APPROXIMATING GEODESICS VIA RANDOM POINTS[1]

BY ERIK DAVIS AND SUNDER SETHURAMAN

University of Arizona

Given a cost functional F on paths γ in a domain $D \subset \mathbb{R}^d$, in the form $F(\gamma) = \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt$, it is of interest to approximate its minimum cost and geodesic paths. Let $X_1, \ldots, X_n$ be points drawn independently from D according to a distribution with a density. Form a random geometric graph on the points where X_i and X_j are connected when $0 < |X_i - X_j| < \varepsilon$, and the length scale $\varepsilon = \varepsilon_n$ vanishes at a suitable rate.

For a general class of functionals F, associated to Finsler and other distances on D, using a probabilistic form of Gamma convergence, we show that the minimum costs and geodesic paths, with respect to types of approximating discrete cost functionals, built from the random geometric graph, converge almost surely in various senses to those corresponding to the continuum cost F, as the number of sample points diverges. In particular, the geodesic path convergence shown appears to be among the first results of its kind.

1. Introduction. Understanding the shortest or geodesic paths between points in a medium is an intrinsic concern in diverse applied problems, from optimal routing in networks and disordered materials to identifying manifold structure in large data sets, as well as in studies of $\mathbb{Z}^d$ and $\mathbb{R}^d$-percolation (cf. recent survey [5] and [17–22]).

There are sometimes abstract formulas for the geodesics, from the calculus of variations, or other differential equation approaches. For instance, with respect to a patch of a Riemannian manifold (M, g), with $M \subset \mathbb{R}^d$ and tensor field $g(\cdot)$, it is known that the distance function $U(\cdot) = d(x, \cdot)$, for fixed x, is a viscosity solution of the Eikonal equation $\|\nabla U(y)\|_{g(y)^{-1}} = 1$ for $y \neq x$, with boundary condition $U(x) = 0$. Here, $\|v\|_A = \sqrt{\langle v, Av \rangle}$, where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^d$. Then a geodesic γ connecting x and z may be recovered from U by solving a descent equation, $\dot{\gamma}(t) = -\eta(t) g^{-1}(\gamma(t)) \nabla U(\gamma(t))$, where $\eta(t)$ is a scalar function controlling the speed.

On the other hand, computing numerically the distances and geodesics may be a complicated issue. One of the standard approaches is the fast marching method to

approximate the distance U, by solving the Eikonal equation on a regular grid of n points. This method has been extended in a variety of ways, including with respect to triangulated domains, as well as irregular samples $\{x_1, \ldots, x_n\}$ of an Euclidean submanifold (cf. [24, 30]). See also [27] in the above contexts for a review.

Alternatively, variants of Dijkstra's or heat flow methods, on graphs approximating the space are sometimes used. In Dijkstra's algorithm, distances and shortest paths are found by successively computing optimal routes to nearest-neighbor edges. In heat flow methods, geodesic distances can be found in terms of the small time asymptotics of a heat kernel on the space. For instance, see [10, 11, 14, 33, 34].

Another idea has been to collect a random sample $\mathcal{X}_n$ of n points from a manifold embedded in $\mathbb{R}^d$, put a network structure on these points, say in terms of a ε-random geometric or k-nearest neighbor graph, and then approximate the continuum geodesics lengths by lengths of discrete geodesic paths found in this network. Presumably, under assumptions on how the points are sampled and how the random graphs are formed, as the number of points diverge, these discrete distances should converge almost surely to the continuum shortest path lengths. Such a statistical consistency result is fundamental in manifold learning [6]. For instance, the popular ISOMAP procedure [7, 32] is based on these notions to elicit manifold structure in data sets.

More specifically, let D be a subset of $\mathbb{R}^d$ corresponding to a patch of the manifold, and consider a kernel $f(x, v) : D \times \mathbb{R}^d \to [0, \infty)$. Define the f-cost of a path $\gamma(t) : [0, 1] \to D$ from $\gamma(0) = a$ to $\gamma(1) = b$ as $F(\gamma) = \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt$. The f-distance from a to b is then the infimum of such costs over paths γ. For example, if $f(x, v) = |v|^p$ and $p \geq 1$, the f-distance is $|b - a|^p$, the pth power of the Euclidean distance. Moreover, when $p = 1$, $F(\gamma)$ is the arc length of the path γ and the f-distance is the length of the line segment from a to b.

With respect to a class of functions f and samples drawn from a distribution on the D with density ρ, papers [7, 28] and [3] address, among other results, how $\varepsilon = \varepsilon_n$ and $k = k_n$ should decrease and increase, respectively, so that various concentration type bounds between types of discrete and continuum optimal distances hold with high probability, leading to consistent estimates.

For instance, in [28], for ε_n-random graphs and smooth ρ, certain density dependent estimators of continuum distances were considered, where $f(x, v) = h(\rho(x))|v|$ and $h(y)$ is decreasing, smooth, constant for $|y|$ small, and bounded away from 0. This work extends [7], which considered $f(x, v) = |v|$ and uniformly distributed samples. On the other hand, in [3], among other results, on k_n-nearest-neighbor graphs, continuum distances, where $f(x, v) = h(\rho(x))|v|$ and h is increasing, Lipschitz, and bounded away from 0, were approximated (see also [15]).

In these contexts, the purpose of this article is twofold. First, we identify a general class of f-distances for which different associated discrete distances, formed from random ε_n-random geometric graphs on a domain $D \subset \mathbb{R}^d$, converge almost

surely to them. Second, we describe when the associated discrete geodesic paths converge almost surely, in uniform and Hausdorff norms, to continuum f-distance geodesic paths, a type of consistency which appears to be among the first contributions of this kind. The main results are Theorems 2.1, 2.2, 2.4, 2.5, 2.6 and Corollary 2.3. See Section 2 and Section 2.3 for precise statements and related remarks.

We consider the following three different discrete costs. The first, d_1, optimizes on paths γ, starting and ending at a and b, respectively, linearly interpolated between points in $\mathcal{X}_n \cup \{a, b\}$, where consecutive points are within ε_n of each other, and the time to traverse each link is the same. The second, d_2, optimizes with respect to quasinormal interpolations between the points, using however the f-geodesic paths. The third, d_3, does not interpolate at all, and optimizes a "Riemann sum" cost $\frac{1}{m} \sum_{i=0}^{m-1} f(v_i, m(v_{i+1} - v_i))$ where m is the number of edges in the discrete path $\{v_0, \dots, v_m\} \subset \mathcal{X}_n \cup \{a, b\}$. We note, discrete distances d_2 and d_3, in the setting $f(x, v) = |v|$ were introduced in [7], and density dependent versions were used in the results in [28]. The discrete distance d_1, although natural, seems not well considered in the literature.

The conditions we impose on f include p-homogeneity in v for $p \geq 1$, convexity and an ellipticity condition with respect to v, and a smoothness assumption away from $v \neq 0$. Such conditions include a large class of kernels f associated to Finsler spaces, as well as those kernels considered in [28] and [7], with respect to ε_n-random graphs. The domain $D \subset \mathbb{R}^d$ is assumed to be bounded and convex. Also, we assume that the rate of decrease of ε_n is such that the graph on $\mathcal{X}_n$ is connected for all large n. Moreover, this choice of ε_n implies that the density of nearby points increases sufficiently quickly so that zig-zags in the optimal path approximations do not accumulate.

While a main contribution of the article is to provide a general setting in which the discrete to continuum convergences hold, we remark our proof method is quite different from that in the literature, where specific features of f, such as $f(x, v) = |v|$ in [7], are important in estimation of distances, not easily generalized. We use a form of Gamma convergence to derive the almost sure limits, which may be of interest itself. This method involves showing liminf, limsup and compactness elements, as in the analysis context, but here on appropriate probability 1 sets. Part of the output of the technique, beyond giving convergence of the distances, is that it yields convergence of the minimizing discrete paths to continuum geodesics in various senses.

The f-costs share different properties depending on if $p = 1$ or $p > 1$, and also when $d \geq 2$ or $d = 1$. For instance, the f-cost is invariant to reparametrization of the path exactly when $p = 1$. Also, when $p > 1$, the form of the f-cost may be seen to be coercive on the modulus of γ, not the case when $p = 1$. In fact, the $p = 1$ case is the most troublesome, and more assumptions on f and ε_n are required in Theorems 2.5 and 2.6 to deal with the linear path cost d_1 and Riemann cost d_3, which are rougher than the quasinormal cost d_2.

At the same time, in $d = 1$, in contrast to $d \geq 2$, all paths lie in the interval $[a, b] \subset \mathbb{R}$. When also $p = 1$, the problem is somewhat degenerate: By invariance to reparametrization, the costs d_1 and d_2 turn out to be nonrandom and to reduce to the integral $\int_a^b f(s, 1) \, ds$. Also, the cost d_3 is a Riemann sum which converges to this integral.

Finally, we comment on a difference in viewpoint with respect to results in continuum percolation. The Riemann sum cost considered here seems related to but is different than the cost optimized in the works [18, 19]. There, for $p > 1$, one optimizes the cost of a path $\{w_0, \dots, w_m\}$, along random points, from the origin 0 to nx, for $x \in \mathbb{R}^d$, given by $\sum_{i=0}^{m-1} |w_{i+1} - w_i|^p$, and infers a scaled distance $d(x) = c(d, p)|x|$, in law of large numbers scale n, where the proportionality constant $c(d, p)$ is not explicit. In contrast, however, in this article, given already an integral f-distance, the viewpoint is to optimize costs of paths of length order 1 (not n as in [18, 19]), where the length scale between points is being scaled of order ε_n, and then to recover the f-distance in the limit. We note also another difference: When $f(x, v) = |v|^p$, as remarked above, the f-distance from the origin to x is $|x|^p$, instead of $\sim |x|$ as in the continuum percolation studies.

In Section 2, the setting, assumptions and results are given with respect to three types of discrete costs. In Section 3, proofs of Theorems 2.1 and 2.2 and Corollary 2.3 on the interpolating costs are given. In Section 4, proofs of Theorems 2.4, 2.5 and 2.6, with respect to Riemann costs, and interpolating costs when $p = 1$, are given. In the Appendix, some technical results, used in the course of the main proofs, are collected.

2. Setting and results. We now introduce the setting of the problem, and standing assumptions, which hold throughout the article.

For $d \geq 1$, we will be working on a subset $D \subset \mathbb{R}^d$,

(2.1)   which is the closure of an open, bounded, convex domain.

Therefore, D is a Lipschitz domain (cf. Corollary 9.1.2 in [2], Section 1.1.8 in [23]).

Consider points $a, b \in D$ and let $\Omega(a, b)$ denote the space of Lipschitz paths $\gamma : [0, 1] \to D$ with $\gamma(0) = a$ and $\gamma(1) = b$. Given $f : D \times \mathbb{R}^d \to [0, \infty)$, we define the cost $F : \Omega(a, b) \to [0, \infty)$ by

$$F(\gamma) = \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt,$$

and associated optimal cost

(2.2) $$d_f(a, b) = \inf_{\gamma \in \Omega(a,b)} F(\gamma).$$

We will make the following assumptions on the integrand f:

(A0) f is continuous on $D \times \mathbb{R}^d$, and C^1 on $D \times (\mathbb{R}^d \setminus \{0\})$,

(A1) $f(x, v)$ is convex in v,

(A2) there exists $p \geq 1$ such that $f(x, v)$ is p-homogenous in v,

$$(2.3) \qquad\qquad f(x, \lambda v) = \lambda^p f(x, v) \qquad \text{for } \lambda > 0,$$

(A3) there exist constants $m_1, m_2 > 0$ such that

$$(2.4) \qquad\qquad m_1 |v|^p \leq f(x, v) \leq m_2 |v|^p \qquad \text{for all } x \in D.$$

We remark, when $p > 1$ and p-homogeneity (A2) holds, that f may be extended to a C^1 function on $D \times \mathbb{R}^d$.

Part of the reasoning for the assumptions (A0)–(A3) is that they include, for $p \geq 1$, the familiar kernel $f(x, v) = |v|^p$, mentioned in the Introduction, for which, when $p = 1$, $F(\gamma)$ is the arclength of the path γ and $d_f(a, b) = |b - a|$.

Also, under these assumptions on f, it is known that the infimum in (2.2) is attained at a path in $\Omega(a, b)$, perhaps nonuniquely (see Proposition A.2 of the Appendix). In addition, we remark, when $p = 1$, under additional differentiability assumptions, d_f represents a Finsler distance (cf. [4, 31] and references therein).

When $p = 1$, there is an interesting scaling property: By 1-homogeneity of f, the cost F is invariant under smooth reparametrization of paths. That is, given a path $\gamma \in \Omega(a, b)$ and smooth, increasing $s : [0, 1] \to [0, 1]$, with $s(0) = 0$ and $s(1) = 1$, one has $F(\tilde{\gamma}) = F(\gamma)$ where $\tilde{\gamma}(t) = \gamma(s(t))$.

This property allows to deduce, when $p = 1$, that d_f satisfies the triangle property (not guaranteed when $p > 1$): Let γ_1 be a path from u to w, and γ_2 be a path from w to z.
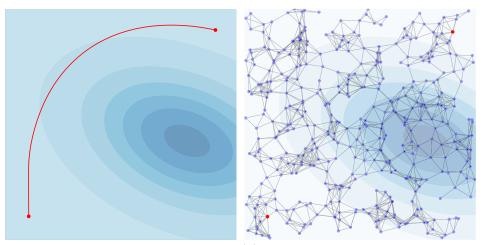
Write

$$
\begin{aligned}
& \int_0^1 f\big(\gamma_1(t), \dot{\gamma}_1(t)\big)\, dt + \int_0^1 f\big(\gamma_2(t), \dot{\gamma}_2(t)\big)\, dt \\
(2.5) \qquad & = \int_0^{1/2} f\big(\gamma_1(2s), 2\dot{\gamma}_1(2s)\big)\, ds + \int_0^{1/2} f\big(\gamma_2(2s), 2\dot{\gamma}_2(2s)\big)\, ds \\
& = \int_0^1 f\big(\gamma_3(t), \dot{\gamma}_3(t)\big)\, dt,
\end{aligned}
$$

where γ_3 is a path from u to z, following $\gamma_1(2\cdot)$ and $\gamma_2(2\cdot)$ on time intervals $[0, 1/2]$ and $[1/2, 1]$, respectively. Optimizing over γ_1, γ_2 and γ_3 gives $d_f(u, w) + d_f(w, z) \geq d_f(u, z)$.

We now construct a random geometric graph on D through which approximations of d_f and its geodesics will be made. Let $\{X_1, X_2, \ldots\} \subset D$ be a sequence of independent points, identically distributed according to a distribution ν with probability density ρ. For each $n \in \mathbb{N}$, let $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ and fix a length scale $\varepsilon_n > 0$. With respect to a realization $\{X_i\}$, we define a graph $\mathcal{G}_n(a, b)$, on the vertex set $\mathcal{X}_n \cup \{a, b\}$, by connecting an edge between u, v in $\mathcal{X}_n \cup \{a, b\}$ iff $0 < |u - v| < \varepsilon_n$, where $|\cdot|$ refers to the Euclidean distance in $\mathbb{R}^d$.

(a) F-minimizing path, with level sets of w indicated.

(b) ϵ_n-graph, on $n = 400$ uniform points with $\epsilon_n = (1/400)^{0.3}$

FIG. 1. *Continuum geodesic and ε_n-graph for $f(x, v) = w(x)|v|$ on the domain $D = [-1, 1] \times [-1, 1]$, where $w(x) = 1 + 8\exp(-2(x_1 - 1/2)^2 + x_1 x_2 + 2x_2^2)$, and $a = (-0.8, -0.8)$, $b = (0.8, 0.8)$.*

For $u, v \in \mathcal{X}_n \cup \{a, b\}$, we say that a finite sequence $(v_0, v_1, \ldots, v_m)$ of vertices is a *path* with m-steps from u to v in $\mathcal{G}_n(a, b)$ if $v_0 = u$, $v_m = v$, and there is an edge from v_i to v_{i+1} for $0 \le i < m$. Let $V_n(a, b)$ denote the set of paths from a to b in $\mathcal{G}_n(a, b)$.

We will assume a certain decay rate on ε_n, namely that $\lim_{n\uparrow\infty} \varepsilon_n = 0$ and

$$(2.6) \qquad \limsup_{n\to\infty} \frac{(\log n)^{1/d}}{n^{1/d}} \frac{1}{\varepsilon_n} = 0.$$

Under this type of decay rate, almost surely, for all large n and $a, b \in D$, points a, b will be connected by a path in the graph $\mathcal{G}_n(a, b)$, in other words, $V_n(a, b)$ will be nonempty. Indeed, under this rate, the degree of a point in the graph will diverge to infinity. See Proposition A.1 in the Appendix, and remarks in Section 2.3.

We will also assume that the underlying probability density ρ is uniformly bounded, that is, there exists a constant $c > 0$ such that

$$(2.7) \qquad c \le \rho(x) \le c^{-1} \qquad \text{for all } x \in D.$$

See Figure 1, parts (a) and (b), which depict a geodesic path with respect to a cost F, and an ε_n-random graph.

Standing assumptions. To summarize, the assumptions, dimension $d \ge 1$, (2.1) on D, items (A0)–(A3) on f when $p \ge 1$, decay rate (2.6) on ε_n, and density bound (2.7) on ρ, denoted as the standing assumptions, will hold throughout the article.

In the next two subsections, we present results on approximation of $d_f(a, b)$ and its geodesics with respect to two types of schemes, where approximating costs are built (1) in terms of interpolations of points in $V_n(a, b)$ and also (2) in terms of Riemann sums.

2.1. *Interpolating costs.* We introduce two types of discrete costs based on linear and quasinormal paths.

Linear interpolations. With respect to a realization $\{X_i\}$, for $u, v \in D$, let $l_{u,v} : [0, 1] \to D$ denote the constant-speed linear path from a to b, given by

$$l_{u,v}(t) = (1 - t)u + tv.$$

Consider now $\mathbf{v} = (v_0, v_1, \dots, v_m) \in V_n(a, b)$. We define $l_{\mathbf{v}} \in \Omega(a, b)$ to be the concatenation of the linear segments $\{l_{v_{i-1}, v_i}\}_{i=1}^m$, where each segment is traversed in the same time $1/m$. More precisely, for $i/m \leq t \leq (i + 1)/m$, define

$$l_{\mathbf{v}}(t) = l_{v_i, v_{i+1}}(mt - i),$$

and note that the resulting piecewise linear path is in $\Omega(a, b)$.

Define now a subset $\Omega_n^l(a, b)$ of $\Omega(a, b)$ by

$$\Omega_n^l(a, b) = \{l_{\mathbf{v}} | \mathbf{v} \in V_n(a, b)\},$$

and define the (random) discrete cost $L_n : \Omega_n^l(a, b) \to [0, \infty]$ by

$$L_n(\gamma) = F(\gamma) \qquad \text{for } \gamma \in \Omega_n^l(a, b).$$

In other words, L_n is the restriction of F to $\Omega_n^l(a, b)$. By the p-homogeneity of f, we may write this as

$$
\begin{aligned}
L_n(l_{\mathbf{v}}) &= \sum_{i=0}^{m-1} \int_{i/m}^{(i+1)/m} f\big(l_{v_i, v_{i+1}}(mt - i), m(v_{i+1} - v_i)\big) \, dt \\
&= m^{p-1} \sum_{i=0}^{m-1} \int_0^1 f\big(l_{v_i, v_{i+1}}(t), v_{i+1} - v_i\big) \, dt.
\end{aligned}
$$
(2.8)

Quasinormal interpolations. Define now a different discrete cost which may nonlinearly interpolate among points in paths of $V_n(a, b)$. We say that a Lipschitz path γ is *quasinormal* with respect to f if there exists a $c > 0$ such that

$$f\big(\gamma(t), \dot{\gamma}(t)\big) = c \qquad \text{for a.e. } t \in [0, 1].$$

It is known, under the standard assumptions on f (see Proposition A.2) that, for $u, v \in D$, there exists a *quasinormal* path $\gamma : [0, 1] \to D$, with $\gamma(0) = u, \gamma(1) = v$, which is optimal, $d_f(u, v) = \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt$. For what follows, when we refer to a *quasinormal* path connecting u and v, we mean such a fixed optimal path denoted by $\gamma_{u,v}$.

Given a path $\mathbf{v} = (v_0, \ldots, v_m) \in V_n(a, b)$, let $\gamma_{\mathbf{v}} \in \Omega(a, b)$ denote the concatenation of $\{\gamma_{v_{i-1}, v_i}\}_{i=1}^m$, where each segment uses the same time $1/m$. More precisely, for $i/m \le t \le (i+1)/m$, define

$$\gamma_{\mathbf{v}}(t) = \gamma_{v_i, v_{i+1}}(mt - i).$$

As with piecewise linear functions, define the subset $\Omega_n^\gamma(a, b)$ of $\Omega(a, b)$ by

$$\Omega_n^\gamma(a, b) = \{\gamma_{\mathbf{v}} | \mathbf{v} \in V_n(a, b)\}.$$

Let $G_n : \Omega_n^\gamma(a, b) \to \mathbb{R}$ denote the restriction of F to $\Omega_n^\gamma(a, b)$.

Then, with respect to a path $\gamma = \gamma_{\mathbf{v}} \in \Omega_n^\gamma(a, b)$, by the p-homogeneity of f, we evaluate that

$$
\begin{aligned}
G_n(\gamma) &= \int_0^1 f\big(\gamma(t), \dot{\gamma}(t)\big)\, dt \\
&= \sum_{i=1}^m \int_{(i-1)/m}^{i/m} f\big(\gamma_{v_{i-1}, v_i}(mt - i), m\dot{\gamma}_{v_{i-1}, v_i}(mt - i)\big)\, dt \\
&= m^{p-1} \sum_{i=1}^m \int_0^1 f\big(\gamma_{v_{i-1}, v_i}(t), \dot{\gamma}_{v_{i-1}, v_i}(t)\big)\, dt \\
&= m^{p-1} \sum_{i=1}^m d_f(v_{i-1}, v_i).
\end{aligned}
$$
(2.9)

Further, by p-homogeneity of f and optimality of $\{\gamma_{v_i, v_{i+1}}\}_{i=1}^m$, the segments of $\gamma = \gamma_{\mathbf{v}}$ are also optimal, in the sense that

$$
\begin{aligned}
&\int_{i/m}^{(i+1)/m} f\big(\gamma(t), \dot{\gamma}(t)\big)\, dt \\
(2.10) \quad &= m^{p-1} \int_0^1 f\big(\gamma_{v_i, v_{i+1}}(t), \dot{\gamma}_{v_i, v_{i+1}}(t)\big)\, dt \\
&= \inf_{\widetilde{\gamma}} m^{p-1} \int_0^1 f\big(\widetilde{\gamma}(t), \dot{\widetilde{\gamma}}(t)\big)\, dt = \inf_{\widehat{\gamma}} \int_{i/m}^{(i+1)/m} f\big(\widehat{\gamma}(t), \dot{\widehat{\gamma}}(t)\big)\, dt,
\end{aligned}
$$

where the infima are over Lipschitz paths $\widetilde{\gamma} : [0, 1] \to D$ and $\widehat{\gamma} : [i/m, (i+1)/m] \to D$ with $\widetilde{\gamma}(0) = v_i$, $\widetilde{\gamma}(1) = v_{i+1}$, $\widehat{\gamma}(i/m) = v_i$ and $\widehat{\gamma}((i+1)/m) = v_{i+1}$.

Relations between G_n and L_n. At this point, we remark there are kernels f for which $G_n = L_n$, namely those such that linear segments are in fact *quasinormal* geodesics. An example is $f(x, v) = |v|$. Identifying these kernels is a question with a long history, going back to Hilbert, whose fourth problem paraphrased asks for which geometries are the geodesics straight lines (cf. surveys [4, 25]). Hamel's criterion, namely $\partial_{x_i} \partial_{v_j} f = \partial_{x_j} \partial_{v_i} f$ for $1 \le i, j \le d$, is a well-known solution to this question (see [4, 13, 25] and references therein).

We also note, as mentioned in the Introduction, that the case $d = p = 1$ is degenerate in that $\min G_n$ and $\min L_n$ are not random. Indeed, let $\gamma_{\mathbf{v}} \in \arg\min G_n$ and suppose $\mathbf{v} = (v_0, \ldots, v_m) \in V_n(a, b)$. We observe that $\gamma_{\mathbf{v}}$ must be nondecreasing, as otherwise, one could build a smaller cost path, from parts of $\gamma_{\mathbf{v}}$ using invariance to reparametrization, violating optimality of $\gamma_{\mathbf{v}}$. In particular, $\dot{\gamma}_{\mathbf{v}} \geq 0$ and $v_i < v_{i+1}$ for $0 \leq i \leq m-1$. Then

$$G_n(\gamma_{\mathbf{v}}) = \sum_{i=0}^{m-1} \int_{i/m}^{(i+1)/m} f\big(\gamma_{\mathbf{v}}(t), \dot{\gamma}_{\mathbf{v}}(t)\big) dt$$

$$= \sum_{i=0}^{m-1} \int_{v_i}^{v_{i+1}} f(s, 1) ds = \int_a^b f(s, 1) ds,$$

using the 1-homogeneity of f and changing variables. The same argument yields that $\min L_n = \int_a^b f(s, 1) ds$. We do not consider this degenerate case further.

The first result is for linearly interpolated paths.

THEOREM 2.1.   *Suppose that $p > 1$. With respect to realizations $\{X_i\}$ in a probability* 1 *set, the following holds. The minimum values of the costs L_n converge to the minimum of F,*

$$\lim_{n \to \infty} \min_{\gamma \in \Omega_n^l(a,b)} L_n(\gamma) = \min_{\gamma \in \Omega(a,b)} F(\gamma).$$

Moreover, consider a sequence of optimal paths $\gamma_n \in \arg\min L_n$. Any subsequence of $\{\gamma_n\}$ has a further subsequence that converges uniformly to a limit path $\gamma \in \arg\min F$,

$$\lim_{k \to \infty} \sup_{0 \leq t \leq 1} \big|\gamma_{n_k}(t) - \gamma(t)\big| = 0.$$

In addition, if γ is the unique minimizer of F, then the whole sequence γ_n converges uniformly to γ.

The case $d \geq 2$ and $p = 1$ requires further development, and is addressed with a few more assumptions in Theorem 2.6.

We now address *quasinormal* interpolations.

THEOREM 2.2.   *Suppose that either* (1) $p > 1$ *or* (2) $d \geq 2$ *and $p = 1$. Then, with respect to realizations $\{X_i\}$ in a probability* 1 *set, the following holds. The minimum values of the energies G_n converge to the minimum of F,*

$$\lim_{n \to \infty} \min_{\gamma \in \Omega_n^\gamma(a,b)} G_n(\gamma) = \min_{\gamma \in \Omega(a,b)} F(\gamma).$$

Moreover, consider a sequence of optimal paths $\gamma_n \in \arg\min G_n$. Any subsequence of $\{\gamma_n\}$ has a further subsequence that converges uniformly to a limit path

$\gamma \in \arg\min F$,

$$\lim_{k\to\infty} \sup_{0\le t\le 1} |\gamma_{n_k}(t) - \gamma(t)| = 0.$$

In addition, if γ is the unique minimizer of F, then the whole sequence γ_n converges uniformly to γ.

We remark, when $d \ge 2$ and $p = 1$, that there is a certain ambiguity in the results of Theorem 2.2, due to the invariance of F under reparametrization of paths. In this case, there is no unique minimizer of F. Consider, for example, the case where $f(x, v) = |v|$ and $F(\gamma) = \int_0^1 |\dot{\gamma}(t)| \, dt$. Any minimizer of this functional is a parametrization of a line, but of course such minimizers are not unique.

One way to address this is to formulate a certain *Hausdorff* convergence with respect to images of the paths. Given $\gamma \in \Omega(a, b)$, we denote the image of γ by

$$S_\gamma = \{\gamma(t) | 0 \le t \le 1\}.$$

Consider the *Hausdorff* metric d_{haus}, defined on nonempty compact subsets A, B of D by

$$d_{\text{haus}}(A, B) = \max\left\{\sup_{x\in A} \inf_{y\in B} d(x, y), \sup_{y\in B} \inf_{x\in A} d(x, y)\right\}.$$

COROLLARY 2.3. *Suppose that either* (1) $d \ge 2$ *and* $p = 1$ *or* (2) $p > 1$. *Consider paths* $\{\gamma_{\mathbf{v}^{(n)}}\}$, *for all large n either in form* $\gamma_{\mathbf{v}^{(n)}} \in \arg\min G_n$, *or* $\gamma_{\mathbf{v}^{(n)}} \in \arg\min L_n$.

Then, with respect to realizations $\{X_i\}$ in a probability 1 set, any subsequence of $\{\mathbf{v}^{(n)}\}$ has a further subsequence which converges in the Hausdorff sense to S_γ, where $\gamma \in \arg\min F$ is an optimal path.

Moreover, if F has a unique (up to reparametrization) minimizer γ, then the whole sequence converges,

$$\lim_{n\to\infty} d_{\text{haus}}(\mathbf{v}^{(n)}, S_\gamma) = 0.$$

2.2. *Riemann sum costs and $p = 1$-linear interpolating costs.* We first introduce a cost which requires knowledge of f only on discrete points and, as a consequence, more applicable. At the end of the subsection, we return to linear interpolated costs when $p = 1$.

Define $H_n : V_n(a, b) \to \mathbb{R}$, for $\mathbf{v} = (v_0, v_1, \ldots, v_m)$, by

(2.11)
$$H_n(\mathbf{v}) = \frac{1}{m} \sum_{i=1}^{m} f(v_i, m(v_{i+1} - v_i)).$$

The functional H_n is, in a sense, a Riemann sum approximation to L_n and G_n and, therefore, its behavior, and the behavior of its minimizing paths, should be similar to that of L_n and G_n. See Figure 2 for an example of an optimal H_n path.
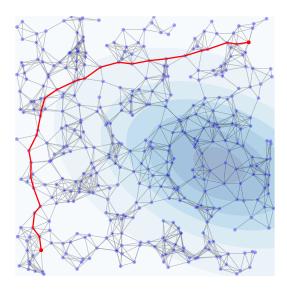
FIG. 2. H_{400}-minimizing discrete path in the setting of Figure 1, linearly interpolated for visual clarity (see Section 2.2).

We make this intuition rigorous by establishing variants of Theorems 2.1 and 2.2 with respect to the cost H_n. Given the rougher nature of H_n, however, additional assumptions on f and ε_n, beyond those in the standing assumptions, will be helpful in this regard. As in the previous subsection, our results differ between the two cases $p = 1$ and $p > 1$.

Define the following smoothness condition:

(Lip) There exists a c such that for all $x, y \in D$ and $v \in \mathbb{R}^d$ we have

$$\left| f(x, v) - f(y, v) \right| \le c |x - y| |v|^p.$$

We note when f satisfies the homogeneity condition (2.3), and $\nabla_x f(x, v)$ is uniformly bounded on $D \times \{y : |y| = 1\}$, that (Lip) holds.

We now consider the behavior of H_n when $p > 1$. The analogue to Theorem 2.1 and Corollary 2.3 in this setting is the following.

THEOREM 2.4. *Suppose $p > 1$, and that f in addition satisfies* (Lip). *With respect to realizations $\{X_i\}$ in a probability 1 set, the minimum values of the energies H_n converge to the minimum of F,*

$$\lim_{n \to \infty} \min_{\mathbf{v} \in V_n(a,b)} H_n(\mathbf{v}) = \min_{\gamma \in \Omega(a,b)} F(\gamma).$$

Further, consider a sequence of optimal discrete paths $\mathbf{w}^{(n)} \in \arg\min H_n$, and their linear interpolations $\{l_{\mathbf{w}^{(n)}}\}$. Then, for any subsequence of $\{l_{\mathbf{w}^{(n)}}\}$ and correspondingly of $\{\mathbf{w}^{(n)}\}$, there is a further subsequence of the linear paths which

converges uniformly to a limit path $\gamma \in \arg\min F$, *and of the discrete paths in the Hausdorff sense to* S_γ.

If F *has a unique minimizer* γ, *the whole sequence of linear paths converges uniformly to* γ, *and the whole sequence of discrete paths converges in the Hausdorff sense to* S_γ.

We will need to impose further assumptions on the integrand f to state results in the case $p = 1$.

(Hilb) We say that f satisfies the "Hilbert condition" if, for each x,

$$\inf_{\gamma \in \Omega(a,b)} \int_0^1 f(x, \dot\gamma(t))\, dt = f(x, b - a),$$

that is, straight lines are geodesics for the kernel $f(x, \cdot)$.

(TrIneq) We say f satisfies the triangle inequality if, for each x,

$$f(x, v - w) \le f(x, v - u) + f(x, u - w)$$

for all $u, v, w \in \mathbb{R}^d$.

(Pythag) Let $\alpha > 1$. Consider points u, v, w where $|uw|, |vw|, |uv| < \eta$ for a $\eta < 1$. Suppose there is a constant c such that, for $0 < r < 1$:

– $\operatorname{dist}(w, \operatorname{line}(u, v)) \ge r$ and
– $|uv| \le cr^{1/\alpha}$.

Then we say f satisfies the Pythagoras α-condition if there is a constant $C = C(\alpha, f, c)$ such that, for all x,

$$f(x, w - u) + f(x, v - w) \ge f(x, v - u) + Cr^\alpha.$$

Here, $\operatorname{line}(u, v)$ is the line segment between u and v, and $|uv| = |u - v|$.

An example of a large class of kernels f, including $f(x, v) = |v|$, satisfying the standing assumptions and the additional conditions above is given in Section 2.4. See also Section 2.3 for further comments.

We will also need to limit the decay properties of ε_n for the next result; see Section 2.3 for comments on this limitation. Namely, we will suppose that ε_n is in form $\varepsilon_n = n^{-\delta}$ where

$$(2.12) \qquad \delta > \max\{[(2 - \alpha^2)\eta + d]^{-1}, [(\alpha(d - 1) + 1]^{-1}\},$$

for an $0 < \eta < 1$ and $1 < \alpha < \sqrt{2}$.

We note that condition (2.6), when ε_n is in form $\varepsilon_n = n^{-\delta}$, yields that $\delta < 1/d$. However, when $d \ge 2$, we have $\max\{[(2 - \alpha^2)\eta + d]^{-1}, [(\alpha(d - 1) + 1]^{-1}\} < 1/d$, and so (2.12), in conjunction with (2.6), limits δ to an interval.

THEOREM 2.5. *Suppose $d \geq 2$ and $p = 1$, and that f also satisfies* (Lip) *and* (Hilb). *With respect to realizations $\{X_n\}_{n \geq 1}$ in a probability 1 set, the minimum values of the cost H_n converge to the minimum of F,*

$$\lim_{n \to \infty} \min_{\mathbf{v} \in V_n(a,b)} H_n(\mathbf{v}) = \min_{\gamma \in \Omega(a,b)} F(\gamma).$$

Moreover, suppose now that f in addition satisfies (TrIneq) *and* (Pythag) *for an $1 < \alpha < \sqrt{2}$, and that ε_n satisfies* (2.12).

Consider a sequence of optimal discrete paths $\mathbf{v}^{(n)} \in \arg \min H_n$, and their linear interpolations $\{l_{\mathbf{v}^{(n)}}\}$. Then, for any subsequence of $\{l_{\mathbf{v}^{(n)}}\}$ and so of $\{\mathbf{v}^{(n)}\}$, there is a further subsequence of the linear paths which converges uniformly to a limit path $\gamma \in \arg \min F$, and of the discrete paths in the Hausdorff sense to S_γ.

If F has a unique (up to reparametrization) minimizer γ, the whole sequence of discrete paths converges, $\lim_{n \to \infty} d_{\mathrm{haus}}(\mathbf{v}^{(n)}, S_\gamma) = 0$.

See Figure 2 for an example of an H_n-cost geodesic path.

As noted in the Introduction, when $d = p = 1$, $H_n(\mathbf{v})$ is a certain Riemann sum. Let $\mathbf{w} \in \arg \min H_n$, and observe by optimality that $\mathbf{w} = (w_0, \ldots, w_m) \in V_n(a,b)$ satisfies $w_i < w_{i+1}$ for $0 \leq i \leq m-1$. Hence, by 1-homogeneity of f, $H_n(\mathbf{w}) = \sum_{i=0}^{m-1} f(w_i, 1)|w_{i+1} - w_i|$, and $H_n(\mathbf{w})$ strongly approximates the integral $\int_a^b f(s, 1)\, ds$, given that the partition length $\max |w_{i+1} - w_i| \leq \varepsilon_n \to 0$. For this reason, this case is not included in the above theorem.

Linear interpolating costs when $p = 1$. Having now introduced (Lip), (Hilb), (TrIneq) and (Pythag), we address the case $d \geq 2$ and $p = 1$ with respect to the cost L_n.

THEOREM 2.6. *Suppose $d \geq 2$ and $p = 1$, and that f also satisfies* (Lip) *and* (Hilb). *With respect to realizations $\{X_n\}_{n \geq 1}$ in a probability 1 set, the minimum values of the cost L_n converge to the minimum of F,*

$$\lim_{n \to \infty} \min_{\gamma \in \Omega_n^l(a,b)} L_n(\gamma) = \min_{\gamma \in \Omega(a,b)} F(\gamma).$$

Moreover, suppose now that f in addition satisfies (TrIneq) *and* (Pythag) *for an $1 < \alpha < \sqrt{2}$, and that ε_n satisfies* (2.12).

Consider a sequence of optimal paths $l_{\mathbf{v}^{(n)}} \in \arg \min L_n$. Then, for any subsequence of $\{l_{\mathbf{v}^{(n)}}\}$ and so of the discrete paths $\{\mathbf{v}^{(n)}\}$, there is a further subsequence of the linear paths which converges uniformly to a limit path $\gamma \in \arg \min F$, and of the discrete paths in the Hausdorff sense to S_γ.

If F has a unique (up to reparametrization) minimizer γ, the whole sequence of discrete paths converges, $\lim_{n \to \infty} d_{\mathrm{haus}}(\mathbf{v}^{(n)}, S_\gamma) = 0$.

2.3. *Remarks.*   We make several comments about the assumptions and related issues.

1. *Domain.* The requirements that D should be closed and connected are needed for the quasinormal path results in [9] and [16] to hold. Also, a bound on the domain allows the Arzelà–Ascoli equicontinuity criterion to be applied in the compactness property, Lemma 3.7. In addition, the proof of Proposition A.1, on the maximum distance to a nearest neighbor vertex, requires that the domain boundary should be Lipschitz, true for convex domains. The convexity of the domain also ensures that all the linearly interpolated paths are within the domain, and allows comparison with quasinormal ones, which by definition are constrained in the domain. In the $p > 1$ case, convexity principally enters in the proof of the limsup inequality, Lemma 3.5, and its requisite propositions. Here, the assumption could perhaps be relaxed in favor of a more complicated construction of the recovery sequence. However, in the $p = 1$ case, convexity plays an important role in our compactness arguments (cf. Proposition 3.6 and Section 4.2.1, which rely on the existence of straight-line paths between nearby points in order to bound the number of vertices in an optimal path), and cannot be immediately removed without some a priori control over the distance from discrete optimal paths to the boundary of the domain.

2. *Ellipticity of ρ.* The bound on ρ is useful to compare ν to the uniform distribution in the nearest-neighbor map result, Proposition A.1, as well as in bounding the number of points in certain sets in Lemma 4.5. We note, as our approximating costs, L_n, G_n, H_n, do not involve density estimators, our results do not depend on the specifics of ρ, unlike for "density based distances" discussed in [28].

3. *Decay of ε_n* (2.6). Intuitively, the rate ε_n cannot vanish too quickly, as then the graph may be disconnected with respect to a positive set of realizations $\{X_i\}$. However, the estimate in (2.6) ensures that the graph $\mathcal{G}_n(a, b)$ is connected for all large n almost surely; see Proposition A.1. This is a version of the δ-sampling condition in [7], and is related to connectivity estimates in continuum percolation (cf. [26]). Moreover, we note, the prescribed rate yields in fact that any vertex X_i will have degree tending to infinity as n grows, as long as ρ is elliptic: One calculates that the mean number of points in the ε_n ball around X_i is of order $n\varepsilon_n^d$ which grows faster than $\log(n)$.

4. *Assumptions* (A0)–(A3) *on f.* These are somewhat standard assumptions to treat parametric variational integrals such as F (cf. [9] and [16]), which include the basic case $f(x, v) = |v|^p$.

5. *Assumption on p.* The assumption $p \geq 1$ is useful to show existence of quasinormal paths in Proposition A.2, and compactness of minimizers. The case $p < 1$ is more problematic in this sense and not discussed here.

6. *Extra assumptions in Theorems 2.5 and 2.6.* The main difficulty is in showing compactness of optimal H_n and L_n paths when $p = 1$. With respect to Theorem 2.4, when $p > 1$, the form of the cost allows a Hölder's inequality argument to deduce equicontinuity of the paths, from which compactness follows using Arzelà–Ascoli's theorem. However, there is no such coercivity when $p = 1$.

Yet, with the additional assumptions, one can approximate a geodesic locally by straight lines. Several geometric estimates on the number of points in small windows around these straight lines are needed to ensure accuracy of the approximation, for which the upper bound on ε_n in (2.12) is useful. In the absence of these assumptions, it is not clear whether one may control the oscillations of the approximating paths.

On the other hand, compactness with respect to the Hausdorff topology, viewing the geodesic as a set of points, might be explored, under less or a different set of assumptions. Perhaps somewhat related, we mention [12], and references therein, which considers Gromov–Hausdorff convergence of the space of probability measures on a random geometric graph in a torus endowed with a discrete Wasserstein metric to the space of probability measures on the torus endowed with the Wasserstein metric.

7. *Unique minimizers of F*. Given that our results achieve their strongest form when arg min F consists of a unique minimizing path, perhaps up to reparametrization, we comment on this possibility. Under suitable smoothness conditions on the integrand f, uniqueness criteria for ordinary differential equations allow to deduce from the Euler–Lagrange equations for F that there is a unique geodesic between points a, b sufficiently close together (cf. Proposition 5.25 in [9] and Chapter 5 in [8]). On the other hand, for general a, b, nonuniqueness may hold depending on the structure of f. For instance, one may construct an f, satisfying the standing assumptions, with several F-minimizing paths, by penalizing portions of D so as to induce forks.

8. *k-nearest neighbor graphs.* It is not clear if our approximation results, say Theorem 2.5, hold with respect to the k-nearest-neighbor graph with k bounded– that is the graph formed by attaching edges from a vertex to the nearest k points. For instance, when $\{X_i\}$ is arranged along a fine regular grid, $f(x, v) \equiv |v|$, $d = 2$, and $k = 4$, the optimal H_n route of moving from the origin to $(1, 1)$ is on a staircase path with length ~ 2, no matter how refined the grid is, yet the Euclidean distance is $\sqrt{2}$. In this respect, the random geometric graph setting of Theorem 2.5 allows enough choices among nearby points, as long as ρ is elliptic, for the optimal path to approximate the straight line from $(0, 0)$ to $(1, 1)$. It would be of interest to investigate the extent to which our results extend to k-nearest neighbor graphs.

9. *Computing optimal paths*. In the $p = 1$ case, minimization of the functionals L_n, G_n and H_n may be naturally posed as a problem of finding shortest paths in a weighted graph [undirected with symmetric weights, in the case of L_n and G_n, and directed with asymmetric weight $f(x, y - x)$ on the edge from x to y for H_n]. Under such a formulation, finding optimal paths is computationally tractable, even when the number of vertices n is large, for example, by means of Dijkstra's algorithm, which requires $O(m \log n)$ operations, where m is the number of weighted edges in the graph, or via the Euclidean heuristic algorithm of [29] which, when the continuum metric is available to prune the search space, requires $O(n)$ operations. We mention, under the assumption (2.6), m grows faster than $n \log(n)$.

10. *More general kernels and domains.* From an applications view, say with respect to the ISOMAP procedure, one presumes that the data lies in a submanifold with a definite, but unknown structure. To estimate distances in the submanifold, one would like to know when approximations via a random sample $\mathcal{X}_n$ are consistent, in the most general situation. Also, from a more theoretical perspective, it is natural to ask about the scope of the consistency of the approximations, both of the distances and the geodesic trajectories.

Although, the class of kernels f that we consider generalize those of [7], where $f(x, v) = |v|$, and [28], where $f(x, v) = h(\rho(x))|v|$ and h satisfies several conditions, in applications where the density ρ is not known in advance it would be useful to formulate density dependent versions, as done in [28].

Also, it would be of interest to consider consistency on a more general domain, such as a compact manifold. In this context, the domain D here represents a single patch of a manifold, and the results here on D may be seen as a basis for further advances.

2.4. *An example class of kernels.* We now describe a class of kernels f satisfying all the assumptions in the article. The following lemma is a case of the Hamel's criterion discussed previously.

LEMMA 2.7. *Given the standing assumptions, suppose also, for fixed $x \in D$, that $v \mapsto f(x, v)$ is C^2 on $\mathbb{R}^d \setminus \{0\}$ with positive definite Hessian. Then* (Hilb) *is satisfied.*

PROOF. Fix an $x_0 \in D$. There is a quasinormal minimizer $\gamma \in C^2$ where both $c = f(x_0, \dot{\gamma}(t)) = \inf_{\gamma \in \Omega(a,b)} \int_0^1 f(x_0, \dot{\gamma}(t))\, dt$ and

$$c^2 = f^2(x_0, \dot{\gamma}(t)) = \inf_{\gamma \in \Omega(a,b)} \int_0^1 f^2(x_0, \dot{\gamma}(t))\, dt$$

for a.e. $0 \le t \le 1$ (cf. Proposition 5.25 in [9]). Let $g(v) = f^2(x_0, v)$. Then γ satisfies the Euler–Lagrange equation $\frac{d}{dt}\nabla_v g(\dot{\gamma}(t)) = \nabla_x g(\dot{\gamma}(t)) = 0$. In other words, $H\ddot{\gamma}(t) = 0$, where H denotes the Hessian of g. By assumption, H is positive definite. Hence, $\ddot{\gamma}(t) \equiv 0$, and so γ is a parametrization of a straight line. $\square$

Recall $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product on $\mathbb{R}^d$.

LEMMA 2.8. *Let $x \mapsto M = M(x)$ be a C^1, strictly elliptic matrix-valued function on D. The kernel $f(x, v) = \langle v, M(x)v \rangle^{1/2}$ satisfies the standing assumptions, and also* (Lip), (Hilb), (TrIneq) *and* (Pythag) *for all $\alpha > 1$.*

PROOF. The kernel clearly satisfies the standing assumptions and (Lip). Next, for fixed x, the map $v \mapsto f(x, v)$ satisfies the conditions of Lemma 2.7, and so satisfies (Hilb). Also, $v \mapsto f(x, v)$ trivially satisfies (TrIneq).
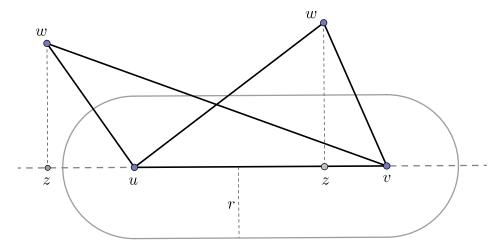
FIG. 3.   *Geometric argument used in proof of Lemma* 2.8 *with respect to* (Pythag).

We show (Pythag) in the case $f(x, v) = h(x)|v|$, that is $M(x) = h^2(x)$Id, as the notation is easier and all the ideas carry over to the more general case. Consider a right triangle joining u, w, z where z is on the line through u, v (cf. Figure 3).

If z is not on the line segment connecting u and v, then either (a) $|uw| \geq |uv|$ and $|wv| \geq r \geq r^\alpha$ or (b) $|wv| \geq |uv|$ and $|uw| \geq r \geq r^\alpha$. In either case, $h(x)[|uw| + |wv|] \geq h(x)|uv| + m_1 r^\alpha$ and (Pythag) is satisfied.

Suppose now z is on the line segment connecting u and v. In the triangle, uw is the hypotenuse, and wz and uz are the legs, such that $|uw|^2 = |wz|^2 + |uz|^2$. Hence, as $|uw| < 1$, all the lengths are less than 1. We are given that $|wz| \geq r$ and $|uz| \leq |uv| \leq cr^{1/\alpha}$. Then, as $r \leq |wz| < 1$ and $2 < \min\{2\alpha, \alpha + 1/\alpha\}$, we have

$$
\begin{aligned}
|uw|^2 &= |uz|^2 + |wz|^2 \\
&\geq |uz|^2 + (1/9)r^{2\alpha} + 2(1/3)r^{\alpha+1/\alpha} \\
&\geq |uz|^2 + (9\max\{c, 1\}^2)^{-1}r^{2\alpha} + 2(3\max\{c, 1\})^{-1}|uz|r^\alpha \\
&= (|uz| + (3\max\{c, 1\})^{-1}r^\alpha)^2,
\end{aligned}
$$

and so $|uw| \geq |uz| + (3\max\{c, 1\})^{-1}r^\alpha$.

A similar inequality, $|wv| \geq |vz| + (3\max\{c, 1\})^{-1}r^\alpha$, holds with the same argument. Hence, $h(x)|uw| + h(x)|wv| \geq h(x)|uv| + C(m_1, c)r^\alpha$.   □

3. Proofs of Theorems 2.1, 2.2 and Corollary 2.3.   As mentioned in the Introduction, the proof of Theorems 2.1 and 2.2 relies on a Gamma convergence argument. After establishing some basic notation and results on quasinormal minimizers, we present three main proof elements, liminf inequality, limsup inequality and compactness, in the following subsections. Proofs of Theorems 2.1 and 2.2, and Corollary 2.3 are the end of the section.

3.1. *Preliminaries.* Define a nearest-neighbor map $T_n : D \to \mathcal{X}_n$ where, for $x \in D$, $T_n(x)$ is the point of $\mathcal{X}_n$ closest to x with respect to the Euclidean distance. In the event of a tie, we adopt the convention that $T_n(x)$ is that nearest neighbor in $\mathcal{X}_n$ with the smallest subscript. Since $\mathcal{X}_n$ is random, we note T_n and the distortion

$$\|T_n - \mathrm{Id}\|_\infty = \sup_{x \in D} |T_n(x) - x| = \|T_n - \mathrm{Id}\|_\infty = \sup_{y \in D} \min_{1 \le i \le n} |X_i - y|$$

are also random. In Proposition A.1 of the Appendix, we show for $a, b \in D$ that, almost surely,

(3.1)   the graph $\mathcal{G}_n(a, b)$ is connected for all $a, b \in D$ and all large n.

Moreover, it is shown that exists a constant C such that almost surely,

$$(3.2) \qquad \limsup_{n \to \infty} \frac{\|T_n - \mathrm{Id}\|_\infty n^{1/d}}{(\log n)^{1/d}} \le C.$$

Throughout, we will be working with realizations where both (3.1) and (3.2) are satisfied. Let

$$A_{\mathrm{NN}} \text{ be the probability 1 event that (3.1) and (3.2) hold.}$$

We observe, when the decay rate (2.6) on ε_n holds, on the set of realizations A_{NN}, "NN" signifying nearest-neighbor, we have $\lim_{n \to \infty} \|T_n - \mathrm{Id}\|_\infty / \varepsilon_n = 0$.

To rule out certain degenerate configurations of points, in $d \ge 2$, let

$$A_Q \text{ be the event that } X_k \notin S_{\gamma_{X_i, X_j}} \ \forall \text{ distinct } i, j, k \in \mathbb{N}.$$

Since the $\{X_i\}$ come from a continuous distribution, and the image of the Lipschitz path γ_{X_i, X_j} in $D \subset \mathbb{R}^d$, when $d \ge 2$, is of lower dimension, A_Q has probability 1; here, "Q" signifies not on quasinormal.

Recall the definitions of quasinormal and linear paths $\gamma_{u,v}$ and $l_{u,v}$.

PROPOSITION 3.1. *Let m_1 and m_2 be the constants in (2.4). Then for $u, v \in D$,*

$$(3.3) \qquad m_1 |u - v|^p \le d_f(u, v) \le m_2 |u - v|^p.$$

Further, the path $\gamma_{u,v}$ satisfies, for $0 \le s, t \le 1$, that

$$(3.4) \qquad |\gamma_{u,v}(s) - \gamma_{u,v}(t)| \le (m_2/m_1)^{1/p} |u - v| |s - t|$$

and

$$(3.5) \qquad \sup_{0 \le t \le 1} |\gamma_{u,v}(t) - l_{u,v}(t)| \le ((m_2/m_1)^{1/p} + 1) |u - v|.$$

PROOF. For a Lipschitz path γ between u and v, we have $m_1|\dot{\gamma}(t)|^p \leq f(\gamma(t), \dot{\gamma}(t)) \leq m_2|\dot{\gamma}(t)|^p$ by (2.4). Also, $\inf_{\gamma \in \Omega(a,b)} \int_0^1 |\dot{\gamma}(t)|^p \, dt = |b - a|^p$ by a standard calculus of variations argument (see also Proposition A.2). So, by taking infimum over γ, we obtain (3.3).

Suppose now $\gamma = \gamma_{u,v}$ is quasinormal, so that $f(\gamma(t), \dot{\gamma}(t)) = c$ for some constant c and a.e. t. Integrating, and noting (3.3), gives $c = \int_0^1 f(\gamma, \dot{\gamma}) \, dt = d_f(u, v) \leq m_2|u - v|^p$. Also, by (2.4), $m_1|\dot{\gamma}(t)|^p \leq f(\gamma(t), \dot{\gamma}(t)) = c$. Hence, $|\dot{\gamma}(t)| \leq (m_2/m_1)^{1/p}|u - v|$ and (3.4) follows.

Finally, to establish (3.5), suppose that there is a t such that $|\gamma_{u,v}(t) - l_{u,v}(t)| > ((m_2/m_1)^{1/p} + 1)|u - v|$. Then, considering that $|l_{u,v}(t) - v| \leq |u - v|$, an application of the triangle inequality gives $|\gamma_{u,v}(t) - v| > (m_2/m_1)^{1/p}|u - v|$. However, by (3.4), $|\gamma_{u,v}(t) - v| = |\gamma_{u,v}(t) - \gamma_{u,v}(1)| \leq (m_2/m_1)^{1/p}|u - v||t - 1| \leq (m_2/m_1)^{1/p}|u - v|$, a contradiction. Thus, inequality (3.5) holds. □

3.2. *Liminf inequality.* A first step in getting some control over the limit cost F in terms of the discrete costs is the following bound.

LEMMA 3.2 (Liminf inequality). *Consider $\gamma \in \Omega(a, b)$, and suppose we have a sequence of paths $\gamma_n \in \Omega(a, b)$ such that*

$$\lim_{n \to \infty} \sup_{0 \leq t \leq 1} |\gamma_n(t) - \gamma(t)| = 0 \quad and \quad \sup_n \int_0^1 |\dot{\gamma}_n|^p \, dt < \infty.$$

Then, $F(\gamma) \leq \liminf_{n \to \infty} F(\gamma_n)$.

PROOF. A sufficient condition for this inequality, a lower semicontinuity property of F, to hold is that $f(x, v)$ be jointly continuous and convex in v. See Theorem 3.5 (and the subsequent Remark 2) of [9] for more discussion on this matter. □

3.3. *Limsup inequality.* To make effective use of the liminf inequality, we need to identify a sufficiently rich set of sequences for which a reverse inequality holds. To this end, we develop certain approximations of Lipschitz paths by piecewise linear or piecewise quasinormal paths.

The following result gives a method for recovering an element of $V_n(a, b)$ from a suitable element of $\Omega(a, b)$.

PROPOSITION 3.3. *Suppose, for constants c, C, that $\gamma \in \Omega(a, b)$ satisfies*

$$(3.6) \qquad\qquad c \leq \frac{|\gamma(s) - \gamma(t)|}{|s - t|} \leq C,$$

for all $0 \leq s < t \leq 1$. Let $N = N(n) = \lceil K/\varepsilon_n \rceil$, where $K = C + 1$ say, and define $\mathbf{v} = (v_0, \ldots, v_N)$ with $v_0 = a$, $v_N = b$, and $v_i = T_n \gamma(i/N)$ for $0 < i < N$.

Then, with respect to realizations $\{X_i\}$ in A_{NN}, we have $\mathbf{v} \in V_n(a, b)$ for all sufficiently large n.

PROOF.   To show that $\mathbf{v} \in V_n(a, b)$, it is sufficient to verify that consecutive vertices v_{i-1} and v_i are connected by an edge in $\mathcal{G}_n(a, b)$, or in other words

$$0 < |v_i - v_{i-1}| < \varepsilon_n \qquad \text{for } i = 1, \dots, N.$$

We first show that $|v_i - v_{i-1}| < \varepsilon_n$. Note that $|\gamma(i/N) - \gamma((i-1)/N)| \le C/N \le (C/K)\varepsilon_n$ and $C/K < 1$. For $1 < i < N$, we have

$$|v_i - v_{i-1}| = \left|T_n\gamma(i/N) - T_n\gamma((i-1)/N)\right|$$
$$\le (C/K)\varepsilon_n + 2\|T_n - \mathrm{Id}\|_\infty.$$

Similarly, for segments incident to an endpoint a or b, we have

$$\max\left(|v_1 - v_0|, |v_N - v_{N-1}|\right) \le (C/K)\varepsilon_n + \|T_n - \mathrm{Id}\|_\infty.$$

In either case, assumption (2.6) on the decay of ε_n implies that, for realizations $\{X_i\}$ in A_{NN}, we have $|v_i - v_{i-1}| < \varepsilon_n$ for all $1 \le i \le N$ and sufficiently large n.

Now, we show that $0 < |v_i - v_{i-1}|$. By the Lipschitz lower bound on γ, we have

$$\left|\gamma(i/N) - \gamma((i-1)/N)\right| \ge c/N > \left(c/(K+1)\right)\varepsilon_n$$

for $1 \le i \le N$. By a triangle inequality argument, the distance between v_i and v_{i-1} is bounded below by $(c/(K+1))\varepsilon_n - 2\|T_n - \mathrm{Id}\|_\infty$, which on the set A_{NN}, as ε_n satisfies (2.6) and, therefore, vanishes slower than $\|T_n - \mathrm{Id}\|_\infty$, is positive for all large n. $\square$

We now establish some approximation properties obtained by interpolating paths between points in $\mathbf{v} = (v_0, \dots, v_N)$.

PROPOSITION 3.4.   *Fix* $\gamma \in \Omega(a, b)$ *satisfying* (3.6), *and a realization* $\{X_i\}$ *in* A_{NN}. *Let* $\gamma_n = \gamma_\mathbf{v}$ *and* $l_n = l_\mathbf{v}$, *where* $N = N(n)$ *and* $\mathbf{v} = (v_0, \dots, v_N)$ *are defined as in Proposition* 3.3. *Then we obtain*

$$\lim_{n\to\infty} \sup_{0\le t\le 1} \left|\gamma_n(t) - \gamma(t)\right| = 0 \quad \text{and} \quad \lim_{n\to\infty} \sup_{0\le t\le 1} \left|l_n(t) - \gamma(t)\right| = 0.$$

In addition,

$$(3.7) \qquad \sup_n \|l_n'\|_\infty < \infty \quad \text{and} \quad l_n'(t) \to \gamma'(t) \qquad \text{for a.e. } t \in [0, 1].$$

PROOF.   We first argue that $\lim_{n\to\infty} \sup_{0\le t\le 1} |l_n(t) - \gamma(t)| = 0$. Let $u_i = \gamma(i/N)$, and let $\tilde{l}_n = l_{\mathbf{u}^{(n)}} \in \Omega(a, b)$ be the piecewise linear interpolation of $\mathbf{u}^{(n)} = (u_0, \dots, u_N)$. As γ is Lipschitz and $\lim_{n\to\infty} N(n) = \infty$, we have $\lim_{n\to\infty} \sup_{0\le t\le 1} |\tilde{l}_n(t) - \gamma(t)| = 0$, and also $\lim_{n\to\infty} \tilde{l}_n'(t) = \gamma'(t)$ for a.e. $t \in [0, 1]$. By construction, $l_n(i/N) = v_i = T_n\gamma(i/N)$ and $\tilde{l}_n(i/n) = u_i = \gamma(i/N)$ so that

$$\max_{0\le i\le N} \left|l_n(i/N) - \tilde{l}_n(i/N)\right| \le \|\mathrm{Id} - T_n\|_\infty.$$

Then, as l_n and $\tilde{l}_n$ are piecewise linear, it follows that $\sup_{0 \le t \le 1} |l_n(t) - \tilde{l}_n(t)| \le \|\mathrm{Id} - T_n\|_\infty$. As $\lim_{n \to \infty} \|\mathrm{Id} - T_n\|_\infty = 0$ on A_{NN}, we obtain $\lim_{n \to \infty} \sup_{0 \le t \le 1} |l_n(t) - \gamma(t)| = 0$.

For $i/N < t < (i+1)/N$, we have

$$l_n'(t) = N(v_{i+1} - v_i).$$

As $|v_{i+1} - v_i| \le \varepsilon_n$ (Proposition 3.3), it follows that $|l_n'(t)| \le N\varepsilon_n \le K + \varepsilon_n$. Hence, $\sup_n \|l_n'\|_\infty < \infty$.

Likewise, $\tilde{l}_n'(t) = N(u_{i+1} - u_i)$, and so

$$\left| l_n'(t) - \tilde{l}_n'(t) \right| \le N\big(|v_{i+1} - u_{i+1}| + |v_i - u_i|\big) \le 2N\|T_n - \mathrm{Id}\|_\infty.$$

For realizations in A_{NN}, since $N = \lceil K/\varepsilon_n \rceil$ and ε_n satisfies (2.6) and, therefore, vanishes slower than $\|T_n - \mathrm{Id}\|_\infty$, we have $\lim_{n \to \infty} N\|T_n - \mathrm{Id}\|_\infty = 0$. Hence, $l_n'(t) \to \gamma'(t)$ for a.e. $t \in [0, 1]$.

Now, considering the bound (3.5), it follows that

$$\sup_{0 \le t \le 1} |\gamma_n(t) - l_n(t)| \le \max\big(C|v_0 - v_1|, \ldots, C|v_{N-1} - v_N|\big) \le C\varepsilon_n,$$

and hence $\|\gamma_n - \gamma\|_\infty \to 0$. $\quad\square$

With the above work in place, we proceed to the main result of this subsection.

LEMMA 3.5 (Limsup inequality). *Suppose $\gamma \in \Omega(a, b)$ satisfies inequality (3.6). Then, with respect to realizations $\{X_i\}$ in A_{NN}, we may find a sequence of paths $\{\gamma_n\}$ taken either in form for all large n as (1) $\gamma_n \in \Omega_n^l(a, b)$ or (2) $\gamma_n \in \Omega_n^\gamma(a, b)$ such that $\lim_{n \to \infty} \sup_{0 \le t \le 1} |\gamma_n(t) - \gamma(t)| = 0$ and*

$$(3.8) \qquad\qquad F(\gamma) \ge \limsup_{n \to \infty} F(\gamma_n).$$

We remark that the sequence $\{\gamma_n\}$ in the last lemma is called the recovery sequence since the liminf inequality in Lemma 3.2 and the limsup inequality in Lemma 3.5 together imply the limit, $\lim_n F(\gamma_n) = F(\gamma)$.

PROOF OF LEMMA 3.5. Let $N = \lceil K/\varepsilon_n \rceil$, where $K = C + 1$ say is a constant greater than C in (3.6). Define $v_0 = a$, $v_N = b$, and $v_i = T_n \gamma(i/N)$ for $0 < i < N$. Then, by Proposition 3.3, $\mathbf{v} = \mathbf{v}^{(n)} = (v_0, \ldots, v_N) \in V_n(a, b)$.

We now consider paths in case (1). By Proposition 3.4, the interpolated paths $l_n = l_\mathbf{v} \in \Omega_n^l(a, b)$ converge uniformly to γ. Consider the bound

$$\left| F(l_n) - F(\gamma) \right| \le \int_0^1 \left| f\big(l_n(t), l_n'(t)\big) - f\big(\gamma(t), \gamma'(t)\big) \right| dt.$$

By Proposition 3.4, l_n' converges almost everywhere to γ', and $\sup_n \|l_n'\|_\infty < \infty$. Hence $(l_n(t), l_n'(t)) \to (\gamma(t), \gamma'(t))$ for almost every t. Also, $\|\dot\gamma\|_\infty < C$ by (3.6).

Since, by (2.4), $f(x, v) \leq m_2 |v|^p$, an application of the bounded convergence theorem yields $\lim_{n \to \infty} |F(l_n) - F(\gamma)| = 0$. Here, $\{l_n\}$ is the desired recovery sequence.

We now consider case (2). Let $\gamma_n = \gamma_{\mathbf{v}} \in \Omega^\gamma(a, b)$. By Proposition 3.4, it follows that $\lim_{n \to \infty} \sup_{0 \leq t \leq 1} |\gamma_n(t) - \gamma(t)| = 0$. To show (3.8) for this sequence, note that $F(\gamma_n) \leq F(l_n)$ as l_n is a possibly more expensive path connecting the vertices in $\mathbf{v}$. Then, by case (1), $\limsup F(\gamma_n) \leq \limsup F(l_n) \leq F(\gamma)$. $\quad\square$

3.4. *Compactness.* In this subsection, we consider circumstances under which a sequence of paths $\{\gamma_n\}$, in the context of Theorems 2.1 and 2.2, has a limit point with respect to uniform convergence. Here, the arguments when $p = 1$ differ from those when $p > 1$.

In particular, consider paths γ_n where $\int_0^1 f(\gamma_n(t), \dot\gamma_n(t))\, dt$ is uniformly bounded. One has $m_1 |v|^p \leq f(x, v)$ and it follows that $\{\gamma_n\}$ is bounded in the $W^{1,p}$ Sobolev space. When $p > 1$, this is sufficient to derive a suitable compactness result. But, when $p = 1$, this is no longer the case.

However, when $p = 1$, our general outlook is that it is enough to establish a compactness result for *sequences of optimal paths*, on which certain eccentric possibilities are ruled out.

We begin by considering such compactness when $p = 1$, when the paths lie in $\Omega_n^\gamma(a, b)$. The setting $p > 1$ is discussed afterwards.

PROPOSITION 3.6. *Suppose $d \geq 2$ and that $p = 1$. Then, with respect to realizations $\{X_i\}$ in $A_{\mathrm{NN}} \cap A_Q$, for all large n, if $\gamma \in \arg\min G_n$ and $0 \leq s, t \leq 1$, we have that*

$$|\gamma(s) - \gamma(t)| \leq (4m_2/m_1^2) G_n(\gamma)|s - t|.$$

PROOF. The path $\gamma \in \Omega_n^\gamma(a, b)$ is a piecewise quasinormal path of the form $\gamma = \gamma_{\mathbf{v}}$ where $\mathbf{v} = (v_0, v_1, \ldots, v_m) \in V_n(a, b)$. We now try to relate m, the number of segments in the path, to $G_n(\gamma)$, the path energy. Recall the formula (2.9).

Let B_i denote the (open) Euclidean ball of radius $\varepsilon_n/2$ around v_i. We claim that $|B_i \cap \{v_0, \ldots, v_m\}| \leq 2$. To see this, suppose that there are at least 3 points of $\{v_0, \ldots, v_m\}$ in B_i. Let v_j and v_l denote the points in B_i with the smallest and largest index, respectively. By minimality of γ, $v_j \neq v_l$. Let v_k denote a third point in B_i.

As $v_k, v_l \in B_i$, we have $|v_k - v_l| < \varepsilon_n$, and so these points are connected in the graph. Applying the triangle inequality for d_f, valid when $p = 1$ [cf. (2.5)], and noting on the event A_Q that $v_k \notin S_{\gamma_{v_j}, v_l}$, we have

$$d_f(v_j, v_l) < d_f(v_j, v_k) + d_f(v_k, v_l) \leq d_f(v_j, v_{j+1}) + \cdots + d_f(v_{l-1}, v_l).$$

Thus, the path $\tilde\gamma = \gamma_{\mathbf{w}}$, where $\mathbf{w} = (v_0, \ldots, v_j, v_l, \ldots, v_m)$, satisfies $G_n(\tilde\gamma) < G_n(\gamma)$. This contradicts the optimality of γ, and so $|B_i \cap \{v_0, \ldots, v_m\}| \leq 2$.

We may thus cover the vertices of γ with balls $\{B_i\}_{i=1}^m$, centered on the vertices $\{v_i\}_{i=1}^m$, and each of these balls contains at most two vertices. It follows that there is a subcover by $s \geq m/2$ balls, $\{B_1', \ldots, B_s'\}$, no two of them containing a common point in $\mathbf{v}$.

A lower bound for $G_n(\gamma)$ is found by considering that part of the G_n-integral contributed to by the portion of the path γ in B_i'. Each such portion, if it does not terminate in B_i', must visit both the center of B_i' and the boundary $\partial B_i'$, and hence has Euclidean length at least $\varepsilon_n/2$. Summing over these portions, we obtain

$$\frac{m\varepsilon_n}{4} \leq \sum_{i=1}^{s} \frac{\varepsilon_n}{2} \leq L,$$

where $L = \int_0^1 |\dot{\gamma}(t)|\, dt$ is the Euclidean arc length of γ.

By (2.4), it follows that

$$(3.9) \qquad m_1 \frac{m\varepsilon_n}{4} \leq m_1 L \leq \int_0^t f(\gamma(t), \dot{\gamma}(t))\, dt = G_n(\gamma).$$

To get a Lipschitz bound for γ, recall the bound (3.4). Then $\gamma_{v_{i-1}, v_i}(m \cdot -i)$ is Lipschitz with constant $(m_2/m_1)m|v_{i-1} - v_i|$. It follows that γ, being the concatenation of these segments, satisfies

$$|\gamma(s) - \gamma(t)| = \left| \sum_{r=0}^{q-1} \gamma(w_r) - \gamma(w_{r+1}) \right|$$

$$\leq (m_2/m_1)m \max(|v_0 - v_1|, \ldots, |v_{m-1} - v_m|) \sum_{r=0}^{q-1} |w_r - w_{r+1}|$$

$$\leq (m_2/m_1)m\varepsilon_n |s - t|,$$

where $s = w_0 < \cdots < w_q = t$, $\{w_r\}_{r=1}^{q-1} \subset \{j/m\}_{j=1}^{m-1}$ so that $|w_r - w_{r+1}| \leq 1/m$ for $0 \leq r \leq q - 1$. Then, with (3.9), we obtain

$$|\gamma(t) - \gamma(s)| \leq (m_2/m_1)m\varepsilon_n |s - t| \leq (4/m_1)(m_2/m_1)G_n(\gamma)|s - t|,$$

completing the proof. □

We now prove our compactness property.

LEMMA 3.7 (Compactness Property).   (I). *Suppose for all large n that either $\gamma_n \in \arg\min L_n$ or $\gamma_n \in \arg\min G_n$. Then, for realizations $\{X_i\}$ in A_{NN}, we have $\sup_n F(\gamma_n) < \infty$.*
(II). *Consider now the following cases:*

  (i) *Suppose paths $\gamma_n \in \arg\min G_n$ for all large n.*
  (ii) *Suppose $p > 1$ and $\gamma_n \in \Omega(a, b)$ for all large n such that $\sup_n F(\gamma_n) < \infty$.*

Then, in case (i) *when* $p > 1$, *and in case* (ii), *with respect to realizations* $\{X_i\}$ *in* A_{NN}, *we have* $\{\gamma_n\}$ *is relatively compact for the topology of uniform convergence. For case* (i) *when* $d \geq 2$ *and* $p = 1$, *the same conclusion holds with respect to realizations* $\{X_i\}$ *in* $A_{NN} \cap A_Q$.

PROOF. We first prove the bound $\sup_n F(\gamma_n) < \infty$ in part (I). Choose a $\tilde{\gamma} \in \Omega(a, b)$, where (3.6) holds, and $F(\tilde{\gamma}) < \infty$. By Lemma 3.5, there is a sequence $\{\tilde{\gamma}_n\}$ of either piecewise linear or quasinormal paths such that $\limsup_{n \to \infty} F(\tilde{\gamma}_n) \leq F(\tilde{\gamma})$. Hence, by minimality of $\{\gamma_n\}$, with respect to paths in either Ω_n^l or Ω_n^γ, we have

$$(3.10) \qquad \sup_n F(\gamma_n) \leq \sup_n F(\tilde{\gamma}_n) < \infty.$$

We now argue the claims for cases (i) and (ii). In both cases, as D is bounded, the paths $\gamma_n : [0, 1] \to D$ are uniformly bounded. To invoke the Arzelà–Ascoli theorem, we must show that $\{\gamma_n\}$ is an equicontinuous family.

In case (i), when $d \geq 2$ and $p = 1$, by Lemma 3.6 on realizations in $A_{NN} \cap A_Q$, we have $|\gamma_n(s) - \gamma_n(t)| \leq C G_n(\gamma_n)|s - t|$, with C independent of n. As $G_n(\gamma_n) = F(\gamma_n)$, combining with (3.10), it follows that $\{\gamma_n\}$ is equicontinuous.

If $p > 1$, with respect to realizations in A_{NN}, (3.10) implies that, if case (i) holds for the sequence, then case (ii) holds.

Without loss of generality, then we focus our attention now on case (ii). Recall, by (2.4), that $m_1|v|^p \leq f(x, v)$. Let q be the conjugate of p, that is, $1/p + 1/q = 1$. Then, for $0 \leq s < t \leq 1$,

$$|\gamma_n(s) - \gamma_n(t)| \leq \int_s^t |\gamma_n'(r)| \, dr \leq |t - s|^{1/q} \left(\int_0^1 |\gamma_n'(t)|^p \, dt \right)^{1/p}$$

$$(3.11) \qquad \leq \left(|t - s|^{1/q} / m_1^{1/p} \right) \left(\int_0^1 f(\gamma_n(t), \gamma_n'(t)) \, dt \right)^{1/p}$$

$$= \left(|t - s|^{1/q} / m_1^{1/p} \right) F(\gamma_n)^{1/p}.$$

Combining (3.11) and the assumption in case (ii) that $\sup_n F(\gamma_n) < \infty$, we have $|\gamma_n(s) - \gamma_n(t)| \leq C|t - s|^{1/q}$ for a constant C independent of n, and hence $\{\gamma_n\}$ is equicontinuous. $\square$

3.5. *Proofs of Theorems* 2.1 *and* 2.2. With the preceding Gamma convergence ingredients in place, the proofs of Theorem 2.1 and 2.2 are similar, and will be given together.

PROOF. Fix a realization $\{X_i\}$ in the probability 1 set A_{NN}. Let $\{\gamma_n\}$ be a sequence of paths such that, for large n, we have either $\gamma_n \in \arg\min L_n$ or $\gamma_n \in \arg\min G_n$. Supposing that $\{\gamma_n\}$ has a subsequential limit $\lim_{k \to \infty} \gamma_{n_k} = \gamma$, with respect to the topology of uniform convergence, we now argue that $\gamma \in \arg\min F$.

By the liminf Lemma 3.2, we have

$$(3.12) \qquad\qquad F(\gamma) \le \liminf_{k\to\infty} F(\gamma_{n_k}).$$

Let $\gamma^* \in \arg\min F$ where, by inequality (A.4) of Proposition A.2, there exist constants c_1, c_2 such that $c_1|s - t| \le |\gamma^*(s) - \gamma^*(t)| \le c_2|s - t|$ for $0 \le s, t \le 1$. Hence, by the limsup Lemma 3.5, there exists a sequence $\{\gamma_n^*\}$, of either piecewise linear or quasinormal paths, converging uniformly to γ^* and $\limsup_{n\to\infty} F(\gamma_n^*) \le F(\gamma^*)$. Recall that $F(\gamma) = L_n(\gamma)$ or $F(\gamma) = G_n(\gamma)$ when γ is piecewise linear or quasinormal, respectively. Combining with (3.12) and minimality of γ_{n_k}, we have

$$(3.13) \qquad F(\gamma) \le \liminf_{k\to\infty} F(\gamma_{n_k}) \le \limsup_{k\to\infty} F(\gamma_{n_k}^*) \le F(\gamma^*) = \min F,$$

and so $\gamma \in \arg\min F$.

In the case the paths $\{\gamma_n\}$ are piecewise linear, since $F(\gamma_n) = \min L_n$, it follows from (3.13) that

$$\lim_{k\to\infty} \min L_{n_k} = \min F.$$

Similarly, when $\{\gamma_n\}$ are piecewise quasinormal, $\lim_{k\to\infty} \min G_{n_k} = \min F$.

Therefore, we have shown that, if a subsequential limit of $\{\gamma_n\}$ exists, it is an optimal continuum path $\gamma \in \arg\min F$.

Consider now Theorem 2.1, where $p > 1$ and $\gamma_n \in \arg\min L_n$, and part (1) of Theorem 2.2 where $p > 1$ and $\gamma_n \in \arg\min G_n$. By the compactness Lemma 3.7, $\sup_n F(\gamma_n) < \infty$ and a subsequential limit exists.

Consider now part (2) of Theorem 2.2 where $d \ge 2$, $p = 1$ and $\gamma_n \in \arg\min G_n$. Suppose that the realization $\{X_i\}$ belongs also to the probability 1 set A_Q. Then subsequential limits follow again from the compactness Lemma 3.7.

Now, consider any subsequence $\{n_k\}$ of $\mathbb{N}$. Then, by the work above, applied to the sequence $\{n_k\}$, there is a further subsequence $\{n_{k_j}\}$, and a $\gamma \in \arg\min F$, with $\gamma_{n_{k_j}} \to \gamma$ uniformly, in the settings of Theorems 2.1 and 2.2. Moreover, $\lim_{j\to\infty} \min L_{n_{k_j}} = \min F$ when the paths $\{\gamma_{n_{k_j}}\}$ are piecewise linear, and $\lim_{j\to\infty} \min G_{n_{k_j}} = \min F$ when the paths are piecewise quasinormal.

Since this argument is valid for any subsequence $\{n_k\}$ of $\mathbb{N}$, we recover that $\min L_n \to \min F$ or $\min G_n \to \min F$ when respectively the paths are piecewise linear or quasinormal. Finally, if F has a unique minimizer γ, by considering subsequences again, the whole sequence $\{\gamma_n\}$ must converges uniformly to γ. $\quad\square$

3.6. *Proof of Corollary* 2.3. Corollary 2.3 is a statement about Hausdorff convergence. In order to adapt the results of Theorems 2.1 and 2.2 to this end, we make the following observation.

PROPOSITION 3.8. *Fix a realization* $\{X_i\}$ *in* A_{NN}, *and consider a sequence of paths* $\{\gamma_n\}$ *such that* γ_n *for all large* n *is either in the form* (1) $\gamma_n = \gamma_{\mathbf{v}^{(n)}}$ *or* (2)

$\gamma_n = l_{\mathbf{v}^{(n)}}$, where $\mathbf{v}^{(n)} \in V_n(a, b)$. Suppose that γ_n converges uniformly to a limit $\gamma \in \Omega(a, b)$. Then

$$\lim_{n \to \infty} d_{\mathrm{haus}}(\mathbf{v}^{(n)}, S_\gamma) = 0.$$

PROOF.   Write $\mathbf{v}^{(n)} = (v_0^{(n)}, \ldots, v_{m(n)}^{(n)})$. Since $\gamma_n \to \gamma$ uniformly and $v_i^{(n)} = \gamma_n(\frac{i}{m(n)})$, it follows that

(3.14) $$\lim_{n \to \infty} \max_{1 \le i \le m(n)} \inf_{x \in S_\gamma} |v_i^{(n)} - x| = 0.$$

On the other hand, consider t with $i/m(n) \le t < (i + 1)/m(n)$. In case (1),

$$|\gamma(t) - v_i^{(n)}| \le |\gamma(t) - \gamma(i/m(n))| + \|\gamma_n - \gamma\|_\infty \le C/m(n) + \|\gamma_n - \gamma\|_\infty,$$

where C is the Lipschitz constant of γ. In case (2), using linearity of the path, $|\gamma(t) - v_i^{(n)}| \le |v_{i+1}^{(n)} - v_i^{(n)}| + \|\gamma_n - \gamma\|_\infty \le \varepsilon_n + \|\gamma_n - \gamma\|_\infty$.

Since $\mathbf{v}^{(n)}$ is a path in $V_n(a, b)$, one may bound $\varepsilon_n m(n) \ge \sum_{i=0}^{m(n)} |v_i^{(n)} - v_{i+1}^{(n)}| \ge |a - b|$, and so $m(n) \ge |a - b|/\varepsilon_n$ diverges. Hence, in both cases,

(3.15) $$\lim_{n \to \infty} \sup_{x \in S_\gamma} \min_{1 \le i \le m(n)} |x - v_i^{(n)}| = 0.$$

Combining (3.14) and (3.15), it follows that $\lim_{n \to \infty} d_{\mathrm{haus}}(\mathbf{v}^{(n)}, S_\gamma) = 0$.   □

We now proceed to prove Corollary 2.3.

PROOF OF COROLLARY 2.3.   We give the argument for the case of piecewise linear optimizers, as the the argument is exactly the same for piecewise quasinormal paths, using Theorem 2.2 instead of Theorem 2.1 below.

Suppose $l_n = l_{\mathbf{v}^{(n)}} \in \arg\min L_n$ is a sequence of paths where $\mathbf{v}^{(n)} \in V_n(a, b)$. By Theorem 2.1, with respect to a probability 1 set of realizations $\{X_i\}$, any subsequence of $\{l_n\}$ has a further subsequence $\{l_{n_k}\}$ which converges uniformly to a $\gamma \in \arg\min F$. By Proposition 3.8, it follows that $\lim_{k \to \infty} d_{\mathrm{haus}}(\mathbf{v}^{(n_k)}, S_\gamma) = 0$.

Suppose now that F has a unique (up to reparametrization) minimizer γ. Note that S_γ is invariant under reparametrization of γ. Then we conclude that all limit points of $\{v^{(n)}\}$ correspond to S_γ, and hence the whole sequence $v^{(n)}$ converges to S_γ, $\lim_{n \to \infty} d_{\mathrm{haus}}(v^{(n)}, S_\gamma) = 0$.   □

4. Proofs of Theorems 2.4, 2.5 and 2.6.   The proofs of Theorems 2.4, 2.5 and 2.6 all make use of Theorems 2.1 and 2.2 in comparing the costs H_n and L_n to G_n. When $p = 1$, as with respect to Theorem 2.2, the arguments in Theorems 2.5 and 2.6 are more involved, especially with respect to the minimal cost H_n-path convergence, where several geometric estimates are used to show a compactness principle.

We begin with the following useful fact.

PROPOSITION 4.1. *Suppose* $U, W, C : X \to \mathbb{R}$ *are functions such that* $|U(x) - W(x)| \leq C(x)$ *for all* $x \in X$. *If* $U(x_U) = \min U$ *and* $W(x_W) = \min W$ *then*

$$-C(x_U) \leq \min U - \min W \leq C(x_W).$$

PROOF. For any $y \in X$, we have $\min U = U(x_U) \leq U(y) \leq W(y) + C(y)$. Taking $y = x_W$ gives $\min U \leq \min W + C(x_W)$. The other inequality follows similarly. □

4.1. *Proof of Theorem* 2.4. Suppose $\mathbf{v} = (v_0, v_1, \ldots, v_m) \in V_n(a, b)$. Then, for $0 \leq i \leq m - 1$, (Lip) and the observation that neighbors v_i, v_{i+1} in the ε_n-graph satisfy $|v_{i+1} - v_i| \leq \varepsilon_n$, implies

$$\left| \int_{\frac{i}{m}}^{\frac{i+1}{m}} f\left(v_i + m\left(s - \frac{i}{m}\right)(v_{i+1} - v_i), m(v_{i+1} - v_i)\right) ds \right.$$

$$\left. - \frac{1}{m} f\left(v_i, m(v_{i+1} - v_i)\right) \right|$$

$$\leq \frac{c}{m} |v_{i+1} - v_i| (m|v_{i+1} - v_i|)^p \leq \frac{c\varepsilon_n}{m} |v_{i+1} - v_i|^p m^p.$$

Recall formulas (2.8) and (2.11). Summing over i gives the following estimate relating L_n and H_n:

$$(4.1) \qquad \left| L_n(l_{\mathbf{v}}) - H_n(\mathbf{v}) \right| \leq \frac{c\varepsilon_n}{m} \sum_{i=0}^{m-1} |v_{i+1} - v_i|^p m^p.$$

Applying (2.4), the right-hand side of (4.1) can be bounded above in terms of both $L_n(l_{\mathbf{v}})$ and $H_n(\mathbf{v})$. Hence, with $c' = cm_1^{-1}$,

$$(4.2) \qquad \left| L_n(l_{\mathbf{v}}) - H_n(\mathbf{v}) \right| \leq c' \varepsilon_n \min\left(L_n(l_{\mathbf{v}}), H_n(\mathbf{v})\right).$$

Suppose $l_{\mathbf{v}^{(n)}} \in \arg\min L_n$ and $\mathbf{w}^{(n)} \in \arg\min H_n$. An immediate consequence of (4.2) and Proposition 4.1 is

$$-c' \varepsilon_n \min L_n \leq -c' \varepsilon_n \min\left(L_n(l_{\mathbf{v}^{(n)}}), H_n(\mathbf{v}^{(n)})\right)$$

$$\leq \min L_n - \min H_n$$

$$\leq c' \varepsilon_n \min\left(L_n(l_{\mathbf{w}^{(n)}}), H_n(\mathbf{w}^{(n)})\right) \leq c' \varepsilon_n \min H_n.$$

By Theorem 2.1, we have $\lim_{n \to \infty} \min L_n = \min F$ for almost all realizations $\{X_i\}$ (those in A_{NN} as the proof shows). Then $\min H_n \leq (1 + c'\varepsilon_n) \min L_n$ and so $\limsup \min H_n \leq \min F$ a.s. In particular, as $\min F < \infty$, we have $\sup_n \min H_n < \infty$ a.s.

On the other hand, $\min H_n \geq \min L_n - c'\varepsilon_n \min H_n$ a.s. As $\sup_n \min H_n < \infty$, we observe that $\liminf \min H_n \geq \min F$ a.s. Hence, $\min H_n \to \min F$ a.s. This finishes one part of Theorem 2.4.

To address the others, consider $l_{\mathbf{w}^{(n)}}$, the piecewise linear interpolation of $\mathbf{w}^{(n)}$. By (4.2), we have

$$L_n(l_{\mathbf{w}^{(n)}}) \leq (1 + c'\varepsilon_n)H_n(\mathbf{w}^{(n)}) = (1 + c'\varepsilon_n)\min H_n.$$

Moreover, noting the optimality of $l_{\mathbf{v}^{(n)}}$ and $\mathbf{w}^{(n)}$ gives

(4.3) $$L_n(l_{\mathbf{v}^{(n)}}) \leq L_n(l_{\mathbf{w}^{(n)}}) \leq (c'\varepsilon_n + 1)H_n(\mathbf{v}^{(n)}).$$

Another application of (4.2) yields

$$H_n(\mathbf{v}^{(n)}) \leq c'\varepsilon_n L_n(l_{\mathbf{v}^{(n)}}) + L_n(l_{\mathbf{v}^{(n)}}).$$

Hence, the left-hand side of (4.3) is bounded as

(4.4)
$$\min L_n = L_n(l_{\mathbf{v}^{(n)}}) \leq L_n(l_{\mathbf{w}^{(n)}}) \leq (c'\varepsilon_n + 1)^2 L_n(\mathbf{v}^{(n)})$$
$$= (c'\varepsilon_n + 1)^2 \min L_n.$$

Hence, as $\min L_n \to \min F$ a.s., we have

(4.5) $$\min F = \lim_{n\to\infty} L_n(l_{\mathbf{v}^{(n)}}) = \lim_{n\to\infty} L_n(l_{\mathbf{w}^{(n)}}) \qquad \text{a.s.}$$

We also observe, as a consequence, that $\sup_n L_n(l_{\mathbf{w}^{(n)}}) < \infty$ a.s.

Given that $p > 1$, by the compactness Lemma 3.7, with respect to realizations $\{X_i\}$ in the probability 1 set A_{NN}, any subsequence of $\{l_{\mathbf{w}^{(n)}}\}$ has a further uniformly convergent subsequence to a limit $\tilde{\gamma} \in \Omega(a, b)$. By the liminf Lemma 3.2, $F(\tilde{\gamma}) \leq \liminf_{n\to\infty} L_n(l_{\mathbf{w}^{(n)}})$ a.s. Finally, by (4.5), it follows that $F(\tilde{\gamma}) = \min F$ and so $\tilde{\gamma} \in \arg\min F$. Consequently, if F has a unique minimizer γ, then the whole sequence $\{l_{\mathbf{w}^{(n)}}\}$ converges uniformly almost surely to it.

The proofs of statements about Hausdorff convergence follow the same arguments as given for Corollary 2.3, and are omitted.

4.2. *Proof of Theorems 2.5 and 2.6.* We prove Theorems 2.5 and 2.6 in two parts.

PROOF. First, we prove in Proposition 4.3 that the minimal costs of H_n and L_n converge to $\min F$, making use of comparisons with quasinormal paths, for which we have control in Theorem 2.2.

Second, in Proposition 4.9 in Section 4.2.2, we show that the minimizing paths converge in the various senses desired. A main tool in this proof is a compactness property (Proposition 4.6), for minimal H_n and L_n-paths when $p = 1$, shown in Section 4.2.1. □

To supply the proofs of the desired propositions, we now obtain an useful estimate between the cost of a quasinormal path and a linear one.

PROPOSITION 4.2.　*Suppose $d \geq 2$, $p = 1$, and that f also satisfies* (Lip) *and* (Hilb). *For $a, b \in D$ such that $|b - a| \leq 1$, there is a constant c_1 such that*

$$\big|d_f(a, b) - f(a, b - a)\big| \leq c_1 |b - a|^2.$$

In particular, as $d_f(a, b) = \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt$ for the quasinormal path $\gamma = \gamma_{a,b}$ connecting a and b, we have

$$\left| \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt - f(a, b - a) \right| \leq c_1 |b - a|^2.$$

PROOF.　By (2.4), for a Lipschitz path β from a to b, we have

$$m_1 \int_0^1 |\dot{\beta}(t)| \, dt \leq \int_0^1 f(\beta(t), \dot{\beta}(t)) \, dt \leq m_2 \int_0^1 |\dot{\beta}(t)| \, dt.$$

Optimizing over β, noting that the infima of the left, right integrals is the arclength $|b - a|$ and the infima of the center one is in terms of the quasinormal path, we recover that $m_1 |b - a| \leq \int_0^1 f(\gamma(t), \dot{\gamma}(t)) \, dt \leq m_2 |b - a|$. By (2.4) again, we have that the arc length of γ satisfies $\int_0^1 |\dot{\gamma}(t)| \, dt \leq (m_2/m_1)|b - a|$. In particular, the path γ is constrained in the Euclidean ball B around a of radius $(m_2/m_1)|b - a|$. Note also that the minimizing Euclidean path $\widetilde{\gamma}$, with constant speed $|b - a|$ on the straight line from a to b in times $0 \leq t \leq 1$, is also constrained in this ball.

Now, for a Lipschitz path β, constrained in the ball B, expand

$$\int_0^1 f(\beta(t), \dot{\beta}(t)) \, dt = \int_0^1 f(a, \dot{\beta}(t)) \, dt + \int_0^1 (f(\beta(t), \dot{\beta}(t)) - f(a, \dot{\beta}(t)) \, dt.$$

As the paths are in B, by (Lip), with respect to a Lipschitz constant C,

$$\big| f(\beta(t), \dot{\beta}(t)) - f(a, \dot{\beta}(t)) \big| \leq C |\beta(t) - a| |\dot{\beta}(t)|$$

$$\leq C(m_2/m_1)|b - a||\dot{\beta}(t)|.$$

Therefore, with respect to Lipschitz paths β constrained in B,

$$\left| \int_0^1 f(\beta(t), \dot{\beta}(t)) \, dt - \int_0^1 f(a, \dot{\beta}(t)) \, dt \right| \leq C(m_2/m_1)|b - a| \int_0^1 |\dot{\beta}(t)| \, dt.$$

Note by condition (Hilb) that, for the cost with respect to $f(a, \cdot)$, straight lines are geodesics, and in particular $\widetilde{\gamma}(t) = (1 - t)a + tb$ is optimal. Hence, the minimal F-cost, with respect to $f(a, \cdot)$, of moving from a to b, given invariance to parametrization when $p = 1$, is $f(a, b - a)$.

Then, by Proposition 4.1 applied to the two functionals of β on the left-hand sides, we obtain

$$\big| d_f(a, b) - f(a, b - a) \big| \leq C(m_2/m_1)|b - a| \max \left[\int_0^1 |\dot{\gamma}(t)| \, dt, \int_0^1 |\dot{\widetilde{\gamma}}(t)| \, dt \right],$$

$$\leq C(m_2/m_1)^2 |b - a|^2,$$

noting the arc length bounds of $\gamma = \gamma_{a,b}$ and $\widetilde{\gamma}$ above.　□

PROPOSITION 4.3. *Suppose $d \geq 2$, $p = 1$, and that f also satisfies* (Lip) *and* (Hilb). *With respect to realizations $\{X_i\}$ in $A_{\mathrm{NN}} \cap A_Q$, the minimum values of H_n and L_n converge to the minimum of F,*

$$\lim_{n \to \infty} \min_{v \in V_n(a,b)} H_n(v) = \lim_{n \to \infty} \min_{\gamma \in \Omega_n^l(a,b)} L_n(\gamma) = \min_{\gamma \in \Omega(a,b)} F(\gamma).$$

PROOF. Consider the energies G_n and H_n in (2.9) and (2.11). For $\gamma = \gamma_{\mathbf{v}}$, the piecewise quasinormal path through the vertices $\mathbf{v} = (v_0, v_1, \ldots, v_m) \in V_n(a,b)$, we have, noting $p = 1$, that

$$G_n(\gamma) = \sum_{i=0}^{m-1} \int_0^1 f(\gamma_i(t), \dot{\gamma}_i(t)) \, dt,$$

where $\gamma_i = \gamma_{v_i, v_{i+1}}$ is a quasinormal path from v_i to v_{i+1}.

An application of Proposition 4.2, noting that $|v_{i+1} - v_i| \leq \varepsilon_n$, gives

$$\left| \int_0^1 f(\gamma_i(t), \dot{\gamma}_i(t)) \, dt - f(v_i, v_{i+1} - v_i) \right| \leq c_1 \varepsilon_n |v_{i+1} - v_i|.$$

Summing this over i gives

(4.6)
$$\left| G_n(\gamma) - H_n(\mathbf{v}) \right| \leq c_1 \varepsilon_n \sum_{i=0}^{m-1} |v_{i+1} - v_i|$$
$$\leq c_1 m_1^{-1} \varepsilon_n \min(G_n(\gamma), H_n(\mathbf{v})),$$

where the last inequality follows from applying (2.4) to both G_n and H_n.

Recall the energy L_n in (2.8). Similarly, and more directly, using (Lip), we have for a linear path $l = l_{\mathbf{v}} \in \Omega_n^l(a,b)$ through vertices $\mathbf{v} = (v_0, \ldots, v_m) \in V_n(a,b)$ that

$$\left| \int_0^1 f(l_i(t), \dot{l}_i(t)) \, dt - f(v_i, v_{i+1} - v_i) \right| \leq c_1 \varepsilon_n |v_{i+1} - v_i|,$$

where $l_i = l_{v_i, v_{i+1}}$ is the linear path from v_i to v_{i+1} with slope $v_{i+1} - v_i$. Summing over i, using (2.4), we obtain

(4.7) $$|L_n(l) - H_n(\mathbf{v})| \leq c_1 \varepsilon_n \sum_{i=0}^{m-1} |v_{i+1} - v_i| \leq c_1 m_1^{-1} \varepsilon_n \min(L_n(l), H_n(\mathbf{v})).$$

We now reprise some of the argument for Theorem 2.4. A consequence of (4.6) and Proposition 4.1 is

$$-c_1 m_1^{-1} \varepsilon_n \min G_n \leq \min G_n - \min H_n \leq c_1 m_1^{-1} \varepsilon_n \min H_n.$$

Hence, $\sup_n \min H_n < \sup_n (1 + c_1 m_1^{-1} \varepsilon_n) \min G_n$. By Theorem 2.2, as seen in its proof, for realizations in the probability 1 set $A_{\mathrm{NN}} \cap A_Q$, we have

$\lim_{n\to\infty} \min G_n = \min F < \infty$. We conclude that also $\lim_{n\to\infty} \min H_n = \min F$ a.s.

Now, we can repeat this same argument with L_n and (4.7) in place of G_n and (4.6), using now $\min H_n \to \min F$ a.s., to conclude that also $\min L_n$ converges to $\min F$ a.s. □

4.2.1. *Compactness property.* When $d \geq 2$ and $p = 1$, analogous to Lemma 3.6, we formulate now a compactness property for minimal paths $w^{(n)} \in \arg\min H_n$ and $l_{\mathbf{v}^{(n)}} \in \arg\min L_n$.

It will be useful to consider a partition of D by a regular grid. Let $z \in \mathbb{Z}^d$ and let $\square_{n,z}$ be the intersection of the box $\prod_{i=1}^{d}[z_i\tau_n, (z_i+1)\tau_n)$ with D, where $\tau_n = \varepsilon_n/\sqrt{d}$. We will refer to these sets as "boxes," with the understanding that the boundary of D results in some of these being irregularly shaped. Regardless, each $\square_{n,z}$ has diameter at most ε_n, and so points of $\{X_i\}_{i=1}^{n}$ in $\square_{n,z}$ are all connected in the random geometric graph.

We first state a bound on the number of boxes visited by an optimal path.

LEMMA 4.4.   *Suppose $d \geq 2$, $p = 1$, and that f also satisfies* (Lip) *and* (Hilb). *Suppose $\mathbf{w} \in \arg\min H_n$ and $l_{\mathbf{v}} \in \arg\min L_n$ are optimal paths. Then, for realizations $\{X_i\}$ in $A_{\mathrm{NN}} \cap A_Q$, for all large n, the number of distinct boxes $\{\square_{n,z}\}_{z\in\mathbb{Z}^d}$ visited by $\mathbf{w}$ and $\mathbf{v}$ is bounded by C/ε_n, where $C = C(d, f)$.*

We also give a bound on the number of points of an optimal path in a box.

LEMMA 4.5.   *Consider the assumptions of the second parts of Theorems 2.5 and 2.6. Suppose $\mathbf{w} \in \arg\min H_n$ is an optimal path. Then, with respect to realizations $\{X_i\}$ in a probability 1 subset of A_{NN}, for all large n, there is a constant $K = K(d, \rho, \alpha)$, such that $|\{w_j\}_{j=1}^{n} \cap \square_{n,z}| \leq K$ for all $z \in \mathbb{Z}^d$.*

Suppose now $l_{\mathbf{v}} \in \arg\min L_n$. Then the same statement above holds with $\mathbf{v}$ in place of $\mathbf{w}$.

With these ingredients, we consider the desired compactness property.

PROPOSITION 4.6.   *Consider the assumptions in the second parts of Theorems 2.5 and 2.6. Suppose $\mathbf{w}^{(n)} \in \arg\min H_n$, and consider the piecewise linear interpolations $l_n = l_{\mathbf{w}^{(n)}}$. Then, with respect to a realizations $\{X_i\}_{i\geq1}$ in a probability 1 subset of $A_{\mathrm{NN}} \cap A_Q$, the sequence $\{l_n\}$ is relatively compact for the topology of uniform convergence.*

Suppose now $l_{\mathbf{v}^{(n)}} \in \arg\min L_n$. Then, the same conclusion holds for the optimal linear interpolations $\{l_{\mathbf{v}^{(n)}}\}$.

PROOF.   We show that the sequence $\{l_n\}$ is equicontinuous for almost all realizations in A_{NN}. As the paths belong to a bounded set D, the proposition would then follow from the Arzelà–Ascoli criterion.

Consider the partition of D by boxes $\{\square_{n,z}\}_{z \in \mathbb{Z}^d}$. By Lemma 4.4, the number of boxes visited by $\mathbf{w}^{(n)}$ and $\mathbf{v}^{(n)}$ is bounded by C/ε_n a.s., for all large n, where $C = C(f, d)$. By Lemma 4.5, the number of vertices in $\mathbf{w}^{(n)}$ and $\mathbf{v}^{(n)}$ in a box is a.s. bounded by a constant $K = K(d, \rho, \alpha)$ for all large n. Thus, the maximum number k_n of points in $\mathbf{w}^{(n)}$ and $\mathbf{v}^{(n)}$ is a.s. bounded,

$$k_n \leq KC/\varepsilon_n.$$

Since $|w_{i+1}^{(n)} - w_i^{(n)}|, |v_{i+1}^{(n)} - v_i^{(n)}| \leq \varepsilon_n$, we obtain $\sup_i k_n |w_{i+1}^{(n)} - w_i^{(n)}| \leq KC$ and $\sup_i k_n |v_{i+1}^{(n)} - v_i^{(n)}| \leq KC$ a.s. for all large n.

This implies a.s. that the piecewise linear paths l_n and $l_{\mathbf{v}^{(n)}}$ are Lipschitz, with respect to the fixed constant KC, for all large n, and so in particular equicontinuous. Indeed, for $l_n = l_{\mathbf{w}^{(n)}}$, where say $\mathbf{w}^{(n)} = (w_0^{(n)}, \ldots, w_{k_n}^{(n)})$, consider the part of the path connecting $w_i^{(n)}$ and $w_{i+1}^{(n)}$ from times i/k_n to $(i+1)/k_n$, namely $l_n(t) = w_i^{(n)}(i+1-k_n t) + w_{i+1}^{(n)}(k_n t - i)$. Then we have $|\dot{l}_n(t)| = k_n |w_{i+1}^{(n)} - w_i^{(n)}| \leq k_n \varepsilon_n \leq KC$. The same argument holds for the paths $l_{\mathbf{v}^{(n)}}$. $\square$

We now prove the lemmas used in the proof Proposition 4.6.

PROOF OF LEMMA 4.4. Any visit of the path $\mathbf{w}$ or $\mathbf{v}$ to $2^d + 1$ distinct boxes has a Euclidean length of at least $\varepsilon_n/\sqrt{d}$, since not all $2^d + 1$ boxes can be adjacent. Recalling (2.4), and the formulas (2.8) and (2.11), such a visitation therefore has a H_n cost or L_n cost of at least $m_1 \varepsilon_n/\sqrt{d}$. So, we may bound the number of boxes visited by $\mathbf{w}$ or $\mathbf{v}$ by $C' H_n(\mathbf{w})/\varepsilon_n$, where $C' = (2^d + 1)\sqrt{d}/m_1$ depends on the dimension and f, but not on the path $\mathbf{w}$ or $\mathbf{v}$. Recalling Proposition 4.3, we have with respect to realizations in $A_{NN} \cap A_Q$ that $\lim \min H_n = \lim \min L_n = \min F < \infty$. The lemma then follows with say $C = 2C' \min F$. $\square$

Before coming to the proof of Lemma 4.5, we need some estimates. The next result shows that optimal paths $\mathbf{w} \in \arg\min H_n$ and $l_{\mathbf{v}} \in \arg\min L_n$ cannot have "long necks," and gives a bound on the number of points nearby an edge in the graph.

LEMMA 4.7. *Suppose $d \geq 2$ and $p = 1$. Fix a realization $\{X_i\}$ in A_{NN}. Suppose $\mathbf{w} \in \arg\min H_n$ is an optimal path. If $i < j$ is such that $|w_i - w_j| < \varepsilon_n$, then*

$$(4.8) \qquad w_k \in B(w_i, C\varepsilon_n) \qquad \text{for } i \leq k \leq j,$$

where $C = 2(m_2/m_1)$.

Further, let $\Theta_n = \sup_{i,j:|w_i - w_j| < \varepsilon_n} |i - j|$, and suppose $\varepsilon_n = n^{-\delta}$, where $\delta > 1/(\beta + d)$ and $\beta > 0$. Then, with respect to realizations in a probability 1 subset of A_{NN}, for all large n, we have

$$(4.9) \qquad \Theta_n \leq \varepsilon_n^{-\beta}.$$

Suppose now $l_{\mathbf{v}} \in \arg\min L_n$. The same conclusions (4.8) and (4.9) hold with $\mathbf{v}$ in place of $\mathbf{w}$.

PROOF. We first show (4.8). If one of the points $\{w_k\}_{k=i}^{j}$ is more than an Euclidean distance $2(m_2/m_1)\varepsilon_n$ away from w_i, we have $\sum_{k=i}^{j-1} |w_{k+1} - w_k| \geq 2(m_2/m_1)\varepsilon_n$. Then, recalling (2.4), we have

$$\sum_{k=i}^{j-1} f(w_k, w_{k+1} - w_k) \geq m_1 \sum_{k=i}^{j-1} |w_{k+1} - w_k|$$
$$\geq m_1(2m_2/m_1)\varepsilon_n$$
$$\geq 2m_2|w_j - w_i| \geq 2f(w_i, w_j - w_i).$$

But, this implies that the path connecting w_i and w_j in one step would be less costly, with respect to H_n, than $\mathbf{w}$. Since $\mathbf{w}$ was taken to be minimal, all points $\{w_k\}_{k=i}^{j}$ therefore must belong to $B(w_i, 2(m_2/m_1)\varepsilon_n)$.

Suppose now $l_{\mathbf{v}} \in \arg\min L_n$ and recall the form of L_n when $p = 1$ in (2.8). Similarly, if one of the points $\{v_k\}_{k=i}^{j}$ is away from v_i by $2(m_2/m_1)\varepsilon_n$, we have

$$\sum_{k=i}^{j-1} \int_0^1 f\big(l_{v_k, v_{k+1}}(t), v_{k+1} - v_k\big) \, dt$$
$$\geq m_1 \sum_{k=i}^{j-1} |v_{k+1} - v_k|$$
$$\geq 2m_2\varepsilon_n \geq 2m_2|v_j - v_i| \geq 2\int_0^1 f\big(l_{v_i, v_j}(t), v_j - v_i\big) \, dt,$$

also a contradiction of minimality of $l_{\mathbf{v}}$.

We now consider (4.9). The proof here is a count bound with respect to $\mathbf{w}$. The argument with respect to $\mathbf{v}$ is exactly the same with $\mathbf{v}$ in place of $\mathbf{w}$.

First, let $N_{i,n}$ be the number of points $\{X_i\}_{i=1}^n$ distinct from X_i in the ball $B(X_i, C\varepsilon_n)$, and let $N_n = \max\{N_{i,n}\}_{i=1}^n$. Then, as $|j - i|$ is bounded by the number of points distinct from w_i, one of the sample points, in the ball $B(w_i, C\varepsilon_n)$, we have $|j - i| \leq N_n$.

Now, $N_{i,n}$ is Binomial$(n-1, p)$ where $p = \nu(B(X_i, C\varepsilon_n))$. For $k \geq 1$, we have

$$\mathbb{P}(N_{i,n} \geq k) \leq \binom{n-1}{k} p^k \leq \frac{(np)^k}{k!}.$$

Recalling that $\nu = \rho dx$ and ρ is bounded, we have $p \leq \|\rho\|_\infty \text{Vol}(B(0,1))\varepsilon_n^d$, and so $P(N_{i,n} \geq k) \leq (C'n\varepsilon_n^d)^k/k!$ for some constant C'.

Let $N_n = \max\{N_{i,n}\}_{i=1}^n$. Then a union bound gives that

$$P(N_n \geq k) \leq \frac{n}{k!} \exp\{k(\log n + d\log\varepsilon_n + \log C')\}.$$

Taking $k = \lceil \varepsilon_n^{-\beta} \rceil$, and noting $k! \geq \sqrt{2\pi} e^{-k} k^{k+1/2}$, yields that

$$(4.10) \quad P\left(N_n \geq \varepsilon_n^{-\beta}\right) \leq \frac{n}{\sqrt{2\pi}} \exp\left\{\varepsilon_n^{-\beta}\left(\log n + (\beta + d)\log \varepsilon_n + \log C' + 1\right)\right\}.$$

If ε_n is in the form $\varepsilon_n = n^{-\delta}$, then the right-hand side of (4.10) is summable when $(\beta + d)\delta > 1$.

Hence, by the Borel–Cantelli lemma, for realizations in the intersection of a probability 1 set and A_{NN} say, we have $\Theta_n \leq N_n \leq \varepsilon_n^{-\beta}$ for all large n, and (4.9) follows. $\square$

We now give a lower bound on the cost of certain "long necks," that is the cost of an optimal H_n-path $\mathbf{w}$ of moving away from two close by vertices.

LEMMA 4.8. *Suppose $d \geq 2$, $p = 1$ and that f also satisfies* (TrIneq), *and* (Pythag) *with $\alpha > 1$. Fix a realization $\{X_i\}$ in A_{NN}. Suppose $\mathbf{w} \in \arg\min H_n$ is an optimal path, and let $i < j$ be indices such that $|w_i - w_j| < \varepsilon_n$. Let ℓ denote the straight line segment from w_i to w_j. Consider a neighborhood $\mathcal{A} = \bigcup_{x \in \ell} B(x, r)$ of ℓ, with $r = \varepsilon_n^{\alpha}$.*

Then, if there is a point $w_k \notin \mathcal{A}$ for some $i < k < j$, there is a constant $C = C(\alpha, f)$ such that

$$(4.11) \qquad \sum_{q=i}^{j-1} f(w_i, w_{q+1} - w_q) \geq f(w_i, w_j - w_i) + C r^{\alpha}.$$

Suppose now $l_{\mathbf{v}} \in \arg\min L_n$. Then (4.11) *holds with $\mathbf{v}$ in place of $\mathbf{w}$.*

PROOF. The argument for $\mathbf{v}$ is the same as for $\mathbf{w}$, which we now present. Suppose a point w_k is at least an Euclidean distance r from ℓ. By the (TrIneq) condition, $\sum_{q=i}^{j-1} f(w_i, w_{q+1} - w_q) \geq f(w_i, w_k - w_i) + f(w_i, w_j - w_k)$.

By Lemma 4.7, as $|w_i - w_j| < \varepsilon_n$, we have $|w_k - w_i| \leq 2(m_2/m_1)\varepsilon_n$, which is strictly less than an $\eta < 1$ for all large n. We also conclude $|w_j - w_k|, |w_i - w_j| < \eta < 1$ for all large n. In addition, $2(m_2/m_1)r^{1/\alpha} = 2(m_2/m_1)\varepsilon_n \geq |w_i - w_j|$. Thus, by (Pythag) with $x = w_i$, we obtain $f(w_i, w_k - w_i) + f(w_i, w_j - w_k) \geq f(w_i, w_j - w_i) + C r^{\alpha}$, where $C = C(\alpha, f)$.

Hence, (4.11) follows by combining the inequalities. $\square$

We now give the proof of Lemma 4.5. The argument is in two steps. In the first step, using a rough count on the number of vertices of the path within a given box, we may approximate the contribution to H_n and L_n from the vertices in the box in terms of a localized cost. Then we use (Pythag), applied to the localized cost, to deduce that the optimal path in the box is trapped in a small set in the box. The second step then is to show that such small sets contain only a constant number of points in $\{X_i\}$.

PROOF OF LEMMA 4.5.   We will give the main argument for $\mathbf{w}$ and indicate modifications with respect to $\mathbf{v}$. Consider a box $\square := \square_{n,z}$. Boxes with at most one point trivially satisfy the claim in the lemma if say $K \geq 2$. Suppose now that there are at least two points in the box.

Step 1. Let w_i and w_j be the first and last points of $\mathbf{w}$ in the box, that is, with the smallest and largest indices respectively. By Lemma 4.7, as $|w_i - w_j| < \varepsilon_n$, we have $w_k \in B(w_i, C'\varepsilon_n)$ for $i \leq k \leq j$. Hence, by (Lip), we have $|\sum_{k=i}^{j-1} f(w_k, w_{k+1} - w_k) - \sum_{k=i}^{j-1} f(w_i, w_{k+1} - w_k)| \leq C'\varepsilon_n^2 |j - i|$. Now, also by Lemma 4.7, when $\delta > (\beta + d)^{-1}$ for $\beta > 0$, we have $|j - i| \leq \varepsilon_n^{-\beta}$. Hence, the following estimate, with respect to a localized energy, where $x = w_i$ is fixed, is obtained:

$$(4.12) \qquad \left| \sum_{k=i}^{j-1} f(w_k, w_{k+1} - w_k) - \sum_{k=i}^{j-1} f(w_i, w_{k+1} - w_k) \right| \leq C'\varepsilon_n^{2-\beta}.$$

Similarly, when $\mathbf{v}$ is considered, by the same reasoning using Lemma 4.7 and (Lip), we may obtain (4.12) with $\mathbf{v}$ in place of $\mathbf{w}$, and moreover,

$$\left| \sum_{k=i}^{j-1} \int_0^1 f(l_{v_k, v_{k+1}}(t), v_{k+1} - v_k) \, dt - \sum_{k=i}^{j-1} f(v_k, v_{k+1} - v_k) \right|$$

$$\leq C'\varepsilon_n^2 |j - i| \leq C'\varepsilon^{2-\beta}.$$

Hence, combining these two estimates, we obtain that

$$(4.13) \quad \left| \sum_{k=i}^{j-1} \int_0^1 f(l_{v_k, v_{k+1}}(t), v_{k+1} - v_k) \, dt - \sum_{k=i}^{j-1} f(v_i, v_{k+1} - v_k) \right| \leq 2C'\varepsilon_n^{2-\beta}.$$

Returning to $\mathbf{w}$, by Lemma 4.8, noting (4.12), any path $(w_i, \ldots, w_j)$ exiting $\mathcal{A}$, the $r = \varepsilon_n^{\alpha}$-neighborhood of the line segment from w_i to w_j, is costlier, with respect to H_n, than the H_n-cost $f(w_i, w_j - w_i)$ of a straight path connecting w_i to w_j in a single hop, as follows:

$$\sum_{q=i}^{j-1} f(w_q, w_{q+1} - w_q) - f(w_i, w_j - w_i) \geq Cr^\alpha - C'\varepsilon_n^{2-\beta}$$

$$(4.14)$$

$$= C\varepsilon_n^{\alpha^2} - C'\varepsilon_n^{2-\beta}.$$

Let us now consider $\mathbf{v}$. Since $|v_i - v_j| \leq C'\varepsilon_n$ by Lemma 4.7, using (Lip), we have that

$$(4.15) \qquad \left| f(v_i, v_j - v_i) - \int_0^1 f(l_{v_i, v_j}(t), v_j - v_i) \, dt \right| \leq C'\varepsilon_n^2.$$

Following the same reasoning given with respect to $\mathbf{w}$, we may obtain (4.14) with $\mathbf{v}$ in place of $\mathbf{w}$. Then, noting (4.13), a path $(v_i, \ldots, v_j)$, exiting the $r = \varepsilon_n^{\alpha}$-neighborhood of the line segment from v_i to v_j, has L_n cost more than the one

step H_n cost $f(v_i, v_j - v_i)$ by the amount $Cr^\alpha - 2C'\varepsilon_n^{2-\beta}$. By (4.15), this H_n cost differs from the one step L_n cost $\int_0^1 f(l_{v_i,v_j}(t), v_j - v_i)\, dt$ of moving from v_i to v_j by $C'\varepsilon_n^2$.

Therefore, the cost savings of moving in one step, in considering $\mathbf{w}$ or $\mathbf{v}$ which exit the r-neighborhood, is bounded below by $Cr^\alpha - 3C'\varepsilon_n^{2-\beta} = O(\varepsilon_n^{\alpha^2} - \varepsilon_n^{2-\beta})$, which is positive, for all large n, when $\alpha^2 < 2 - \beta$. This is the case when we fix $\beta = (2 - \alpha^2)\eta > 0$, for an $0 < \eta < 1$, since $1 < \alpha < \sqrt{2}$.

Hence, with this choice of β, such exiting paths are not optimal, and all the points $\{w_i, \ldots, w_j\}$ or $\{v_i, \ldots, v_j\}$ in the box must belong to the $r = \varepsilon_n^\alpha$-neighborhood of the line segment connecting the ith and jth points.

We note, given the value of β, to use Lemma 4.7 above, the exponent δ should satisfy $\delta > [(2 - \alpha^2)\eta + d]^{-1}$, afforded by our assumption (2.12).

Step 2. We now focus on $\mathbf{w}$ as the following counting argument is the same with respect to $\mathbf{v}$. We will count the points in the small set $\mathcal{A}$.

To this end, given two samples X_i and X_j, consider the line segment connecting them. Let $\ell_{i,j}$ be that portion of the line segment within distance ε_n of X_i. Define $N_{n,i,j}$ as the number of distinct points in the $r = \varepsilon_n^\alpha$ neighborhood of $\ell_{i,j}$. Then $N_{n,i,j}$ is Binomial$(n - 2, p)$ where p is the ν measure of the neighborhood. As the neighborhood is nearly covered by a cylinder with length ε_n and radius ε_n^α, and the density ρ is bounded, $p \leq C(\rho)\varepsilon_n^{\alpha(d-1)+1}$. Let also N_n be the maximum of such counts $N_{n,i,j}$ over the sample points.

Now, by Step 1 and the above definitions, the cardinality $|\{w_k\}_{k=i}^j| = |j - i| =: M_z$ of points in the path in $\square_z$ is bounded by $|\mathcal{X}_n \cap \mathcal{A}| \leq 2 + N_n$, which does not depend on z. Observe that

$$P(N_{n,i,j} \geq K') \leq (n^{K'}/K'!)\nu(\mathcal{A})^{K'} \leq C(\rho)^{K'} n^{K'} \varepsilon_n^{K'\alpha d + K'(1-\alpha)}.$$

Hence, by a union of events bound, as the total number of pairs (X_i, X_j) is $n(n-1)$ in the sample, we have

$$P(N_n \geq K') \leq n^2 \times C(\rho)^{K'} n^{K'} \varepsilon_n^{K'\alpha d + K'(1-\alpha)}.$$

Therefore, we have that

$$(4.16) \quad P\left(\max_z M_z \geq K' + 2\right) \leq P(N_n \geq K') \leq C(\rho)^{K'} n^{K'+2} \varepsilon_n^{K'\alpha d + K'(1-\alpha)}.$$

Suppose ε_n is of the form $\varepsilon_n = n^{-\delta}$ for $0 < \delta < 1/d$. If $d < K' < \infty$ and

$$(4.17) \quad \delta > (K' + 3)/[K'(\alpha(d-1)+1)],$$

the display (4.16) is summable in n. In particular, when $\delta > [\alpha(d-1)+1]^{-1}$, part of our assumptions (2.12), a large but fixed K' can be chosen so that (4.17) holds.

Hence, by Borel–Cantelli lemma, on the intersection of a probability 1 set and A_{NN} say, we recover the claim for all large n that the path visits at most $K = K' + 2$ points between the first and last visit to a visited box. $\square$

4.2.2. *Convergence of optimal paths.* We now consider the behavior of the optimal paths, in analogy to Theorem 2.2, for the energy H_n.

PROPOSITION 4.9. *Consider the assumptions for the second parts of Theorems 2.5 and 2.6. Consider a discrete path $\mathbf{w}^{(n)} \in \arg\min H_n$ and its linear interpolation, $l_{\mathbf{w}^{(n)}}$. Then, with respect to realizations in a probability 1 subset of $A_{NN} \cap A_Q$, for any subsequence of $\{l_{\mathbf{w}^{(n)}}\}$, and correspondingly of $\{\mathbf{w}^{(n)}\}$, there is a further subsequence of the linear paths which converges uniformly to a limit path $\gamma \in \arg\min F$, and of the discrete paths in the Hausdorff sense to S_γ.*

If F has a unique (up to reparametrization) minimizer γ, then the whole sequence $\{\mathbf{w}^{(n)}\}$ converges, $\lim_{n\to\infty} d_{\text{haus}}(\mathbf{w}^{(n)}, S_\gamma) = 0$.

Consider now a path $l_{\mathbf{v}^{(n)}} \in \arg\min L_n$. The same conclusions holds for $\{\mathbf{v}^{(n)}\}$ in place of $\{\mathbf{w}^{(n)}\}$.

PROOF. Consider first $\mathbf{w}^{(n)} \in \arg\min H_n$. By the compactness criterion, Proposition 4.6, almost surely, any subsequence of the paths $\{l_{\mathbf{w}^{(n)}}\}$ has a further subsequence $\{l_{\mathbf{w}^{(n_k)}}\}$ converging uniformly to a limit γ. By the liminf Lemma 3.2, $F(\gamma) \leq \liminf_{k\to\infty} F(l_{\mathbf{w}^{(n_k)}})$.

The same argument and conclusion holds with $l_{\mathbf{v}^{(n)}} \in \arg\min L_n$ and $\mathbf{v}^{(n)}$ in place of $l_{\mathbf{w}^{(n)}}$ and $\mathbf{w}^{(n)}$.

We now show that $\gamma \in \arg\min F$. With respect to optimal L_n paths, as $F(l_{\mathbf{v}^{(n)}}) = \min L_n$, and $\min L_n \to \min F$ a.s. by Proposition 4.3, we obtain $F(\gamma) \leq \min F$, and so the desired conclusion.

For H_n optimal paths, we recall an argument in the proof of Theorem 2.4. Using only the standing assumptions (allowing $p = 1$) and (Lip), we derived (4.2), namely, for $\mathbf{u} \in V_n(a, b)$, that

$$\left| L_n(l_{\mathbf{u}}) - H_n(\mathbf{u}) \right| \leq cm_1^{-1}\varepsilon_n \min\left(L_n(l_{\mathbf{u}}), H_n(\mathbf{u})\right),$$

where c is the constant in (Lip). Then, as a consequence of Proposition 4.1, we saw in (4.4) that $\min L_n \leq L_n(l_{\mathbf{w}^{(n)}}) \leq (cm_1^{-1}\varepsilon_n + 1)^2 \min L_n$. Since, by Proposition 4.3, $\min L_n \to \min F$ a.s., we conclude that $\gamma \in \arg\min F$.

Finally, we remark that the Hausdorff convergences are argued as in the proof of Corollary 2.3. $\square$

APPENDIX

Here, we collect some results which we had previously assumed.

A.1. Nearest-neighbor rate.

PROPOSITION A.1. *Let $\{X_i\}$ be i.i.d. samples from a probability measure $v = \rho(x)\,dx$ on a Lipschitz domain D, and let*

$$R_n = \sup_{y \in D} \min_{1 \leq i \leq n} |X_i - y|.$$

Suppose $\rho(x)$ is uniformly bounded below by a positive constant. Then there exists a constant C, independent of n, such that, for almost all realizations $\{X_i\}$,

$$\limsup_{n\to\infty} \frac{R_n n^{1/d}}{(\log n)^{1/d}} \leq C.$$

In particular, when ε_n satisfies (2.6), *for $a, b \in D$, almost surely for all large n, the graph $\mathcal{G}_n(a, b)$ is connected.*

PROOF.   We first address the claim with respect to R_n. Let $B(y, r)$ be the Euclidean ball of radius r centered at $y \in D$. Since D is Lipschitz, there is a constant c such that $m(B(y, r) \cap D) \geq cm(B(y, r))$ for all small $r > 0$, where m denotes Lebesgue measure (cf. the discussion about cone conditions in Section 4.11 of [1]). It follows that there is a constant c such that $m(B(y, r) \cap D)/m(D) \geq cr^d$ for all $y \in D$ and all small $r > 0$. Since ν has density ρ bounded below by a positive constant, there exists a constant c such that $\nu(B(y, r)) \geq cr^d$ for all $y \in D$ and $0 < r < r_0$, where r_0 is a sufficiently small constant. Therefore, recalling $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$, we have

$$\mathbb{P}(|B(y, r) \cap \mathcal{X}_n| = 0) = (1 - \nu(B(y, r)))^n$$

(A.1)

$$\leq (1 - cr^d)^n \leq e^{-cnr^d}.$$

Let $\{y_1, \ldots, y_k\} \subset D$ be a collection of points so that $\sup_{x \in X} \min_{1 \leq i \leq k} |x - y_i| \leq r$. We may take the number of points k to satisfy $k \leq c/r^d$ for some constant c independent of r, say, by choosing $\{y_i\}$ to be a regular grid, with grid length $\sim r$.

Let E_i denote the event that $|B(y_i, r) \cap \mathcal{X}_n| = 0$, and consider the event $\{R_n > 2r\}$ that there exists a $y \in D$ with $\min_{1 \leq n \leq n} |X_i - y| > 2r$. Then, by a triangle inequality argument, we have $\{R_n > 2r\} \subset \bigcup_{i=1}^{k} E_i$. Hence, together with (A.1), we have

$$\text{(A.2)} \qquad\qquad \mathbb{P}(R_n > 2r) \leq \sum_{i=1}^{k} \mathbb{P}(E_i) \leq \frac{c}{r^d} e^{-cnr^d}.$$

Let $r^d = (3 \log n)/(cn)$. Then (A.2) gives a summable term,

$$\mathbb{P}\left(R_n > 2\frac{(3 \log n)^{1/d}}{n^{1/d}}\right) \leq \frac{c^2}{3n^2}.$$

By the Borel–Cantelli lemma, $R_n \leq 2(3 \log n)^{1/d}/n^{1/d}$ for all large n.

We now show that $\mathcal{G}_n(a, b)$ is connected when ε_n satisfies (2.6). Let v_1, v_2 be any vertices in $\mathcal{X}_n \cup \{a, b\}$, and consider the line $\ell(t) = v_1(1 - t) + v_2(t)$ between them for $t \in [0, 1]$. By convexity of D, the path ℓ is contained in D. Consider points on the path $v_1 = \ell(0), \ell(R_n), \ell(2R_n), \ldots, \ell(kR_n), \ell(1) = v_1$, where $k = \lfloor |v_2 - v_1|/R_n \rfloor$ so that $|1 - kR_n| \leq R_n$. Each point $y = \ell(jR_n)$ is within Euclidean distance R_n of a point $u_j \in \mathcal{X}_n$. By construction, for $1 \leq j \leq k - 1$, the Euclidean

distance between u_j and u_{j+1} is less than sum of the distances, from u_j to $\ell(jR_n)$, from $\ell(jR_n)$ to $\ell((j+1)R_n)$, and from $\ell((j+1)R_n)$ to u_{j+1}, which is bounded by $3R_n$. Similarly, the endpoints v_1, v_2 are within Euclidean distance $2R_n$ of u_1 and u_k, respectively. Since, by (2.6), $3R_n/\varepsilon_n < 1$ for all large n, the path along vertices $\{v_1, u_1, \ldots, u_k, v_2\}$ belongs to $V_n(a, b)$ and so v_1 and v_2 are connected in $G_n(a, b)$ a.s. for all large n. $\square$

A.2. Existence of quasinormal minimizers.   We discuss a conservation law for F-minimizing paths, and existence of F-minimizing Lipschitz paths, following the treatment in [9].

PROPOSITION A.2.   *Consider the integral functional $F(\gamma) = \int_0^1 f(\gamma, \dot\gamma)\, dt$, where f satisfies* (A0)–(A3). *Then, F attains a minimum on the set $\Omega(a, b)$ of Lipschitz paths from a to b. In other words, there exists a $\gamma^* \in \Omega(a, b)$ with $F(\gamma^*) = \inf_{\gamma \in \Omega(a,b)} F(\gamma)$.*

In the case that $p = 1$, there exists a $\gamma \in \arg\min F$ and constants c, c_1, c_2 such that

$$(\text{A.3}) \qquad\qquad f(\gamma(t), \dot\gamma(t)) = c \qquad \text{for a.e. } t \in [0, 1]$$

and

$$(\text{A.4}) \qquad\qquad c_1 \leq |\dot\gamma(t)| \leq c_2 \qquad \text{for a.e. } t \in [0, 1].$$

If $p > 1$, then there are constants c, c_1, c_2 such that (A.3) *and* (A.4) *hold for any $\gamma \in \arg\min F$.*

PROOF.   We first give an argument in the case where $p > 1$. Note, by assumption, the integrand f is continuous and C^1 on $D \times (\mathbb{R}^d \setminus \{0\})$, convex and p-homogenous in the second argument, and satisfies (2.4). As $p > 1$, f may be extended continuously to a C^1 function on $D \times \mathbb{R}^d$.

When the domain of F is extended to all Sobolev paths $\gamma \in W^{1,p}([0, 1]; D)$ with $\gamma(0) = a$, $\gamma(1) = b$, the existence of a minimizer follows from Remark 2 of Section 3.2 in [9]. In particular, the continuity and convexity assumptions (A0) and (A1) imply that F is lower-semicontinuous with respect to weak convergence of Sobolev functions. The existence of a minimizer then follows from a standard compactness argument.

Let γ denote such a Sobolev minimizer. Consider now an inner variation $\omega(t, \varepsilon) = \gamma(\xi(t, \varepsilon))$ of γ, where ξ is C^1 on $[0, 1] \times (-\varepsilon_0, \varepsilon_0)$ for some $\varepsilon_0 > 0$ and $\xi(\cdot, \varepsilon)$ is a C^1 diffeomorphism of the interval $[0, 1]$ to itself, with $\gamma(t, 0) = t$. It may be shown (see the discussion on pages 19–21, Proposition 1.16 and Remark 3 in Section 1.1 of [9]) that the optimality condition $\frac{d}{d\varepsilon} F(\omega(\cdot, \varepsilon))|_{\varepsilon=0}$ over the class of inner variations, together with Euler's identity for homogenous functions, $v \cdot \nabla_v f(x, v) = pf(x, v)$, together imply, for some constant c, that

$$(\text{A.5}) \qquad\qquad (p - 1)f(\gamma(t), \dot\gamma(t)) = c \qquad \text{for a.e. } t.$$

Finally, by assumption (A3) on f, it follows that $c > 0$ and there exist constants $c_1, c_2 > 0$ with $c_1 \leq |\dot{\gamma}(t)| \leq c_2$ for almost every t. In particular, $\gamma \in \Omega(a, b)$, and the proposition is proved for $p > 1$.

The argument for the $p = 1$ case is complicated by a lack of compactness with respect to weak convergence in the Sobolev space $W^{1,1}([0, 1]; D)$, as well as difficulty in establishing an analogue of (A.5). By a more involved argument, relating optimizers of F to optimizers of the quadratic functional $Q(\gamma) := \int_0^1 f^2(\gamma, \dot{\gamma}) \, dt$, the existence of a Lipschitz path $\gamma \in \arg\min F$ satisfying (A.3) is established in Theorem 1 of [16] (see also Theorem 5.22 of [9] which gives an alternative argument). From this and assumption (A3), inequality (A.4) follows. $\quad\square$

REFERENCES

[1] ADAMS, R. A. and FOURNIER, J. J. F. (2003). *Sobolev Spaces*, 2nd ed. *Pure and Applied Mathematics* (*Amsterdam*) **140**. Elsevier/Academic Press, Amsterdam. MR2424078

[2] AGRANOVICH, M. S. (2015). *Sobolev Spaces*, *Their Generalizations and Elliptic Problems in Smooth and Lipschitz Domains*. Springer, Cham. MR3287270

[3] ALAMGIR, M. and VON LUXBURG, U. (2012). Shortest path distance in random k-nearest neighbor graphs. In *Proc. 29th Int. Conf. on Machine Learning*. Available at arXiv:1206.6381.

[4] ÁLVAREZ PAIVA, J. C. (2006). Some problems on Finsler geometry. In *Handbook of Differential Geometry. Vol. II* 1–33. Elsevier/North-Holland, Amsterdam. MR2194667

[5] AUFFINGER, A., DAMRON, M. and HANSON, J. (2017). 50 *Years of First-Passage Percolation*. *University Lecture Series* **68**. Amer. Math. Soc., Providence, RI. MR3729447

[6] BERN, M. et al. (1999). Emerging Challenges in Computational Topology. NSF report. Available at arXiv:cs/9909001.

[7] BERNSTEIN, M., DE SILVA, V., LANGFORD, J. and TENENBAUM, J. (2000). Graph approximations to geodesics on embedded manifolds. Technical Report, Dept. Psychology, Stanford Univ., Stanford, CA.

[8] BURAGO, D., BURAGO, Y. and IVANOV, S. (2001). *A Course in Metric Geometry*. *Graduate Studies in Mathematics* **33**. Amer. Math. Soc., Providence, RI. MR1835418

[9] BUTTAZZO, G., GIAQUINTA, M. and HILDEBRANDT, S. (1998). *One-Dimensional Variational Problems*: *An Introduction*. Oxford Lecture Series in Mathematics and Its Applications **15**. Oxford Univ. Press, Oxford. MR1694383

[10] CABELLO, S. and JEJCIC, M. (2015). Shortest paths in intersection graphs of unit disks. *Comput. Geom.* **48** 360–367. MR3296073

[11] CRANE, K., WEISCHEDEL, C. and WARDETZKY, M. (2013). Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Trans. Graph.* **28** 106.

[12] GARCIA TRILLOS, N. (2017). Gromov–Hausdorff limit of Wasserstein spaces on point clouds. Available at arXiv:1702.03464.

[13] GELFAND, I. M. and SMIRNOV, M. M. (1994). Lagrangians satisfying Crofton formulas, Radon transforms, and nonlocal differentials. *Adv. Math.* **109** 188–227. MR1304752

[14] GIESEN, J. and WAGNER, U. (2004). Shape dimension and intrinsic metric from samples of manifolds. *Discrete Comput. Geom.* **32** 245–267. MR2074340

[15] HASHIMOTO, T. B., SUN, Y. and JAAKKOLA, T. S. (2015). Metric recovery from directed unweighted graphs. In *Proc. 18th Int. Conference on AI and Stat* (*AISTATS*) **38**, JMLR: W&CP, San Diego, CA.

[16] HILDEBRANDT, S. (2015). Minimizers of one-dimensional parametric variational integrals. *St. Petersburg Math. J.* **27** 569–576.

[17] HIRSCH, C., NEUHÄUSER, D., GLOAGUEN, C. and SCHMIDT, V. (2015). First passage percolation on random geometric graphs and an application to shortest-path trees. *Adv. in Appl. Probab.* **47** 328–354. MR3360380

[18] HOWARD, C. D. and NEWMAN, C. M. (1997). Euclidean models of first-passage percolation. *Probab. Theory Related Fields* **108** 153–170. MR1452554

[19] HOWARD, C. D. and NEWMAN, C. M. (2001). Geodesics and spanning trees for Euclidean first-passage percolation. *Ann. Probab.* **29** 577–623. MR1849171

[20] HWANG, S. J., DAMELIN, S. B. and HERO, A. O. III (2016). Shortest path through random points. *Ann. Appl. Probab.* **26** 2791–2823. MR3563194

[21] LAGATTA, T. and WEHR, J. (2010). A shape theorem for Riemannian first-passage percolation. *J. Math. Phys.* **51** 053502, 14. MR2666980

[22] LAGATTA, T. and WEHR, J. (2014). Geodesics of random Riemannian metrics. *Comm. Math. Phys.* **327** 181–241. MR3177936

[23] MAZ'YA, V. (2011). *Sobolev Spaces with Applications to Elliptic Partial Differential Equations*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* **342**. Springer, Heidelberg. MR2777530

[24] MÉMOLI, F. and SAPIRO, G. (2005). Distance functions and geodesics on submanifolds of $\mathbb{R}^d$ and point clouds. *SIAM J. Appl. Math.* **65** 1227–1260. MR2147326

[25] PAPADOPOULOS, A. (2014). Hilbert's fourth problem. In *Handbook of Hilbert Geometry* (A. Papadopoulos and M. Troyanov, eds.). *IRMA Lect. Math. Theor. Phys.* **22** 391–431. Eur. Math. Soc., Zürich. MR3329889

[26] PENROSE, M. (2003). *Random Geometric Graphs. Oxford Studies in Probability* **5**. Oxford Univ. Press, Oxford. MR1986198

[27] PEYRÉ, G., PÉCHAUD, M., KERIVEN, R., COHEN, L. et al. (2010). Geodesic methods in computer vision and graphics. *Foundations and Trends in Computer Graphics and Vision* **5** 197–397.

[28] SAJAMA and ORLITZKY, A. (2005). Estimating and computing density based distance metrics. In *Proc. 22nd Int. Cont. on Machine Learning*, Bonn, Germany.

[29] SEDGEWICK, R. and VITTER, J. S. (1986). Shortest paths in Euclidean graphs. *Algorithmica* **1** 31–48. MR0833117

[30] SETHIAN, J. A. (1999). Fast marching methods. *SIAM Rev.* **41** 199–235. MR1684542

[31] TAMÁSSY, L. (2008). Relation between metric spaces and Finsler spaces. *Differential Geom. Appl.* **26** 483–494. MR2458274

[32] TENENBAUM, J. B., DE SILVA, V. and LANGFORD, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* **290** 2319–2323.

[33] VON DEYLEN, S. W., GLICKENSTEIN, D. and WARDETZKY, M. (2016). Distortion estimates for barycentric coordinates on Riemannian simplices. Available at arXiv:1610.01168.

[34] YU, H., ZHANG, J. J. and JIAO, Z. (2014). Geodesics on point clouds. *Math. Probl. Eng.* Art. ID 860136, 12. MR3206968

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721
USA
E-MAIL: edavis@math.arizona.edu
          sethuram@math.arizona.edu