



Lower Bounding the Gromov–Hausdorff Distance in Metric Graphs

Henry Adams  

University of Florida, Gainesville, FL, USA

Sushovan Majhi  

George Washington University, Washington, DC, USA

Fedor Manin  

University of Toronto, Canada

Žiga Virk  

University of Ljubljana, Slovenia

Institute IMFM, Ljubljana, Slovenia

Nicolò Zava  

Institute of Science and Technology, Klosterneuburg, Austria

Abstract

Let G be a finite, connected metric graph and let X be a subset of G . If X is sufficiently dense in G , we show that the Gromov–Hausdorff distance matches the Hausdorff distance, namely $d_{\text{GH}}(G, X) = d_{\text{H}}(G, X)$. When the metric graph is the circle $G = S^1$ with circumference 2π , a recent study established the equality $d_{\text{GH}}(S^1, X) = d_{\text{H}}(S^1, X)$ whenever $d_{\text{GH}}(S^1, X) < \frac{\pi}{6}$. Our results relax this hypothesis to $d_{\text{GH}}(S^1, X) < \frac{\pi}{3}$, and furthermore, we show that the constant $\frac{\pi}{3}$ is the best possible. We lower bound the Gromov–Hausdorff distance $d_{\text{GH}}(G, X)$ by the Hausdorff distance $d_{\text{H}}(G, X)$ via a simple topological obstruction: the existence of a possibly discontinuous function $f: G \rightarrow X$ with too small distortion contradicts the connectedness of G .

2012 ACM Subject Classification Mathematics of computing \rightarrow Topology; Mathematics of computing \rightarrow Graph theory

Keywords and phrases Gromov–Hausdorff distance, distortion, connectedness, Borsuk–Ulam theorem

Digital Object Identifier 10.4230/LIPIcs.SoCG.2026.3

Related Version *Full Version*: <https://arxiv.org/abs/2411.09182> [3]

Funding *Henry Adams*: Simons Foundation Travel Support for Mathematicians.

Žiga Virk: Slovene research agency grant P1-0292.

Nicolò Zava: FWF Grant, Project number I4245-N35.

1 Introduction

The Gromov–Hausdorff distance between two abstract metric spaces (X, d_X) and (Y, d_Y) , denoted $d_{\text{GH}}(X, Y)$, provides a dissimilarity measure quantifying how far the two metric spaces are from being isometric [13, 14, 24]. In the past two decades, this distance has found applications in topological data analysis (TDA) as a theoretical framework for shape and dataset comparison [21], which motivated the study of its quantitative aspects. However, precise computations of the Gromov–Hausdorff distance between even simple spaces are mostly unknown, with certain exceptions such as the Gromov–Hausdorff distance between a line segment and a circle [16] and between spheres of certain dimensions [19, 1, 15]. Even approximating the Gromov–Hausdorff distance by a factor of 3 in the case of trees with unit-edge length is known to be NP-hard [4, 23]. A better approximation factor can only be achieved in very special cases; for example, [20] provides a $\frac{5}{4}$ -approximation scheme if



© Henry Adams, Sushovan Majhi, Fedor Manin, Žiga Virk, and Nicolò Zava; licensed under Creative Commons License CC-BY 4.0

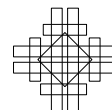
42nd International Symposium on Computational Geometry (SoCG 2026).

Editors: Hee-Kap Ahn, Michael Hoffmann, and Amir Nayyeri; Article No. 3; pp. 3:1–3:16

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



X and Y are finite subsets of the real line. Developing tools and pipelines to estimate the Gromov–Hausdorff distance is a central research direction in the applied topology community. To estimate the Gromov–Hausdorff distance between two metric spaces X, Y , one may:

- (A) Find nice samples $X' \subseteq X$ and $Y' \subseteq Y$ approximating the geometry of X and Y , and
- (B) Efficiently bound the Gromov–Hausdorff distance between the subsets X' and Y' .

Indeed, when X and Y are infinite continuous objects – e.g., Riemannian manifolds, metric graphs – and one wants to use computational machinery, it is often essential to resort to finite subsets approximating the geometry of the original spaces.

We reformulate the first task (A) as follows: how dense does the sample X' have to be in X to capture the geometry of the ambient space X ? To systematize this question, we use the Hausdorff distance as a measure of density and ask if we can provide a threshold ε and constant $0 < C \leq 1$ such that $d_{\text{GH}}(X, X') \geq C \cdot d_{\text{H}}(X, X')$ whenever $d_{\text{H}}(X, X') < \varepsilon$. In particular, $d_{\text{GH}}(X, X') = d_{\text{H}}(X, X')$ if $C = 1$. In [2], the authors show that when X is a manifold and $d_{\text{GH}}(X, X') < \varepsilon$ (which is implied by $d_{\text{H}}(X, X') < \varepsilon$) for some constant ε , then $d_{\text{GH}}(X', X) \geq C \cdot d_{\text{H}}(X, X')$ for some $\frac{1}{2} \leq C \leq 1$. Our study continues that investigation and proves improved results when X is a metric graph, hence relaxing the manifold assumption.

A fruitful approach to provide bounds on the Gromov–Hausdorff distance in task (B) is to use stable metric invariants. One associates each metric space X with a value $\psi(X)$ in a simpler metric space so that $\psi(X)$ and $\psi(Y)$ are close whenever X and Y are. An immediate example is the diameter; we have $d_{\text{GH}}(X, Y) \geq \frac{1}{2}|\text{diam}(X) - \text{diam}(Y)|$. For further stable metric invariants, we refer the reader to curvature sets [22], to Vietoris–Rips persistence diagrams [11], and to hierarchical clustering [10]. In [25], dimension theory is used to limit the precision of metric invariants with values in Hilbert spaces.

One of the precursors of algebraic topology was Brouwer’s proof of *invariance of dimension*, namely that \mathbb{R}^m is not homeomorphic to \mathbb{R}^n for $n > m$. Proving \mathbb{R}^1 is not homeomorphic to \mathbb{R}^n for $n > 1$ relies only on connectedness: removing a single point from \mathbb{R}^1 leaves a disconnected space. To prove invariance of dimension for $n > m \geq 2$, Brouwer relied on concepts related to the *degree* of a map between spheres [8, 7], which is often now formalized using higher-dimensional homology groups. The topological obstructions to small Gromov–Hausdorff distances in [2] relied on homological tools, namely the fundamental class of a manifold. One of the key contributions of our paper is to go “back in time” and produce strong lower bounds on Gromov–Hausdorff distances using simpler topological tools: we show that small Gromov–Hausdorff distances would contradict the connectedness of a space.

Overview of our results

Let G be a connected, finite metric graph, and let $X \subseteq G$ be a subset. We show how to lower bound the Gromov–Hausdorff distance $d_{\text{GH}}(G, X)$ in terms of the Hausdorff distance $d_{\text{H}}(G, X)$, achieving $d_{\text{GH}}(G, X) = d_{\text{H}}(G, X)$. A unifying theme in our proofs is that we show that a possibly discontinuous function $f: G \rightarrow X$ induces a continuous and non-surjective map $\bar{f}: G \rightarrow G$. We then use non-surjectivity (and related properties) to lower bound the distortion of \bar{f} and f , hence lower bounding $d_{\text{GH}}(G, X)$. The simplest metric graph is a tree T . Let ∂T denote the leaves of T , and let $\vec{d}_{\text{H}}(\partial T, X)$ denote the directed Hausdorff distance from ∂T to the sample X (see Section 2). We show:

Theorem 8. Let T be a finite metric tree and let $X \subseteq T$. If $d_{\text{H}}(T, X) > \vec{d}_{\text{H}}(\partial T, X)$, then $d_{\text{GH}}(T, X) = d_{\text{H}}(T, X)$. The interpretation is as follows: if the biggest gaps in the sampling $X \subseteq T$ are not near the leaves of T (which by Proposition 6 is typical behavior), then $d_{\text{GH}}(T, X) = d_{\text{H}}(T, X)$. Some control on the leaf vertices is needed:

Theorem 5. For any $\varepsilon > 0$ there exists a finite metric tree T and a closed subset $X \subseteq T$ such that $d_{\text{GH}}(T, X) \leq \varepsilon \cdot d_{\text{H}}(T, X)$.

When $G = S^1$ is the circle of circumference 2π , [2] shows that if $d_{\text{GH}}(S^1, X) < \frac{\pi}{6}$, then $d_{\text{GH}}(S^1, X) = d_{\text{H}}(S^1, X)$. The following theorem improves the constant $\frac{\pi}{6}$ to $\frac{\pi}{3}$.

Theorem 11. For any subset $X \subseteq S^1$, we have $d_{\text{GH}}(S^1, X) \geq \min \{d_{\text{H}}(S^1, X), \frac{\pi}{3}\}$.

In **Theorem 12** we show that this constant $\frac{\pi}{3}$ is the best possible, namely that for any $\varepsilon \in (0, \frac{\pi}{6})$, there exists a nonempty compact subset $X \subseteq S^1$ with $d_{\text{H}}(S^1, X) = \frac{\pi}{3} + \varepsilon$ and $d_{\text{GH}}(S^1, X) = \frac{\pi}{3} < d_{\text{H}}(S^1, X)$. We generalize the above results for the circle and trees to general metric graphs. Let G be a finite metric graph with set of leaf vertices ∂G . We prove that if X is dense enough relative to $e(G_0)$, which is the smallest edge length between vertices in the core G_0 of G (see Section 2), then the Gromov–Hausdorff distance matches the Hausdorff distance.

Theorem 23. Let G be a finite connected metric graph and let $X \subseteq G$. If $\partial G \neq \emptyset$, suppose $d_{\text{H}}(G, X) > \bar{d}_{\text{H}}(\partial G, X)$. If $d_{\text{GH}}(G, X) < \frac{e(G_0)}{16}$, then $d_{\text{GH}}(G, X) = d_{\text{H}}(G, X)$. Theorems 11 and 23 are obtained using low-dimensional versions of the Borsuk–Ulam theorem for maps from the circle into \mathbb{R} or into a tree, which are proven using connectedness. Our results have corollaries allowing us to lower bound the Gromov–Hausdorff distance between two subsets $X, Y \subseteq G$, using the triangle inequality. That is, for two sufficiently dense subsets $X, Y \subseteq G$, we provide new lower bounds on $d_{\text{GH}}(X, Y)$ in terms of the Hausdorff distance $d_{\text{H}}(X, Y)$, which is $\mathcal{O}(n^2)$ -time computable if the subsets have at most n points. We demonstrate such a corollary explicitly in the case of metric trees in Section 4.

Organization of the paper

We begin in Section 2 with background. In Section 3, we show that some control over the boundary of the metric graph is necessary. We begin with metric trees in Section 4, proceed to the circle in Section 5, and consider general metric graphs in Section 6.

2 Background and notation

We refer the reader to [9, 6] for more information on metric spaces, metric graphs, the Hausdorff distance, and the Gromov–Hausdorff distance.

2.1 Metric spaces

Let (Z, d) be a metric space. For $z \in Z$, we let $B(z; r) = \{z' \in Z \mid d(z, z') < r\}$ denote the open metric ball of radius r about z . For any $r \geq 0$ and $A \subseteq Z$, we denote the r -thickening of A in Z , namely, the union of metric balls of radius r centered at the points of A , by

$$A^r = \cup_{a \in A} B(a; r) = \left\{ z \in Z \mid \inf_{a \in A} d(a, z) < r \right\}.$$

The *diameter* of a subset $A \subseteq Z$ is the supremum of all distances between pairs of points of A . More formally, the diameter is defined as $\text{diam}(A) = \sup\{d(a, a') \mid a, a' \in A\}$.

2.2 Metric graphs

To define metric graphs, we follow Definitions 3.1.12 (quotient metric) and 3.2.9 (metric graph) of [9]. A *metric segment* is a metric space isometric to a real line segment $[0, l]$ for $l \geq 0$. Our metric graphs will be connected and have a finite number of vertices and edges.

► **Definition 1 (Metric Graph).** A finite metric graph (G, d) is the metric space obtained by gluing a finite collection E of metric segments along their boundaries to a finite collection V of vertices, and equipping the result $(\bigsqcup_{e \in E} e) / \sim$ with the quotient metric. We assume this equivalence relation results in a connected space.

The metric segments $e \in E$ are the *edges* and the equivalence classes of their endpoints in V are the *vertices* of G . Since G is connected, we have $d(y, y') < \infty$ for all $y, y' \in G$. Since E is finite, the quotient semi-metric [9, Definition 3.1.12] is in fact a metric.¹ For an edge $e \in E$, its length $|e|$ is the length of the corresponding metric segment. Note that the length can be different from the distance between the endpoints of e according to the metric d . Extending the definition of length for a continuous path $\gamma: [a, b] \rightarrow G$, we define the length $|\gamma|$ as the sum of the (full or partial) lengths of the edges that γ traverses. For a vertex v of a graph, its *degree*, denoted $\deg(v)$, is the number of edges incident to v . A self-loop at v contributes 2 to the degree of v . As an example, the circle S^1 can be given a metric graph structure with a single vertex v of degree 2 at the north pole, along with a self-loop. We denote the set of leaf vertices, i.e. the vertices of degree one, by ∂G . A metric graph G is therefore endowed with two layers of information: a metric structure turning it into a length space [9, Definition 2.1.6] and a combinatorial structure (V, E) as an abstract graph. Using the above definition, we are also allowing G to have single-edge cycles and multiple edges between a pair of vertices. In this paper, we only consider path connected metric graphs (G, d) with finitely many edges. Therefore, the metric graphs we consider are compact, and since a compact length space is geodesic, they are also geodesic (meaning for every $x, y \in G$ there is an isometric embedding $\gamma: [0, d(x, y)] \rightarrow G$ such that $\gamma(0) = x$ and $\gamma(d(x, y)) = y$). This justifies why for any $y, y' \in G$, there is always at least one *shortest path* or *geodesic path* γ in G joining y, y' whose length satisfies $|\gamma| = d(y, y')$. We define an undirected *simple loop* γ of G to be a simple path that intersects itself only at the endpoints. We denote by $\mathcal{L}(G)$ the finite set of all simple loops of the finite graph G . Since a *metric tree* is a metric graph with no loops, $\mathcal{L}(G) = \emptyset$ for a metric tree G . For a metric graph G , let $e(G)$ denote the length of the shortest edge in G . If for $A \subseteq G$ there exists a unique shortest geodesic in G between any points $a, a' \in A$, then we define the (*geodesic*) *convex hull* of A in G to be the union of geodesics between all a and a' . If $\text{diam}(A) < \frac{e(G)}{2}$, then the convex hull of A is well-defined. Given a (connected and finite) metric graph G , let G_0 be the smallest connected subgraph of G containing the union of all simple loops of G . We call G_0 the *core* of G . If G has no leaf edges, then $G = G_0$. Note that G_0 is just G with its “hanging trees” removed, or in other words, the minimal subspace onto which G deformation retracts. We will use the following two properties about G_0 . First, if G (and hence G_0) contains at least two simple loops, then at least one vertex of G_0 has degree at least 3. We can then consider the canonical graph representation of G_0 consisting of only vertices of degree at least 3. Second, we note that if a path connected subset $Z \subseteq G$ has first Betti number equal to the first Betti number of G , then $G_0 \subseteq Z$.

¹ One needs to be careful if E is infinite. Suppose one has two vertices v and v' , and one edge e_n of length $\frac{1}{n}$ with endpoints v and v' for each $n \in \mathbb{N}$. The quotient semi-metric gives $d(v, v') = \inf_{n \in \mathbb{N}} \frac{1}{n} = 0$ though $v \neq v'$; this semi-metric is not a metric. These subtleties disappear since we assume E is finite.

2.3 Hausdorff distances

Let (Z, d) be a metric space. If X and Y are two subsets of Z , then the *directed Hausdorff distance* from X to Y , denoted $\vec{d}_H(X, Y)$, is defined by

$$\vec{d}_H(X, Y) = \sup_{x \in X} \inf_{y \in Y} d(x, y) = \inf \left\{ r \geq 0 \mid X \subseteq Y^r \right\}.$$

Note that $\vec{d}_H(X, Y)$ is not symmetric in general. To retain symmetry, the *Hausdorff distance*, denoted $d_H(X, Y)$, is defined as $d_H(X, Y) = \max \left\{ \vec{d}_H(X, Y), \vec{d}_H(Y, X) \right\}$. In other words, the Hausdorff distance finds the infimum over all real numbers r such that if we thicken Y by r it contains X , and if we thicken X by r it contains Y . If no such r exists then the Hausdorff distance is infinity.

2.4 Gromov–Hausdorff distances

Let (X, d_X) and (Y, d_Y) be metric spaces. In equations (1)–(3) we give three equivalent definitions of the Gromov–Hausdorff distance between X and Y . The first definition follows [14]. The *Gromov–Hausdorff distance* $d_{GH}(X, Y)$ is the infimum of the Hausdorff distances between $F(X)$ and $G(Y)$ in Z , with the infimum being over all metric spaces Z and all isometric embeddings $F: X \hookrightarrow Z$ and $G: Y \hookrightarrow Z$:

$$d_{GH}(X, Y) = \inf_{F, G, Z} d_H(F(X), G(Y)). \quad (1)$$

The second definition uses correspondences. A relation $\mathcal{R} \subseteq X \times Y$ is a *correspondence* if for every $x \in X$ there is some $y \in Y$ with $(x, y) \in \mathcal{R}$, and for every $y \in Y$ there is some $x \in X$ exists with $(x, y) \in \mathcal{R}$. The *distortion* of a correspondence \mathcal{R} is

$$\text{dis}(\mathcal{R}) = \sup_{(x, y), (x', y') \in \mathcal{R}} |d_X(x, x') - d_Y(y, y')|.$$

The *Gromov–Hausdorff distance* between two (arbitrary) metric spaces X and Y is

$$d_{GH}(X, Y) = \frac{1}{2} \inf_{\mathcal{R} \subseteq X \times Y} \text{dis}(\mathcal{R}), \quad (2)$$

where the infimum is taken over all correspondences \mathcal{R} between X and Y [9, 17]. For metric spaces (X, d_X) and (Y, d_Y) , let $f: X \rightarrow Y$ and $g: Y \rightarrow X$ be two (possibly discontinuous) functions between them. We define the *distortion* of f as

$$\text{dis}(f) = \sup_{x, x' \in X} |d_X(x, x') - d_Y(f(x), f(x'))|.$$

The distortion of g is defined similarly. The *codistortion* of f and g is defined as

$$\text{codis}(f, g) = \sup_{x \in X, y \in Y} |d_X(x, g(y)) - d_Y(f(x), y)|.$$

Kalton and Ostrovskii [17] show that

$$d_{GH}(X, Y) = \frac{1}{2} \inf_{f, g} \max \{ \text{dis}(f), \text{dis}(g), \text{codis}(f, g) \}, \quad (3)$$

where $f: X \rightarrow Y$ and $g: Y \rightarrow X$ are any functions. It follows that if $d_{GH}(X, Y) < C$, then there is some function $f: X \rightarrow Y$ with $\frac{1}{2} \text{dis}(f) < C$.

2.5 Gromov–Hausdorff distances and compactness

The following result is folklore; see [3] for the proof.

► **Lemma 2.** *Let X be a metric space and Y be a subset thereof. Let us denote by \bar{Y} the closure of Y in X . Then*

$$d_{\text{H}}(X, Y) = d_{\text{H}}(X, \bar{Y}) \quad \text{and} \quad d_{\text{GH}}(X, Y) = d_{\text{GH}}(X, \bar{Y}).$$

Furthermore, for every two subsets Y and Z of X , we have $\vec{d}_{\text{H}}(Z, Y) = \vec{d}_{\text{H}}(Z, \bar{Y})$.

We will consistently apply Lemma 2 when the ambient metric space is compact. Thus, when we are studying the Hausdorff and the Gromov–Hausdorff distances between a compact metric space and a subspace thereof, we can assume that the subspace is compact as well.

2.6 Gromov–Hausdorff distances and connectedness

► **Lemma 3.** *Let $f: X \rightarrow Y$ be a function between metric spaces, with X path-connected and Y geodesic. Then for every $r > \frac{\text{dis}(f)}{2}$, the thickening $f(X)^r$ is path connected.*

Proof. Let $y_1, y_2 \in f(X)^r$ and $x_1, x_2 \in X$ be points such that $d_Y(y_i, f(x_i)) < r$ for $i = 1, 2$. Since Y is geodesic, the entire geodesic connecting y_i to $f(x_i)$ is contained in $B(f(x_i), r) \subseteq f(X)^r$. Pick $\varepsilon > 0$ sufficiently small to satisfy $r > \frac{\text{dis}(f)}{2} + \varepsilon$. Since X is path connected, there are $z_1 = x_1, z_2, \dots, z_{n+1} = x_2$ such that $d(z_j, z_{j+1}) < \varepsilon$ for every $j = 0, \dots, n$. Thus, $d(f(z_j), f(z_{j+1})) < \text{dis}(f) + \varepsilon$. Since Y is geodesic, there are $m_j \in Y$ such that $d(f(z_j), m_j) = d(f(z_{j+1}), m_j) = \frac{1}{2}d(f(z_j), f(z_{j+1})) < r$ and the geodesics connecting m_j to $f(z_j)$ and $f(z_{j+1})$ are contained in $f(X)^r$. ◀

2.7 Gromov–Hausdorff distances and graphs

Let G be a finite metric graph and let $X \subseteq G$ be a compact subset. To use functions $f: G \rightarrow X$ to prove bounds on the Gromov–Hausdorff distance, we need to be a bit careful. Indeed, even if the sample X is close to the leaves of G , it may be that $f(G)$ is not close to the leaves. Luckily, we have a lot of freedom to construct our function f :

► **Lemma 4.** *Let G be a finite metric graph and let $X \subseteq G$ be compact. If $d_{\text{GH}}(G, X) < r$, then there is a function $f: G \rightarrow X$ with $\frac{1}{2}\text{dis}(f) < r$ and $\vec{d}_{\text{H}}(\partial G, f(G)) = \vec{d}_{\text{H}}(\partial G, X)$.*

Proof. Let $\mathcal{R} \subseteq G \times X$ be a correspondence satisfying $\text{dis}(\mathcal{R}) < 2r$. Fix $\varepsilon > 0$ such that $\text{dis}(\mathcal{R}) + \varepsilon < 2r$. For every $v \in \partial G$, pick $x_v \in X$ as a point satisfying $d(v, x_v) = \min_{x \in X} d(v, x)$. We first recursively construct a surjective map \tilde{f} from a subset of G to $\{x_v \mid v \in \partial G\}$. Let us arbitrarily order the leaves v_1, \dots, v_n and denote the point x_{v_i} by x_i for the sake of simplicity. **Base step.** Pick any $y_1 \in G$ such that $(y_1, x_1) \in \mathcal{R}$ and set $\tilde{f}(y_1) = x_1$. **Recursive step.** Assume we have defined $y_1, \dots, y_m \in G$ distinct and set $\tilde{f}(y_i) = x_i$. Pick $y \in G$ satisfying $(y, x_{m+1}) \in \mathcal{R}$. Let us distinguish two cases. If $y \notin \{y_1, \dots, y_m\}$, then set $y_{m+1} = y$ and $\tilde{f}(y_{m+1}) = x_{m+1}$. Otherwise, fix an arbitrary point $y_{m+1} \in G \setminus \{y_1, \dots, y_m\}$ satisfying $d(y_{m+1}, y) < \frac{\varepsilon}{2}$ and define $\tilde{f}(y_{m+1}) = x_{m+1}$. Consider an arbitrary extension $f: G \rightarrow X$ of $\tilde{f}: \{y_1, \dots, y_n\} \rightarrow \{x_1, \dots, x_n\}$ with the property that the graph of $f|_{G \setminus \{y_1, \dots, y_n\}}$ is contained in \mathcal{R} . Then $\text{dis}(f) \leq \text{dis}(\mathcal{R}) + 2 \cdot \frac{\varepsilon}{2} < 2r$ and $\vec{d}_{\text{H}}(\partial G, f(G)) = \vec{d}_{\text{H}}(\partial G, X)$ by construction. ◀

It is worth noticing that in the proof above we did not use any property of ∂G , and the result can be shown with ∂G replaced by any finite subset Y of G .

3 Ratio of d_{GH} vs d_H can be arbitrarily small for metric trees

In this section we prove Theorem 5, a variant of Theorem 5 of [2], now adapted to metric trees. As we show in Theorems 8 and 23, the Gromov–Hausdorff distance $d_{GH}(T, X)$ between a metric graph T and a subset $X \subseteq T$ can be bounded below by their Hausdorff distance $d_H(T, X)$ when X is close enough to the leaves of T , i.e., when $d_H(T, X) > \vec{d}_H(\partial T, X)$. When $d_H(T, X) < \vec{d}_H(\partial T, X)$, however, the ratio of $d_{GH}(T, X)$ over $d_H(T, X)$ can become arbitrarily small as we show below. Furthermore, Theorem 5 suggests that ∂T may not map close to itself via correspondences with small distortion; see Figure 1.

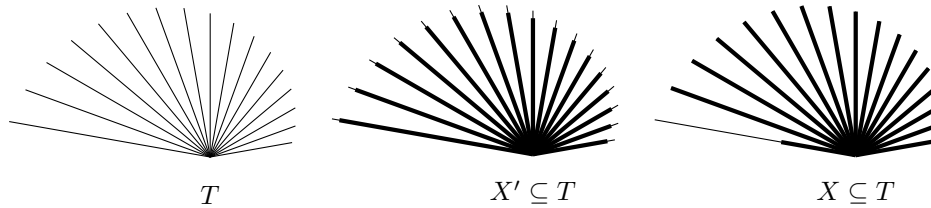


Figure 1 A sketch of the construction of Theorem 5: (left) The metric tree T , (middle) the bold superimposed $X' \subseteq T$ with small $d_H(T, X')$, and (right) a different isometric embedding of X' into T indicated by $X \subseteq T$ with large $d_H(T, X)$.

► **Theorem 5.** *For any $\varepsilon > 0$ there exists a finite metric tree T and a closed subset $X \subseteq T$ such that $d_{GH}(T, X) \leq \varepsilon \cdot d_H(T, X)$.*

Proof. Choose $n \in \mathbb{N}$ and let $\varepsilon = \frac{1}{n}$. For $t \geq 0$, let $I(t)$ be a copy of the interval $[0, t]$ with a designated basepoint 0. Define a metric tree T as the wedge of n intervals of lengths $1 + \varepsilon, 1 + 2\varepsilon, \dots, 1 + n\varepsilon = 2$, and define a metric tree X_0 as the wedge of intervals of lengths $1, 1 + \varepsilon, 1 + 2\varepsilon, \dots, 1 + (n - 1)\varepsilon = 2 - \varepsilon$:

$$T = \bigvee_{t=1}^n I(1 + t \cdot \varepsilon) \quad \text{and} \quad X_0 = \bigvee_{t=0}^{n-1} I(1 + t \cdot \varepsilon).$$

Recall from (1) that $d_{GH}(T, X)$ is the infimum of the Hausdorff distances between $F(T)$ and $G(X)$ in Z , with the infimum being over all metric spaces Z and all isometric embeddings $F: T \hookrightarrow Z$ and $G: X \hookrightarrow Z$. In Figure 1 we have two different isometric embeddings X', X of the same tree X_0 into $Z = T$: one with a small Hausdorff distance $d_H(T, X') = \varepsilon$ as X' is T with each ray being shortened by ε , and one with a large Hausdorff distance $d_H(T, X) = 1$ as X is T with the longest ray $I(2)$ of length 2 being replaced with a subray $I(1)$ of length 1. By (1), we have $d_{GH}(T, X) \leq d_H(T, X') = \varepsilon d_H(T, X)$. ◀

4 Metric trees with coefficient 1

Section 3 demonstrates that results claiming $d_{GH}(G, X) = d_H(G, X)$ for X a subset of a metric graph G do not hold in general, even if $d_H(G, X)$ is arbitrarily small. In this section, as well as in Section 6, we prove related results relying on bounds on $\vec{d}_H(\partial G, X)$. In particular, we will assume the condition $d_H(G, X) > \vec{d}_H(\partial G, X)$ which rules out the example constructed in Theorem 5. The following proposition shows that condition has a high probability of being satisfied if X is a large uniform sample of G ; see [3] for the proof.

► **Proposition 6.** *If G is a metric graph and X is a sample of n uniformly random points, then with high probability as $n \rightarrow \infty$, we have $d_H(G, X) > \vec{d}_H(\partial G, X)$.*

We prove the following results for metric trees.

► **Proposition 7.** *Let T be a finite tree, and let $f: T \rightarrow T$ be a function. If $d_{\text{H}}(T, f(T)) > \vec{d}_{\text{H}}(\partial T, f(T))$, then $\frac{1}{2}\text{dis}(f) \geq d_{\text{H}}(T, f(T))$.*

Proof. Let $\frac{1}{2}\text{dis}(f) < r$ and let $\bar{x} \in T$ be a point such that $B(\bar{x}, d_{\text{H}}(T, f(T))) \cap f(T) = \emptyset$. The assumption $d_{\text{H}}(T, f(T)) > \vec{d}_{\text{H}}(\partial T, f(T))$ implies that $T \setminus \{\bar{x}\}$ is disconnected and at least two distinct components thereof have non-empty intersection with $f(T)$ (since a tree has at least two leaves). Since $f(T)^r$ is connected by Lemma 3, we have $\bar{x} \in f(T)^r$, which implies $d_{\text{H}}(T, f(T)) < r$. Since r can be taken arbitrarily, $d_{\text{H}}(T, f(T)) \leq \frac{1}{2}\text{dis}(f)$. ◀

► **Theorem 8.** *Let T be a finite metric tree and let $X \subseteq T$. If $d_{\text{H}}(T, X) > \vec{d}_{\text{H}}(\partial T, X)$, then $d_{\text{GH}}(T, X) = d_{\text{H}}(T, X)$.*

Proof. By Lemma 2, we may assume that X is compact. Suppose for a contradiction that $d_{\text{GH}}(T, X) < d_{\text{H}}(T, X)$. Then by Lemma 4, there is a function $f: T \rightarrow X \subseteq T$ with $\frac{1}{2}\text{dis}(f) < d_{\text{H}}(T, X) \leq d_{\text{H}}(T, f(T))$ and $\vec{d}_{\text{H}}(\partial T, f(T)) = \vec{d}_{\text{H}}(\partial T, X) < d_{\text{H}}(T, X) \leq d_{\text{H}}(T, f(T))$, contradicting Proposition 7. ◀

In other words, $d_{\text{H}}(T, X) = \max\{d_{\text{GH}}(T, X), \vec{d}_{\text{H}}(\partial T, X)\}$. All of the results in our paper have simple corollaries for two subsets of the tree or graph; we give one such example.

► **Corollary 9.** *Let T be a finite metric tree and let $X, Y \subseteq T$. For any $\varepsilon \geq d_{\text{H}}(T, Y)$, if $d_{\text{H}}(X, Y) > \vec{d}_{\text{H}}(\partial T, X) + \varepsilon$, then $d_{\text{GH}}(X, Y) \geq d_{\text{H}}(X, Y) - 2\varepsilon$.*

Proof. We need only apply the triangle inequality:

$$\begin{aligned}
d_{\text{H}}(X, Y) &\leq d_{\text{H}}(T, X) + d_{\text{H}}(T, Y) \\
&= \max\left\{d_{\text{GH}}(T, X), \vec{d}_{\text{H}}(\partial T, X)\right\} + d_{\text{H}}(T, Y) && \text{by Theorem 8} \\
&\leq \max\left\{d_{\text{GH}}(X, Y) + d_{\text{GH}}(T, Y), \vec{d}_{\text{H}}(\partial T, X)\right\} + d_{\text{H}}(T, Y) \\
&\leq \max\left\{d_{\text{GH}}(X, Y) + d_{\text{H}}(T, Y), \vec{d}_{\text{H}}(\partial T, X)\right\} + d_{\text{H}}(T, Y) && \text{since } d_{\text{GH}}(T, Y) \leq d_{\text{H}}(T, Y) \\
&= \max\left\{d_{\text{GH}}(X, Y) + 2d_{\text{H}}(T, Y), \vec{d}_{\text{H}}(\partial T, X) + d_{\text{H}}(T, Y)\right\} \\
&\leq \max\left\{d_{\text{GH}}(X, Y) + 2\varepsilon, \vec{d}_{\text{H}}(\partial T, X) + \varepsilon\right\} && \text{since } d_{\text{H}}(T, Y) \leq \varepsilon. \quad \blacktriangleleft
\end{aligned}$$

As a particular case of Theorem 8, we can provide a closed formula for the Gromov–Hausdorff distance between a segment and a sample of it; see [3] for the proof.

► **Corollary 10.** *Let X be a compact non-empty subset of an interval $I = [a, b]$. Denote by $c = \min X$ and $d = \max X$. Then $d_{\text{GH}}(I, X) = \max\left\{\frac{c-a+b-d}{2}, d_{\text{H}}([c, d], X)\right\}$.*

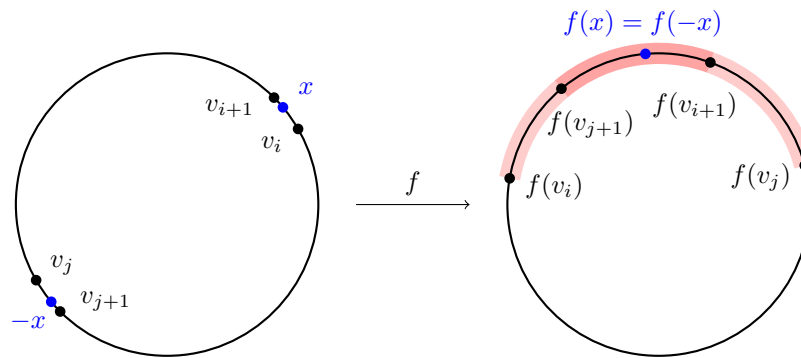
5 The case of the circle

Before considering general metric graphs in the following section, in this section we first understand the simplest connected metric graph which is not a tree: the circle $G = S^1$. The authors of [2] showed that $d_{\text{GH}}(S^1, X) = d_{\text{H}}(S^1, X)$ for a compact subset $X \subseteq S^1$ with $d_{\text{GH}}(S^1, X) < \frac{\pi}{6}$. A curious question regarding the optimality of the density constant $\frac{\pi}{6}$ was also posed therein [2, Question 2]. While the constant $\frac{\pi}{6}$ provides a sufficient Gromov–Hausdorff density of a subset X to guarantee the equality of its Hausdorff and Gromov–Hausdorff distances to S^1 , it turns out this constant is suboptimal. In this section,

we show that $\frac{\pi}{3}$ is the optimal constant in the case of the circle. That is, for a subset $X \subseteq S^1$, the density condition $d_{GH}(S^1, X) < \frac{\pi}{3}$ implies $d_{GH}(S^1, X) = d_H(S^1, X)$ (Theorem 11), and there is a compact $X \subseteq S^1$ with $\frac{\pi}{3} = d_{GH}(S^1, X) < d_H(S^1, X)$ (Theorem 12).

► **Theorem 11.** For any subset $X \subseteq S^1$, we have $d_{GH}(S^1, X) \geq \min \{d_H(S^1, X), \frac{\pi}{3}\}$.

Proof. This proof is closely related to [18, 12]. It suffices to show that if $d_{GH}(S^1, X) < \frac{\pi}{3}$, then $d_{GH}(S^1, X) \geq d_H(S^1, X)$. Let $f: S^1 \rightarrow X$ be a (possibly discontinuous) function. By (3), it suffices to show that if $\frac{1}{2}\text{dis}(f) < \frac{\pi}{3}$, then $\frac{1}{2}\text{dis}(f) \geq d_H(S^1, X)$, or equivalently that $\frac{\varepsilon + \text{dis}(f)}{2} \geq d_H(S^1, X)$ for all $\varepsilon > 0$ sufficiently small. Suppose $\frac{1}{2}\text{dis}(f) < \frac{\pi}{3}$. Fix $\varepsilon > 0$ with $5\varepsilon < 2\pi - 3\text{dis}(f)$. Form a finite triangulation K of S^1 with edges of length less than ε . Let $K^{(0)} \subseteq S^1$ be the vertex set of K . Since adjacent vertices $v_i, v_{i+1} \in K^{(0)}$ satisfy $d(v_i, v_{i+1}) < \varepsilon$, we have $d(f(v_i), f(v_{i+1})) < \varepsilon + \text{dis}(f) < \pi$. Hence we can map the edge between v_i and v_{i+1} to the unique shortest path in S^1 between $f(v_i)$ and $f(v_{i+1})$, obtaining a continuous map $\bar{f}: K \rightarrow S^1$. Let $r = \frac{\varepsilon + \text{dis}(f)}{2}$. Suppose for a contradiction that X^r is



■ **Figure 2** Proof of Theorem 11: Since $\bar{f}(x) = \bar{f}(-x)$, the shortest paths in S^1 between $f(v_i), f(v_{i+1})$ and $f(v_j), f(v_{j+1})$ intersect.

not equal to all of S^1 . Then there is an arc of length at least $2r$ in S^1 that contains no points from X . So \bar{f} is non-surjective, since the image of each edge in K is an edge of length less than $2r$. Hence $\bar{f}: K \rightarrow S^1$ is a map of degree zero. The Borsuk–Ulam theorem [5, Page 206, Exercise 21] then implies there are antipodal points $x, -x \in K$ with $\bar{f}(x) = \bar{f}(-x)$. Let x be on the closed edge between $v_i, v_{i+1} \in K^{(0)}$, and let $-x$ be on the closed edge between $v_j, v_{j+1} \in K^{(0)}$. Since $\bar{f}(x) = \bar{f}(-x)$, the shortest path in S^1 between $f(v_i), f(v_{i+1})$ intersects the shortest path between $f(v_j), f(v_{j+1})$. Since these paths in S^1 are of length less than $2r$, one of $f(v_i), f(v_{i+1})$ is at a distance less than r from one of $f(v_j), f(v_{j+1})$, i.e. $d(f(v), f(v')) < r$ for some $v \in \{v_i, v_{i+1}\}$ and $v' \in \{v_j, v_{j+1}\}$; see Figure 2. We obtain

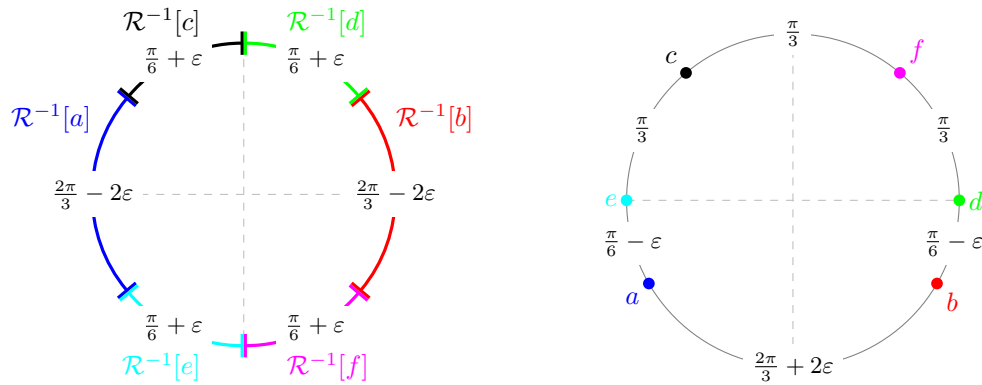
$$\text{dis}(f) > d(v, v') - d(f(v), f(v')) > (\pi - 2\varepsilon) - \frac{\varepsilon + \text{dis}(f)}{2} = \pi - \frac{\text{dis}(f)}{2} - \frac{5\varepsilon}{2}.$$

So $\frac{5\varepsilon}{2} > \pi - \frac{3\text{dis}(f)}{2}$, contradicting the choice of ε . Hence it must be that $X^r = S^1$. Since $\frac{\varepsilon + \text{dis}(f)}{2} = r \geq d_H(S^1, X)$ for all $\varepsilon > 0$ sufficiently small, we are done. ◀

We now prove the optimality of $\frac{\pi}{3}$ for the circle, i.e. that $\frac{\pi}{3}$ is also a necessary condition.

► **Theorem 12.** For any $\varepsilon \in (0, \frac{\pi}{6})$, there exists a nonempty compact subset $X \subseteq S^1$ with $d_H(S^1, X) = \frac{\pi}{3} + \varepsilon$ and $d_{GH}(S^1, X) = \frac{\pi}{3} < d_H(S^1, X)$.

Proof. We construct X to be a six-element subset of S^1 . Both X and a particular correspondence $\mathcal{R} \subseteq S^1 \times X$ are shown in Figure 3. One can check that $d_H(S^1, X) = \frac{\pi}{3} + \varepsilon$. Also, one can check that $\text{dis}(\mathcal{R}) = \frac{2\pi}{3}$, giving $d_{GH}(S^1, X) \leq \frac{\pi}{3}$. So $d_{GH}(S^1, X) < d_H(S^1, X)$. Moreover, Theorem 11 implies that $d_{GH}(S^1, X) = \frac{\pi}{3}$. ◀



(a) Correspondence $\mathcal{R} \subseteq S^1 \times X$ has a distortion of $\frac{2\pi}{3}$. So, $d_{\text{GH}}(S^1, X) \leq \frac{\pi}{3}$. (b) $X = \{a, b, c, d, e, f\} \subseteq S^1$ with $d_{\text{H}}(S^1, X) = \frac{\pi}{3} + \epsilon$.

■ **Figure 3** The configuration of points $X \subseteq S^1$ is depicted to show the optimal constant for the circle.

6 Metric graphs with loops with coefficient 1

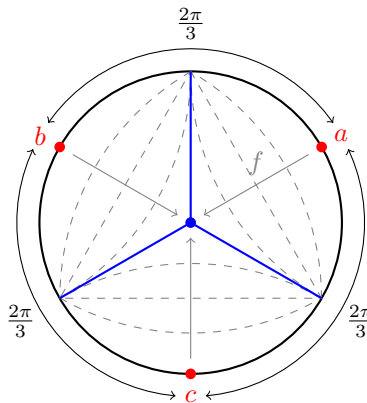
In this section, we extend our results for the circle to general metric graphs; see Theorem 23.

6.1 Distortion of continuous maps between graphs

We first need a Borsuk–Ulam theorem for maps into trees; see [3] for the proof.

► **Lemma 13.** *If T is a tree and $f: S^1 \rightarrow T$ is continuous, then there exist $x, x' \in S^1$ with $d(x, x') \geq \frac{2\pi}{3}$ and $f(x) = f(x')$.*

Note that the lower bound $\frac{2\pi}{3}$ in Lemma 13 cannot be improved, as shown in Figure 4. Recall

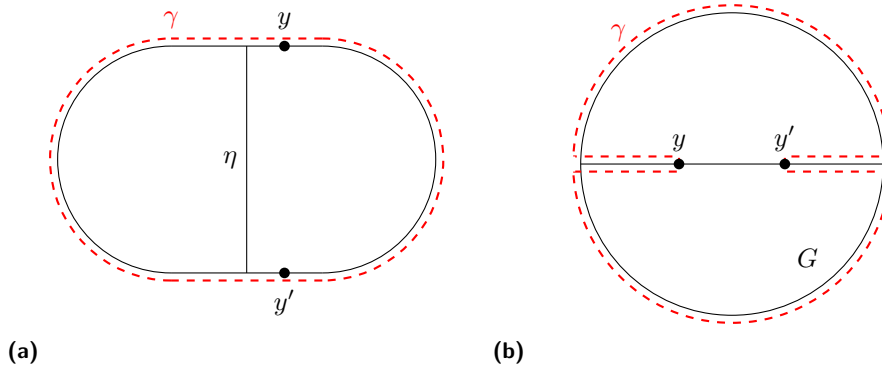


■ **Figure 4** The lower bound in Lemma 13 is tight. The function f sends the black circle to the blue tripod. The points a, b and c are at distance $\frac{2\pi}{3}$ apart and in the same fiber of f .

that G_0 is the core of a metric graph G (see Section 2), and that $e(G_0)$ is the length of the shortest edge in G_0 .

► **Lemma 14.** *Let (G, d) be a metric graph and γ be a simple loop. If $y, y' \in \gamma$ satisfy $d(y, y') < \frac{e(G_0)}{2}$, then there is a unique geodesic joining y and y' and it lies on γ .*

Proof. Let η be a simple path connecting y and y' . If η is not entirely in γ , its length is at least $e(G_0)$. So since $d(y, y') < \frac{e(G_0)}{2}$, a geodesic connecting the two points must lie on γ ; see Figure 5. This geodesic is unique since the length of a loop is at least $e(G_0)$. ◀



■ **Figure 5** The assumptions in Lemma 14 are necessary. Loop γ is represented with a dashed red line. (a) If there is no bound on $e(G_0)$, then there may exist a shorter path connecting y and y' using a shortcut η . (b) If γ is not simple, then y and y' may be arbitrarily close and joined by geodesics that are not in γ .

The following lemma is the crucial step showing that a continuous map between finite graphs must send simple loops to simple loops, provided that the distortion is sufficiently small compared to the length of the shortest internal edge.

► **Lemma 15.** *Let $\bar{f}: (G, d_G) \rightarrow (G', d_{G'})$ be a continuous map between finite metric graphs. Assume that $\frac{\text{dis}(\bar{f})}{2} < \frac{\min\{e(G_0), e(G'_0)\}}{8}$. Then \bar{f} induces an injection $\Phi: \mathcal{L}(G) \rightarrow \mathcal{L}(G')$ between the simple loops of G and the simple loops of G' as follows: given a simple loop $\gamma \in \mathcal{L}(G)$, $\Phi(\gamma)$ is the only simple loop of G' contained in $\bar{f}(\gamma)$.*

Proof of Lemma 15 ◻. For readability purposes, let us divide the proof into several claims. First, let us show that the map Φ is well-defined, i.e., for every simple loop $\gamma \in \mathcal{L}(G)$ there is a simple loop of G' contained in $\bar{f}(\gamma)$ and that simple loop is unique.

▷ **Claim 16 (Existence).** If γ is a simple loop of G , then $\bar{f}(\gamma)$ contains a simple loop of G' .

Proof of Claim 16. Suppose for a contradiction that $\bar{f}(\gamma)$ does not contain a loop. Since $\bar{f}(\gamma)$ is connected, it is a tree. So Lemma 13 applies to give points $x, x' \in \gamma$ with $\bar{f}(x) = \bar{f}(x')$ and $d_\gamma(x, x') \geq \frac{\text{length}(\gamma)}{3}$ where d_γ is the geodesic distance on γ . We must have $d_G(x, x') \geq \frac{e(G_0)}{3}$, since if $d_G(x, x') < \frac{e(G_0)}{3} < \frac{e(G_0)}{2}$, then Lemma 14 would give that a geodesic connecting x and x' lies on γ , and so $d_G(x, x') = d_\gamma(x, x') \geq \frac{\text{length}(\gamma)}{3} \geq \frac{e(G_0)}{3}$. We obtain the contradiction

$$\text{dis}(\bar{f}) \geq |d_G(x, x') - d_{G'}(\bar{f}(x), \bar{f}(x'))| = d_G(x, x') \geq \frac{e(G_0)}{3} > \frac{e(G_0)}{4} > \text{dis}(\bar{f}).$$

Hence $\bar{f}(\gamma)$ must contain a loop. ◀

▷ **Claim 17 (Uniqueness).** If γ is a simple loop, then $\bar{f}(\gamma)$ contains at most one simple loop.

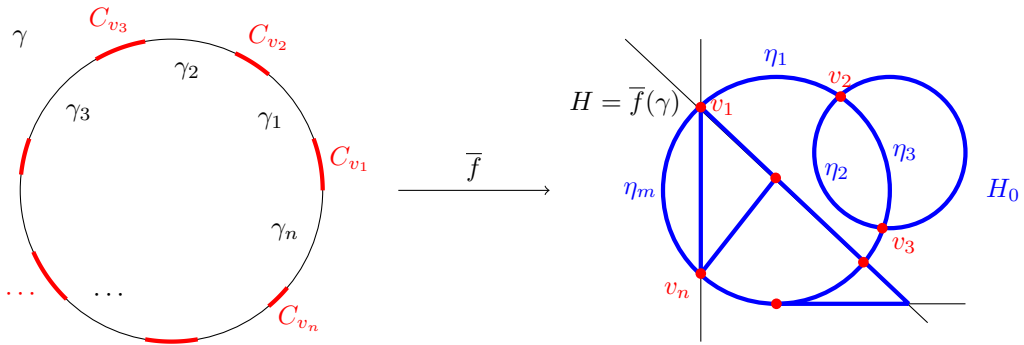
Proof of Claim 17. Note that $\bar{f}(\gamma)$ is a closed subset of G' since it is the continuous image of a compact subset. Let $H = \bar{f}(\gamma)$, which is a connected subgraph of G' . Furthermore, let H_0 be the smallest connected subgraph of G' containing the union of all simple loops of H . Suppose for a contradiction that $\bar{f}(\gamma) = H$ contains at least two simple loops. Therefore,

3:12 Lower Bounding the Gromov–Hausdorff Distance in Metric Graphs

at least one vertex of H_0 has degree at least 3. We can then consider the canonical graph representation of H_0 consisting of only vertices of degree at least 3. For every $v \in V(H_0)$, the closed subset $(\bar{f})^{-1}(v)$ has diameter at most $\text{dis}(\bar{f}) < \frac{e(G'_0)}{2}$. Hence, the convex hull C_v of $(\bar{f})^{-1}(v) \cap \gamma$ is well-defined. Furthermore, for every $v, v' \in V(H_0)$ with $v \neq v'$, we have $C_v \cap C_{v'} = \emptyset$. Indeed, for $x \in (\bar{f})^{-1}(v)$ and $x' \in (\bar{f})^{-1}(v')$ with $v \neq v'$, we have

$$d_G(x, x') \geq d_{G'}(v, v') - \text{dis}(\bar{f}) \geq e(G'_0) - \text{dis}(\bar{f}) > \text{dis}(\bar{f}) \geq \max\{\text{diam}C_v, \text{diam}C_{v'}\}.$$

Let $|V(H_0)| = n$ and let $\gamma_1, \dots, \gamma_n$ be the non-empty connected components of $\gamma \setminus \bigcup_{v \in V(H_0)} C_v$. For every $i \in \{1, \dots, n\}$, note that $\bar{f}(\gamma_i)$ is contained in a connected component of $H \setminus V(H_0)$. Let η_1, \dots, η_m be the connected components of $H_0 \setminus V(H_0)$. Note that each η_j is an open-



■ **Figure 6** A representation of the objects constructed in the proof of Claim 17.

ended path in H_0 . We claim that $\eta_j \cap \bar{f}(\gamma_1 \cup \dots \cup \gamma_n) \neq \emptyset$ for every $j \in \{1, \dots, m\}$. Let y be the midpoint of η_j . Since $H = \bar{f}(\gamma)$, there is $z \in \gamma$ such that $\bar{f}(z) = y$. For every $v \in V(H_0)$ and every $x \in (\bar{f})^{-1}(v)$,

$$d_G(z, x) \geq d_{G'}(y, v) - \text{dis}(\bar{f}) \geq \frac{e(G'_0)}{2} - \text{dis}(\bar{f}) > 2\text{dis}(\bar{f}) - \text{dis}(\bar{f}) = \text{dis}(\bar{f}) \geq \text{diam}C_v.$$

Hence, $z \notin C_v$ and so $z \in \gamma_1 \cup \dots \cup \gamma_n$. So far, we have shown that each $\bar{f}(\gamma_i)$, for $i \in \{1, \dots, n\}$, is contained in some connected component of $H \setminus V(H_0)$. Furthermore, among these connected components, each η_j with $j \in \{1, \dots, m\}$ contains the \bar{f} -image of at least one γ_i . It will be a contradiction if we prove that $n < m$. Note that the set $\{\eta_j\}_{j=1}^m$ is in bijection with the edge set $E(H_0)$ of H_0 . Hence

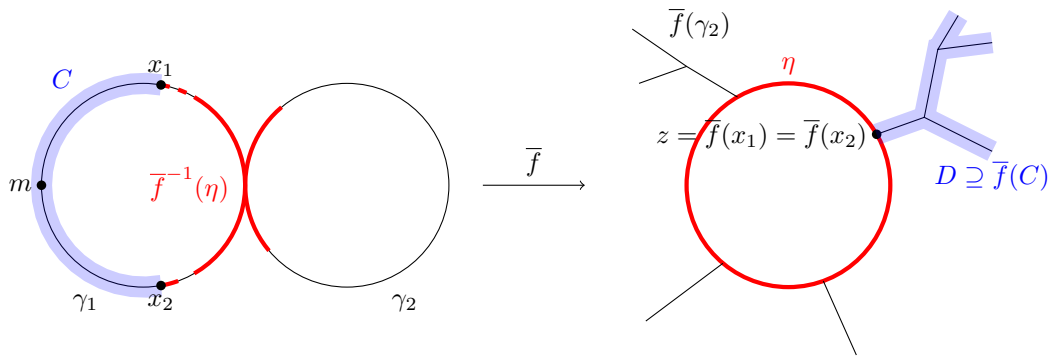
$$2m = 2|E(G'_0)| = \sum_{v \in V(G'_0)} \deg_{G'_0}(v) \geq 3|V(G'_0)| = 3n,$$

giving the contradiction $n < m$. It must be that $\bar{f}(\gamma)$ contains at most one simple loop. \triangleleft

According to Claims 16 and 17, map Φ is well-defined. Lastly, we show that it is injective.

▷ **Claim 18.** The map $\Phi: \mathcal{L}(G) \rightarrow \mathcal{L}(G')$ is injective.

Proof of Claim 18. Let γ_1, γ_2 , and η be three simple loops such that $\eta \subseteq \bar{f}(\gamma_1) \cap \bar{f}(\gamma_2)$. Suppose, by contradiction, that the simple loops γ_1 and γ_2 are distinct, meaning they differ by at least one edge. Without loss of generality, let us fix an edge $e \in E(G)$ that is contained in γ_1 but whose interior is disjoint from γ_2 . Fix the midpoint m of e . In particular,



■ **Figure 7** A representation of the objects constructed in the proof of Claim 18.

$d_G(m, \gamma_2) \geq \frac{e(G_0)}{2}$. Let us estimate $d_G(m, (\bar{f})^{-1}(\eta))$. Suppose that $x \in (\bar{f})^{-1}(\eta)$, and fix $y \in \gamma_2$ such that $\bar{f}(x) = \bar{f}(y)$. Note that $d_G(x, y) \leq \text{dis}(\bar{f})$. Using the triangle inequality, we obtain

$$\begin{aligned} d_G(m, x) &\geq d_G(m, y) - d_G(x, y) \geq d_G(m, \gamma_2) - \text{dis}(\bar{f}) \\ &\geq \frac{e(G_0)}{2} - \text{dis}(\bar{f}) > 2\text{dis}(\bar{f}) - \text{dis}(\bar{f}) = \text{dis}(\bar{f}). \end{aligned} \tag{4}$$

Hence $d(m, (\bar{f})^{-1}(\eta)) \geq \text{dis}(\bar{f})$, and in particular, $\bar{f}(m) \notin \eta$. Consider the connected component C of $\gamma_1 \setminus (\bar{f})^{-1}(\eta)$ containing m . Since $(\bar{f})^{-1}(\eta)$ is closed by continuity, C is an open arc in γ_1 . Let x_1 and x_2 be the two points of γ_1 on the boundary of C . Note that they both belong to $(\bar{f})^{-1}(\eta)$, and so $d_G(m, x_i) > \text{dis}(\bar{f})$ by (4). Let us estimate $d_G(x_1, x_2)$. A geodesic connecting x_1 and x_2 satisfies one of the following:

- (i) either it passes through m ,
- (ii) or it crosses $(\bar{f})^{-1}(\eta) \cap \gamma_1$,
- (iii) or it uses an edge that is not in γ_1 .

In all three cases, we have $d_G(x_1, x_2) > \text{dis}(\bar{f})$. For (i), this follows since $d_G(m, x_i) > \text{dis}(\bar{f})$. For (ii), let $A := (\bar{f})^{-1}(\eta) \cap \gamma_1$. Since $\bar{f}(\gamma_1)$ contains η , we have $\bar{f}(A) = \eta$. The diameter of η is at least $\frac{e(G'_0)}{2}$, and hence $\text{diam}(A) \geq \frac{e(G'_0)}{2} - \text{dis}(\bar{f}) > 2\text{dis}(\bar{f}) - \text{dis}(\bar{f}) = \text{dis}(\bar{f})$.

Finally, we also have $d_G(x_1, x_2) > \text{dis}(\bar{f})$ in case (iii) since $e(G_0) > \text{dis}(\bar{f})$. Since \bar{f} is continuous, $\bar{f}(C)$ is contained in a connected component D of $\bar{f}(\gamma_1) \setminus \eta$. Since $\bar{f}(\gamma_1)$ contains only one loop, D is a tree. Furthermore, $\bar{f}(\gamma_1)$ is a closed subset of G' since it is the continuous image of a compact subset, and it contains only one simple loop. Thus, the closure of D in G' intersects η in only one point z . We have already observed that $\bar{f}(x_i) \in \eta$ for $i \in \{1, 2\}$. Furthermore, $\bar{f}(x_i)$ belongs to the closure of $\bar{f}(C)$ in G' , which is contained in the closure of D . Therefore, $\bar{f}(x_i) = z$ since the latter is the only point in the closure of D that lies on η . This gives $\bar{f}(x_1) = z = \bar{f}(x_2)$, which contradicts the above established inequality $d_G(x_1, x_2) > \text{dis}(\bar{f})$. ◀

This completes the proof of Lemma 15. ◀

6.2 Distortion of arbitrary functions between graphs

To apply Lemma 15 in the context of Gromov–Hausdorff distances, we will need a way to turn arbitrary functions into nearby continuous ones with controlled distortion, as described in the following result.

▶ **Lemma 19.** *Let $f: (G, d_G) \rightarrow (G', d_{G'})$ be a function between finite metric graphs with distortion $\frac{\text{dis}(f)}{2} < r < \frac{\min\{e(G_0), e(G'_0)\}}{4}$. Then there is a continuous map $\bar{f}: G \rightarrow G'$ such that $\frac{\text{dis}(\bar{f})}{2} < 2r$ and $\bar{f}(G) \subseteq (f(G))^r$.*

3:14 Lower Bounding the Gromov–Hausdorff Distance in Metric Graphs

Proof. We organize the proof into three steps, to help its readability. Let $\varepsilon > 0$ satisfy $\varepsilon + \text{dis}(f) < 2r$.

Step 1: Construct a continuous map \bar{f} . Fix a finite triangulation K of G with edges of length at most ε such that $V(G) \subseteq K^{(0)}$. Let us construct a continuous map $\bar{f}: K \rightarrow G'$ as follows. For every vertex $x \in K^{(0)}$, set $\bar{f}(x) = f(x)$. Consider now two adjacent vertices x and y in $K^{(0)}$. Send the segment from x to y to the geodesic segment connecting $f(x)$ and $f(y)$ at uniform speed. The geodesic between the two points $f(x)$ and $f(y)$ is unique since $d_{G'}(f(x), f(y)) \leq \varepsilon + \text{dis}(f) < \frac{\varepsilon(G'_0)}{2}$.

Step 2: Estimate the distortion of \bar{f} . We claim that $\text{dis}(\bar{f}) \leq 2(\text{dis}(f) + \varepsilon) < 4r$. The second inequality is trivial. Let $x, y \in K$. Choose a closest point $x_i \in K^{(0)}$ to x . Let $x_{i+1} \in K^{(0)}$ be such that x is contained in the edge $\{x_i, x_{i+1}\}$. Note that $d_G(x, x_i) \leq \frac{\varepsilon}{2}$. Thus, by construction of \bar{f} ,

$$d_{G'}(\bar{f}(x), f(x_i)) \leq \frac{d_{G'}(f(x_i), f(x_{i+1}))}{2} \leq \frac{\varepsilon + \text{dis}(f)}{2} \quad (5)$$

Similarly, if $y_j \in K^{(0)}$ is a closest point to y , $d_G(y, y_j) \leq \frac{\varepsilon}{2}$ and $d_{G'}(\bar{f}(y), f(y_j)) \leq \frac{\varepsilon + \text{dis}(f)}{2}$. By applying the triangle inequality, we obtain the following chain of inequalities:

$$\begin{aligned} d_{G'}(\bar{f}(x), \bar{f}(y)) &\leq d_{G'}(\bar{f}(x), \bar{f}(x_i)) + d_{G'}(f(x_i), f(y_j)) + d_{G'}(\bar{f}(y_j), \bar{f}(y)) \\ &\leq \frac{\text{dis}(f) + \varepsilon}{2} + d_G(x_i, y_j) + \text{dis}(f) + \frac{\text{dis}(f) + \varepsilon}{2} \leq 2\text{dis}(f) + \varepsilon + d_G(x_i, x) + d_G(x, y) + d_G(y, y_j) \\ &\leq 2(\text{dis}(f) + \varepsilon) + d_G(x, y) < 4r + d_G(x, y). \end{aligned}$$

With similar computations, we can show that $d_G(x, y) - d_{G'}(\bar{f}(x), \bar{f}(y)) \leq 2(\text{dis}(f) + \varepsilon) < 4r$. This proves $\frac{\text{dis}(\bar{f})}{2} < 2r$.

Step 3: Show $\bar{f}(K) \subseteq (f(G))^r$. For every point $\bar{f}(x) \in \bar{f}(K)$, choose a closest point $x_i \in K^{(0)}$ to x as in **Step 2**. According to (5), $d_{G'}(\bar{f}(x), f(x_i)) < r$ since $\varepsilon + \text{dis}(f) < 2r$. ◀

We finally obtain that the core G_0 of a graph G is preserved by a function with distortion bounded by a fraction of the shortest edge length.

► **Proposition 20.** *Let G be a finite metric graph and let $f: G \rightarrow G$ be a function with $\frac{\text{dis}(f)}{2} < r < \frac{\varepsilon(G_0)}{16}$. Then $G_0 \subseteq f(G_0)^r$.*

Proof. Let $\bar{f}: G \rightarrow G$ be the continuous map described in Lemma 19. Then $\frac{\text{dis}(\bar{f})}{2} < 2r < \frac{\varepsilon(G_0)}{8}$. The restriction $\bar{f}|_{G_0}: G_0 \rightarrow G$ is also continuous and has the same bound on its distortion. We can then apply Lemma 15 to $\bar{f}|_{G_0}$ and obtain an injection $\Phi: \mathcal{L}(G_0) \rightarrow \mathcal{L}(G)$. Since $\mathcal{L}(G_0) = \mathcal{L}(G)$ by construction and these sets are finite, Φ is a bijection. Hence, every simple loop of G is contained in $\bar{f}(G_0) \subseteq f(G_0)^r$ according to Lemma 19. Given that $f(G_0)^r$ is connected, it contains G_0 . ◀

As an immediate consequence of Proposition 20 we obtain that the Hausdorff and the Gromov–Hausdorff distances between a graph with no leaves and a subset thereof coincide, provided that the latter is sufficiently small.

► **Corollary 21.** *Let G be a finite metric graph with no leaves, and let $X \subseteq G$. If $d_{\text{GH}}(G, X) < \frac{\varepsilon(G_0)}{16}$, then $d_{\text{GH}}(G, X) = d_{\text{H}}(G, X)$.*

Proof. Let $r > 0$ satisfy $d_{\text{GH}}(G, X) < r < \frac{\epsilon(G_0)}{16}$. Then there is a map $f: G \rightarrow X$ whose distortion satisfies $\frac{\text{dis}(f)}{2} < r < \frac{\epsilon(G_0)}{16}$. We can apply Proposition 20 to obtain that $G = G_0 \subseteq f(G_0)^r \subseteq X^r$. Hence $d_{\text{H}}(G, X) \leq r$. Since r can be taken arbitrarily, we have $d_{\text{H}}(G, X) \leq d_{\text{GH}}(G, X)$. \blacktriangleleft

We now discuss how to lower bound the Gromov–Hausdorff distance with respect to the Hausdorff distance for graphs with leaves. The first result relies on hypotheses around the subset X of the graph G . More precisely, similarly to the tree case, we request that the subset X is sufficiently close to the leaves of G .

► **Proposition 22.** *Let (G, d) be a finite metric graph and let $f: G \rightarrow G$ be a (possibly discontinuous) function with $\frac{\text{dis}(f)}{2} < \frac{\epsilon(G_0)}{16}$. If $\partial G \neq \emptyset$, suppose that $d_{\text{H}}(G, f(G)) > \vec{d}_{\text{H}}(\partial G, f(G))$. Then $d_{\text{H}}(G, f(G)) \leq \frac{\text{dis}(f)}{2}$.*

Proof. Let $r > 0$ satisfy $\frac{\text{dis}(f)}{2} < r < \frac{\epsilon(G_0)}{16}$. Let $\bar{x} \in G$ be a point such that the intersection $B(\bar{x}, d_{\text{H}}(G, f(G))) \cap f(G)$ is empty. If $\bar{x} \in G_0$, then $\bar{x} \in f(G)^r$ according to Proposition 20, and so $d_{\text{H}}(G, f(G)) < r$. Suppose otherwise that $\bar{x} \notin G_0$. Thus, $G \setminus \{\bar{x}\}$ is disconnected, and at least two connected components thereof have non-empty intersection with X by the assumption. We can conclude as in Proposition 7. \blacktriangleleft

► **Theorem 23.** *Let G be a finite connected metric graph and let $X \subseteq G$. If $\partial G \neq \emptyset$, suppose $d_{\text{H}}(G, X) > \vec{d}_{\text{H}}(\partial G, X)$. If $d_{\text{GH}}(G, X) < \frac{\epsilon(G_0)}{16}$, then $d_{\text{GH}}(G, X) = d_{\text{H}}(G, X)$.*

Proof. As in the proof of Theorem 8, by Lemma 2 we can assume that X is compact by replacing it with its closure \bar{X} if necessary. Suppose, by contradiction, that $d_{\text{GH}}(G, X) < d_{\text{H}}(G, X)$. By Lemma 4 there is a function $f: G \rightarrow X$ such that $\frac{1}{2}\text{dis}(f) < \min\{d_{\text{H}}(G, X), \frac{\epsilon(G_0)}{16}\}$ and $\vec{d}_{\text{H}}(\partial G, X) = \vec{d}_{\text{H}}(\partial G, f(G))$. Note that $d_{\text{H}}(G, f(G)) \geq d_{\text{H}}(G, X) > \vec{d}_{\text{H}}(\partial G, X) = \vec{d}_{\text{H}}(\partial G, f(G))$. We can apply Proposition 22 to obtain a contradiction with the following chain of inequalities: $d_{\text{H}}(G, X) \leq d_{\text{H}}(G, f(G)) \leq \frac{\text{dis}(f)}{2} \leq d_{\text{GH}}(G, X)$. \blacktriangleleft

7 Conclusion and open questions

We have furthered the study of lower bounding the Gromov–Hausdorff distance between a space and a subset by their Hausdorff distance, particularly for metric graphs. This investigation sparks several open questions and new research directions, especially for spaces beyond graphs: (1) Is the density constant $\frac{\epsilon(G_0)}{16}$ in Theorem 23 optimal? (2) Do our results generalize to general length spaces? If so, under what density assumptions? (3) Are there versions of our results that hold for manifolds with boundary? (4) Are there classes of metric spaces (other than graphs) where the Gromov–Hausdorff distance between the space and a dense enough subset equals their Hausdorff distance?

References

- 1 Henry Adams, Johnathan Bush, Nate Clause, Florian Frick, Mario Gómez, Michael Harrison, R. Amzi Jeffs, Evgeniya Lagoda, Sunhyuk Lim, Facundo Mémoli, Michael Moy, Nikola Sadovek, Matt Superdock, Daniel Vargas, Qingsong Wang, and Ling Zhou. Gromov–Hausdorff distances, Borsuk–Ulam theorems, and Vietoris–Rips complexes. *arXiv preprint*, 2023. [arXiv: 2301.00246](#).
- 2 Henry Adams, Florian Frick, Sushovan Majhi, and Nicholas McBride. Hausdorff vs Gromov–Hausdorff distances. *Discrete & Computational Geometry*, 2025.

- 3 Henry Adams, Sushovan Majhi, Fedor Manin, Žiga Virk, and Nicolò Zava. Lower bounding the Gromov–Hausdorff distance in metric graphs. [arXiv:2411.09182](#).
- 4 Pankaj K. Agarwal, Kyle Fox, Abhinandan Nath, Anastasios Sidiropoulos, and Yusu Wang. Computing the Gromov–Hausdorff distance for metric trees. *ACM Trans. Algorithms*, 14(2), 2018. doi:10.1145/3185466.
- 5 Mark Anthony Armstrong. *Basic topology*. Springer Science & Business Media, 2013.
- 6 Martin R Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319. Springer Science & Business Media, 2011.
- 7 L. E. J. Brouwer. Beweis der Invarianz der Dimensionenzahl. *Mathematische Annalen*, 70(2):161–165, 1911.
- 8 L. E. J. Brouwer. Über Abbildung von Mannigfaltigkeiten. *Mathematische Annalen*, 71(1):97–115, 1911.
- 9 Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33. American Mathematical Society, Providence, 2001.
- 10 Gunnar Carlsson and Facundo Mémoli. Characterization, stability and convergence of hierarchical clustering methods. *Journal of Machine Learning Research*, 11(47):1425–1470, 2010. doi:10.5555/1756006.1859898.
- 11 Frédéric Chazal, Vin de Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 174:193–214, 2014.
- 12 Lester Dubins and Gideon Schwarz. Equidiscontinuity of Borsuk–Ulam functions. *Pacific Journal of Mathematics*, 95(1):51–59, 1981.
- 13 David A Edwards. The structure of superspace. In *Studies in topology*, pages 121–133. Elsevier, 1975.
- 14 Mikhael Gromov. Structures métriques pour les variétés Riemanniennes. *Textes Mathématiques*, 1, 1981.
- 15 Michael Harrison and R Amzi Jeffs. Quantitative upper bounds on the Gromov–Hausdorff distance between spheres. *arXiv preprint*, 2023. [arXiv:2309.11237](#).
- 16 Yibo Ji and Alexey A Tuzhilin. Gromov–Hausdorff distance between segment and circle. *arXiv preprint*, 2021. [arXiv:2101.05762](#).
- 17 Nigel J Kalton and Mikhail I Ostrovskii. Distances between Banach spaces. *Forum Math.*, 11:1–17, 1999.
- 18 Alejandro Leon. The Gromov–Hausdorff distance between spheres and their boundaries. To appear, 2026.
- 19 Sunhyuk Lim, Facundo Mémoli, and Zane Smith. The Gromov–Hausdorff distance between spheres. *Geometry & Topology*, 27:3733–3800, 2023.
- 20 Sushovan Majhi, Jeffrey Vitter, and Carola Wenk. Approximating Gromov–Hausdorff distance in Euclidean space. *Computational Geometry*, 116:102034, 2024. doi:10.1016/j.comgeo.2023.102034.
- 21 Facundo Mémoli. On the use of Gromov–Hausdorff distances for shape comparison, 2007.
- 22 Facundo Mémoli. Some properties of Gromov–Hausdorff distances. *Discrete & Computational Geometry*, 48(2):416–440, 2012. doi:10.1007/S00454-012-9406-8.
- 23 Felix Schmiedl. Computational aspects of the Gromov–Hausdorff distance and its application in non-rigid shape matching. *Discrete & Computational Geometry*, 57(4):854, 2017.
- 24 Alexey A Tuzhilin. Who invented the Gromov–Hausdorff distance? *arXiv preprint*, 2016. [arXiv:1612.00728](#).
- 25 Nicolò Zava. Coarse and bi-Lipschitz embeddability of subspaces of the Gromov–Hausdorff space into Hilbert spaces. *Algebr. Geom. Topol.*, 25(8):5153–5174, 2024.