ABSOLUTELY MINIMAL LIPSCHITZ EXTENSION OF TREE-VALUED MAPPINGS

ASSAF NAOR AND SCOTT SHEFFIELD

Abstract. We prove that every Lipschitz function from a subset of a locally compact length space to a metric tree has a unique absolutely minimal Lipschitz extension (AMLE). We relate these extensions to a stochastic game called **Politics** — a generalization of a game called **Tug of War** that has been used in [42] to study real-valued AMLEs.

1. Introduction

For a pair of metric spaces (X, d_X) and (Z, d_Z), a mapping $h : X \to Z$, and a subset $S \subseteq X$, the Lipschitz constant of h on S is denoted

$$\mathrm{Lip}_S(h) \stackrel{\mathrm{def}}{=} \sup_{\substack{x,y\in S\\ x\neq y}} \frac{d_Z(h(x),h(y))}{d_X(x,y)}.$$

Given a closed subset $Y \subseteq X$ and a Lipschitz mapping $f : Y \to Z$, a Lipschitz mapping $\widetilde{f} : X \to Z$ is called an **absolutely minimal Lipschitz extension** (AMLE) of f if its restriction to Y coincides with f, and for every open subset $U \subseteq X \setminus Y$ and every Lipschitz mapping $h : X \to Z$ that coincides with $\widetilde{f}$ on $X \setminus U$ we have

$$\mathrm{Lip}_U(h) \geqslant \mathrm{Lip}_U\left(\widetilde{f}\right). \tag{1}$$

In other words, $\widetilde{f}$ extends f, and it is not possible to modify $\widetilde{f}$ on an open set in a way that decreases the Lipschitz constant on that set.

Our main result is:

Theorem 1. *Let X be a locally compact length space and let T be a metric tree. For every closed subset $Y \subseteq X$, every Lipschitz mapping $f : Y \to T$ has a unique AMLE $\widetilde{f} : X \to T$.*

Recall that a metric space (X, d_X) is a **length space** if for all $x, y \in X$, the distance $d_X(x, y)$ is the infimum of the lengths of curves in X that connect x to y. By a **metric tree** we mean the one-dimensional simplicial complex associated to a finite graph-theoretical tree with arbitrary edge lengths (i.e., a finite graph-theoretical tree whose edges are present as actual intervals of arbitrary length, equipped with the graphical shortest path metric). We did not investigate here the greatest possible generality in which Theorem 1 holds true; in particular, we conjecture that the assumption that X is locally compact can be dropped, and that T need not correspond to a finite graph-theoretical tree, but rather can belong to the more general class of bounded $\mathbb{R}$-trees (see [14, 15]). The requirement that X be locally compact is not used in our proof of the uniqueness assertion of Theorem 1.

In the special case when T is an interval $[a, b] \subseteq \mathbb{R}$ and $X = \mathbb{R}^n$, Theorem 1 was proved in [17]; see also [6, 4, 2] for different proofs of the uniqueness part of Theorem 1 in this special case. The existence part of Theorem 1 was generalized to arbitrary length spaces X and $T = [a, b]$ in [37]; see also [21] and [29] for different proofs of this existence result with the additional assumptions that the length space X is separable or compact, respectively. The uniqueness part of Theorem 1 was proved in [42] for X a general length space and $T = [a, b]$. Additionally, [42] contains a new (game theoretic) proof of the existence part of Theorem 1 when $T = [a, b]$ and X is a general length space.

The purpose of the present article is to initiate the study of absolutely minimal Lipschitz extensions of mappings that are not necessarily real-valued, the tree-valued case being the first non-trivial setting of this type where such theorems can be proved. Our proofs overcome various difficulties that arise since we can no longer use the order structure of the real line, which was crucially used in [17, 37, 21, 29, 42]. We also introduce a stochastic game called **Politics**, related to tree-valued AMLE, that generalizes the stochastic game called **Tug of War** that was introduced and related to real-valued AMLE in [42].

In the remainder of this introduction we explain the relevant background from the classical theory of Lipschitz extension and ∞-harmonic functions, and also describe the main steps of our proof.

1.1. **Background on the Lipschitz extension problem.** The classical Lipschitz extension problem asks for conditions on a pair of metric spaces (X, d_X) and (Z, d_Z) which ensure that there exists $K \in [1, \infty)$ such that for all $Y \subseteq X$ and all Lipschitz mappings $f : Y \to Z$, there exists $\widetilde{f} : X \to Z$ with $\widetilde{f}\big|_Y = f$ and

$$\mathrm{Lip}_X\left(\widetilde{f}\right) \leqslant K \cdot \mathrm{Lip}_Y(f). \tag{2}$$

Stated differently, in the Lipschitz extension problem we are interested in geometric conditions ensuring the existence of $\widetilde{f} : X \to Z$ such that the diagram in (3) commutes, where $\iota : Y \to X$ is the formal inclusion, and the Lipschitz constant of $\widetilde{f}$ is guaranteed to be at most a fixed multiple (depending only on the geometry of the spaces X, Z) of the Lipschitz constant of f.

$$\begin{tikzcd} X \arrow[dr, dashed, "\widetilde{f}"] & \\ Y \arrow[u, "\iota"] \arrow[r, "f"'] & Z \end{tikzcd} \tag{3}$$

Note that if (Z, d_Z) is complete then we can trivially extend f to the closure of Y.

When $K = 1$ in (2), i.e., when one can always extend functions while preserving their Lipschitz constant, the pair (X, Z) is said to have the **isometric extension property**. When $K \in (1, \infty)$ the corresponding extension property is called the **isomorphic extension property**. The present article is devoted to the isometric extension problem, though we will briefly discuss questions related to its isomorphic counterpart in Section 1.4. We refer to the books [47, 7] and the references therein, as well as the introductions of [30, 40] (and the references therein), for more background on the Lipschitz extension problem.

It is rare for a pair of metric spaces (X, Z) to have the isometric extension property. A famous instance when this does happen is Kirszbraun's extension theorem [24], which asserts that if X and Z are Hilbert spaces then (X, Z) have the isometric extension property.

Another famous example is the non-linear Hahn-Banach theorem [34], i.e., when $Z = \mathbb{R}$ and X is arbitrary; this (easy) fact follows from the same proof as the proof of the classical Hahn-Banach theorem (i.e., by extending to one additional point at a time; alternatively, one can construct the maximal and minimal isometric extensions explicitly).

More generally, one may consider metric spaces Z such that for every metric space X the pair (X, Z) has the isometric extension property (i.e., Z is an injective metric space in the isometric category). This is equivalent to the fact that there is a 1-Lipschitz retraction from any metric space containing Z onto Z (see [7, Prop. 1.2]); such spaces are called in the literature **absolute 1-Lipschitz retracts**. It is a well known fact (see [7, Prop. 1.4]) that (Z, d_Z) is an absolute 1-Lipschitz retract if and only if (a) Z is *metrically convex*, i.e., for every $x, y \in Z$ and $\lambda \in [0, 1]$ there is $z \in Z$ such that $d_Z(x, z) = \lambda d_Z(x, y)$ and $d_Z(y, z) = (1 - \lambda)d_Z(x, y)$, and (b) Z has the *binary intersection property*, i.e., if every collection of pairwise intersecting closed balls in Z has a common point. Examples of absolute 1-Lipschitz retracts are ℓ_∞ and metric trees (see [23, 19]). Additional examples are contained in [16] (see also [7, Ch. 1]).

If (X, d_X) is path-connected and the pair (X, Z) has the isometric extension property, then the AMLE condition (1) is equivalent to the requirement:

$$\forall \text{ open } U \subseteq X \setminus Y, \quad \mathrm{Lip}_U\left(\widetilde{f}\right) = \mathrm{Lip}_{\partial U}\left(\widetilde{f}\right). \tag{4}$$

When $Z = \mathbb{R}$ and $X = \mathbb{R}^n$, the AMLE formulation (4) was first introduced by Aronsson [3], in connection with the theory of ∞-harmonic functions. Specifically, it was shown in [3] that if $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}$ is smooth then the validity of (4) is equivalent to the requirement that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial \widetilde{f}}{\partial x_i} \cdot \frac{\partial \widetilde{f}}{\partial x_j} \cdot \frac{\partial^2 \widetilde{f}}{\partial x_i \partial x_j} = 0 \quad \text{on } \mathbb{R}^n \setminus Y. \tag{5}$$

If one interprets (5) in terms of viscosity solutions, then it was proved in [17] that the equivalence of (4) and (5) (when $Z = \mathbb{R}$ and $X = \mathbb{R}^n$) holds for general Lipschitz $\widetilde{f}$. We refer to the survey article [4] and the references therein for more information on the many works that investigate this remarkable connection between the classical Lipschitz extension problem and PDEs.

Existence of isometric and isomorphic Lipschitz extensions has a wide variety of applications in pure and applied mathematics. Despite this rich theory, the issue raised by Aronsson's seminal paper [3] is that even when isometric Lipschitz extension is possible, many such extensions usually exist, and it is therefore natural to ask for extension theorems ensuring that the extended function has additional desirable properties. In particular, the notion of AMLE is an isometric Lipschitz extension which is locally the "best possible" extension. In this context, one can ask for (appropriately defined) "AMLE versions" of known Lipschitz extension theories. As a first step, in light of Theorem 1 it is tempting to ask the following:

Question 1. *Let Z be an absolute 1-Lipschitz retract. Is it true that for every length space X and every closed subset $Y \subseteq X$, any Lipschitz $f : Y \to Z$ admits an AMLE $\widetilde{f} : X \to Z$?*

Note that unlike the situation when Z is a metric tree, in the setting of Question 1 one cannot expect in general that the AMLE will be unique: consider for example $Z = \ell_\infty^2$, i.e., the absolute 1-Lipschitz retract $\mathbb{R}^2$, equipped with the ℓ_∞ norm. Let $X = \mathbb{R}$ and $Y = \{0, 1\}$. The 1-Lipschitz mapping $f : Y \to \ell_\infty^2$ given by $f(0) = (0, 0)$, $f(1) = (1, 1)$ has many AMLEs

$\widetilde{f} : \mathbb{R} \to \ell_\infty^2$, since for every 1-Lipschitz function $g : \mathbb{R} \to \mathbb{R}$ with $g(0) = 0$, $g(1) = 1$, the mapping $x \mapsto (x, g(x))$ will be an AMLE of f. At the same time, by using the existence of real-valued AMLEs coordinate-wise, the answer to Question 1 is trivially positive when $Z = \ell_\infty(\Gamma)$ for any set Γ.

While we do not give a general answer to Question 1, we show here that general absolute 1-Lipschitz retracts Z do enjoy a stronger Lipschitz extension property: Z-valued functions defined on subsets of vertices of 1-dimensional simplicial complexes associated to unweighted finite graphs admit isometric Lipschitz extensions which are ∞-harmonic. This issue, together with the relevant definitions, is discussed in Section 1.2 below. In addition to being crucially used in our proof of Theorem 1, this result indicates that absolute 1-Lipschitz retracts do admit enhanced Lipschitz extension theorems that go beyond the simple existence of isometric Lipschitz extensions (which is the definition of absolute 1-Lipschitz retracts). At the same time, we describe below a simple example indicating inherent difficulties in obtaining a positive answer to Question 1 beyond the class of metric trees (and their ℓ_∞-products).

1.2. **∞-harmonic functions and AMLEs on finite graphs.** Let $G = (V, E)$ be a finite connected (unweighted) graph. We shall consider G as a 1-dimensional simplicial complex, i.e., the edges of G are present as intervals of length 1 joining their endpoints. This makes G into a length space, where the shortest-path metric is denoted by d_G. Given a vertex $v \in V$ denote its neighborhood in G by $N_G(v)$, i.e., $N_G(v) = \{u \in V : \ uv \in E\}$.

Let (Z, d_Z) be a metric space. We shall say that a function $f : V \to Z$ is ∞-harmonic at $v \in V$ if there exist $u, w \in N_G(v)$ such that

$$d_Z(f(u), f(v)) = d_Z(f(w), f(v)) = \max_{z \in N_G(v)} d_Z(f(z), f(v)), \tag{6}$$

and

$$d_Z(f(u), f(w)) = 2 \max_{z \in N_G(v)} d_Z(f(z), f(v)). \tag{7}$$

(Here we allow for $u = w$, so that if $N_G(v)$ has only one element u, then (6) and (7) imply that $f(u) = f(v)$.) We say $f : V \to Z$ is ∞-harmonic on $W \subseteq V$ if it is ∞-harmonic at every $v \in W$.

The connection to AMLEs is simple: for $\Omega \subseteq V$ and $f : \Omega \to Z$, if $\widetilde{f} : G \to Z$ is an AMLE of f then $\widetilde{f}$ must be geodesic on edges, i.e., for $u, v \in V$ with $uv \in E$, if $x \in G$ is a point on the edge uv at distance $\lambda \in [0, 1]$ from u, then $d_Z(f(x), f(u)) = \lambda d_Z(f(u), f(v))$ and $d_Z(f(x), f(v)) = (1 - \lambda) d_Z(f(u), f(v))$ (apply (4) to the open segment joining u and v). Moreover, if G is triangle-free, then $\widetilde{f}$ is ∞-harmonic on $V \smallsetminus \Omega$. This follows from considering in (4) the open set $U \subseteq G$ consisting of the union of the half-open edges incident to $v \in V \smallsetminus \Omega$ (including v itself). The vertices $u, w \in N_G(v) = \partial U$ in (6) will be the points at which $\mathrm{Lip}_{\partial U}\left(\widetilde{f}\right)$ is attained. The restriction that G is triangle-free implies that $d_G(u, w) = 2$, using which (7) follows from (4).[1]

The converse to the above discussion is true for mappings into metric trees. This is contained in Theorem 2 below, whose simple proof appears in Section 4. A local-global

[1]In the above reasoning the assumption that G is triangle-free can be dropped if Z is a metric tree. But, this is not important for us: we only care about G as a length space, and therefore we can replace each edge of G by a path of length 2, resulting in a triangle-free graph whose associated 1-dimensional simplicial complex is the same as the original simplicial complex, with distances scaled by a factor of 2.

statement analogous to Theorem 2 *fails* when the target (geodesic) metric space is not a metric tree, as we explain in Remark 1 below.

Given a metric tree T, a finite graph $G = (V, E)$ and a function $f : V \to T$, the *linear interpolation* of f is the T-valued function defined on the 1-dimensional simplicial complex associated to G as follows: given an edge $e = uv \in E$ and $x \in e$ with $d_G(x, u) = \lambda d_G(u, v)$ and $d_G(x, v) = (1-\lambda)d_G(u, v)$, the image $f(x) \in T$ is the point on the geodesic joining $f(u)$ and $f(v)$ in T with $d_T(f(x), f(u)) = \lambda d_T(f(u), f(v))$ and $d_T(f(x), f(v)) = (1-\lambda)d_T(f(u), f(v))$.

Theorem 2. *Let T be a metric tree and $G = (V, E)$ a finite connected (unweighted) graph. Assume that $\Omega \subseteq V$ and that $f : V \to T$ is ∞-harmonic on $V \smallsetminus \Omega$. Then the linear interpolation of f is an AMLE of $f|_\Omega$.*

Remark 1. Consider the example depicted in Figure 1, viewed as a 12 vertex graph G with vertices

$$V = \{A, B, C, X, Y, Z\} \cup \{S_i\}_{i=1}^6$$

and edges

$$E = \{XS_3, S_3A, AS_2, S_2B, AS_4, S_4C, BS_6, S_6C, ZS_5, S_5C, YS_1, S_1B\}.$$

(The role of the vertices $\{S_i\}_{i=1}^6$ is just to subdivide edges so that the graph will be triangle-free.) The picture in Figure 1 can also be viewed as a mapping $f : V \to \mathbb{R}^2$. Denoting $\Omega = \{X, Y, Z\}$, this mapping is by construction ∞-harmonic on $V \smallsetminus \Omega$. In spite of this fact, the linear interpolation of f is not an AMLE of $f|_\Omega$. Indeed, consider the open set $U = G \smallsetminus \Omega$. Since the planar Euclidean distance between any two of the points $f(X), f(Y), f(Z)$ is strictly less than 3 (=the distance between any two of the vertices $\{X, Y, Z\}$ in G), we have $\mathrm{Lip}_{\partial U}(f) = \mathrm{Lip}_{\{X,Y,Z\}}(f) < 1$. At the same time, by considering the vertices A, B, C we see that $\mathrm{Lip}_U(f) = 1$.

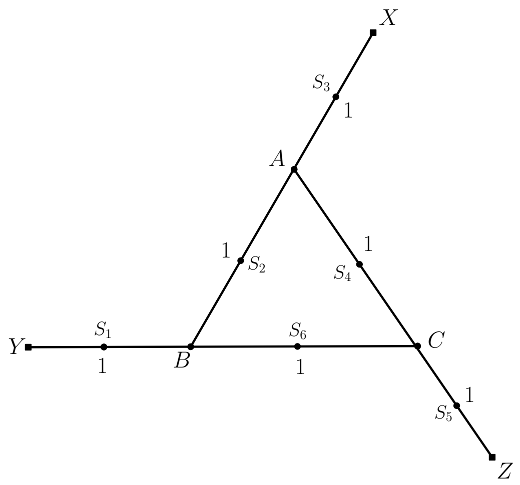

FIGURE 1. An example of an ∞-harmonic function which isn't an AMLE.

In Section 4 we show that absolute 1-Lipschitz retracts have a stronger Lipschitz extension property, namely they admit ∞-harmonic extensions for functions from finite graphs:

Theorem 3. *Assume that (Z, d_Z) is an absolute 1-Lipschitz retract and that $G = (V, E)$ is a finite connected (unweighted) graph. Fix $\Omega \subseteq V$ and $f : \Omega \to Z$. Then there exists a mapping $\widetilde{f} : V \to Z$ which is ∞-harmonic on $V \setminus \Omega$ such that*

$$\widetilde{f}\Big|_{\Omega} = f \quad \text{and} \quad \mathrm{Lip}_V\left(\widetilde{f}\right) = \mathrm{Lip}_{\Omega}(f).$$

The existence part of Theorem 1 is deduced in Section 5 from Theorem 3 via a compactness argument that relies on a comparison-based characterization of AMLE that we establish in Section 2. The uniqueness part of Theorem 1 is proved via a topological argument (and the results of Section 2) in Section 3.

1.3. **Tug of War and Politics.** In the special case when $T \subseteq \mathbb{R}$ is an interval, Theorem 1 was proved in [42] without the local compactness assumption using a two-player, zero-sum stochastic game called Tug of War. We expect that one could adapt the arguments in [42] and the game called Politics (introduced below) to give a proof of Theorem 1 that does not use local compactness; however, this would involve rewriting large sections of [42] in a significantly more complicated way, and we will not attempt to do this here.

Tug of War is a two-player, zero-sum stochastic game. In this game, one starts with an initial point $x_0 \in X \setminus Y$; then at the kth stage of the game, a fair coin is tossed and the winner gets to choose any $x_k \in X$ with $|x_k - x_{k-1}| < \varepsilon$. Informally, the winning player "tugs" the game position up to ε units in a direction of her choice. The game ends the first time K that $x_K \in Y$, and player one collects a payoff of $f(x_K)$ from player two. It was shown that as $\varepsilon \to 0$, the value of the game (informally, the amount the first player wins in expectation when both players play optimally; see Section 6) tends to $\widetilde{f}(x_0)$. In addition to its usefulness in proofs, the game theory provides a deeper understanding of what an AMLE is. Although AMLEs are often difficult to compute explicitly, one can always provide upper and lower bounds by giving explicit strategies for the game and showing that they guarantee a certain expected payoff for one player or the other. It is therefore natural to ask for an analog of Tug of War that makes sense when T is not an interval.

Since $\widetilde{f}(x_0)$ is a point in T, however, and not in $\mathbb{R}$, it is not immediately obvious how $\widetilde{f}(x_0)$ can represent a value for either player. We will solve this problem by augmenting the state space of the game to include declared "targets" $t_k, o_k \in T$ as well as "game positions" $x_k \in X$. Before explaining this, we remark that one obtains a slight generalization of Tug of War by letting x_k be vertices of any (possibly infinite) graph with vertex set X and $Y \subseteq X$. One then requires that x_k and x_{k-1} be adjacent in that graph (instead of requiring $|x_k - x_{k-1}| < \varepsilon$). We now introduce the game of **Politics** in a similar setting.

Let $G = (V, E)$ be an unweighted undirected graph which may have self loops. Fix $Y \subseteq V$ and a mapping $f : Y \to T$. Begin with an initial game position $x_0 \in V \setminus Y$ and an initial "target" $t_0 \in T$. At the kth round of the game, the players determine the values (x_k, t_k) as follows:

(1) Player I chooses an "opposition target" $o_k \in T$ and collects $d_T(o_k, t_{k-1})$ units from player II.
(2) Player II chooses a new target $t_k \in T$ and collects $d_T(o_k, t_k)$ units from player I.
(3) A fair coin is tossed and the winner of the toss chooses a new game position $x_k \in X$ with $\{x_{k-1}, x_k\} \in E$.

The total amount player I gains at each round is $d_T(o_k, t_{k-1}) - d_T(o_k, t_k)$. Similarly, player II gains $d_T(o_k, t_k) - d_T(o_k, t_{k-1})$ at each round. The game ends after round K, where K is the smallest value of k for which $x_k \in Y$. At this point player I collects an additional $d_T(f(x_k), t_k)$ units from player II. (If the game never ends, we declare the total payout for each player to be zero.)

The game is called "Politics" because we may view it as a model for a rather cynical zero-sum political struggle in which $f(x_K)$ represents a "political outcome," but both parties care only about their own perceived political strength, and not about the actual outcome. We think of the target as representing the "declared political objective" of player II; the terminal payoff rule, makes it clear that player II would prefer $f(x_K)$ be close to this declared target (in order to "appear successful"). Player II is allowed to adjust the target during each round, but loses points for moving her target closer to the declared opposition target o_k (because "making a concession" makes her appear weak) and gains points for moving her target further from the opposition target because "taking a harder line" makes her appear strong). [2]

We will prove the following for finite graphs:

Proposition 4. *Fix a finite graph $G = (V, E)$, some $Y \subseteq V$, a metric tree T, and a function $f : Y \to T$. View G as a length space (with all edges having length one) and let $\widetilde{f} : G \to T$ be the AMLE of f. Then the value of the game of Politics with these parameters and initial vertex $x_0 \in V \setminus Y$ is given by*

$$d_T\left(\widetilde{f}(x_0), t_0\right).$$

Proposition 4 will be proved in Section 6. An extension of Proposition 4 to infinite graphs (via the methods of [42]) is probably possible, but we will not attempt it here.

1.4. **Some open questions and directions for future research.** It would be of interest to understand known isometric extension theorems in the context of the AMLE problem. Specifically, we ask:

Question 2. *Is there an AMLE version of Kirszbraun's extension theorem, i.e, is it true that for every pair of Hilbert spaces H_1, H_2 and every closed subset $Y \subseteq H_1$, any Lipschitz mapping $f : Y \to H_2$ admits an AMLE $\widetilde{f} : H_1 \to H_2$?*

We refer to the manuscript [44] for a discussion of subtleties related to Question 2, as well as some partial results in this direction. Examples of additional isometric extension theorems that might have AMLE versions are contained in [45, 46, 47, 27, 39].

The study of isomorphic extensions in the context of the AMLE problem is wide open. Since when Lipschitz extension is possible a constant factor loss is usually necessary, and since isomorphic extensions suffice for many applications, it would be of interest if some isomorphic extension theorems had "almost locally optimal" counterparts. For example, one

[2] There is a more player-symmetric variant of this game in which each player, upon moving a target, earns the net change in the distance from the opponent's target. That is, player II earns $d_T(o_k, t_k) - d_T(o_k, t_{k-1})$ when choosing t_k (so player I earns $d_T(o_k, t_{k-1}) - d_T(o_k, t_k)$) and player I earns $d_T(o_k, t_{k-1}) - d_T(o_{k-1}, t_{k-1})$ when choosing o_k. In fact, by combining like terms, modifying the end-of-game payout function, and defining $o_0 = t_0$, one can make this game *equivalent* to the one described above but with twice the total payout.

might ask for the existence of a constant $K > 0$ such that one can extend any mapping $f : Y \to Z$ to a mapping $\widetilde{f} : X \to Z$ so that for every open $U \subseteq X \smallsetminus Y$ we have

$$\mathrm{Lip}_U\left(\widetilde{f}\right) \leqslant K \cdot \mathrm{Lip}_{\partial U}\left(\widetilde{f}\right). \tag{8}$$

Examples of isomorphic extension results that could be studied in the context of the AMLE problem include [31, 32, 18, 20, 5, 43, 25, 9, 30, 40, 8, 36, 22, 26]. Unlike isometric extension theorems, isomorphic extension theorems cannot be done "one point at time", since naïvely the constant factor losses at each step would accumulate. For this reason, isomorphic extension theorems usually require methods that are very different from their isometric counterparts. One would therefore expect that entirely new approaches are necessary in order to prove AMLE versions of isomorphic extension.

1.5. **Possible applications.** The image processing literature makes use of real-valued AMLEs as a technique for image inpainting and surface reconstruction — see [11, 1, 35, 10]. Since many data sets in areas ranging from computer science to biology have a natural tree structure, it stands to reason that problems involving reconstruction/interpolation of missing tree-valued data could be similarly approached using tree valued AMLEs.

Tree-valued AMLEs may also be useful for problems that do not involve trees *a priori.* To give a simple illustration of this, suppose we have a two-dimensional surface S embedded in $\mathbb{R}^3$ that separates an "inside" from an "outside," but such that on some open $W \subseteq \mathbb{R}^3$ the shape of the surface is not known. Let $d(x)$ be the signed distance of x from S (i.e., the actual distance if x is on the outside and minus that distance if x is on the inside). If we can compute or approximate $d(x)$ outside of W, then the extension of $d(x)$ to W has a zero set that can be interpreted as a "reconstructed" approximation to S. This approach and related methods are explored in [10].

If instead of a single "inside" and "outside" there were three or more regions of space meeting at a point v, and the union S of the interfaces between these regions was unknown in a neighborhood W of v, then we could use the same approach but replace $\mathbb{R}$ with the metric tree $\bigcup \omega_i[0,\infty) \subseteq \mathbb{C}$ for some complex roots of unity ω_i, and let $d(x)$ be ω_i (when x is in the ith region) times the distance from x to S. A similar technique could be used for inpainting a two-dimensional image comprised of a small number of monochromatic regions. Indeed, for such problems, it is not clear how one could apply the AMLE method without using trees.

2. Comparison formulation of absolute minimality

We take the following definition from [12] (see also [17, 13] for the case $X = \mathbb{R}^n$). Let U be an open subset of a length-space (X, d_X) and let $f : \overline{U} \to \mathbb{R}$ be continuous. Then f is said to satisfy **comparison with distance functions from above** on U if for every open $W \subseteq U$, $z \in X \smallsetminus W$, $b \geqslant 0$ and $c \in \mathbb{R}$ we have the following:

$$\Big(\forall\, x \in \partial W \quad f(x) \leqslant b\, d_X(x,z) + c\Big) \implies \Big(\forall\, x \in W \quad f(x) \leqslant b\, d_X(x,z) + c\Big). \tag{9}$$

The function f is said to satisfy **comparison with distance functions from below** on U if the function $-f$ satisfies comparison with distance functions from above on U, i.e., for every open $W \subseteq U$, $z \in X \setminus W$, $b \geqslant 0$ and $c \in \mathbb{R}$ we have the following:

$$\Big(\forall\, x \in \partial W \quad f(x) \geqslant -b\, d_X(x,z) + c\Big) \implies \Big(\forall\, x \in W \quad f(x) \geqslant -b\, d_X(x,z) + c\Big). \tag{10}$$

Finally, f satisfies **comparison with distance functions** on U if it satisfies comparison with distance functions from above and from below on U. We cite the following:

Proposition 5 ([12]). *Let U be an open subset of a length space. A continuous $f : \overline{U} \to \mathbb{R}$ satisfies comparison with distance functions on U if and only if it is an AMLE of $f|_{\partial U}$.*

Remark 2. The definition of comparison with distance functions from above would not change if we added the requirement that $z \notin \partial W$; if (9) or (10) fails and $z \in \partial W$, then it will fail (with a modified c) when W is modified to include some neighborhood of z. The definition would also not change if we required $b > 0$. If (9) or (10) fails with $b = 0$, then it fails for some sufficiently small $b' > 0$.

We will need to have an analog of the above definition with the real line $\mathbb{R}$ replaced with T. The definition makes sense when T is any metric space, but we will only use it in the case when T is a metric tree. We say $f : \overline{U} \to T$ satisfies **T-comparison** on U if for every $t \in T$, the function $x \mapsto d_T(t, f(x))$ satisfies comparison with distance functions from above on U. This generalizes comparison with distance functions:

Proposition 6. *If T is the closed interval $[t_1, t_2] \subseteq \mathbb{R}$, then $f : \overline{U} \to T$ satisfies T-comparison on U if and only if it satisfies comparison with distance functions on U.*

Proof. If f satisfies T-comparison on U, then the mappings $x \mapsto d_T(t_1, f(x)) = f(x) - t_1$ and $x \mapsto d_T(t_2, f(x)) = t_2 - f(x)$ satisfy comparison with distance functions from above on U, hence f and $-f$ both satisfy comparison with distance functions from above. Conversely, if f satisfies comparison with distance functions on U, then for all $t \in [t_1, t_2]$ the mapping $x \mapsto d_T(f(x), t) = (f(x) - t) \vee (t - f(x))$ satisfies comparison with distance functions from above because it is a maximum of two functions with this property. $\square$

Proposition 5 also has a natural generalization, which is contained in Proposition 7 below. Note that the proof of this generalization uses the assumption that T is a metric tree in the "only if" direction; for the "if" direction T can be any metric space.

Proposition 7. *Let U be an open subset of a length space (X, d_X), and let (T, d_T) be a metric tree. A continuous function $f : \overline{U} \to T$ satisfies T-comparison on U if and only if it is an AMLE of $f|_{\partial U}$.*

Proof. We will first suppose, to obtain a contradiction, that f is not an AMLE of $f|_{\partial U}$, but satisfies T-comparison. Then there is an open $W \subseteq U$ such that $\mathrm{Lip}_W(f) > \mathrm{Lip}_{\partial W}(f)$. That is, there is a path P in X connecting points x and y in W whose length L satisfies

$$\frac{d_T(f(x), f(y))}{L} > \mathrm{Lip}_{\partial W}(f). \tag{11}$$

If y_1 and y_2 are the first and last times P hits ∂W, $d_T(f(y_1), f(y_2)) \leqslant \mathrm{Lip}_{\partial W}(f) \cdot d_X(y_1, y_2)$; hence the property (11) holds for either the portion of P between x and y_1 or the portion between y_2 and y. Thus, we may take P to be entirely contained in W; replacing P with a slightly shorter sub-path of P, we may assume the endpoints of P are both in W as well, and that P is some positive distance δ from ∂W. Set

$$m \stackrel{\text{def}}{=} \frac{d_T(f(x), f(y))}{L} > \mathrm{Lip}_{\partial W}(f).$$

We may then find x_1 arbitrarily close to some fixed point x_0 along P satisfying

$$\frac{d_T(f(x_0), f(x_1))}{d_X(x_0, x_1)} \geqslant m > \mathrm{Lip}_{\partial W}(f).$$

Now we consider the distance function $md_X(x_0, \cdot)$. We will compare it to the function $d_T(f(x_0), f(\cdot))$. Since the latter is at least as large as the former at the point x_1, T-comparison implies that it must be at least as large at some point on ∂W. This implies that for any $\varepsilon > 0$ we may find a $z \in \partial W$ where

$$m' \stackrel{\text{def}}{=} \frac{d_T(f(x_0), f(z))}{d_X(x_0, z)} \geqslant m - \varepsilon.$$

In particular, we may assume $m' > \mathrm{Lip}_{\partial W}(f)$. Next choose $m'' \in \big(\mathrm{Lip}_{\partial W}(f), m'\big)$. Consider the distance function $m'' d_X(z, \cdot)$ and compare it to $d_T(f(z), f(\cdot))$. Since the functions are equal at z and the latter is larger than the former at x_0, the latter must be larger than the former at some point $w \in (\partial W) \smallsetminus \{z\}$. But this implies

$$\frac{d_T(f(z), f(w))}{d_X(z, w)} > \mathrm{Lip}_{\partial W}(f),$$

a contradiction.

We now proceed to the converse. Note that since T is a bounded metric space, by intersecting U with a large ball it suffices to prove the converse when U is bounded. Suppose, to obtain a contradiction, that f is an AMLE of $f|_{\partial U}$ and does not satisfy T-comparison on U. Since f does not satisfy T-comparison on U, there exists an open $W \subseteq U$, a point $x_0 \notin W$ and $c \in \mathbb{R}$, $b \geqslant 0$, such that for some $t \in T$ we have $d_T(t, f(x)) \leqslant bd_X(x_0, x) + c$ for all $x \in \partial W$, yet $d_T(t, f(y)) > bd_X(x_0, y) + c$ for some $y \in W$. Write $F(z) = bd_X(x_0, z) + c$. We may replace W with the connected component of $\{x \in W : \ d_T(t, f(x)) > F(x)\}$ containing y, so that one has $d_T(t, f(x)) = F(x)$ at the boundary of W. By looking at a nearly-shortest path from y to x_0, we deduce that $\mathrm{Lip}_W(f) > b$. If we could also show that $\mathrm{Lip}_{\partial W}(f) = b$ (which is trivially the case when $T \subseteq \mathbb{R}$, but not for a more general metric tree T) we would have a contradiction to the AMLE property of f. Instead of proving this for the particular W constructed above, we will show that there exists a smaller W for which the analogous statement holds.

Consider the function

$$G(s) \stackrel{\text{def}}{=} \sup_{\substack{x \in \overline{W} \\ d_X(x_0, x) = s}} d_T(t, f(x)),$$

which is defined on the interval (s_1, s_2), where s_1 and s_2 are the infimum and supremum of the set $\{d_X(x_0, x) : \ x \in W\}$, respectively. By assumption $G(s)$ lies above the line $bs + c$ for some $s \in (s_1, s_2)$, though not for s_1 and s_2. Hence, if we define

$$M \stackrel{\text{def}}{=} \sup \big\{G(s) - bs - c : \ s \in (s_1, s_2)\big\}$$

then $M > 0$. Write

$$S \stackrel{\mathrm{def}}{=} \left\{ \sigma \in [s_1, s_2] : \limsup_{\substack{s \to \sigma \\ s \in (s_1, s_2)}} \big(G(s) - bs - c \big) = M \right\},$$

and note that S is a nonempty closed subset of $[s_1, s_2]$, so that $s_0 \stackrel{\mathrm{def}}{=} \inf S \in S$.

For $\varepsilon > 0$ and $x \in X$ define

$$F_\varepsilon(x) \stackrel{\mathrm{def}}{=} (b + \varepsilon) d_X(x_0, x) + M + c - \varepsilon s_0 - \varepsilon^2,$$

and

$$W_\varepsilon \stackrel{\mathrm{def}}{=} \big\{ x \in W : \; d_T(t, f(x)) > F_\varepsilon(x) \big\}.$$

Observe that $W_\varepsilon \neq \emptyset$ for all $\varepsilon > 0$. To see this fix $\delta > 0$. Since $s_0 \in S$ there exists $s \in (s_1, s_2)$ such that $|s - s_0| \leqslant \delta$ and $G(s) - bs - c \geqslant M - \delta$. By the definition of $G(s)$, there is $z_0 \in \overline{W}$ satisfying $d_X(z_0, x_0) = s$ and $G(s) \leqslant d_T(t, f(z_0)) + \delta$. Since f is continuous at z_0, there is $\eta \in (0, \delta)$ such that if $d_X(z, z_0) < \eta$ then $d_T(f(z), f(z_0)) < \delta$. Take $z \in W$ with $d_X(z, z_0) < \eta$. Then,

$$\begin{aligned}
d_T(t, f(z)) &> d_T(t, f(z_0)) - \delta \\
&\geqslant G(s) - 2\delta \\
&\geqslant M + bs + c - 3\delta \\
&\geqslant M + bs_0 + c - (3 + b)\delta \\
&= F_\varepsilon(z) - (b + \varepsilon) d_X(x_0, z) + bs_0 + \varepsilon s_0 + \varepsilon^2 - (3 + b)\delta \\
&> F_\varepsilon(z) - (b + \varepsilon)(s + \eta) + bs_0 + \varepsilon s_0 + \varepsilon^2 - (3 + b)\delta \\
&\geqslant F_\varepsilon(z) - (b + \varepsilon)(s_0 + \delta + \eta) + bs_0 + \varepsilon s_0 + \varepsilon^2 - (3 + b)\delta \\
&> F_\varepsilon(z) + \varepsilon^2 - (3\delta + 2\varepsilon\delta + 3b\delta).
\end{aligned}$$

Thus for δ small enough we have $z \in W_\varepsilon$. The following claim contains additional properties of the sets W_ε that we will use later.

Claim 8. *The open sets $\{W_\varepsilon\}_{\varepsilon > 0}$ have the following properties:*

(1) *If $0 < \varepsilon_1 < \varepsilon_2$ then $\overline{W_{\varepsilon_1}} \subseteq W_{\varepsilon_2}$,*
(2) $\lim_{\varepsilon \to 0} \sup_{x \in W_\varepsilon} \big| d_X(x_0, x) - s_0 \big| = 0$,
(3) $\lim_{\varepsilon \to 0} \sup_{x \in W_\varepsilon} \big| d_T(t, f(x)) - (M + bs_0 + c) \big| = 0$.

Proof. Fix $0 < \varepsilon_1 < \varepsilon_2$ and $x \in \overline{W_{\varepsilon_1}}$. Write $s = d_X(x_0, x)$. Since $d_T(t, f(x)) \geqslant F_{\varepsilon_1}(x)$, we have $G(s) - bs - c \geqslant \varepsilon_1 s + M - \varepsilon_1 s_0 - \varepsilon_1^2$. By the definition of M, this implies that $s \leqslant s_0 + \varepsilon_1$. Hence,

$$\begin{aligned}
d_T(t, f(x)) &\geqslant (b + \varepsilon_1)s + M + c - \varepsilon_1 s_0 - \varepsilon_1^2 = F_{\varepsilon_2}(x) + (\varepsilon_2 - \varepsilon_1)s_0 + \varepsilon_2^2 - \varepsilon_1^2 - (\varepsilon_2 - \varepsilon_1)s \\
&\geqslant F_{\varepsilon_2}(x) + (\varepsilon_2 - \varepsilon_1)s_0 + \varepsilon_2^2 - \varepsilon_1^2 - (\varepsilon_2 - \varepsilon_1)(s_0 + \varepsilon_1) = F_{\varepsilon_2}(x) + \varepsilon_2(\varepsilon_2 - \varepsilon_1) > F_{\varepsilon_2}(x).
\end{aligned}$$

Thus $x \in W_{\varepsilon_2}$, proving the first assertion of Claim 8.

To prove the second assertion of Claim 8, note that we have already proved above that if $x \in W_\varepsilon$ then $d_X(x_0, x) \leqslant s_0 + \varepsilon$. Thus, if the second assertion of Claim 8 fails there is some $\delta > 0$ and a sequence $\{\varepsilon_n\}_{n=1}^\infty \subseteq [0, 1]$ with $\lim_{n \to \infty} \varepsilon_n = 0$, such that for each $n \in \mathbb{N}$ there is

$z_n \in W_{\varepsilon_n}$ with $d_X(z_n, x_0) \leqslant s_0 - \delta$. Write $\sigma_n = d_X(z_n, x_0)$, and by passing to a subsequence assume that $\lim_{n\to\infty} \sigma_n = \sigma_\infty$ exists. Then $\sigma_\infty \leqslant s_0 - \delta$ and,

$$\begin{aligned}\limsup_{n\to\infty} \big(G(\sigma_n) - b\sigma_n - c\big) &\geqslant \limsup_{n\to\infty} d_T(t, f(z_n)) - b\sigma_\infty - c \geqslant \limsup_{n\to\infty} F_{\varepsilon_n}(z_n) - b\sigma_\infty - c \\ &= \limsup_{n\to\infty} \big((b+\varepsilon_n)\sigma_n + M + c - \varepsilon_n s_0 - \varepsilon_n^2\big) - b\sigma_\infty - c = M.\end{aligned}$$

Thus $\sigma_\infty \in S$. But since $\sigma_\infty \leqslant s_0 - \delta$, this contradicts the choice of s_0 as the minimum of S. The proof of the second assertion of Claim 8 is complete. The third assertion of Claim 8 now follows, since if $x \in W_\varepsilon$ then by writing $s = d_X(x, x_0)$ we see that

$$\begin{aligned}b(s - s_0) &\geqslant [(G(s) - bs - c) - M] + b(s - s_0) \geqslant d_T(t, f(x)) - (M + bs_0 + c) \\ &\geqslant F_\varepsilon(x) - (M + bs_0 + c) = b(s - s_0) + \varepsilon s - \varepsilon s_0 - \varepsilon^2.\end{aligned}$$

Thus

$$\sup_{x\in W_\varepsilon} \big|d_T(t, f(x)) - (M + bs_0 + c)\big| \leqslant b \sup_{x\in W_\varepsilon} \big|d_X(x_0, x) - s_0\big| + \varepsilon s_0 + \varepsilon^2,$$

and therefore the third assertion of Claim 8 follows from the second assertion of Claim 8. $\square$

We are now in position to conclude the proof of Proposition 7. Let V be the the set of vertices of the metric tree T. We claim that for all $\varepsilon > 0$ such that $\varepsilon s_0 + \varepsilon^2 \leqslant M$ we have $f(W_\varepsilon) \cap V \neq \emptyset$ (recall that by our assumption we have $M > 0$). Indeed, if $x \in \partial W_\varepsilon$ then either $d_T(t, f(x)) = F_\varepsilon(x)$ or $x \in \partial W$. In the latter case, by assumption we have $d_T(t, f(x)) \leqslant b d_X(x_0, x) + c \leqslant F_\varepsilon(x)$, where the last inequality follows from $\varepsilon s_0 + \varepsilon^2 \leqslant M$. Thus, by the definition of W_ε, the function f does not satisfy T-comparison on W_ε. Since f is an AMLE of $f|_{\partial U}$, Proposition 5, combined with Proposition 6, now implies that $f|_{W_\varepsilon}$ must take values in V.

Due to part (1) of Claim 8, there exists $v \in V$ such that $v \in \bigcap_{\varepsilon>0} W_\varepsilon$. Let W'_ε be the connected component of W_ε whose image under f contains v. By part (3) of Claim 8, for ε small enough we have $f(W'_\varepsilon) \cap V = \{v\}$. Since, by the definition of W_ε and the connectedness of W'_ε, for $x \in \partial W'_\varepsilon$ we have $d_T(t, f(x)) = F_\varepsilon(x)$, by considering a nearly-shortest path from a point in W'_ε to x_0 we see that $\mathrm{Lip}_{W'_\varepsilon}(f) > b + \varepsilon$. Since f is an AMLE of $f|_{\partial U}$, it follows that $\mathrm{Lip}_{\partial W'_\varepsilon}(f) > b + \varepsilon$. This implies that there are distinct $x_\varepsilon, y_\varepsilon \in \partial W'_\varepsilon$ such that $d_T(f(x_\varepsilon), f(y_\varepsilon)) > (b+\varepsilon) d_X(x_\varepsilon, y_\varepsilon)$. But, since $d_T(t, f(x_\varepsilon)) = F_\varepsilon(x_\varepsilon)$ and $d_T(t, f(y_\varepsilon)) = F_\varepsilon(y_\varepsilon)$, it must be the case that the distance from t of both $f(x_\varepsilon)$ and $f(y_\varepsilon)$ is at least their distance from v. Indeed, if $t \in \bigcap_{\varepsilon>0} f(W'_\varepsilon)$ then it would follow from part (3) of Claim 8 that $t = v$, and there is nothing to prove. Otherwise, for ε small enough $t \notin f(W'_\varepsilon)$, and therefore, since T is a tree, if at least one of the points $f(x_\varepsilon), f(y_\varepsilon)$ is closer to t than to v then the points $t, f(x_\varepsilon), f(y_\varepsilon)$ all lie on the same geodesic in T, implying that:

$$\begin{aligned}d_T(f(x_\varepsilon), f(y_\varepsilon)) &= \big|d_T(t, f(x_\varepsilon)) - d_T(t, f(y_\varepsilon))\big| = |F_\varepsilon(x_\varepsilon) - F_\varepsilon(y_\varepsilon)| \\ &= (b+\varepsilon)\big|d_X(x_0, x_\varepsilon) - d_X(x_0, y_\varepsilon)\big| \leqslant (b+\varepsilon) d_X(x_\varepsilon, y_\varepsilon),\end{aligned}$$

a contradiction to the choice of $x_\varepsilon, y_\varepsilon$.

Having proved that both $f(x_\varepsilon)$ and $f(y_\varepsilon)$ lie further away from t than v, if we consider a nearly shortest-path between x_ε and y_ε, it must include points $x_1, x_2 \in \overline{W'_\varepsilon}$ such that $f(x_1)$

and $f(x_2)$ lie on the same component I of $T \setminus V$ — on the other side of v from t — and satisfy

$$\frac{d_T(f(x_1), f(x_2))}{d_X(x_1, x_2)} > (b + \varepsilon). \tag{12}$$

Suppose that $f(x_1)$ is closer to t than $f(x_2)$. Note that

$$\begin{aligned} d_T(f(x_1), f(x_\varepsilon)) = d_T(t, f(x_\varepsilon)) - d_T(t, f(x_1)) \leqslant F_\varepsilon(x_\varepsilon) - F_\varepsilon(x_1) \\ = (b+\varepsilon)\big(d_X(x_0, x_\varepsilon) - d_X(x_0, x_1)\big) \leqslant (b+\varepsilon) d_X(x_\varepsilon, x_1). \end{aligned}$$

Moreover, if $x = x_1$ then trivially $d_T(f(x_1), f(x)) \leqslant (b+\varepsilon) d_X(x_1, x)$. Thus, if we let $J \subseteq I$ be the open interval joining $f(x_\varepsilon)$ and $f(x_1)$, then $d_T(f(x_1), f(\cdot)) \leqslant (b+\varepsilon) d_X(x_1, \cdot)$ on $\partial\left(f^{-1}(J) \cap W'_\varepsilon\right)$. By (12) we now have a violation of T-comparison on $f^{-1}(J)$, which contradicts Proposition 5. □

3. Uniqueness

In this section we prove the uniqueness half of Theorem 1 (which does not require the locally compact assumption) as Lemma 12 below. Before doing so, we prove some preliminary lemmas.

Lemma 9. *Suppose that X is a length space, that $Y \subseteq X$ is closed, that $f : Y \to \mathbb{R}$ is Lipschitz and bounded, and that $\widetilde{f} : X \to \mathbb{R}$ is the AMLE of f. (Existence and uniqueness of $\widetilde{f}$ are proved in* [42]*.) Suppose that $g : X \to \mathbb{R}$ is another bounded and continuous extension of f, and that for some fixed $\delta > 0$, this g satisfies comparison with distance functions from above on every radius δ ball centered in $X \setminus Y$. Then $g \leqslant \widetilde{f}$ on X.*

Proof. This is proved (though not explicitly stated) in [42]. Precisely, it is shown there that for $\varepsilon > 0$, the comparison with distance functions from above on balls of radius larger than 2ε implies that the first player in a modified tug of war game (with game position v_k and step size ε) can make $g(v_k)$ a submartingale until the termination of the game, which in turn implies that $g \leqslant f_\varepsilon$ where f_ε is the value of this game. It is also shown that $\lim_{\varepsilon \to 0} f_\varepsilon = \widetilde{f}$ holds on X. Taking $\varepsilon \to 0$ (and noting $2\varepsilon < \delta$ for small enough ε) gives $g \leqslant \widetilde{f}$. □

The following was proved in [2, Lem. 5]. The statement in [2] was only made for the special case $X \subseteq \mathbb{R}^n$, but the (short) proof was not specific to $\mathbb{R}^n$. For completeness, we copy the proof from [2], adapted to our notation. We will vary the presentation just slightly — using suprema over open balls instead of maxima over closed balls — because in our context (since we do not assume any kind of local compactness) maxima of continuous functions on closed balls are not necessarily obtained.

Lemma 10. *Let (X, d_X) be a length space, $x_0 \in X$ and $\varepsilon > 0$. Suppose that $f : X \to \mathbb{R}$ satisfies comparison with distance functions from above on a domain containing $B(x_0, 2\varepsilon)$. Write*

$$f^\varepsilon(x) \stackrel{\mathrm{def}}{=} \sup_{B(x,\varepsilon)} f, \quad f_\varepsilon(x) \stackrel{\mathrm{def}}{=} \inf_{B(x,\varepsilon)} f$$

and

$$S_\varepsilon^+ f(x) \stackrel{\mathrm{def}}{=} \sup_{y \in B(x,\varepsilon)} \frac{f(y) - f(x)}{\varepsilon}, \quad S_\varepsilon^- f(x) \stackrel{\mathrm{def}}{=} \sup_{y \in B(x,\varepsilon)} \frac{f(x) - f(y)}{\varepsilon}.$$

Then

$$S_\varepsilon^- f^\varepsilon(x_0) \leqslant S_\varepsilon^+ f^\varepsilon(x_0).$$

Proof. For $\delta > 0$ we may select $y_0 \in B(x_0, \varepsilon)$ and $z_0 \in B(x_0, 2\varepsilon)$ such that $|f(y_0) - f^\varepsilon(x_0)| \leqslant \delta$ and $|f(z_0) - f^{2\varepsilon}(x_0)| \leqslant \delta$. Then,

$$\begin{aligned} \varepsilon\left(S_\varepsilon^- f^\varepsilon(x_0) - S_\varepsilon^+ f^\varepsilon(x_0)\right) &= 2f^\varepsilon(x_0) - (f^\varepsilon)^\varepsilon(x_0) - (f^\varepsilon)_\varepsilon(x_0) \\ &\leqslant 2f^\varepsilon(x_0) - f^{2\varepsilon}(x_0) - f(x_0) \\ &\leqslant 2f(y_0) - f(z_0) - f(x_0) + 2\delta, \end{aligned} \tag{13}$$

where we used the fact that $(f^\varepsilon)^\varepsilon(x_0) = f^{2\varepsilon}(x_0)$ (since X is a length space), and that by definition $(f^\varepsilon)_\varepsilon(x_0) \geqslant f(x_0)$.

Note that if $d_X(w, x_0) = 2\varepsilon$ then

$$f(w) \leqslant f^{2\varepsilon}(x_0) = f(x_0) + \frac{f^{2\varepsilon}(x_0) - f(x_0)}{2\varepsilon} d_X(w, x_0).$$

Hence for all $w \in \partial\big(B(x_0, 2\varepsilon) \smallsetminus \{x_0\}\big)$ we have

$$f(w) \leqslant f(x_0) + \frac{f^{2\varepsilon}(x_0) - f(x_0)}{2\varepsilon} d_X(w, x_0). \tag{14}$$

Since $f^{2\varepsilon}(x_0) - f(x_0) \geqslant 0$, we may apply the fact that f satisfies comparison with distance functions from above to deduce that (14) holds for every $w \in B(x_0, 2\varepsilon) \smallsetminus \{x_0\}$, and thus for every $w \in B(x_0, 2\varepsilon)$. Substituting $w = y_0$, we see that

$$\begin{aligned} 2f(y_0) - f(x_0) - f(z_0) &\leqslant f(x_0) - f(z_0) + \frac{f^{2\varepsilon}(x_0) - f(x_0)}{\varepsilon} d_X(y_0, x_0) \\ &\leqslant -\left(1 - \frac{d_X(y_0, x_0)}{\varepsilon}\right)\left(f^{2\varepsilon}(x_0) - f(x_0)\right) + \delta \\ &\leqslant \delta, \end{aligned} \tag{15}$$

where we used the fact that $d_X(y_0, x_0) \leqslant \varepsilon$. Since (15) holds for all $\delta > 0$, the required result follows from a combination of (13) and (15). □

Lemma 11. *Assume the following structures and definitions:*

(1) *A length space X, a closed $Y \subseteq X$, a metric tree T, and a Lipschitz $f : Y \to T$.*
(2) *A fixed $x_0 \in X \smallsetminus Y$ and the set $\widehat{X}$ defined as the space of finite-length closed paths in X (parameterized at unit speed) that begin at x_0 and remain in $X \smallsetminus Y$ except possibly at right endpoints.*
(3) *A metric on $\widehat{X}$ defined as follows: given $\widehat{u}, \widehat{v} \in \widehat{X}$, $d_{\widehat{X}}(\widehat{u}, \widehat{v})$ is the sum of the lengths of the portions of the two paths that occur* after *the largest time at which $\widehat{u}$ and $\widehat{v}$ agree. (Note that $\widehat{X}$ is an $\mathbb{R}$-tree under this metric.)*
(4) *The "covering map" $M : \widehat{X} \to X$ that sends a path in $\widehat{X}$ to its right endpoint.*

If $\widetilde{f} : X \to T$ is an AMLE of f, then $\widehat{f} \stackrel{\mathrm{def}}{=} \widetilde{f} \circ M : \widehat{X} \to T$ is an AMLE of $f \circ M$ (which is defined on $\widehat{Y} \stackrel{\mathrm{def}}{=} M^{-1}(Y)$).

Proof. First we claim that M is path-length preserving. That is, if $\widehat{\gamma}$ is any rectifiable path in $\widehat{X}$ then $\gamma \stackrel{\mathrm{def}}{=} M \circ \widehat{\gamma}$ is a rectifiable path of the same length in X. This is true by definition if $L \circ \widehat{\gamma}$ (here $L(\cdot)$ denotes path length) is strictly increasing, and similarly if $L \circ \widehat{\gamma}$ is strictly decreasing. Since the length of $\widehat{\gamma}$ is the total variation of $L \circ \widehat{\gamma}$ (and the latter is finite), the general statement can be derived by approximating $\widehat{\gamma}$ with paths for which $L \circ \widehat{\gamma}$ is piecewise monotone. To do this, first note that one can take an increasing set of times $0 = t_0, t_1, \ldots, t_k$ such that the total variation of $L \circ \widehat{\gamma}$ restricted to those times is arbitrarily close to the unrestricted total variation. Then the length of γ traversed between times t_j and t_{j+1} is at least $r \stackrel{\mathrm{def}}{=} |L \circ \widehat{\gamma}(t_j) - L \circ \widehat{\gamma}(t_{j+1})|$ — this is because the longer of the two paths $\widehat{\gamma}(t_j)$ and $\widehat{\gamma}(t_{j+1})$ contains a segment of length at least r that is not part of the other, and γ must traverse all the points of that segment in order (or in reverse order) somewhere between times t_j and t_{j+1}.

Consider an open subset $\widehat{W} \subseteq \widehat{X} \setminus \widehat{Y}$ and note that $W \stackrel{\mathrm{def}}{=} M\left(\widehat{W}\right)$ is also open. We need to show that if $\widetilde{f}$ is an AMLE of f then we cannot have $\mathrm{Lip}_{\widehat{W}}\left(\widehat{f}\right) > \mathrm{Lip}_{\partial\widehat{W}}\left(\widehat{f}\right)$.

Indeed, suppose we had $\mathrm{Lip}_{\widehat{W}}\left(\widehat{f}\right) > \mathrm{Lip}_{\partial\widehat{W}}\left(\widehat{f}\right)$. Then we could find a path $\widehat{\gamma}$ within $\widehat{W}$ connecting points $a, b \in \widehat{W}$ such that

$$\frac{d_{\widehat{X}}\left(\widehat{f}(a), \widehat{f}(b)\right)}{L(\gamma)} > m$$

for some

$$m > \mathrm{Lip}_{\partial\widehat{W}}\left(\widehat{f}\right).$$

Since $d_{\widehat{X}}\left(\widehat{f}(a), \widehat{f}(\widehat{\gamma}(s))\right)$ is Lipschitz (hence a.e. differentiable) in s, we can find an s_0 at which its derivative is greater than m. Thus, for all sufficiently small ε_0, writing $x_0 = \gamma(s_0)$ (where $\gamma = M \circ \widehat{\gamma}$), we can find points $x_1, x_{-1} \in W$ such that $d_X(x_1, x_0) = d_X(x_{-1}, x_0) = \varepsilon_0$ and $\widetilde{f}(x_1), \widetilde{f}(x_{-1})$ are both at distance greater than $m\varepsilon_0$ from $\widetilde{f}(x_0)$ and lie in distinct components of $T \setminus \left\{\widetilde{f}(x_0)\right\}$.

Now consider some $\varepsilon > 0$ much smaller than ε_0. Since $\widetilde{f}$ is an AMLE of f, the T-comparison property implies that the function

$$g(\cdot) \stackrel{\mathrm{def}}{=} d_T\left(\widetilde{f}(\cdot), \widetilde{f}(x_0)\right)$$

satisfies comparison with distance functions from above. Moreover, along any near-geodesic from x_0 to x_1, the function g increases at an average speed greater than m. Write, as before, $g^\varepsilon(x) \stackrel{\mathrm{def}}{=} \sup_{B(x,\varepsilon)} g$. Let C be the Lipschitz constant of $\widetilde{f}$, so that $|g - g^\varepsilon| \leqslant C\varepsilon$. Note that $g(x_0) = 0$ and $g(x_1) > m\varepsilon_0$, and therefore $g^\varepsilon(x_1) - g^\varepsilon(x_0) > m\varepsilon_0 - 2C\varepsilon$. By considering a near-geodesic from x_0 to x_1, this implies that when ε is small enough, we can find points $y_1, y_2 \in B(x_0, \varepsilon_0)$ with $g^\varepsilon(y_2) - g^\varepsilon(y_1) > m\varepsilon$ and $d_T(y_1, y_2) \leqslant \varepsilon$, such that $\widetilde{f}(y_1)$ and $\widetilde{f}(y_2)$ lie in the same component of $T \setminus \left\{\widetilde{f}(x_0)\right\}$ as $\widetilde{f}(x_1)$, and both $\widetilde{f}(y_1)$ and $\widetilde{f}(y_2)$ are at distance at least $3C\varepsilon$ from $\widetilde{f}(x_0)$.

Fix $\delta > 0$. Applying Lemma 10 inductively we obtain a sequence of points $\{y_i\}_{i=1}^k$ such that for all $i \in \{1, \ldots, k-1\}$ we have $d_X(y_i, y_{i+1}) \leqslant \varepsilon$ and for all $i \in \{1, \ldots, k-2\}$,

$$g^\varepsilon(y_{i+2}) - g^\varepsilon(y_{i+1}) \geqslant g^\varepsilon(y_{i+1}) - g^\varepsilon(y_i) - \frac{\delta}{2^i}. \tag{16}$$

This iterative construction can continue until the first k for which y_k has distance at most ε from ∂W. It follows from (16) that

$$g^\varepsilon(y_{i+1}) - g^\varepsilon(y_i) \geqslant g^\varepsilon(y_2) - g^\varepsilon(y_1) - \sum_{j=1}^{\infty} \frac{\delta}{2^j} > m\varepsilon - \delta.$$

Thus, assuming δ is small enough, we have $g^\varepsilon(y_{i+1}) - g^\varepsilon(y_i) > m\varepsilon$ for all $i \geqslant 1$. It follows that for all $j > i \geqslant 1$ we have

$$g^\varepsilon(y_j) > g^\varepsilon(y_i) + (j-i)m\varepsilon. \tag{17}$$

A consequence of (17) is that, since g^ε is bounded, the above construction cannot continue indefinitely without reaching a point within ε distance from ∂W.

Another consequence of (17) and the fact that $\widetilde{f}$ is C-Lipschitz is that for all $j > i$,

$$g(y_j) > g(y_i) + (j-i)m\varepsilon - 2C\varepsilon. \tag{18}$$

In particular, since $g(y_1) \geqslant 3C\varepsilon$, it follows from (18) that for all $i \geqslant 1$ we have $g(y_i) \geqslant C\varepsilon$. This implies that the points $\left\{\widetilde{f}(y_i)\right\}_{i=1}^{\infty}$ are all in the same component of $T \smallsetminus \left\{\widetilde{f}(x_0)\right\}$ as $\widetilde{f}(x_1)$, since otherwise if i is the first index such that $\widetilde{f}(y_i)$ is not in this component, then $d_T\left(\widetilde{f}(y_i), \widetilde{f}(y_{i-1})\right) \geqslant 2C\varepsilon$, contradicting the fact that $\widetilde{f}$ is C-Lipschitz.

We similarly construct the sequence $\{z_i\}_{i=1}^{\ell}$, starting with the point x_{-1} instead of the point x_1, such that for all $j > i$,

$$g(z_j) > g(z_i) + (j-i)m\varepsilon - 2C\varepsilon.$$

As before, the entire sequence $\left\{\widetilde{f}(z_i)\right\}_{i=1}^{\ell}$ must remain in the same component of $T \smallsetminus \left\{\widetilde{f}(x_0)\right\}$ as $f(x_{-1})$ and z_ℓ is within ε distance from ∂W.

Now consider some M pre-image $\widehat{x}_0$ of x_0 that is contained in $\widehat{W}$, and let D denote the the distance from $\widehat{x}_0$ to $\partial\widehat{W}$. Among the rectifiable paths in W from one boundary point of W to another that pass through all the z_k in reverse order and the subsequently the y_k in order, let γ_0 be one which is near the shortest. Take any path $\widehat{\gamma}_0$ through $\widehat{x}_0$ such that $M \circ \widehat{\gamma}_0 = \gamma_0$. Then consider a maximal arc of this path contained in $\widehat{W}$ and containing $\widehat{x}_0$ (which necessarily has length at least D and connects two points on $\partial\widehat{W}$). The length of this maximal arc is at least D and the change in $\widehat{f}$ from one endpoint of the arc to the other is (in distance) at least m times the length of the arc, plus an $O(\varepsilon_0)$ error, which contradicts the definition of m. □

Lemma 12. *Given a length space (X, d_X), a closed $Y \subseteq X$, a metric tree (T, d_T), and a Lipschitz function $f : Y \to T$, there exists at most one $\widetilde{f} : X \to T$ which is an AMLE of f.*

Proof. By Proposition 7 we must show that if $g, h : X \to T$ are continuous functions such that $g = h$ on Y, and g, h both satisfy T-comparison on $X \smallsetminus Y$, then $g = h$ throughout X. By Lemma 11, it is enough to prove this in the case that X is an $\mathbb{R}$-tree: in particular, we may assume that X and $X \smallsetminus Y$ are simply connected (note that the connected components of an open subset of a metric tree are simply connected).

For the sake of obtaining a contradiction, suppose that g, h satisfy these hypotheses, but $g(x) \neq h(x)$ for some $x \in X \smallsetminus Y$. Then the hypotheses still hold if we replace Y with complement of the connected component of $\{y \in X : \; g(y) \neq h(y)\}$ containing x, and redefine g to be equal to h on all other connected components of $\{y \in X : \; g(y) \neq h(y)\}$. In other words, we lose no generality in assuming $g(x) \neq h(x)$ for all $x \in X \setminus Y$. The pair $(g(\cdot), h(\cdot))$ may then be viewed as a map from $X \setminus Y$ to

$$\mathcal{T} \stackrel{\text{def}}{=} \{(t_1, t_2) \in T \times T : \; t_1 \neq t_2\}.$$

We will use this map to define a certain pair of real-valued functions on X.

To this end, consider an arbitrary continuous function $P : [0,1] \to \mathcal{T}$. Define functions $t_1, t_2 : [0,1] \to T$ by writing $(t_1(s), t_2(s)) = P(s)$. Let $I((t_1, t_2))$ denote the geodesic joining t_1 and t_2 in T. For every $s \in [0,1]$, we will define an isometry Ψ_s^P from $I(P(s))$ to an interval $(a_1(s), a_2(s))$ of $\mathbb{R}$, sending t_1 to $a_1(s)$ and t_2 to $a_2(s)$. Clearly, the values of $a_1(s)$ and $a_2(s)$ determine the isometry, and for each s, we must have

$$a_2(s) - a_1(s) = d_T(t_1(s), t_2(s)).$$

However, the latter observation only determines $a_1(s)$ and $a_2(s)$ up to the addition of a single constant to both values. This constant is determined by the following requirements:

(1) $a_1(0) = 0$.
(2) For each fixed $t \in T$, the function $s \mapsto \Psi_s^P(t)$ is constant on every connected interval of the (open) set $\{s \in [0,1] : \; t \in I(P(s))\}$.

Informally, at each time s, the geodesic $I(P(s))$ is "glued" isometrically to the interval $(a_1(s), a_2(s)) \subseteq \mathbb{R}$. As s increases, if t_1 and t_2 move closer to each other, then points are being removed from the ends of $I(P(s))$, and these points are "unglued" from $\mathbb{R}$. As t_1 and t_2 move further from each other, new points are added to the geodesic and these new points are glued back onto $\mathbb{R}$.

When P' and P are paths in $\mathcal{T}$ as above, write $P \sim P'$ if $P(0) = P'(0)$ and $P(1) = P'(1)$ and for each $s \in [0,1]$ we have $I(P(s)) \cap I(P'(s)) \neq \emptyset$. We claim that in this case $\Psi_1^P = \Psi_1^{P'}$. Note that given s and $t \in I(P(s)) \cap I(P'(s))$, requirement (2) above implies that Ψ_s^P and $\Psi_s^{P'}$ agree up to additive constant on a neighborhood of s; namely, they must agree on the connected component of $\{s' : \; t \in I(P(s')) \cap I(P'(s'))\}$ containing s. Since Ψ_s^P and $\Psi_s^{P'}$ agree up to additive constant on an open neighborhood of every $s \in [0,1]$, they must be equal up to additive constant throughout the interval, and requirement (1) above implies that this constant is zero.

A corollary of this discussion is that $\Psi_1^P = \Psi_1^{P'}$ whenever P and P' are homotopically equivalent paths in $\mathcal{T}$. To obtain this, it is enough to observe that if P^r is a homotopy, with $r \in [0,1]$ and $P = P^0, P' = P^1$, then for each $r \in [0,1]$ we have $P^r \sim P^{r'}$ for all r' in some neighborhood of r. This implies that $\Psi_1^{P^r}$ (as a function of r) is constant on a neighborhood of each point in $[0,1]$, hence constant throughout $[0,1]$.

Since we are assuming $X \setminus Y$ is simply connected, we can fix a point $x_0 \in X \setminus Y$ and define the pair $(a_1(x), a_2(x))$ to be the value $(a_1(1), a_2(1))$ obtained above by taking $P(s) = (g(p(s)), h(p(s)))$ where $p : [0,1] \to X$ is any path from x_0 to x. For y in some neighborhood of each $x \in X \setminus Y$, and for some $t \in T$, we have that $a_1(y)$ is an affine function of $d_T(t, g(y))$. For this, it suffices to take a neighborhood and a t such that $t \in I(P(s))$ throughout that neighborhood and use requirement (2) above. In this neighborhood, T-comparison implies that a_1 satisfies comparison with distance functions from below and a_2 satisfies comparison with distance functions from above.

Now let C be the Lipschitz constant of f, and for $\delta > 0$ write

$$U_\delta \stackrel{\text{def}}{=} \{x : \; a_1(x) < a_2(x) - 2C\delta\}.$$

We claim that the argument in the paragraph just above implies that a_2 satisfies comparison with distance functions from above on $B(x,\delta)$ for any $x \in U_\delta$. To see this, first observe that the fact that a_2 is Lipschitz with constant at most C implies that $B(x,\delta) \subseteq X \smallsetminus Y$, and then take t to be the midpoint in T between $g(x)$ and $h(x)$, noting that since g and h are both Lipschitz with constant C, we have that t is on the geodesic between $g(x')$ and $h(x')$ for all $x' \in B(x,\delta)$.

We may now apply Lemma 9 (where the Y of the lemma statement is chosen so that $U_\delta = X \smallsetminus Y$) to see that $a_2 \leqslant \widetilde{a}_2$ on U_δ, where $\widetilde{a}_2$ is the AMLE of the restriction of a_2 to ∂U_δ. By symmetry, we may apply the same arguments to $-a_1$ to obtain that $a_1 \geqslant \widetilde{a}_1$ on U_δ. Since $a_1 < a_2$ by construction, we now have

$$\widetilde{a}_1 \leqslant a_1 < a_2 \leqslant \widetilde{a}_2. \tag{19}$$

Now it is a standard fact (see [4]) about AMLE that the suprema and infima of a difference of AMLEs (in this case $\widetilde{a}_2 - \widetilde{a}_1$) is obtained on the boundary set (in this case ∂U_δ). This implies that $\widetilde{a}_1 \leqslant \widetilde{a}_2 \leqslant \widetilde{a}_1 + 2C\delta$ throughout the set U_δ. By (19), we now have that

$$\sup_{x \in U_\delta} |a_1(x) - a_2(x)| \leqslant 2C\delta$$

This implies $\sup_{x\in X} |a_1(x) - a_2(x)| \leqslant 2C\delta$, and since this holds for all $\delta > 0$, we have $a_1 = a_2$ throughout X, a contradiction. □

4. Proofs of Theorem 2 and Theorem 3

We first present the simple proof of Theorem 2, i.e, the local-global result for tree-valued ∞-harmonic functions. The proof is a modification of an argument in [42] from the setting of real-valued mappings to the setting of tree-valued mappings.

Proof of Theorem 2. Let $U \subseteq G \smallsetminus \Omega$ be an open subset of G. Denote

$$L = \mathrm{Lip}_{\overline{U}}(f) = \max_{\substack{x,y\in\overline{U}\\ x\neq y}} \frac{d_T(f(x), f(y))}{d_G(x,y)}. \tag{20}$$

Let $x, y \in \overline{U}$ be points at which the maximum in (20) is attained and $d_G(x,y)$ is maximal among all such points. We will be done if we show that $x, y \in \partial U$. Assume for the sake of contradiction that $x \in U$ (the case $y \in U$ being similar).

If x is in the interior of an edge of G, we could move x slightly along the edge and increase $d_G(x,y)$ without decreasing $d_T(f(x), f(y))/d_G(x,y)$. So, assume that $x \in V$. The fact that f is ∞-harmonic on $V \smallsetminus \Omega \supseteq \{x\}$ means that there exist $u, v \in N_G(x)$ with $d_T(f(u), f(v)) = 2L$ and $f(x)$ is a midpoint between $f(u)$ and $f(v)$ in T. Since $x \in U$ (and f is linear on the edges of G) there exists $\varepsilon > 0$ and $z_u, z_v \in U$ such that $z_u \in xu \in E$, $z_v \in xv \in E$, $d_G(x, z_u) = d_G(x, z_v) = \varepsilon$ and $f(z_u)$ (resp. $f(z_v)$) is the point on the geodesic in T joining $f(x)$ and $f(u)$ (resp. $f(v)$) at distance $L\varepsilon$ from $f(x)$. Because T is a metric tree, either $d_T(f(y), f(z_u)) = d_T(f(x), f(y)) + L\varepsilon$ or $d_T(f(y), f(z_v)) = d_T(f(x), f(y)) + L\varepsilon$. Assume without loss of generality that $d_T(f(y), f(z_u)) = d_T(f(x), f(y)) + L\varepsilon = L(d_G(x,y)+\varepsilon)$. Then

$d_G(z_u,y) = d_G(x,y)+\varepsilon$ and $d_T(f(z_u),f(y))/d_G(z_u,y) = L$, contradicting the maximality of $d_G(x,y)$. $\square$

As in the classical proof that metric convexity and the binary intersection property implies the isometric extension property (see [7, Prop. 1.4]), for the proof of Theorem 3 we shall construct $\widetilde{f}$ by extending to one additional point at a time. The proof of Theorem 3 relies on a specific choice of the ordering of the points for the purpose of such a point-by-point construction. Our argument uses a variant of an algorithm from [28].

Proof of Theorem 3. Write $|V \setminus \Omega| = n$. We shall construct inductively a special ordering $w_1,\ldots,w_n$ of the points of $V\setminus\Omega$, and extend f to these points one by one according to this ordering. Assume that $w_1,\ldots,w_k$ have been defined, as well as the values $\widetilde{f}(w_1),\ldots,\widetilde{f}(w_k)\in Z$ (if $k=0$ this assumption is vacuous).

Write $\Omega_0=\Omega$ and $\Omega_k=\Omega\cup\{w_1,\ldots,w_k\}$. Given distinct $x,y\in V$ we shall say that $x_0,x_1,\ldots x_\ell\in V$ is a path joining x and y which is external to Ω_k if $x_0=x$, $x_\ell=y$, and for all $i\in\{0,\ldots,\ell-1\}$ we have $x_ix_{i+1}\in E$ and $\{x_i,x_{i+1}\}\not\subseteq\Omega_k$. Let $d_k(x,y)$ be the minimum over $\ell\in\mathbb{N}$ such that there exists a path $x_0,x_1,\ldots x_\ell\in V$ joining x and y which is external to Ω_k. If no such path exists we set $d_k(x,y)=\infty$. We also set $d_k(x,x)=0$ for all $x\in V$. Then $d_k:V\times V\to\{0\}\cup\mathbb{N}\cup\{\infty\}$ clearly satisfies the triangle inequality and $d_k(\cdot,\cdot)\geqslant d_G(\cdot,\cdot)$ pointwise.

We distinguish between two cases:

Case 1. For all distinct $x,y\in\Omega_k$ we have $d_k(x,y)=\infty$. In this case order the points of $V\setminus\Omega_k$ arbitrarily, i.e., $V\setminus\Omega_k=\{w_{k+1},\ldots,w_n\}$. If $w\in\{w_{k+1},\ldots,w_n\}$ then by the connectedness of G, there exists a path in G joining w and some point $x_w\in\Omega_k$. Note that x_w is uniquely determined by w, since if there were another path joining w and some point $y_w\in\Omega_k$ which isn't x_w then $d_k(x_w,y_w)<\infty$, contradicting our assumption in Case 1. We can therefore define in this case $\widetilde{f}(w)=\widetilde{f}(x_w)$.

Case 2. For some distinct $x,y\in\Omega_k$ we have $d_k(x,y)<\infty$. In this case define

$$L_k=\max_{\substack{x,y\in\Omega_k\\ x\neq y}}\frac{d_Z\left(\widetilde{f}(x),\widetilde{f}(y)\right)}{d_k(x,y)}. \tag{21}$$

Our assumption implies that $L_k>0$. Choose $x,y\in\Omega_k$ that are distinct and satisfy $L_kd_k(x,y)=d_Z\left(\widetilde{f}(x),\widetilde{f}(y)\right)$. Write $\ell=d_k(x,y)$ and let $x_0,x_1,\ldots x_\ell\in V$ be a path joining x and y which is external to Ω_k. Then $x_1\notin\Omega_k$, so we may define $w_{k+1}=x_1$. We claim that

$$\bigcap_{a\in\Omega_k}B_Z\left(\widetilde{f}(a),L_kd_k(a,w_{k+1})\right)\neq\emptyset. \tag{22}$$

To prove (22), by the fact that Z has the binary intersection property, it suffices to show that for all $a,b\in\Omega_k$ we have

$$B_Z\left(\widetilde{f}(a),L_kd_k(a,w_{k+1})\right)\cap B_Z\left(\widetilde{f}(b),L_kd_k(b,w_{k+1})\right)\neq\emptyset. \tag{23}$$

If either $d_k(a, w_{k+1}) = \infty$ or $d_k(b, w_{k+1}) = \infty$ then (23) is trivial. Assume therefore that $d_k(a, w_{k+1})$ and $d_k(b, w_{k+1})$ are finite. Define $\lambda \in [0, 1]$ by

$$\lambda = \frac{d_k(a, w_{k+1})}{d_k(a, w_{k+1}) + d_k(b, w_{k+1})}. \tag{24}$$

Since Z is metrically convex, there exists a point $z \in Z$ such that

$$d_Z\left(z, \widetilde{f}(a)\right) = \lambda d_Z\left(\widetilde{f}(a), \widetilde{f}(b)\right) \quad \text{and} \quad d_Z\left(z, \widetilde{f}(b)\right) = (1-\lambda) d_Z\left(\widetilde{f}(a), \widetilde{f}(b)\right). \tag{25}$$

The definition of L_k implies

$$d_Z\left(\widetilde{f}(a), \widetilde{f}(b)\right) \leqslant L_k d_k(a, b) \leqslant L_k\big(d_k(a, w_{k+1}) + d_k(b, w_{k+1})\big). \tag{26}$$

Using (24) and (25), we deduce from (26) that

$$d_Z\left(z, \widetilde{f}(a)\right) \leqslant L_k d_k(a, w_{k+1}) \quad \text{and} \quad d_Z\left(z, \widetilde{f}(b)\right) \leqslant L_k d_k(b, w_{k+1}),$$

proving (23). Having proved (22), we let $\widetilde{f}(w_{k+1})$ be an arbitrary point satisfying

$$\widetilde{f}(w_{k+1}) \in \bigcap_{a \in \Omega_k} B_Z\left(\widetilde{f}(a), L_k d_k(a, w_{k+1})\right). \tag{27}$$

The above inductive construction produces a function $\widetilde{f} : V \to Z$ that extends f. We claim that $\widetilde{f}$ is ∞-harmonic on $V \setminus \Omega$. To see this note that for all $x, y \in V$ the sequence $\{d_k(x, y)\}_{k=1}^n \subseteq \{0\} \cup \mathbb{N} \cup \{\infty\}$ is non-decreasing. We shall next show that the sequence $\{L_k\}_{k=1}^n$, defined in (21), is non-increasing. Indeed, assume that $L_{k+1} > 0$ and take distinct $a, b \in \Omega_{k+1}$ such that $L_{k+1} d_{k+1}(a, b) = d_Z\left(\widetilde{f}(a), \widetilde{f}(b)\right)$. If $a, b \in \Omega_k$ then it follows from the definition of L_k that $L_{k+1} \leqslant L_k$, since $d_{k+1}(a, b) \geqslant d_k(a, b)$. By symmetry, it remains to deal with the case $a \in \Omega_k$ and $b = w_{k+1}$. In this case, since by our construction we have $\widetilde{f}(w_{k+1}) \in B_Z\left(\widetilde{f}(a), L_k d_k(a, w_{k+1})\right) \subseteq B_Z\left(\widetilde{f}(a), L_k d_{k+1}(a, w_{k+1})\right)$, it follows once more that $L_{k+1} \leqslant L_k$.

Fix $k \in \{0, \ldots, n-1\}$. If $\widetilde{f}(w_{k+1})$ was defined in Case 1 of our inductive construction, then $\widetilde{f}$ is constant on $N_G(w_{k+1}) \cup \{w_{k+1}\}$, in which fact the ∞-harmonic conditions (6), (7) for $\widetilde{f}$ at w_{k+1} hold trivially. If, on the other hand, $\widetilde{f}(w_{k+1})$ was defined in Case 2 of our inductive construction, then there exist distinct $x, y \in \Omega_k$ with $L_k d_k(x, y) = d_Z\left(\widetilde{f}(x), \widetilde{f}(y)\right)$, such that for $\ell = d_k(x, y)$ there are $x_0, x_1, \ldots x_\ell \in V$ which form a path joining x and y which is external to Ω_k, and $x_1 = w_{k+1}$. For every $i \in \{1, \ldots, \ell - 1\}$ either $x_i \notin \Omega_k$ or $x_{i+1} \notin \Omega_k$, and therefore at least one of the values $\widetilde{f}(x_i), \widetilde{f}(x_{i+1})$ was defined after stage $k+1$ of our inductive construction. This means that for some $j \geqslant k$ we have $|\Omega_j \cap \{x_i, x_{i+1}\}| = 1$ and

$$d_Z\left(\widetilde{f}(x_i), \widetilde{f}(x_{i+1})\right) \leqslant L_j d_j(x_i, x_{i+1}) = L_j \leqslant L_k, \tag{28}$$

where we used the fact that $x_ix_{i+1} \in E$, and therefore, since $|\Omega_j \cap \{x_i, x_{i+1}\}| = 1$, the path x_i, x_{i+1} is external to Ω_j. Thus

$$\begin{aligned}
L_k\ell &= d_Z\left(\widetilde{f}(x), \widetilde{f}(y)\right) \\
&\leqslant d_Z\left(\widetilde{f}(x_0), \widetilde{f}(x_2)\right) + d_Z\left(\widetilde{f}(x_2), \widetilde{f}(x_\ell)\right) \\
&\leqslant d_Z\left(\widetilde{f}(x_0), \widetilde{f}(x_2)\right) + \sum_{i=2}^{\ell-1} d_Z\left(\widetilde{f}(x_i), \widetilde{f}(x_{i+1})\right) \\
&\overset{(28)}{\leqslant} d_Z\left(\widetilde{f}(x_0), \widetilde{f}(x_1)\right) + d_Z\left(\widetilde{f}(x_1), \widetilde{f}(x_2)\right) + L_k(\ell-2) \\
&\overset{(28)}{\leqslant} L_k\ell. \qquad (29)
\end{aligned}$$

It follows that all the inequalities in (29) actually hold as equality. Therefore we have $d_Z\left(\widetilde{f}(x), \widetilde{f}(w_{k+1})\right) = d_Z\left(\widetilde{f}(w_{k+1}), \widetilde{f}(x_2)\right) = L_k$ and $d_Z\left(\widetilde{f}(x), \widetilde{f}(x_2)\right) = 2L_k$. Since by construction $x, x_2 \in N_G(w_{k+1})$, in order to show that $\widetilde{f}$ is ∞-harmonic at w_{k+1} it remains to check that for all $u \in N_G(w_{k+1})$ we have $d_Z\left(\widetilde{f}(u), \widetilde{f}(w_{k+1})\right) \leqslant L_k$. But, our construction ensures that for some $j \geqslant k$ we have $d_Z\left(\widetilde{f}(u), \widetilde{f}(w_{k+1})\right) \leqslant L_j d_j(u, w_{k+1}) = L_j$ (using $uw_{k+1} \in E$), and the required result follows since $L_j \leqslant L_k$.

Denote $L = \mathrm{Lip}_\Omega(f)$. In order to complete the proof of Theorem 3 it remains to argue that $\mathrm{Lip}_V\left(\widetilde{f}\right) = L$. We will do so by proving by induction on k that $\mathrm{Lip}_{\Omega_k}\left(\widetilde{f}\right) = L$. For $k = 0$ this is the definition of L, so assuming that $\mathrm{Lip}_{\Omega_k}\left(\widetilde{f}\right) = L$ we need to deduce that if $x \in \Omega_k$ then $d_Z\left(\widetilde{f}(x), \widetilde{f}(w_{k+1})\right) \leqslant L d_G(x, w_{k+1})$. Denote $\ell = d_G(x, w_{k+1})$ and let $w_{k+1} = y_0, y_1, \ldots, y_\ell = x \in V$ satisfy $y_iy_{i+1} \in E$ for all $i \in \{0, \ldots, \ell-1\}$. Let $j \in \{1, \ldots, \ell\}$ be the smallest index such that $y_j \in \Omega_k$. Since $y_0, y_1, \ldots, y_\ell$ is a shortest path in G joining w_{k+1} and x, and the path $w_{k+1} = y_0, \ldots, y_j$ is external to Ω_k by the definition of j, it follows that $d_k(w_{k+1}, y_j) = d_G(w_{k+1}, y_j) = j$ and $d_G(y_j, x) = \ell - j$. Since $y_j, x \in \Omega_k$, by the inductive hypothesis we have $d_Z\left(\widetilde{f}(y_j), \widetilde{f}(x)\right) \leqslant L(\ell - j)$. Also, it follows from (27) that $d_Z\left(\widetilde{f}(w_{k+1}), \widetilde{f}(y_j)\right) \leqslant L_k d_k(y_j, w_{k+1}) = L_k j \leqslant Lj$, where we used the fact that since $d_k(\cdot,\cdot) \geqslant d_G(\cdot,\cdot)$, it follows from (21) and the inductive hypothesis that $L_k \leqslant \mathrm{Lip}_{\Omega_k}\left(\widetilde{f}\right) = L$. So, in conclusion,

$$d_Z\left(\widetilde{f}(x), \widetilde{f}(w_{k+1})\right) \leqslant d_Z\left(\widetilde{f}(w_{k+1}), \widetilde{f}(y_j)\right) + d_Z\left(\widetilde{f}(y_j), \widetilde{f}(x)\right) \leqslant Lj + L(\ell - j) = L_\ell. \quad \square$$

5. Existence

Here we prove the existence part of Theorem 1, i.e., we establish the following:

Theorem 13. *Let (X, d_X) be a locally compact length space and (T, d_T) a metric tree. Then for every closed $Y \subseteq X$, every Lipschitz mapping $f : Y \to T$ admits an AMLE.*

Proof. Assume first that X is compact. We will construct an AMLE $\widetilde{f}$ of f as a limit of discrete approximations.

For each $\varepsilon \in (0, 1/4)$, let Λ_ε be a finite subset of X such that

$$X \subseteq \bigcup_{x \in \Lambda_\varepsilon} B_X(x, \varepsilon) \quad \text{and} \quad Y \subseteq \bigcup_{y \in \Lambda_\varepsilon \cap Y} B_X(y, \varepsilon). \tag{30}$$

Let G_ε be the graph whose vertices are the elements of Λ_ε, with $x, y \in \Lambda_\varepsilon$ adjacent when $d_X(x, y) \leqslant \sqrt{\varepsilon}$.

For any x and y in Λ_ε, we can find an arbitrarily-close-to-minimal length path between them and a sequence of points $x = x_0, x_1, x_2, \ldots, x_k = y$ spaced at intervals of $\sqrt{\varepsilon} - 2\varepsilon$ along the path, where $k - 1$ is the integer part of $d_X(x, y)/(\sqrt{\varepsilon} - 2\varepsilon)$, and can then find points $\tilde{x}_i \in B(x_i, \varepsilon) \cap \Lambda_\varepsilon$. Since $d(x_i, x_{i+1}) \leqslant \sqrt{\varepsilon}$ we conclude that $d_{G_\varepsilon}(x, y) \leqslant k$. It is also clear that $d_{G_\varepsilon}(x, y) \geqslant d(x, y)/\sqrt{\varepsilon}$. Hence,

$$\left| d_{G_\varepsilon}(x, y)\sqrt{\varepsilon} - d_X(x, y) \right| \leqslant C\sqrt{\varepsilon}, \tag{31}$$

where C depends only on the diameter of X.

Let $\widetilde{f}_\varepsilon$ be an ∞-harmonic extension of $f|_{Y \cap \Lambda_\varepsilon}$ to all of G_ε, the existence of which is due to Theorem 3 (since T is a 1-absolute Lipschitz retract). Note that on Λ_ε we have the pointwise inequality $d_X(\cdot, \cdot) \leqslant \sqrt{\varepsilon} d_{G_\varepsilon}(\cdot, \cdot)$. It follows that the Lipschitz constant of $f|_{Y \cap \Lambda_\varepsilon}$ with respect to the metric $\sqrt{\varepsilon} d_{G_\varepsilon}$ is bounded above by $\mathrm{Lip}_Y(f)$, and hence the Lipschitz constant of $\widetilde{f}_\varepsilon$ with respect to the metric $\sqrt{\varepsilon} d_{G_\varepsilon}$ is also bounded above by $\mathrm{Lip}_Y(f)$.

Let $\mathcal{N}_\varepsilon \subseteq \Lambda_\varepsilon$ be a $\sqrt{\varepsilon}$-net in $(\Lambda_\varepsilon, d_X)$, i.e., a maximal subset of Λ_ε, any two elements of which are separated in the metric d_X by at least $\sqrt{\varepsilon}$. For any distinct $x, y \in \mathcal{N}_\varepsilon$ we have

$$\begin{aligned} d_T\left(\widetilde{f}_\varepsilon(x), \widetilde{f}_\varepsilon(y)\right) &\leqslant \mathrm{Lip}_Y(f)\sqrt{\varepsilon} d_{G_\varepsilon}(x, y) \\ &\overset{(31)}{\leqslant} \mathrm{Lip}_Y(f)\left(d_X(x, y) + C\sqrt{\varepsilon}\right) \leqslant \mathrm{Lip}_Y(f)(1 + C) d_X(x, y). \end{aligned} \tag{32}$$

It follows that we can extends $\widetilde{f}\Big|_{\mathcal{N}_\varepsilon}$ to a function $f_\varepsilon^* : X \to T$ that is Lipschitz with constant $\mathrm{Lip}_Y(f)(1 + C)$ (this extension can be done in an arbitrary way, using the fact that T is a 1-absolute Lipschitz retract). Since the functions f_ε^* are equicontinuous, the Arzela-Ascoli Theorem [38, Thm. 6.1] says that there exists a subsequence $\{\varepsilon_n\}_{n=1}^\infty \subseteq (0, 1/4)$ tending to zero such that $f_{\varepsilon_n}^*$ converges uniformly to $f^* : X \to T$. We aim to show that f^* is an AMLE of f.

By Proposition 7 it is enough to show that for each $t \in T$ and open $W \subseteq X \smallsetminus Y$, $z \in X \smallsetminus W$, $b \geqslant 0$ and $c \in \mathbb{R}$, we have the following:

$$\begin{aligned} &\forall x \in \partial W \quad d_T\left(t, f^*(x)\right) \leqslant b d_X(x, z) + c \\ &\qquad\qquad \implies \quad \forall x \in W \quad d_T\left(t, f^*(x)\right) \leqslant b d_X(x, z) + c. \end{aligned} \tag{33}$$

By uniform convergence, for every $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that for every $n \geqslant n_0$ if for every $x \in \partial W$ we have

$$d_T\left(t, f^*(x)\right) \leqslant b d_X(x, z) + c, \tag{34}$$

then for every $x \in \partial W$ we have

$$d_T\left(t, f_{\varepsilon_n}^*(x)\right) \leqslant b d_X(x, z) + c + \delta. \tag{35}$$

Assume from now on that (34) holds for all $x \in \partial W$. Let $V_{\varepsilon_n} \subseteq \mathcal{N}_{\varepsilon_n}$ be the set of $u \in \mathcal{N}_{\varepsilon_n} \subseteq \Lambda_{\varepsilon_n}$ for which there exists $w \in W$ such that $d_X(u,w) \leqslant \sqrt{\varepsilon_n}$. Define W_{ε_n} to be the open subset of the 1-dimensional simplicial complex corresponding to the graph G_{ε_n} consisting of the union of all the half-open intervals $[u,v)$, where $u,v \in \Lambda_{\varepsilon_n}$, uv is an edge of G_{ε_n} and $u \in V_{\varepsilon_n}$. Any point $v \in \partial W_{\varepsilon_n}$ of the boundary of W_{ε_n} in G_{ε_n} is at d_X-distance greater than $\sqrt{\varepsilon_n}$ from W, but at d_X-distance at most $\sqrt{\varepsilon_n}$ from some point of $\mathcal{N}_{\varepsilon_n}$ whose d_X-distance from W is at most $\sqrt{\varepsilon_n}$. Thus

$$v \in \partial W_{\varepsilon_n} \implies d_X(u, \partial W) \leqslant 2\sqrt{\varepsilon_n}. \tag{36}$$

Let z_{ε_n} be any one of the d_X-closest points of z in $\mathcal{N}_{\varepsilon_n}$. By (30) and the definition of $\mathcal{N}_{\varepsilon_n}$ we have

$$d_X(z_{\varepsilon_n}, z) \leqslant \sqrt{\varepsilon_n} + \varepsilon_n \leqslant 2\sqrt{\varepsilon_n}. \tag{37}$$

Since $f^*_{\varepsilon_n}$ is Lipschitz with constant $\mathrm{Lip}_Y(f)(1+C)$, for every $v \in \partial W_{\varepsilon_n}$ we have

$$\begin{aligned}
d_T\left(t, f^*_{\varepsilon_n}(v)\right) &\overset{(35)\wedge(36)}{\leqslant} b d_X(v,z) + c + \delta + 2\mathrm{Lip}_Y(f)(1+C)\sqrt{\varepsilon_n} \\
&\overset{(37)}{\leqslant} b d_X\left(v, z_{\varepsilon_n}\right) + c + \delta + 2(\mathrm{Lip}_Y(f)(1+C) + b)\sqrt{\varepsilon_n} \\
&\overset{(31)}{\leqslant} b\sqrt{\varepsilon_n} d_{G_{\varepsilon_n}}\left(v, z_{\varepsilon_n}\right) + c + \delta + (2\mathrm{Lip}_Y(f)(1+C) + 2b + bC)\sqrt{\varepsilon_n} \\
&\leqslant b\sqrt{\varepsilon_n} d_{G_{\varepsilon_n}}\left(v, z_{\varepsilon_n}\right) + c + \delta + K\sqrt{\varepsilon_n},
\end{aligned} \tag{38}$$

where $K > 0$ is independent of n. Observe that if $z_{\varepsilon_n} \in W_{\varepsilon_n}$ then since $z \notin W$ we have $d_X(z_{\varepsilon_n}, \partial W) \leqslant 6\sqrt{\varepsilon_n}$. In this case the same argument as above shows that (38) holds for $v = z_{\varepsilon_n}$ as well (with a different value of K). Thus, the bound (38) holds for all $v \in \partial(W_{\varepsilon_n} \smallsetminus \{z_{\varepsilon_n}\})$. By Theorem 2 and Proposition 7, it follows that for every $v \in W_{\varepsilon_n} \smallsetminus \{z_{\varepsilon_n}\}$, and hence also for all $v \in V_{\varepsilon_n}$, we have

$$\begin{aligned}
d_T\left(t, f^*_{\varepsilon_n}(v)\right) &\leqslant b\sqrt{\varepsilon_n} d_{G_{\varepsilon_n}}\left(v, z_{\varepsilon_n}\right) + c + \delta + K\sqrt{\varepsilon_n} \\
&\overset{(31)\wedge(37)}{\leqslant} b d_X\left(v, z\right) + c + \delta + (K + Cb + 2b)\sqrt{\varepsilon_n}.
\end{aligned} \tag{39}$$

Since any point of W is at d_X-distance at most $\varepsilon_n + \sqrt{\varepsilon_n} \leqslant 2\sqrt{\varepsilon_n}$ from V_{ε_n}, and since $f^*_{\varepsilon_n}$ is Lipschitz with constant independent of n, we see from (39) that for some $K' > 0$ independent of n, for all $x \in W$ we have:

$$d_T\left(t, f^*_{\varepsilon_n}(x)\right) \leqslant b d_X\left(x, z\right) + c + \delta + K'\sqrt{\varepsilon_n}. \tag{40}$$

Letting n tend to ∞ in (40), it follows that

$$d_T\left(t, f^*(x)\right) \leqslant b d_X\left(x, z\right) + c + \delta. \tag{41}$$

Since (41) holds for all $\delta > 0$, we have proved the desired implication (33).

When X is locally compact but not necessarily compact, the proof of Theorem 13 follows from a direct reduction to the compact case. Indeed, by Remark 2 it suffices to prove (33) when $b > 0$. In this case, since T is bounded, the upper bounds in (33) are trivial if $d_X(x,z)$ is sufficiently large. Thus, it suffices to prove (33) for the intersection of W with a large enough ball centered at z. $\square$

6. Politics

In this section we prove Proposition 4. We require some notation (in particular, a definition of value) to make the statement of Proposition 4 precise.

A **strategy** for a player is a way of choosing the player's next move as a function of all previously played moves and all previous coin tosses. It is a map from the set of partially played games to moves (or in the case of a **random strategy**, a probability distribution on moves). We might expect a good strategy to be Markovian, i.e., a map from the current state to the next move, but it is useful to allow more general strategies that take into account the history.

Given two strategies $\mathcal{S}_{\mathrm{I}}, \mathcal{S}_{\mathrm{II}}$, let $\mathcal{F}(\mathcal{S}_{\mathrm{I}}, \mathcal{S}_{\mathrm{II}})$ be the expected total payoff (including the running payoffs received) when the players adopt these strategies. We define $\mathcal{F}$ to be some fixed constant C if the game does not terminate with probability one, or if this expectation does not exist.

The **value of the game for player** I is defined as $\sup_{\mathcal{S}_{\mathrm{I}}} \inf_{\mathcal{S}_{\mathrm{II}}} \mathcal{F}(\mathcal{S}_{\mathrm{I}}, \mathcal{S}_{\mathrm{II}})$. The **value for player** II is $\inf_{\mathcal{S}_{\mathrm{II}}} \sup_{\mathcal{S}_{\mathrm{I}}} \mathcal{F}(\mathcal{S}_{\mathrm{I}}, \mathcal{S}_{\mathrm{II}})$. The game has a **value** when these two quantities are equal. It turns out that Politics always has a value for any choice of initial states $x_0 \in V$ and $t_0 \in T$; this is a consequence of a general theorem (since the payoff function is a zero-sum Borel-measurable function of the infinite sequence of moves [33]; see also [41] for more on stochastic games).

Proof of Proposition 4. First we introduce some notation: when the game position is at x_k, we let y_k and z_k denote two of the vertices adjacent to x_k that maximize $d_T\left(\widetilde{f}(x_k), \widetilde{f}(\cdot)\right)$, chosen so that $\widetilde{f}(x_k)$ is the midpoint of $\widetilde{f}(y_k)$ and $\widetilde{f}(z_k)$. Write for $x \in V$,

$$\delta(x) \stackrel{\mathrm{def}}{=} \sup_{y \in N_G(x)} d_T\left(\widetilde{f}(x), \widetilde{f}(y)\right),$$

and

$$M_k \stackrel{\mathrm{def}}{=} \delta(x_k) = d_T\left(\widetilde{f}(x_k), \widetilde{f}(y_k)\right) = d_T\left(\widetilde{f}(x_k), \widetilde{f}(z_k)\right).$$

Using this notation, we now give a strategy for player II that makes $d\left(\widetilde{f}(x_k), t_k\right)$ plus the total payoff thus far for Player I a *supermartingale.* Player II always chooses t_k to be the element in $\left\{\widetilde{f}(y_{k-1}), \widetilde{f}(z_{k-1})\right\}$ on which $d_T(\cdot, o_k)$ is largest; if she wins the coin toss, she then chooses x_k to be so that $\widetilde{f}(x_k)$ is that element. To establish the supermartingale property, we must show that, regardless of player I's strategy, we have

$$\mathbb{E}\Big[d_T(o_k, t_k) - d_T(o_k, t_{k-1})\Big] \geqslant \mathbb{E}\left[d_T\left(\widetilde{f}(x_k), t_k\right) - d_T\left(\widetilde{f}(x_{k-1}), t_{k-1}\right)\right]. \tag{42}$$

It is not hard to see that we have deterministically

$$d_T(o_k, t_k) - d_T(o_k, t_{k-1}) \geqslant d_T\left(\widetilde{f}(x_{k-1}), t_k\right) - d_T\left(\widetilde{f}(x_{k-1}), t_{k-1}\right). \tag{43}$$

Indeed, if o_k and t_{k-1} are in distinct components of $T \smallsetminus \left\{\widetilde{f}(x_{k-1})\right\}$, then the same will be true of o_k and t_k, and (43) holds as equality; if o_k and t_{k-1} are in the same component of $T \smallsetminus \left\{\widetilde{f}(x_{k-1})\right\}$ then o_k and t_k will be in opposite components of $T \smallsetminus \left\{\widetilde{f}(x_{k-1})\right\}$, and the left hand side minus the right hand side of (43) becomes twice the distance from $\widetilde{f}(x_{k-1})$ of the least common ancestor of o_k and and t_{k-1} in the tree rooted at $\widetilde{f}(x_{k-1})$.

Due to (43), in order to prove (42) it is enough to show that

$$\mathbb{E}\left[d_T\left(\widetilde{f}(x_k), t_k\right) - d_T\left(\widetilde{f}(x_{k-1}), t_k\right)\right] \leqslant 0,$$

which is clear since if player II wins the coin toss this quantity will be $-M_k$ and if player I wins the coin toss it will be at most M_k.

Next we give a very similar strategy for player I that makes $d_T\left(\widetilde{f}(x_k), t_k\right)$ plus the total payoff thus far for Player I a *submartingale*. In this strategy, Player I always chooses o_k to be the element in $\left\{\widetilde{f}(y_{k-1}), \widetilde{f}(z_{k-1})\right\}$ on which $d_T(\cdot, t_{k-1})$ is largest; if he wins the coin toss, he then chooses x_k to be so that $\widetilde{f}(x_k)$ is that element. To establish the submartingale property, we must now show that

$$\mathbb{E}\Big[d_T(o_k, t_k) - d_T(o_k, t_{k-1})\Big] \leqslant \mathbb{E}\left[d_T\left(\widetilde{f}(x_k), t_k\right) - d_T\left(\widetilde{f}(x_{k-1}), t_{k-1}\right)\right]. \tag{44}$$

Note that by strategy definition o_k is on the opposite side of $\widetilde{f}(x_{k-1})$ from t_{k-1}, so we may write $d_T(o_k, t_{k-1}) = M_k + d_T\left(\widetilde{f}(x_{k-1}), t_{k-1}\right)$. Plugging this into (44), what we seek to show becomes

$$\mathbb{E}\Big[d_T(o_k, t_k) - M_k\Big] \leqslant \mathbb{E}\left[d_T\left(\widetilde{f}(x_k), t_k\right)\right], \tag{45}$$

which we see by noting that the right hand side of (45) is equal to $d_T(o_k, t_k)$ when player I wins the coin toss (and makes $\widetilde{f}(x_k) = o_k$) and at least $d_T(o_k, t_k) - 2M_k$ when player II wins the coin toss, since $d_T\left(\widetilde{f}(x_k), o_k\right) \leqslant 2M_k$ for any valid choice of x_k.

To conclude the proof, we need to modify the strategy in such a way that forces the game to terminate without sacrificing the payoff expectation. If both players adopt the above strategy, it is clear that the increments $d_T\left(\widetilde{f}(x_{k-1}), \widetilde{f}(x_k)\right)$ are non-decreasing, and that the distance from any fixed endpoint of the tree has at least probability $1/2$ of increasing at each step; from this, it follows that the length of game play is a random variable with exponential decay. If the other player makes other moves, which are not optimal from the point of view of optimizing the payoff, then we can wait until the cumulative amount the other player has "given up" is greater than twice the diameter of T, and then force the game to end by placing a target at a single point and subsequently always moving x_k closer to that point when winning a coin toss. (The loss from the sub-optimality of this strategy is less than the gain from the amount the other player gave up.) If a player adopts this strategy, then the total time duration of the game is a random variable whose law decays exponentially; this yields the uniform integrability necessary for the sub-martingale optional stopping theorem, which implies Proposition 4. □

Acknowledgements. A. N. is supported by NSF grants CCF-0635078 and CCF-0832795, BSF grant 2006009, and the Packard Foundation. S. S. is supported by NSF grants DMS-0645585 and OISE-0730136. We are grateful to the anonymous referee for many helpful comments and suggestions.

References

[1] A. Almansa, F. Cao, Y. Gousseau, and B. Rouge. Interpolation of digital elevation models using AMLE and related methods. *IEEE Transactions on Geoscience and Remote Sensing*, 40(2):314–325, 2002.

[2] S. N. Armstrong and C. K. Smart. An easy proof of Jensen's theorem on the uniqueness of infinity harmonic functions. *Calc. Var. Partial Differential Equations*, 37(3-4):381–384, 2010.

[3] G. Aronsson. Extension of functions satisfying Lipschitz conditions. *Ark. Mat.*, 6:551–561 (1967), 1967.

[4] G. Aronsson, M. G. Crandall, and P. Juutinen. A tour of the theory of absolutely minimizing functions. *Bull. Amer. Math. Soc. (N.S.)*, 41(4):439–505 (electronic), 2004.

[5] K. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2(2):137–172, 1992.

[6] G. Barles and J. Busca. Existence and comparison results for fully nonlinear degenerate elliptic equations without zeroth-order term. *Comm. Partial Differential Equations*, 26(11-12):2323–2337, 2001.

[7] Y. Benyamini and J. Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[8] A. Brudnyi and Y. Brudnyi. Linear and nonlinear extensions of Lipschitz functions from subsets of metric spaces. *Algebra i Analiz*, 19(3):106–118, 2007.

[9] Y. Brudnyi and P. Shvartsman. Stability of the Lipschitz extension property under metric transforms. *Geom. Funct. Anal.*, 12(1):73–79, 2002.

[10] V. Caselles, G. Haro, G. Sapiro, and J. Verdera. On geometric variational models for inpainting surface holes. *Computer Vision and Image Understanding*, 111(3):351–373, 2008.

[11] V. Caselles, J.-M. Morel, and C. Sbert. An axiomatic approach to image interpolation. *IEEE Trans. Image Process.*, 7(3):376–386, 1998.

[12] T. Champion and L. De Pascale. Principles of comparison with distance functions for absolute minimizers. *J. Convex Anal.*, 14(3):515–541, 2007.

[13] M. G. Crandall, L. C. Evans, and R. F. Gariepy. Optimal Lipschitz extensions and the infinity Laplacian. *Calc. Var. Partial Differential Equations*, 13(2):123–139, 2001.

[14] A. Dress. Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces. *Adv. in Math.*, 53(3):321–402, 1984.

[15] A. Dress, V. Moulton, and W. Terhalle. T-theory: an overview. *European J. Combin.*, 17(2-3):161–175, 1996. Discrete metric spaces (Bielefeld, 1994).

[16] J. R. Isbell. Six theorems about injective metric spaces. *Comment. Math. Helv.*, 39:65–76, 1964.

[17] R. Jensen. Uniqueness of Lipschitz extensions: minimizing the sup norm of the gradient. *Arch. Rational Mech. Anal.*, 123(1):51–74, 1993.

[18] W. B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

[19] W. B. Johnson, J. Lindenstrauss, D. Preiss, and G. Schechtman. Lipschitz quotients from metric trees and from Banach spaces containing l_1. *J. Funct. Anal.*, 194(2):332–346, 2002.

[20] W. B. Johnson, J. Lindenstrauss, and G. Schechtman. Extensions of Lipschitz maps into Banach spaces. *Israel J. Math.*, 54(2):129–138, 1986.

[21] P. Juutinen. Absolutely minimizing Lipschitz extensions on a metric space. *Ann. Acad. Sci. Fenn. Math.*, 27(1):57–67, 2002.

[22] N. J. Kalton. Extending Lipschitz maps into $\mathscr{C}(K)$-spaces. *Israel J. Math.*, 162:275–315, 2007.

[23] N. V. Khuê and N. T. Nhu. Lipschitz extensions and Lipschitz retractions in metric spaces. *Colloq. Math.*, 45(2):245–250, 1981.

[24] M. D. Kirszbraun. Über die zusammenziehenden und Lipschitzchen Transformationen. *Fundam. Math.*, 22:77–108, 1934.

[25] U. Lang, B. Pavlović, and V. Schroeder. Extensions of Lipschitz maps into Hadamard spaces. *Geom. Funct. Anal.*, 10(6):1527–1553, 2000.

[26] U. Lang and T. Schlichenmaier. Nagata dimension, quasisymmetric embeddings, and Lipschitz extensions. *Int. Math. Res. Not.*, (58):3625–3655, 2005.

[27] U. Lang and V. Schroeder. Kirszbraun's theorem and metric spaces of bounded curvature. *Geom. Funct. Anal.*, 7(3):535–560, 1997.

[28] A. J. Lazarus, D. E. Loeb, J. G. Propp, W. R. Stromquist, and D. H. Ullman. Combinatorial games under auction play. *Games Econom. Behav.*, 27(2):229–264, 1999.

[29] E. Le Gruyer. On absolutely minimizing Lipschitz extensions and PDE $\Delta_\infty(u) = 0$. *NoDEA Nonlinear Differential Equations Appl.*, 14(1-2):29–55, 2007.

[30] J. R. Lee and A. Naor. Extending Lipschitz functions via random metric partitions. *Invent. Math.*, 160(1):59–95, 2005.

[31] J. Lindenstrauss. On nonlinear projections in Banach spaces. *Michigan Math. J.*, 11:263–287, 1964.

[32] M. B. Marcus and G. Pisier. Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. *Acta Math.*, 152(3-4):245–301, 1984.

[33] D. A. Martin. The determinacy of Blackwell games. *J. Symbolic Logic*, 63(4):1565–1581, 1998.

[34] E. J. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.*, 40(12):837–842, 1934.

[35] F. Mémoli, G. Sapiro, and P. Thompson. Geometric surface and brain warping via geodesic minimizing Lipschitz extensions. In *MFCA-2006 International Workshop on Mathematical Foundations of Computational Anatomy (MICCAI)*, pages 58–67, 2006.

[36] M. Mendel and A. Naor. Some applications of Ball's extension theorem. *Proc. Amer. Math. Soc.*, 134(9):2577–2584 (electronic), 2006.

[37] V. A. Mil′man. Absolutely minimal extensions of functions on metric spaces. *Mat. Sb.*, 190(6):83–110, 1999.

[38] J. R. Munkres. *Topology: a first course*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.

[39] A. Naor. A phase transition phenomenon between the isometric and isomorphic extension problems for Hölder functions between L_p spaces. *Mathematika*, 48(1-2):253–271, 2001.

[40] A. Naor, Y. Peres, O. Schramm, and S. Sheffield. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134(1):165–197, 2006.

[41] A. Neyman and S. Sorin, editors. *Stochastic games and applications*, volume 570 of *NATO Science Series C: Mathematical and Physical Sciences*, Dordrecht, 2003. Kluwer Academic Publishers.

[42] Y. Peres, O. Schramm, S. Sheffield, and D. B. Wilson. Tug-of-war and the infinity Laplacian. *J. Amer. Math. Soc.*, 22(1):167–210, 2009.

[43] K. Przesławski and D. Yost. Lipschitz retracts, selectors, and extensions. *Michigan Math. J.*, 42(3):555–571, 1995.

[44] S. Sheffield and C. Smart. Vector-valued optimal Lipschitz extensions. Preprint, 2010.

[45] F. A. Valentine. Contractions in non-Euclidean spaces. *Bull. Amer. Math. Soc.*, 50:710–713, 1944.

[46] F. A. Valentine. A Lipschitz condition preserving extension for a vector function. *Amer. J. Math.*, 67:83–93, 1945.

[47] J. H. Wells and L. R. Williams. *Embeddings and extensions in analysis*. Springer-Verlag, New York, 1975. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84.

Courant Institute, New York University
E-mail address: `naor@cims.nyu.edu`

Massachusetts Institute of Technology
E-mail address: `sheffield@math.mit.edu`