B(z; R)$. When $y_n \in B(z; R/4)$, $L_n(x_n, y_n)$ is conditionally independent of the outside $B(z; R)$ on the event $L_n(x_n, y_n) < L_n(x_n; 5R/8)$, due to the annulus buffer region $\{u: R/4 < |z - u| < R\}$.

hence for $n$ sufficiently large,

$$\mathbb{P}\{\sup_{x,y} \left| \frac{L_n(x, y)}{(nf(z))^{(1-p)/d}} - C(d, p) \right| > \varepsilon \} \leq (1 - \mathbb{P}(F_n)) + \sum_{i,j} \left\{ \left| \frac{L_n(w_i, w_j)}{(nf(z))^{(1-p)/d}} - C(d, p) \right| > \frac{\varepsilon}{2} \right\},$$

where the sum is over $i$ and $j$ that $|w_i - w_j| \geq c(nf(z))^{-ab} - 2(nf(z))^{(a-1)/d}$. The theorem now follows from (13) and Lemma 3 as $m \leq (nf(z)R)^d$ is of polynomial order of $nf(z)$.

**Corollary 8.** Let $f_m, f_M$ be constants such that $0 < f_m \leq f(u) \leq f_M < \infty$ for all $u \in B(z; R)$. Let $0 < b < 1$ and $c > 0$ be constants and let $\alpha = (d + 2p)^{-1}$. Fix $\varepsilon > 0$. Let $H_n$ denote the event that

$$\sup_{x,y} \frac{L_n(x, y)}{(nf_m)^{(1-p)/d}|x - y|} \leq C(d, p) + \varepsilon,$$

where the supremum is taken over $x, y \in B(z; R/4)$ and $(nf_m)^{\alpha b}|x - y| \geq c$. Then

$$\lim_{n \to \infty} \sup_{x} \frac{1}{(nf_m)^{(1-b)\alpha}} \log(1 - \mathbb{P}(H_n \cap F_n)) < 0,$$

where $F_n$ is the event defined in Lemma 3.
Similarly, let $K_n$ denote the event that

$$\inf_{x,y} \frac{L(x, y; X_n)}{(n f_M)^{(1-p)/d}} |x - y| \geq C(d, p) - \varepsilon,$$

where the infimum is taken over $x, y \in B(z; R/4)$ and $(n f_M)^{ba}|x - y| \geq c$. Assume that $\int_{B(z; R)} f_M \, du \leq 1$. Then

$$\limsup_{n \to \infty} \frac{1}{(n f_M)^{(1-b)\alpha}} \log(1 - P(K_n \cap F_n)) < 0. \tag{16}$$

Proof. Consider the claim (15) involving $H_n$ first. For each point $X_i \in X_n$, if $X_i$ is in $B(z; R)$ then discard $X_i$ with probability $1 - f_m/f(X_i)$. Let $\tilde{X}_n$ denote the filtered collection from $X_n$. Then, obviously, $L_n(x, y) = L(x, y; X_n) \leq L(x, y; \tilde{X}_n)$ for all $x, y \in \mathbb{R}^d$, and if $L_n(x, y) \leq (C(d, p) + \varepsilon) (n f_m)^{(1-p)/d} |x - y|$ holds, then (14) holds. Note that $\tilde{X}_n$ is an i.i.d. sample with uniform density $f_m$ when restricted to $B(z; R)$. Apply Theorem 7 with the Chernoff bound, and we obtain (15).

For the claim (16) involving $K_n$ repeat a similar argument to prove (16). \qed

### 3.2 Convergence in Riemannian manifolds

We adapt Corollary 8 to the case when the probability distribution is supported on a Riemannian manifold $M$ instead of on a Euclidean space.

**Lemma 9.** Let $(M, g_1)$ be a Riemannian manifold equipped with metric $g_1$. Let $b \in (0, 1)$, $c > 0$, $\varepsilon > 0$ be fixed constants. Let $\alpha = (d + 2p)^{-1}$. Assume that $f(z) > 0$ and that $f$ is smooth at $z$, i.e., $f$ is infinitely differentiable at $z$. We denote by $E_n = E_n(z, R)$, for $z \in M$ and $R > 0$, the event that

(i) every shortest path link in $B(z; R)$ has length less than $(n f(z))^{(\alpha - 1)/d}$, and

(ii) the supremum

$$\sup_{x,y} \frac{L_n(x, y; X_n)}{n^{(1-p)/d} \text{dist}_p(x, y)} - C(d, p) < \varepsilon,$$

for $x, y \in M$ such that $\text{dist}_1(x, z) < R$, $\text{dist}_1(y, z) < R$, and $\text{dist}_1(x, y) \geq c(n f(z))^{-ab}$.

For every $z \in M$, there exists $R > 0$ such that

$$\limsup_{n \to \infty} (n f(z))^{\alpha(b-1)} \log(1 - P(E_n)) < 0.$$

Proof. The event condition (i) may be easily proved with the same arguments from Lemma 3. We concentrate on condition (ii).
Define \( U = B(z; 4R) = \{ u \in M : \text{dist}_1(u, z) < 4R \} \) for \( R > 0 \). For every \( \delta > 0 \) we may choose \( R > 0 \) small enough so that there exists a normal chart map \( \varphi : U \subset M \rightarrow V \subset \mathbb{R}^d \),

\[
(1 - \delta)^d \sup_U f \leq f(z) \leq (1 + \delta)^d \inf_U f,
\]

and

\[
1 - \delta \leq \frac{\text{dist}_1(u, v)}{|\varphi(u) - \varphi(v)|} \leq 1 + \delta,
\]

for all \( u \neq v \in U \). Note that the denominator in (19) is a Euclidean distance.

Let \( x,y \in B(z; R) \subset U \). Let \( L(\varphi(x), \varphi(y); \varphi(\mathcal{X}_n)) \) denote the shortest path length between \( \varphi(x), \varphi(y) \in V \) in Euclidean space \( \mathbb{R}^d \). Suppose that

\[
C(d, p) - \frac{\varepsilon}{2} \leq \frac{L_n(\varphi(x), \varphi(y); \varphi(\mathcal{X}_n))}{(nf(z))^{(1-p)/d}|\varphi(x) - \varphi(y)|} \leq C(d, p) + \frac{\varepsilon}{2},
\]

holds. Then by the assumptions (18) and (19),

\[
(C(d, p) - \varepsilon)(\inf_U f)^{(1-p)/d} \leq \frac{L_n(x, y)}{n^{(1-p)/d}\text{dist}_1(x, y)} \leq (C(d, p) + \varepsilon)(\sup_U f)^{(1-p)/d},
\]

when \( \delta \) is sufficiently small. We next show that (17) follows from (21).

Since \( x,y \in B(z; R) \), the minimal \( g_1 \)-geodesic curve from \( x \) to \( y \) is contained in \( U \) by assumption (19). It follows from the definition of Riemannian distance that

\[
\text{dist}_p(x, y) \leq \text{dist}_1(x, y)(\inf_U f)^{(1-p)/d}.
\]

Furthermore, if the minimal \( g_p \)-geodesic curve from \( x \) to \( y \) were contained in \( U \), then

\[
\text{dist}_p(x, y) \geq \text{dist}_1(x, y)(\sup_U f)^{(1-p)/d}.
\]

If a (piece-wise) smooth curve from \( x \) exits outside \( U \), then the curve length under \( g_p \) must be at least \((3R)(\sup_U f)^{(1-p)/d}\) by the assumptions \( \text{dist}_1(x, z) < R \) and \( U = B(z; 4R) \). Therefore it follows from (18) that (23) holds since

\[
\text{dist}_p(x, y) \leq \text{dist}_1(x, y)(\inf_U f)^{(1-p)/d} \leq (2R)(\sup_U f)^{(1-p)/d} \left(\frac{1 - \delta}{1 + \delta}\right)^{(1-p)/d} < (3R)(\sup_U f)^{(1-p)/d},
\]

if \( \delta \) were sufficiently small, and the minimal \( g_p \)-geodesic must be contained in \( U \).

After we combine (21), (22), and (23),

\[
(C(d, p) - \varepsilon) \text{dist}_p(x, y) \leq n^{(p-1)/d}L_n(x, y)
\]

\[
\leq (C(d, p) + \varepsilon) \text{dist}_p(x, y),
\]

holds provided that (20) is true. Apply Corollary 8 to (20), and the lemma is proved. \( \square \)
Our main result Theorem 1 can be now obtained by applying Lemma 9 to a finite open cover of the manifold.

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 9 holds in open cover elements, then the local convergences may be assembled together to yield global convergence of the curve length.

For each \( w_i \in M \), we may associate positive \( R_i > 0 \) such that Lemma 9 holds within the region \( V_i = \{ v \in M : \text{dist}_1(v, w_i) < 3R_i \} \). Let \( U_i = \{ u \in M : \text{dist}_1(u, w_i) < R_i \} \). Since \( M \) is assumed to be compact, we may choose finite \( m > 0 \), \( \{ w_i \in M \}_{i=1}^m \), and corresponding \( \{ R_i > 0 \}_{i=1}^m \) such that corresponding \( \{ U_i \} \) is a finite open cover of \( M \).

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 \( z_2 \notin U_1 \). Then \( \text{dist}_1(z_1, z_2) \geq R_1 \). Note that \( z_2 \in M \) need not be in \( X_n \). 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 in, say, \( V_k \). Set \( z_{k+1} = y \). Then points \( x = z_1, z_2, \ldots, z_k, z_{k+1} = y \) satisfy the conditions \( z_i, z_{i+1} \in V_i \) for \( i = 1, 2, \ldots, k \), and \( \text{dist}_1(z_i, z_{i+1}) \geq R_i \geq R \) for \( i = 1, 2, \ldots, k - 1 \), where \( R = \min_i R_i \). The last edge length \( \text{dist}_1(z_k, z_{k+1}) \) may 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 it is easily checked that \( z_k \) is still in \( V_k \). See Figure 2 for illustration.
Suppose that
\begin{equation}
(C(d, p) - \varepsilon) \text{dist}_p(z_i, z_{i+1}) \leq n^{(p-1)/d} L_n(z_i, z_{i+1}),
\end{equation}
holds for all \(i = 1, 2, \ldots, k\). Then by the triangle inequality and the property \(\nu^p + \omega^p \leq (\nu + \omega)^p\) for \(\nu, \omega \geq 0\),
\[(C(d, p) - \varepsilon) \text{dist}_p(x, y) \leq (C(d, p) - \varepsilon) \sum_{i=1}^{k} \text{dist}_p(z_i, z_{i+1}) \leq \sum_{i=1}^{k} n^{(p-1)/d} L_n(z_i, z_{i+1}) \leq n^{(p-1)/d} L_n(x, y).\]
Since \(m\) is finite and Lemma 9 applies in \(V_1, V_2, \ldots, V_m\),
\[
\limsup_{n \to \infty} \left( n \inf_{f} (1/(d+2p)) \mathbb{P} \left\{ \sup_{x, y} \frac{L_n(x, y)}{d, p \text{dist}_p(x, y)} \leq C(d, p) - \varepsilon \right\} < 0.\]
For the upper tail, we follow a similar strategy to Bernstein et al. (2000). If \(z_1 = x, z_{k+1} = y\), and \(z_i\) are points on the minimal geodesic curve from \(x\) to \(y\) under \(g_p\), 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 \(\text{dist}(z_i, z_{i+1}) \geq R\) for all \(i = 1, 2, \ldots, k\). Since Lemma 9 applies in \(V_1, \ldots, V_m\), every edge length of the shortest path from \(z_i\) to \(z_{i+1}\) is at most \((n \sup f)^{(\alpha-1)/d}\) for \(i = 1, \ldots, k\), where \(\alpha = (d + 2p)^{-1}\) from Lemma 9. Therefore each paste procedure may incur additional cost of at most \(2p(n \sup f)^{\alpha-1/d}\) so that
\begin{equation}
L_n(x, y) \leq \sum_{i=1}^{k} L_n(z_i, z_{i+1}) + k2^p(n \sup f)^{(\alpha-1)/d},
\end{equation}
where \(\alpha = (d + 2p)^{-1}\). Therefore if Lemma 9 holds in \(V_1, V_2, \ldots, V_m\), then
\[
n^{(p-1)/d} L_n(x, y) \leq \left( C(d, p) + \frac{\varepsilon}{2} \right) \text{dist}_p(x, y) + k2^p n^{(\alpha p-1)/d} (\sup f)^{(\alpha-1)/d},
\]
and if \(n\) is large enough, \(n^{(p-1)/d} L_n(x, y) \leq (C(d, p) + \varepsilon) \text{dist}_p(x, y)\) since \(n^{(\alpha p-1)/d} n^{\alpha b}\) shrinks to zero as \(n \to \infty\). Therefore Theorem 1 is established by applications of Lemma 9 to \(V_1, V_2, \ldots, V_m\).

\section{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.
As we mentioned, the proofs require some results in Poisson processes. 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 uniform i.i.d. over $B$. As in the i.i.d. case, we use a shorthand notation $\mathcal{L}_\lambda(x, y) = L(x, y; \mathcal{H}_\lambda)$ for $x, y \in \mathbb{R}^d$.

Let $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^d$ be a fixed 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
\begin{equation}
T(u, v; b) = \bigcup_{0 \leq s \leq 1} B(su + (1 - s)v; b).
\end{equation}
Note that $\bigcup_{b > 0} T(u, v; b) = \mathbb{R}^d$. For convenience, define
\begin{equation}
\mathcal{L}_\lambda(u, v; b) = L(u, v; \mathcal{H}_\lambda \cap T(u, v; b)).
\end{equation}

4.1 Percolation lemma

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

**Lemma 10.** 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 strategy of this proof is similar to that of Meester and Roy (1996, Theorem 6.1). We first define a Galton-Watson process $X_n$. Let $X_0 = \{x_0 = 0 \in \mathbb{R}^d\}$ be the ancestor of the family, and associate the parameter $r_0 > 0$. Then define the offsprings $X_1(r_0)$ to be $\mathcal{H}_\lambda \cap B(x_0, r_0^{1/p})$. $X_1(r_0)$ is the set of points in $\mathcal{H}_\lambda$ that may be reached from $x_0$ with exactly single edge with path length at most $r_0$ in power-weighted sense. Note that $\mathbb{E}|X_1(r_0)| = \lambda V_d r_0^{d/p}$ where $|X_1(r_0)|$ denotes the cardinality of $X_1(r_0)$, and $V_d$ denotes the volume of $B(0; 1)$.

For each offspring $x_{1,k} \in X_1(r_0)$, 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 $X_2(r_0)$ to be the collection of all the second generation offsprings, and define recursively the $n$-th generation offsprings $X_n(r_0)$. Then $X_n(r_0)$ 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 3. We prove by induction that
\begin{equation}
\mathbb{E}|X_n(r_0)| \leq \left(\lambda V_d r_0^{d/p}\right)^n \frac{\Gamma(1 + d/p)^n}{\Gamma(1 + nd/p)}.
\end{equation}
Figure 3: A run through the family tree generated by $X_n$ 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 $r_1^{1/p}/r_0 = 3$. Here $x_{1,1} \in X_1$ is among the first generations since it is within $B(x_0; r_0^{1/p})$, and $x_{2,1} \in X_2$ is among the second generations since it is within the balls centered at the first generation offsprings, e.g., $x_{1,1}$. 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.

We mentioned above that $E|X_1(r_0)| = \lambda V_d r_0^{d/p}$, and (28) is true for $n = 1$. For general $n$, apply the Campbell-Mecke formula, (Baddeley 2007, Theorem 3.2, p.48)

$$E|X_n(r_0)| \leq \lambda \int_{B(x_0; r_0^{1/p})} E|X_{n-1}(r_0 - |x - x_0|^p)| dx$$

$$\leq \lambda^n V_d^{n-1} \frac{\Gamma(1 + d/p)^{n-1}}{\Gamma(1 + (n-1)d/p)} \int_{B(x_0; r_0^{1/p})} (r_0 - |x - x_0|^p)^{(n-1)d/p} dx$$

$$= (\lambda V_d r_0^{d/p})^n \frac{\Gamma(1 + d/p)^{n}}{\Gamma(1 + nd/p)}.$$ 

We have established (28).

Using the Markov inequality and Stirling’s approximation,

$$\log P\{X_n(r_0) \neq \emptyset\} \leq n \log \left( V_d \Gamma\left(1 + \frac{d}{p}\right) \left( \frac{c_0}{c_1} \cdot \frac{pc}{a} \right)^{d/p} \right) + o(n)$$

as $n \to \infty$. Note that if a path starting at $x$ passes through more than $n = c_1 \lambda^{1/d}$ nodes and has path length less than $r_0 = c_0 \lambda^{(1-p)/d}$, then the $n$-th generation set
Lemma 10 follows since, if the ratio $c_1/c_0$ is sufficiently large, the logarithm term above is negative.

4.2 Mean convergence for Poisson point processes

Lemma 11. Consider the shortest path length $L_1(0, te_1)$ from $0 \in \mathbb{R}^d$ to $te_1 \in \mathbb{R}^d$ in $\mathcal{H}_1$ for $t > 0$.

\begin{equation}
\lim_{t \to \infty} \frac{1}{t} E L_1(0, te_1) = C(d, p).
\end{equation}

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

\begin{equation}
\lim_{t \to \infty} \frac{1}{t} E L_1(0, te_1; b_t) = C(d, p).
\end{equation}

Recall that $L_1(0, te_1; b_t)$ denotes $L(0, te_1; \mathcal{H}_1 \cap T(0, te_1; b_t))$ from (26) and (27). When $b = +\infty$, (29) is a consequence of, e.g., 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 the event $L_1(0, te_1) \neq L_1(0, te_1; b_t)$ occurs is exponentially small of order at least $t^{3p/4}$ for some $\varepsilon > 0$ when $b_t \geq t^{3/4+\varepsilon}$. Lemma 11 is weaker in the sense that it only asserts closeness in the mean. On the other hand, Lemma 11 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 be even sub-polynomial.

Proof of Lemma 11. 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 $\Theta(x,y)$ in the proof of Lemma 3. Let

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

and let

$$\xi_t(\lambda, b) = \sup \left\{|u - te_1| : u \in T(b), \mathcal{H}_\lambda \cap T(b) \cap \Theta(u, te_1) = \emptyset\right\}.$$ 

In other words, $\xi_t(\lambda, b)$ denotes the maximum 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 $h$ which is used to define $\Theta(x,y)$ in the proof of Lemma 3, 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 $t \in \mathbb{R}$,

\begin{equation}
E \xi_t(\lambda, b)^p \leq \frac{k\Gamma(1 + p/d)}{(\lambda A)^{p/d}} + \frac{m2p\Gamma(1 + p)}{\lambda^p(\delta b)^{p(d-1)}}.
\end{equation}

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 $L_1(0, se_1; b)$ between $se_1$ and $(s + t)e_1$, and let $\gamma_-$ denote the node that directly connects to $se_1$. Similarly consider the shortest path for $L_1(se_1, (s + t)e_1; b)$ 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 $\gamma_0$ 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 4. This new path has length that is an upper bound of $L_1(0, (s + t)e_1; b)$.

\[
L_1(0, (s + t)e_1; b) \leq L_1(0, se_1; b) + L_1(se_1; (s + t)e_1; b) + (2^{p-1} - 1)(|\gamma_0 - \gamma_-| + |\gamma_+ - \gamma_0|)^p,
\]

by the convex property of the power function for $p \geq 1$. Note that both $|\gamma_0 - \gamma_-|$ and $|\gamma_+ - \gamma_0|$ are bounded above by $\xi_s(1, b)$. Therefore $\mathbb{E}L_1(0, (s + t)e_1; b)$ is bounded above by

\[
\mathbb{E}L_1(0, se_1; b) + \mathbb{E}L_1(se_1; (s + t)e_1; b) + \mathbb{E}(2\xi_s(1, b))^p
= \mathbb{E}L_1(0, se_1; b) + \mathbb{E}L_1(0; te_1; b) + \mathbb{E}(2\xi_0(1, b))^p.
\]

The equality holds by the translation invariant property of the distribution of $H_1$. Therefore $\mathbb{E}L_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p$ is a sub-additive function of $t$. Note that $L_1(0, te_1; b) \leq t^p$, and $t \mapsto t^p$ is Lipschitz in compact intervals. Therefore $\mathbb{E}L_1(0, te_1; b)$ is continuous for $t \geq 0$. A standard proof of the Fekete’s lemma (for example, see Steele 1997, Lemma 1.2.1) may be easily adapted to continuous sub-additive functions, and the limit of $t^{-1}\mathbb{E}L_1(0, te_1; b)$ exists. The limiting value is given by

\[
\lim_{t \to \infty} \frac{1}{t} \mathbb{E}L_1(0, te_1; b) = \inf_{t > 0} \frac{\mathbb{E}L_1(0, te_1; b) + \mathbb{E}(2\xi_0(1, b))^p}{t},
\]

and we denote the limit by $C(d, p; b)$. In the special case $b = \infty$, $C(d, p; \infty) = C(d, p)$ establishing (29).
We now show that \( C(d, p; b) \) converges to \( C(d, p; \infty) = C(d, p) \) when \( b \to \infty \). Choose an arbitrary \( \varepsilon > 0 \). By (31) and by the fact that \( C(d, p) \) is the limit of \( t^{-1}L_1(0, te_1) \), there exists \( T > 0 \) such that
\[
\frac{1}{T}E_{L_1(0, Te_1)} < C(d, p) + \frac{\varepsilon}{3},
\]
and
\[
\frac{1}{T}E(2\xi_0(1, b))^p < \frac{\varepsilon}{3},
\]
for all \( b > 1 \). For this fixed \( T \), note that \( \lim_{b \to \infty} L_1(0, Te_1; b) = L_1(0, Te_1) \) monotonically from above almost surely, and by the monotone convergence theorem, there exists \( B > 1 \) such that for all \( b > B \) and fixed \( T \),
\[
\frac{1}{T}E_{L_1(0, Te_1; b)} \leq \frac{1}{T}E_{L_1(0, Te_1)} + \frac{\varepsilon}{3}.
\]
Combine the three inequalities above with (32),
\[
C(d, p; b) \leq \frac{1}{T} (E_{L_1(0, Te_1; b)} + E(2\xi_0(1, b))^p) \leq C(d, p) + \varepsilon,
\]
for all \( b > B \). Therefore \( C(d, p; b) \) converges to \( C(d, p) \) as \( b \to \infty \).

Finally, suppose \( b = b_t \) is a function of \( t \) rather than a constant. If \( \lim \inf_t b_t = \infty \) then
\[
C(d, p) \leq \lim_{t \to \infty} \frac{1}{t}E_{L_1(0, te_1; b_t)} \leq \frac{1}{t}E_{L_1(0, te_1; B)} = C(d, p; B),
\]
for every fixed \( B > 0 \). (30) follows as \( B \to \infty \) on the right side.

### 4.3 Shortest path size

In order to prove Lemma 4, we need an exponential probability bound on an upper tail of \( L \).

**Lemma 12.** Let \( z \in \mathbb{R}^d \), \( R_2 > R_1 > 0 \), and \( \alpha = (d+2p-1)^{-1} \). Let \( x_\lambda, y_\lambda \in B(z; R_1) \) be functions of \( \lambda \) satisfying \( \lim \inf_{\lambda \to \infty} \lambda^\alpha |x_\lambda - y_\lambda| = +\infty \). Fix \( \varepsilon > 0 \) and let \( E_\lambda \) denote the event that
\[
L(x_\lambda, y_\lambda; \mathcal{H}_\lambda \cap B(z; R_2)) \leq \frac{\lambda^{1-p}/d |x_\lambda - y_\lambda| C(d, p) + \varepsilon.}
\]
Then
\[
\lim \sup_{\lambda \to \infty} \frac{1}{\lambda^\alpha |x_\lambda - y_\lambda|} \log(1 - P(E_\lambda)) < 0.
\]
Proof. Let \( b_\lambda \) be a positive-valued function of \( \lambda \) such that \( b_\lambda \to 0 \) and \( \lambda^{1/d} b_\lambda \to \infty \) as \( \lambda \to \infty \). Recall the definitions (26) and (27). Since \( b_\lambda \to 0 \), for all sufficiently large \( \lambda \),

\[
\mathcal{H}_\lambda \cap T(x_\lambda, y_\lambda; b_\lambda) \subset \mathcal{H}_\lambda \cap B(z; R_2),
\]

and hence

\[
L(x_\lambda, y_\lambda; \mathcal{H}_\lambda \cap B(z; R_2)) \leq L(x_\lambda, y_\lambda; b_\lambda).
\]

Let \( E'_\lambda \) denote the event that (33) holds with \( \mathcal{L}_\lambda(x_\lambda, y_\lambda; b_\lambda) \) in place of \( L(x_\lambda, y_\lambda; \mathcal{H}_\lambda \cap B(z; R_2)) \). By the inequality above, if \( E'_\lambda \) occurs then \( E_\lambda \) occurs. Therefore it is sufficient to show that \( 1 - P(E'_\lambda) \), which dominates \( 1 - P(E_\lambda) \), is exponentially small. Fix sufficiently large \( \lambda \) so that the inequalities above hold, and set \( b = b_\lambda \).

As in Lemma 11, by the convex property of the power functions, \( \mathcal{L}_\lambda(0, b e_1; b) \) may be bounded above by \( \mathcal{L}_\lambda(0, b e_1; b) + \mathcal{L}_\lambda(b e_1, 2 b e_1; b) + (2^{p-1} - 1)(Z^p_k + Y^p_0) \), where \( Z_k \) and \( Y_k \) are the first and the last edge lengths in \( \mathcal{L}_\lambda(k b e_1, (k + 1) b e_1; b) \), respectively. In Figure 4, when \( s = b \) and \( s + t = 2 b \), \( Z_1 \) and \( Y_0 \) correspond to \( |\gamma_+ - \gamma_0| \) and \( |\gamma_0 - \gamma_-| \), respectively.

Note that the shortest path for \( \mathcal{L}_\lambda(k b e_1, (k + 1) b e_1; b) \) is not likely to be the direct path \( k b e_1 \to (k + 1) b e_1 \). That is, if it were the direct path, then \( \mathcal{H}_\lambda \) is empty in \( \Theta(k b e_1, (k + 1) b e_1) \) where \( \Theta \) was defined in Lemma 3, and it happens with probability at most \( \exp(-\lambda \theta_0 b^d) \), where \( \theta_0 = \theta_0(d, p) > 0 \) was also defined in Lemma 3. If none of the shortest paths for \( \mathcal{L}_\lambda(k b e_1, (k + 1) b e_1; b) \) is a direct path, then the previous paste procedure may be repeated so that

\[
\mathcal{L}_\lambda(0, m b e_1; b) \leq \sum_{k=0}^{m-1} \left( \mathcal{L}_\lambda(k b e_1, (k + 1) b e_1; b) + (2^{p-1} - 1)(Z^p_k + Y^p_0) \right),
\]

with probability at least \( 1 - m \exp(-\lambda \theta_0 b^d) \).

If \( k, l \) are integers and \( l - k \geq 3 \), then \( T(k b e_1, (k + 1) b e_1; b) \) and \( T(l b e_1, (l + 1) b e_1; b) \) are disjoint, hence \( \mathcal{L}_\lambda(k b e_1, (k + 1) b e_1; b) \) and \( \mathcal{L}_\lambda(l b e_1, (l + 1) b e_1; b) \) are mutually independent, and so are \( Z_k \) and \( Z_l \), as well as \( Y_k \) and \( Y_l \). Then the sum in (35) may split into \( K \geq 3 \) sums of independent variables, and each sum has at most \( K^{-1}m \) summands. Note that each summand is almost surely bounded since \( Z^p_k + Y^p_k \leq \mathcal{L}_\lambda(k b e_1, (k + 1) b e_1; b) \leq b^p \). Apply the Azuma’s inequality for \( K = 4 \) separate sequences,

\[
P\left( \frac{\mathcal{L}_\lambda(0, m b e_1; b)}{\lambda^{(1-p)/d} m b} \geq \mu_b + \varepsilon \right) \leq m e^{-\lambda \theta_0 b^d} + 4 \exp\left( -\frac{(m - 3)\varepsilon^2}{21 + 2p(\lambda^{1/d} b)^2(p-1)} \right),
\]

where \( \mu_b \) is the expectation \( \mathbb{E}\mathcal{L}_\lambda(0, b e_1; b) + (2^{p-1} - 1)(\mathcal{E}_Z^p + \mathcal{E}_Y^p) \) divided by \( \lambda^{(1-p)/d} b \).

Set \( m b = |x_\lambda - y_\lambda| \) and \( m = |\lambda^{(1-\alpha)/d}| |x_\lambda - y_\lambda| |. By the definition, both \( \mathbb{E}Z_k^p \) and \( \mathbb{E}Y_{k-1}^p \) are bounded above by \( \mathbb{E}\xi_{kk}(\lambda, b)^p = \mathbb{E}\xi_0(\lambda, b)^p \) defined in (31), and a direct computation shows that \( \lambda^{(p-1)/d} b^{-1} \mathbb{E}\xi_0(\lambda, b)^p \) shrinks to zero when \( \lambda^{1/d} b \geq \lambda^{\alpha/d} \to \infty \).
∞. See Hwang (2012, Lemma 2.5) for more details. Apply Lemma 11 after scaling to see that \( \mu_0 \) converges to \( C(d, p) \) as \( \lambda \to \infty \). Then (36) becomes

\[
\Pr \left\{ \frac{L_\lambda(x_\lambda, y_\lambda; b)}{\lambda^{(1-p)/d}|x-y|} \geq C(d, p) + 2\varepsilon \right\} \leq \lambda^{(1-\alpha)/d}|x_\lambda - y_\lambda| e^{-\theta_0 \lambda^\alpha} + 4 \exp \left( -\frac{\lambda^\alpha|x_\lambda - y_\lambda| \varepsilon^2}{2^{1+2p}} \left( 1 + o \left( \frac{1}{\lambda^\alpha|x_\lambda - y_\lambda|} \right) \right) \right),
\]

as \( \lambda \) and \( \lambda^\alpha|x_\lambda - y_\lambda| \) tends to infinity. The claim follows since \( |x_\lambda - y_\lambda| \) is bounded above by \( 2R_1 \) and the first upper bound term has exponential decay in \( \lambda^{1/d}|x_\lambda - y_\lambda| \) as well.

**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 \( H_n \) denote the event that \( N \geq n \) and \( N' \leq n \). Let \( a = A/(V_d R^d) \) and \( a' = A'/(V_d R^d) \) where \( V_d \) denotes the volume of unit ball \( B(z; 1) \). In other words, \( a \) (or \( a' \)) is \( A \) (or \( A' \)) divided by the volume of \( B(z; R_2) \), respectively. Let \( K_n \) denote the event that

\[
\frac{L(x_n, y_n; H_{na'}) \cap B(z; R_2)}{(na')^{(1-p)/d}|x_n - y_n|} \leq C(d, p) + \frac{\varepsilon}{2}.
\]

We first show that if both \( H_n \) and \( K_n \) occur, then the following conditions are satisfied.

(i) \( L_n(x_n, y_n) \) is a path in \( \mathcal{H}_{na} \).

(ii) \( L_n(x_n, y_n) \leq (C(d, p) + \varepsilon/2)(na')^{(1-p)/d}|x_n - y_n| \).

Note that restriction of \( \mathcal{H}_{na} \) to \( B(z; R_2) \) may be realized as \( \mathcal{X}_N \) since \( na = nA/(V_d R^d) \). Since \( H_n \) is assumed to occur, it follows that \( N \geq n \), and \( \mathcal{X}_N = \mathcal{H}_{na} \cap B(z; R_2) \subset H_{na} \). Therefore \( L_n(x_n, y_n) = L(x_n, y_n; \mathcal{X}_n) \) is a path in \( \mathcal{H}_{na} \), and condition (ii) holds.

For condition (ii), \( H_n \) is assumed to occur, so we have \( N' \leq n \). Then similar to the previous argument, \( \mathcal{H}_{na'} \cap B(z; R_2) = \mathcal{X}_{N'} \subset \mathcal{X}_n \) and it follows that \( L(x_n, y_n; \mathcal{X}_n) \leq L(x_n, y_n; \mathcal{H}_{na'} \cap B(z; R_2)) \). Condition (ii) follows by (38).

\( 1 - \Pr(H_n) \) is exponentially small in \( n \) by the Chernoff bound, and \( 1 - \Pr(K_n) \) is exponentially small in \( (na')^{\alpha}|x_n - y_n| \) by Lemma 12. Note that \( f(z) = 1/(V_d R^d) \) and \( A' \) is a fixed constant in this proof, hence \( 1 - \Pr(K_n) \) is exponentially small in \( (nf(z))^{\alpha}|x_n - y_n| \). Lastly, when \( H_n \) and \( K_n \) occur, we have shown that (i) and (ii) hold, and an application of Lemma 10 shows that \( 1 - \Pr(G_n) \), the probability that \( \#L_n(x_n, y_n) \) is greater than \( C_*(nf(z))^{1/d}|x_n - y_n| \), is exponentially small in \( (nf(z))^{1/d}|x_n - y_n| \) when \( C_* > (C(d, p) + \varepsilon/2)/\rho_0 \). (See Lemma 10 for \( \rho_0 \).) The decay rate of \( 1 - \Pr(G_n) \) is determined by the slower one, i.e., \( 1 - \Pr(K_n) \), and the lemma holds.

\( \square \)
4.4 Mean convergence in i.i.d. cases

Proof of Proposition 5. Fix \(x_n\) and \(y_n\), so that \(L_k\) denotes \(L_k(x_n, y_n)\) (not \(L_k(x, y)\)) for all \(k \geq 0\). Let \(C_{\ast} > 0\) as in Lemma 4 and suppose that the number of nodes \#\(L_n\) in the shortest path \(L_n\) is less than \(C_{\ast}(n f(z))^{-1/d}|x_n - y_n|\). Suppose that the event \(F_n\) from Lemma 3 occurred and all the shortest path edge lengths are at most \((n f(z))^{(\alpha - 1)/d}\). When a sample point from \(X_n\) is discarded, \(L_{n-1}\) remains the same as \(L_n\) if the discarded sample point were not a node in \(L_n\). Furthermore since edge lengths are at most \((n f(z))^{(\alpha - 1)/d}\), \(L_{n-1}\) and \(L_n\) may differ at most by \(2^p(n f(z))^{(\alpha - 1)/d}\). Therefore

\[
\mathbb{E}L_{n-1} - \mathbb{E}L_n \leq C_{\ast}(n f(z))^{1/d} |x_n - y_n|(1 - 2^p(n f(z))^{(\alpha - 1)/d} + h_n \mathbb{E}L_0),
\]

where \(h_n\) denotes the probability that either \#\(L_n\) > \(C_{\ast}(n f(z))^{-1/d}|x_n - y_n|\), or the event \(F_n\) 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

\[
\mathbb{E}L_N = \sum_{k \geq 0} \mathbb{E}L_k \mathbb{P}\{N = k\}.
\]

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

\[
\sum_{k \geq 0} |\mathbb{E}L_n - \mathbb{E}L_k| \mathbb{P}\{N = k\} \\
\leq \mathbb{E}L_0 \mathbb{P}\{N < 2^{-1}n\} + \sum_{k \geq 2^{-1}n} |\mathbb{E}L_n - \mathbb{E}L_k| \mathbb{P}\{N = k\}.
\]

Note that the first term on the right of (39) has monotonic decrease for fixed \(x_n\) and \(y_n\). Therefore if \(k \geq 2^{-1}n\),

\[
|\mathbb{E}L_n - \mathbb{E}L_k| \leq 2^p C_{\ast} |x_n - y_n| n - k \left(\frac{n}{2}\right)^{-1} \left(\frac{n f(z)}{2}\right)^{1+\alpha/2} \mathbb{E}L_0 \sum_{l \geq 2^{-1}n} h_l,
\]

and since \(\mathbb{E}|N - n| \leq \sqrt{n}\) and \(\mathbb{E}L_0 = |x_n - y_n|^p\),

\[
\frac{|\mathbb{E}L_n - \mathbb{E}L_N|}{(n f(z))^{1-p/d}|x_n - y_n|} \leq O((n f(z))^{\alpha p/d} n^{-1/2}) + \frac{\mathbb{P}\{N < 2^{-1}n\} + \sum_{l > 2^{-1}n} h_l}{(n f(z))^{1-p/d}|x_n - y_n|^{1-p}},
\]

where the summation \(\sum h_l\) is still for \(l > 2^{-1}n\). The first term on the right decays to zero since \(\alpha < d/(2p)\). The second term also decays to zero since, while the denominator has at most polynomial decay in \((n f(z))^{\alpha}|x_n - y_n|\), the numerator has exponential decay by the Chernoff bound, Lemma 4, and Lemma 3. Note that \(X_N\) is identically distributed as \(H_{n/(V_d R_2)} \cap B(z; R_2)\) where \(V_d\) is the volume of unit ball \(B(z; 1)\), and the proposition follows by Lemma 11 after a scale adjustment. \(\Box\)
References

ADRIAN BADDELEY (2007). “Spatial point processes and their applications”. In: Stochastic geometry. Ed. by WOLFGANG WEIL. Lecture Notes in Math. 1892. Lectures given at the C.I.M.E. Summer School held in Martina Franca, Italy, September 13–18, 2004. Springer. doi: 10.1007/3-540-38174-0.

MIRA BERNSTEIN et al. (Dec. 20, 2000). “Graph approximations to geodesics on embedded manifolds”. url: http://isomap.stanford.edu/BdSLT.pdf.

PATRICK BILLINGSLEY (1995). Probability and measure. 3rd ed. Wiley.

THOMAS H. CORMEN et al. (Sept. 2009). Introduction to Algorithms. 3rd ed. The MIT Press.

JOSE A. COSTA and ALFRED O. HERO (2004). “Geodesic entropic graphs for dimension and entropy estimation in manifold learning”. In: IEEE Trans. Signal Process. 52.8, pp. 2210–2221. doi: 10.1109/TSP.2004.831130.

E. W. DIJKSTRA (1959). “A note on two problems in connexion with graphs”. In: Numer. Math. 1, pp. 269–271. doi: 10.1007/BF01386390.

J. M. HAMMERSLEY (1966). “First-passage percolation”. In: J. R. Stat. Soc. Ser. B Stat. Methodol. 28.3, pp. 491–496. JSTOR: 2984443.

C. DOUGLAS HOWARD and CHARLES M. NEWMAN (1997). “Euclidean models of first-passage percolation”. In: Probab. Theory Related Fields 108.2, pp. 153–170. doi: 10.1007/s004400050105.

— (2001). “Geodesics and spanning trees for Euclidean first-passage percolation”. In: Ann. Probab. 29.2, pp. 577–623. doi: 10.1214/aop/1008956686.

SUNG JIN HWANG (2012). “Geometric representations of high dimensional random data”. PhD thesis. University of Michigan. url: http://web.eecs.umich.edu/~hero/Preprints/sjh_thesis_final.pdf.

HARRY KESTEN (1987). “Percolation theory and first-passage percolation”. In: Ann. Probab. 15.4, pp. 1231–1271. doi: 10.1214/aop/1176991975.

T. LA GATTA and J. WEHR (2010). “A shape theorem for Riemannian first-passage percolation”. In: J. Math. Phys. 51.5, p. 053502. doi: 10.1063/1.3409344.

RONALD MEESTER and RAHUL ROY (1996). Continuum percolation. Cambridge University Press.

MATHEW D. PENROSE and J. E. YUKICH (2003). “Weak laws of large numbers in geometric probability”. In: Ann. Appl. Probab. 13.1, pp. 277–303. doi: 10.1214/aoap/1042765669.

— (Apr. 2011). Limit theory for point processes in manifolds. arXiv:1104.0914v1 [math.PR].

J. MICHAEL STEELE (1997). Probability theory and combinatorial optimization. CBMS-NSF Regional Conf. Ser. in Appl. Math. 69. Soc. Ind. Appl. Math. doi: 10.1137/1.9781611970029.
Michel Talagrand (1995). “Concentration of measure and isoperimetric inequalities in product spaces”. In: *Publ. Math. Inst. Hautes Études Sci.* 81.1, 73:205. doi: 10.1007/BF02699376. arXiv:math/9406212v1 [math].

J. E. Yukich (2000). “Asymptotics for weighted minimal spanning trees on random points”. In: *Stochastic Process. Appl.* 85.1, pp. 123–138. doi: 10.1016/S0304-4149(99)00068-X.

Joseph E. Yukich (1998). *Probability theory of classical Euclidean optimization problems*. Lecture Notes in Math. 1675. Springer. doi: 10.1007/BFb0093472.