Received: 7 June 2022 | Revised: 6 February 2023 | Accepted: 28 February 2023

DOI: 10.1112/mtk.12198

RESEARCH ARTICLE

Mathematika

Sums of distances on graphs and embeddings into Euclidean space

Stefan Steinerberger

Department of Mathematics, University of Washington, Seattle, Washington, USA

Correspondence
Stefan Steinerberger, Department of Mathematics, University of Washington, Seattle, WA 98195, USA.
stefan.steinerberger@gmail.com

Abstract

Let $G = (V, E)$ be a finite, connected graph. We consider a greedy selection of vertices: given a list of vertices $x_1, \dots, x_k$, take x_{k+1} to be any vertex maximizing the sum of distances to the vertices already chosen and iterate, keep adding the "most remote" vertex. The frequency with which the vertices of the graph appear in this sequence converges to a set of probability measures with nice properties. The support of these measures is, generically, given by a rather small number of vertices $m \ll |V|$. We prove that this suggests that the graph G is, in a suitable sense, "m-dimensional" by exhibiting an explicit 1-Lipschitz embedding $\phi : V \to \ell^1(\mathbb{R}^m)$ with good properties.

MSC 2020
05C69, 31E05, 51K99 (primary)

1 | INTRODUCTION AND RESULTS

1.1 | Motivation

Our original motivation was trying to understand the curious behavior of a simple procedure: given a finite, connected graph and given a list of vertices $x_1, \dots, x_k$ (note that this is a sequence: the same vertex may appear multiple times), one could try to extend the list by adding the vertex that is the furthest away in the sense of maximizing the sum over the distances to the vertices already in the list (counted with multiplicity). This vertex may not be unique and if there is more

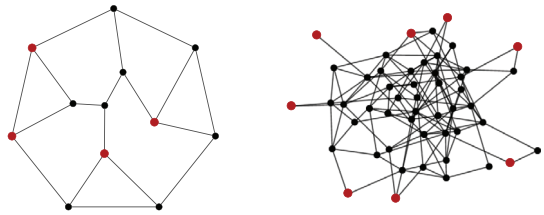

FIGURE 1 The Frucht graph (left) and an Erdős–Renyi random graph on $n = 50$ vertices (right). The rule ends up only selecting the red vertices (not necessarily with equal frequency).

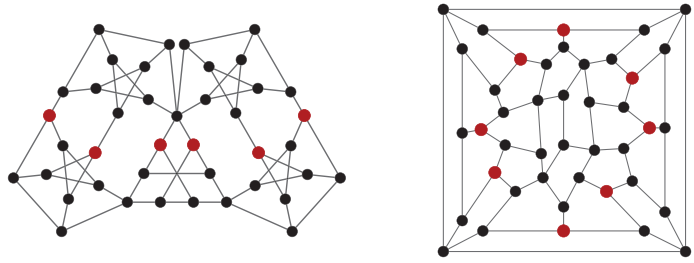

FIGURE 2 Thomassen 41 and Wiener–Araya. The rule ends up only selecting the red vertices (not necessarily equally often).

than one, any can be chosen. This means that we ask that

$$\sum_{i=1}^{k} d(x_{k+1}, x_i) = \max_{v \in V} \sum_{i=1}^{k} d(v, x_i).$$

The long-term behavior of this greedy procedure turns out to be connected to probability measures on V with nice properties.

Some experimentation suggests that the frequency with which vertices arise seems to quickly stabilize to a type of limiting distribution. We also observe that, typically, any such limiting measure μ seems to be supported on a fairly small subset of vertices. This is not always the case and it is possible that the limiting measure is supported on all vertices (but this seems rare and it seems that the graph needs to have a lot of special properties for this to happen, see §2.1). Indeed, for most graphs (both highly structured and random), the limiting measure seems to be supported on a small subset. One could think of this as a procedure favoring subsets of vertices that are at great distance from each other. The purpose of this paper is to demonstrate that the procedure has a number of interesting properties and applications (including that of inducing an interesting graph embedding) (Figure 2).

1.2 | Long-time behavior

Given an initial list $x_1, \dots, x_k \in V$, we use the rule

$$\sum_{i=1}^{k} d(x_{k+1}, x_i) = \max_{v \in V} \sum_{i=1}^{k} d(v, x_i).$$

to generate an infinite sequence of vertices. Note that an initial list does not necessarily specify a unique sequence but rather a family of sequences (since the maximum might be assumed in multiple vertices). We will not distinguish between different elements of the same family and always refer to them as "a sequence." The rest of the paper is concerned with results that are independent of the initial list $x_1, \dots, x_k \in V$. The empirical density of vertices of such a sequence will be shown to approach the set of balanced probability distributions which we now define.

Definition. A probability measure μ on the vertices V is *balanced* if

$$\forall w \in V \qquad \mu(w) > 0 \Rightarrow \sum_{u \in V} d(w,u)\mu(u) = \max_{v \in V} \sum_{u \in V} d(v,u)\mu(u).$$

One interpretation is as follows: for any probability measure μ on the vertices, we may introduce the (transport cost) function $T : V \rightarrow \mathbb{R}$

$$T(w) = \sum_{u \in V} d(w,u)\mu(u).$$

$T(w)$ is the minimal transport cost of sending all the mass of μ to the vertex w under the assumption that transporting $\varepsilon > 0$ units of mass across one edge has transport cost ε. This is also known as the Wasserstein W^1 cost or Earth Mover Distance. A probability measure μ on the vertices V is said to be *balanced* if it has the property that whenever $\mu(w) > 0$, then T assumes a global maximum in w. Poetically put, a measure is balanced if the points where probability mass can actually be found are simultaneously the vertices for which global transport of the entire measure to a single vertex is the most expensive. We can now describe the asymptotic behavior of the greedy vertex selection: it is assumed that we start with an arbitrary initial list of vertices and then turn the list into an infinite sequence by picking vertices maximizing the sum of distances.

Theorem 1. *Let $(x_k)_{k=1}^{\infty}$ denote an infinite sequence of vertices obtained by the greedy procedure and let μ_m denote the empirical probability measure of the first m vertices. There exists a constant* $\mathrm{diam}(G)/2 \leqslant \alpha \leqslant \mathrm{diam}(G)$

(1) *such that*

$$\lim_{m \rightarrow \infty} \sum_{v,w \in V} \mu_m(v) d(v,w) \mu_m(w) \rightarrow \alpha,$$

(2) *the maximal transport cost converges to α*

$$\lim_{m \rightarrow \infty} \max_{v \in V} \sum_{w \in V} d(v,w)\mu_m(w) = \alpha,$$

(3) *and for any $\varepsilon > 0$ and all m sufficiently large (depending on ε)*

$$\forall v \in V \qquad \mu(v) > \varepsilon \quad \Rightarrow \quad \sum_{w \in V} d(v,w)\mu_m(w) \geqslant \alpha - \varepsilon.$$

This implies that any convergent subsequence of μ_m (which exist due to compactness) has to converge to a balanced measure. One consequence is that at least one balanced probability measure always exists (though this is not difficult to show by other means, see Proposition 1). Another

FIGURE 3 Balanced measure on a path graph.

consequence is that if one wishes to create a measure that is close to a balanced measure, one can simply use the greedy selection procedure and is guaranteed to end up close to such a measure.

1.3 | Embedding $G = (V, E)$ into $\mathbb{R}^m$

If a balanced measure μ is supported on a small number $m \ll n$ of vertices, then this implies that the graph behaves approximately like an m-dimensional object in a way that we will make now precise. More precisely, it will now be argued that if the balanced measure μ is supported in $w_1, \dots, w_m \in V$, then keeping track of the distances to these m vertices and weighting them according to μ leads to a good embedding $\phi : V \to \mathbb{R}^m$ via

$$\phi(v) = (\mu(w_i)d(w_i, v))_{i=1}^m \in \mathbb{R}^m.$$

We note that this embedding is completely explicit, there are no hidden constants or computations. We start with a couple of examples.

The first natural example is that of *path graph* on n vertices (see Figure 3). Unsurprisingly, the balanced measure puts half of its mass on the two endpoints 1 and n. The embedding $\phi : \{1, 2, \dots, n\} \to \mathbb{R}^2$ is then simply

$$\phi(i) = \left(\frac{i-1}{2}, \frac{n-i}{2}\right),$$

which maps the graph to a line in $\mathbb{R}^2$. This is certainly a reasonable representation of the path graph. Another simple example is the *complete graph* K_n: here, the balanced measure is the uniform measure and the arising embedding $\phi : V \to \mathbb{R}^n$ sends the vertices to a simplex since

$$\phi(i) = \frac{1}{n}(1, 1, \dots, 1, \underbrace{0}_{i\text{th coordinate}}, 1, \dots, 1),$$

which is a reasonable representation of the complete graph. In particular, we conclude that K_n is effectively n-dimensional (or $(n-1)$-dimensional by projecting on the common hyperplane), which is perhaps well aligned with intuition.

There can be multiple balancing measures that can lead to different embeddings for the same graph. The following example was constructed by Noah Kravitz and is included with his kind permission. We take m path graphs of length $2\ell + 1$ and glue them together at the two endpoints (see Figure 4). This graph supports two very different balanced measures. One is concentrated at the two end points: the arising embedding is a line in $\mathbb{R}^2$ with diameter $\sim \mathrm{diam}(G)$ that collapses the m paths all into a single path, it identifies the graph as being predominantly one-dimensional. The other is supported on the m central points of the m paths and leads to a very different embedding emphasizing a somewhat different aspect of the graph.

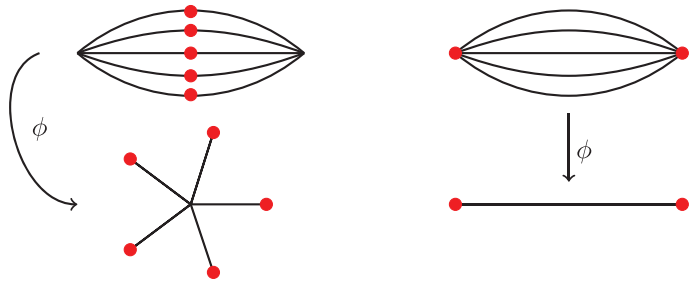

FIGURE 4 m path graphs of length $2\ell + 1$ glued together at the endpoints. Top: two different balanced measures leading to (bottom) two different embeddings emphasizing different aspects.

Theorem 2 (Graph embedding). *Let $G = (V, E)$ be connected and let μ be a balanced measure on m vertices $w_1, \dots, w_m$. Then $\phi : V \to \mathbb{R}^m$ given by*

$$\phi(v) = (\mu(w_1)d(w_1, v), \ \dots, \ \mu(w_m)d(w_m, v))$$

(1) *is 1-Lipschitz from the Graph Distance to $\ell^1(\mathbb{R}^m)$ and for all $u, v \in V$*

$$\|\phi(u) - \phi(v)\|_{\ell^1(\mathbb{R}^m)} \leqslant d(u, v).$$

(2) *ϕ sends $\operatorname{supp}\mu$ to a hyperplane: for some $\operatorname{diam}(G)/2 \leqslant \alpha \leqslant \operatorname{diam}(G)$*

$$\phi(\operatorname{supp}\mu) \subset \left\{ x \in \mathbb{R}^m_{\geqslant 0} \ : \ x_1 + x_2 + \cdots + x_m = \alpha \right\}.$$

(3) *Points in $\phi(\operatorname{supp}\mu)$ are, on average, not too close: for all $v \in \operatorname{supp}\mu$*

$$\frac{1}{m} \sum_{w \in \operatorname{supp}\mu} \|\phi(v) - \phi(w)\|_{\ell^\infty(\mathbb{R}^m)} \geqslant \frac{\operatorname{diam}(G)}{2m}.$$

We emphasize the interplay between property (1) and (3). It is very easy to produce a Lipschitz map $\phi : V \to \mathbb{R}^m$ by sending all vertices to the same point or to points that are very, very close (and this can be done independently of the overall structure of the graph and is not particularly instructive). Property (3) says that this does not happen here: the vertices in $\operatorname{supp}\mu$ end up being mapped to points that are at least on average not too close to each other. However, any mapping that is both 1-Lipschitz and whose image $\phi(V)$ has substantial diameter necessarily needs to capture at least some major global structures of the graph.

Remarks. Several more remarks are in order.

(1) *Small entries.* The embedding can be thought of as a triangulation by $\operatorname{supp}\mu$ weighted by μ. We observe from the explicit closed form of ϕ that if $\mu(w_j)$ is small, then the embedding does not have a lot of variation in its j-th coordinate. One could thus obtain another embedding of comparable quality in a lower dimension by omitting this coordinate (see below for an example). This suggests that in practice the "effective dimensionality" is perhaps closer related to the number of vertices in $\operatorname{supp}\mu$ whose weight is not too small when compared with $1/(\#\operatorname{supp}\mu)$.

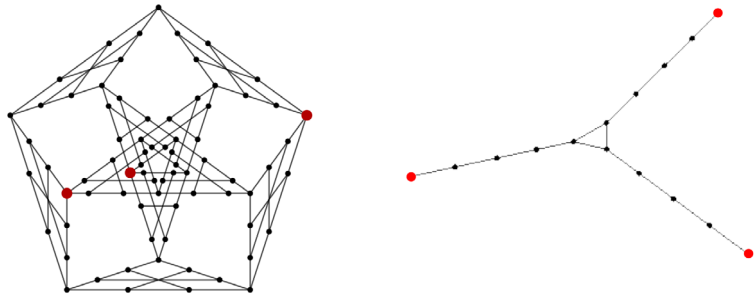

FIGURE 5 Zamfirescu 75 (left) and the embedding $\phi : V \to \mathbb{R}^3$ (right, with lines connecting $\phi(v)$ and $\phi(w)$ for all $(v, w) \in E$).

(2) *Variable dimensionality.* There is no sense in which this m, the notion of "effective dimension," is unique: it is certainly conceivable that the graph embedding $\phi(V)$ is contained in a lower dimensional subset of $\mathbb{R}^m$ and it is easy to construct examples for which this happens: by property (2) of the embedding, this will always happen if μ is supported on all of V, for example. The size of $\mathrm{supp}(\mu)$ provides an upper bound on the dimension. We also note that graphs can support balanced measures whose supports have a different cardinality (see also the example above).

(3) *Sharpness.* Theorem 2 is sharp up to constants: considering the example of glued paths (see Figure 4) and the balanced measure supported on the m midpoints of the m paths. Then, for all $v \in \mathrm{supp}\,\mu$,

$$\frac{1}{m} \sum_{w \in \mathrm{supp}\,\mu} \|\phi(v) - \phi(w)\|_{\ell^1(\mathbb{R}^m)} \sim \frac{\mathrm{diam}(G)}{m}.$$

In particular, the graph is being folded into a rather small region of $\ell^1(\mathbb{R}^m)$ whose diameter shrinks as m increases. Nonetheless, the embedding itself is certainly a good representation of the structure of the graph.

Graph embedding: Examples. We conclude with several more intricate graphs and a discussion of the behavior of the graph embedding for these cases.

1. Combinatorial example. The Zamfirescu 75 graph [33] looks rather complicated (see Figure 5). However, creating the sequence of vertices, we quickly find that the game ends up jumping between only three vertices. Computing the arising embedding into $\mathbb{R}^3$ demonstrates that there is a rather simple structure underlying the graph. Note that the embedding is not injective and does, in fact, collapse different vertices onto the same point in $\mathbb{R}^3$. This is an interesting example where the graph is actually combinatorially somewhat simpler than it may at first glance appear; this is reflected in particularly simple long-term behavior of the game and a correspondingly simple embedding.

2. Gaussian point cloud. Another type of example is shown in Figure 6: these points are obtained from taking nearest-neighbor connections between Gaussian point clouds. What we observe is that for these types of examples, it seems that the embedding is fairly close to an isometry for a majority of vertices: up to a universal constant $c_G > 0$, it seems that for the vast majority of pairs of vertices $(u, v) \in V$ (say 95% of $V \times V$), we have

$$\frac{1}{2} \leqslant c_G \frac{d(u, v)}{\|\phi(u) - \phi(v)\|_{\ell^1}} \leqslant \frac{3}{2}.$$

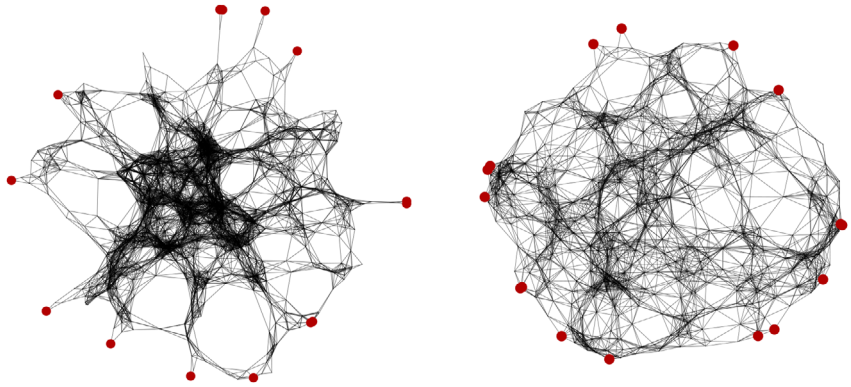

FIGURE 6 Two Gaussian point clouds with supp μ in red.

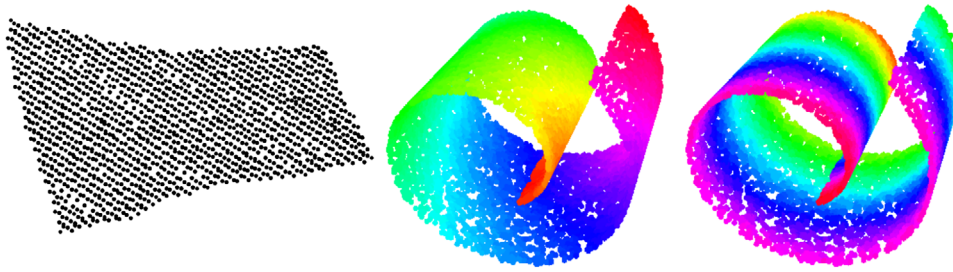

FIGURE 7 Left: embedding of a Swiss Roll into 2D using the $m = 3$ followed by PCA. Original Swiss roll colored by x-coordinate (middle) and y-coordinate (right) of the embedding.

It would be interesting if such a statement could be made precise for, say, certain types of random graphs. The purpose of this example is to illustrate that the theorem does appear to actually lead to bi-Lipschitz embeddings with good constants even in somewhat rough settings.

3. Swiss roll. In practice, the following two variations on the graph embedding might be reasonable (the first has already been discussed in the remarks above).

(1) What is important for the properties of ϕ is not only the size of the support of μ but also the weights $\mu(w_i)$. We see from the explicit form of the embedding ϕ that if $\mu(w_i)$ is rather small, then the corresponding entries in the embedding will vary very little. One could thus consider omitting and only focus on coordinates for which $\mu(w_i)$ is large.
(2) We know that ϕ sends supp μ to a hyperplane and it might, in practice, make sense to move the embedding from $\mathbb{R}^m$ to $\mathbb{R}^{m-1}$ either by PCA (principal component analysis) or by projecting onto the plane $x_1 + \cdots + x_m = 0$.

We will now apply both these ideas to an example of a manifold embedding: we generate a swiss roll in $\mathbb{R}^3$ using 10,000 points and building a graph connecting each points to its k-nearest neighbors (where $k = 40$, this is not tremendously important). Out of these 10,000 points, the measure ends up being supported on 50 points. Taking the vertices where μ is the largest, we realize that $1/3$ of the total probability mass is contained in only three vertices. We can take these three vertices to create an embedding into $\mathbb{R}^3$ which then, via PCA, we map to $\mathbb{R}^2$. Figure 7 shows

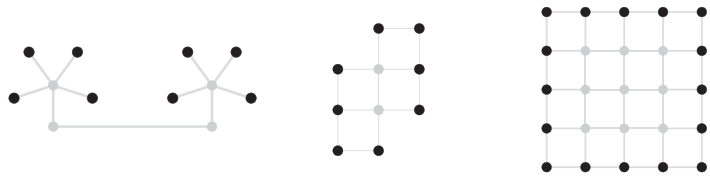

FIGURE 8 Three graphs with their boundary vertices highlighted.

that the approximate shape is recovered by the embedding and that the embedding coordinates are smooth variables in the original space.

2 | COMMENTS, EXAMPLES, AND RELATED RESULTS

2.1 | Balanced measures as critical points

We recall that for any given graph $G = (V, E)$, we call a probability measure μ on the set of vertices balanced if

$$\mu(u) > 0 \Rightarrow \sum_{v \in V} d(u, v)\mu(v) = \max_{w \in W} \sum_{v \in V} d(w, v)\mu(v).$$

There is a simple variational characterization of balanced measures. The distance matrix $D \in \mathbb{R}^{n \times n}$ defined via $D_{ij} = d(v_i, v_j)$ has integer entries and is symmetric. It has many interesting properties [4, 29] and balanced measures might represent an interesting new aspect. We identify the probability measures on V with

$$\Delta = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n_{\geqslant 0} \, : \, x_1 + \cdots + x_n = 1 \right\}.$$

Proposition 1. *Each critical point of the functional $J : \Delta \to \mathbb{R}_{\geqslant 0}$ given by*

$$J(\mu) = \langle \mu, D\mu \rangle$$

is a balanced measure. A balanced measure μ has the property that the directional derivative in all admissible directions is nonpositive.

If μ lies on the boundary of Δ, we say that it is a critical point if all directional derivatives in admissible directions vanish: for any signed measure ν such that $\mu + \varepsilon\nu$ is a probability measure for all $0 \leqslant \varepsilon \leqslant \varepsilon_0$ for some $\varepsilon_0 > 0$ (depending on μ), we require

$$J(\mu + \varepsilon\nu) = J(\mu) + o(\varepsilon) \qquad \text{as } \varepsilon \to 0^+.$$

Since Δ is a compact set and J is continuous, there is at least one maximum, and thus, there is always at least one balanced measure. We note that the global minima are given by the Dirac measures in a single vertex: these are *not* critical points (possible because they are assumed at the boundary of Δ) (Figure 8).

2.2 | Balancing measures and boundary

One could wonder whether there is any way of deciding *a priori* where the measure μ can be supported. We recall a definition of *boundary* of a graph given in [27]: for any connected graph $G = (V, E)$, we define the boundary $\partial G \subseteq V$ to be the set of all vertices $u \in V$ for which there exists another vertex v such that the *average* neighbor of u is closer to v than $d(u, v)$. Formally, we define the boundary as

$$\partial G = \left\{ u \in V \Big| \exists v \in V \ : \ \frac{1}{\deg(u)} \sum_{(u,w)\in E} d(w, v) < d(u, v) \right\}.$$

This notion of boundary is motivated by the fact that it satisfies a type of isoperimetric inequality, stating that large graphs have to have a large boundary: if the maximal degree of G is given by Δ, then

$$\#\partial G \geqslant \frac{1}{2\Delta} \frac{\#V}{\operatorname{diam}(G)}.$$

We refer to [27] for more details. As it turns out, there are balanced measures that are supported purely in the boundary.

Proposition 2. *Let $x_1, \dots, x_k \in V$ be a list of vertices. Then $f(v) = \sum_{i=1}^{k} d(v, x_i)$ assumes its maximum in ∂G. If f assumes its maximum in $V \setminus \partial G$, then f is constant. In particular, there exists a balanced measure supported in ∂G.*

One might assume that this means that "typically" balanced measures can only be supported in the boundary — it would be interesting to understand this better.

2.3 | Remarks

In many of the examples shown throughout the paper, we see that $\#\operatorname{supp}(\mu) = m \ll n$. This need not always be the case: examples are given by the dodecahedral graph or the Desargue graphs for which the uniform measure is balanced. Both examples have distance matrices $(d(v_i, v_j))_{i,j=1}^{n}$ with one positive eigenvalue whose corresponding eigenvector is constant which, in light of Proposition 1, is perhaps not a coincidence (see also [29]). There are many other examples, indeed, graphs for which the uniform measure is balanced are precisely the graphs with constant curvature in the sense of [28]. In practice, it seems very difficult for graphs to have $\operatorname{supp}\mu$ contain a large number of vertices and it would be interesting to have a more quantitative understanding of this. Are there upper bounds on $\#\operatorname{supp}\mu$ depending on some graph parameters?

Graphs with symmetries will naturally present with balanced measures inheriting these symmetries. However, the support of a balanced measure is *not* a graph invariant: graphs can have balanced measures with very different cardinality (see Figure 9). It would be interesting to understand how many different balanced measures a graph can support. A graph having a balanced measure supported in two vertices indicates that the graph is elongated and behaves, mostly, like a one-dimensional interval — there are many graphs like this. Conversely, a graph supporting

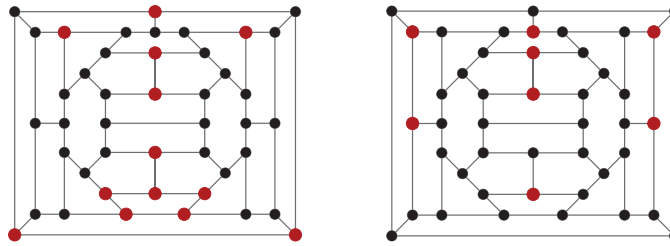

FIGURE 9 Support of two balanced measures on Grinberg 44.

a balanced measure in $m \geqslant n/100$ vertices seems to require a tremendous amount of symmetries and these are probably rare. It might also be interesting to study the properties of balanced measures on Erdős–Renyi random graphs.

We conclude with an interesting phenomenon that appears to be rare. Recall that the definition of a balanced measures is given by

$$\mu(u) > 0 \Rightarrow \sum_{v \in V} d(u,v)\mu(v) = \max_{w \in W} \sum_{v \in V} d(w,v)\mu(v).$$

One could wonder whether any type of inverse result of this flavor is true: are the maxima *only* assumed in vertices where the measure is supported? The Frucht graph (see Figure 1) is a counterexample: there is a balanced measure on four vertices (with weights $0.1, 0.2, 0.3, 0.4$) such that $\sum d(v,w)\mu(w)$ assumes its maximum in six vertices. However, this seems to be exceedingly rare.

2.4 | Related results

Potential theory. Let (X,d) be a compact metric space and let μ be a probability measure on the space. The study of the energy integral

$$J(\mu) = \int_{X \times X} d(x,y) d\mu(x) d\mu(y)$$

is classical and dates back to a 1956 paper of Björck [7]. Björck proved that if X is a compact subset of $\mathbb{R}^n$, then the maximizing measure is unique and supported on the boundary ∂X. This has inspired a lot of work in Euclidean space: we refer to Alexander [1, 2], Alexander and Stolarsky [3], Carando, Galicer and Pinasco [8], Hinrichs, Nickolas, and Wolf [12], and Wolf [31]. A continuous version of the greedy procedure was proposed by Larcher, Schmid, and Wolf [15] to approximate the (unique) maximizing measure. Outside the Euclidean setting, there has been substantial work in quasi-hypermetric spaces, we refer to Nickolas and Wolf [18–21, 32]. The Euclidean space is an example of a quasi-hypermetric space; in our setting, a graph is quasi-hypermetric if the distance matrix D is negative semidefinite when restricted to vectors whose sum of entries is 0. Partial motivation for many of these questions is a fascinating 1964 result of Gross [11]: in any compact, connected metric space (X,d), there exists a unique, real number $r(X,d) \in \mathbb{R}$ such that for any

set of points $\{x_1, \dots, x_n\} \subset X$, there exists a point $x \in X$ with

$$\frac{1}{n}\sum_{i=1}^{n} d(x, x_i) = r(X, d).$$

A wonderful introduction to the subject is given by Cleary, Morris, and Yost [9]. No such results can be true on graphs, however, see Thomassen [30] for a substitute result (see also [28] for a connection to the von Neumann Minimax Theorem). The result of Gross can be extended to probability measures, which has the following implication: if there exists a probability measure μ on X so that

$$\int_X d(x, y)d\mu(y) \qquad \text{is independent of } x,$$

then $r(X, d)$ has to equal the value of the integral evaluated at an arbitrary point x. This allows for an explicit computation of $r(X, d)$, which, in general, is very difficult. Simultaneously, if we consider the problem of maximizing

$$\mu \to \int_{X\times X} d(x, y)d\mu(x)d\mu(y),$$

then an Euler–Lagrange ansatz shows that $\int_X d(x, y)\mu(y)$ is constant on $\mathrm{supp}(\mu)$. If $\mathrm{supp}(\mu) = X$, the maximizing measure can be used to determine $r(X, d)$. Our paper is somewhat opposite to the main direction in the literature since (besides the graph setting that is a metric space but *not* a connected metric space) we are interested in cases where $\mathrm{supp}(\mu)$ is much smaller than the space X.

Beacons. Another related area is in the theory of networks. A *beacon-based* approach is to select, often randomly, a number of vertices (the "beacons"). One is then given the corresponding rows in the Graph Distance Matrix, the distance between the beacons and all other vertices and tries to reconstruct all distances. An example of such a result is given by Ng and Zhang [22] who showed that a beacon-based approach could embed all but a small fraction of Internet distances; see also Kleinberg, Slivkins, and Wexler [13]. These results seem are somewhat different in flavor insofar as the beacons can be freely chosen, whereas in our setting, they arise from balanced measures that are determined by the graph.

Greedy sequences. Another motivation of the present paper is the recent introduction of greedy sequences for the purpose of generating uniformly distributed sequences with good regularity properties [5, 14, 23–26]. This is in contrast to the present paper where the limiting measure μ tends to be highly irregular; however, this is due to the particular type of kernel being used; kernels with the purpose of generating a more uniform distribution on graphs have been studied by Brown [6], Cloninger and Mhaskar [10], and Linderman and Steinerberger [16].

3 | PROOFS

3.1 | Proof of Theorem 1

The proof decouples into the following steps.

(1) We first show that the sequence of real numbers

$$E_m = \frac{1}{m(m-1)} \sum_{i,j=1}^{m} d(x_i, x_j) \qquad \text{converges as } m \to \infty$$

(2) and then that

$$\frac{\operatorname{diam}(G)}{2} \leqslant \lim_{m\to\infty} E_m \leqslant \operatorname{diam}(G).$$

(3) The next step consists in proving that a suitable weighted average of $T_m(v) = \sum_{w\in V} d(v,w)\mu_m(w)$ is also close to E_m.
(4) A stability estimate: if $T_m(v)$ is large, then so is $T_{m+\ell}(v)$ for small $\ell \in \mathbb{N}$.
(5) From this, we deduce an upper bound on $\|T_m\|_{\ell^\infty}$.
(6) Steps (3) and (5) imply that $T_m(v)$ is nearly constant on $\operatorname{supp} \mu_m$.

Proof. **Step 1.** We start by considering the behavior of the rescaled energy

$$E_m = \frac{1}{m(m-1)} \sum_{i,j=1}^{m} d(x_i, x_j).$$

The first step of the argument consists in showing that E_m converges. This sequence is trivially bounded since

$$E_m = \frac{1}{m(m-1)} \sum_{i,j=1}^{m} d(x_i, x_j) \leqslant \frac{1}{m(m-1)} \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \operatorname{diam}(G) = \operatorname{diam}(G).$$

We observe the algebraic identity

$$(m+1)m \cdot E_{m+1} = m(m-1) \cdot E_m + 2\sum_{i=1}^{m} d(x_{m+1}, x_i).$$

We also see that, because of the definition of the greedy sequence, the sum always attains it largest possible value (assuming $x_1, \dots, x_m$ to be given) and

$$\sum_{i=1}^{m} d(x_{m+1}, x_i) = \max_{v\in V} \sum_{i=1}^{m} d(v, x_i).$$

One can always bound a maximum from below by *any* weighted average of the values that appear. For any probability measure ν on V, we have

$$\max_{v\in V} \sum_{i=1}^{m} d(v, x_i) \geqslant \sum_{v\in V} \sum_{i=1}^{m} d(v, x_i)\nu(v).$$

Our choice will be the empirical measure of the first m points

$$\nu = \frac{1}{m}\sum_{j=1}^{m} \delta_{x_j}.$$

For this particular choice of probability measure, the lower bound simplifies further

$$\max_{v \in V} \sum_{i=1}^{m} d(v, x_i) \geqslant \sum_{v \in V} \sum_{i=1}^{m} d(v, x_i)\nu(v) = \sum_{i=1}^{m} \sum_{v \in V} d(v, x_i)\nu(v)$$
$$= \frac{1}{m} \sum_{i,j=1}^{m} d(x_i, x_j) = (m-1) \cdot E_m.$$

Plugging in shows that this implies

$$(m+1)mE_{m+1} \geqslant (m+2)(m-1)E_m,$$

and thus,

$$E_{m+1} \geqslant \frac{(m+2)(m-1)}{m(m+1)} \cdot E_m.$$

A simple computation shows that

$$\frac{(m+2)(m-1)}{m(m+1)} \geqslant 1 - \frac{2}{m^2}.$$

This implies, in particular, that for all $\ell \in \mathbb{N}$,

$$E_{m+\ell} \geqslant E_m \cdot \prod_{k=m}^{m+\ell-1} \left(1 - \frac{2}{k^2}\right) \geqslant E_m \cdot \prod_{k=m}^{\infty} \left(1 - \frac{2}{k^2}\right).$$

The next step is to argue that this product converges to a number that is arbitrarily close to 1 as $m \to \infty$. We make this quantitative. Using standard Taylor series estimates, we have for m sufficiently large

$$\prod_{k=m}^{\infty} \left(1 - \frac{2}{k^2}\right) = \exp\left[\log\left(\prod_{k=m}^{\infty} \left(1 - \frac{2}{k^2}\right)\right)\right]$$
$$= \exp\left[\sum_{k=m}^{\infty} \log\left(1 - \frac{2}{k^2}\right)\right]$$
$$\geqslant \exp\left(\sum_{k=m}^{\infty} -\frac{4}{k^2}\right) \geqslant \exp\left(-\frac{8}{m}\right) \geqslant 1 - \frac{16}{m}.$$

This implies that, for m sufficiently large,

$$\inf_{k \geqslant m} E_k \geqslant \left(1 - \frac{16}{m}\right) E_m.$$

Thus, given E_m, subsequent values cannot be much smaller. Let

$$\alpha = \limsup_{m \to \infty} E_m.$$

Then, by definition of lim sup, for any $\varepsilon > 0$, there exists a subsequence $(E_{m_\ell})_{\ell=1}^{\infty}$ such that all elements of the subsequence satisfy

$$E_{m_\ell} \geqslant \alpha - \varepsilon/2.$$

Picking ℓ sufficiently large so that $16/m_\ell \leqslant \varepsilon/(2\alpha)$, we deduce that

$$\liminf_{m \to \infty} E_m \geqslant \alpha - \varepsilon$$

and since ε was arbitrary, we deduce the convergence of $(E_m)_{m=1}^{\infty}$.

Step 2. We denote the limit by

$$\alpha = \lim_{m \to \infty} E_m$$

and will now establish $\operatorname{diam}(G)/2 \leqslant \alpha \leqslant \operatorname{diam}(G)$. The upper bound is obvious, it remains to prove the lower bound. We start by proving a lemma that will also be useful in the proof of Theorem 2. It appears implicitly in [27].

Lemma 3. *Let $G = (V, E)$ be a connected, finite graph and let ν be a probability measure on the vertices. Then there exists a vertex $w \in V$ such that*

$$\sum_{v \in V} d(w, v)\nu(v) \geqslant \frac{\operatorname{diam}(G)}{2}.$$

Proof. We use the von Neumann Minimax Theorem [17] in the special case of symmetric matrices (see also [28, 30]) and apply it to the case where the matrix is given by the distance matrix $D_{ij} = d(v_i, v_j) \in \mathbb{R}^{n \times n}$. The Minimax Theorem implies that there exists a constant β depending only on the matrix D such that for all probability vectors $\nu \in \Delta$ representing probability measures on V

$$\min_{1 \leqslant i \leqslant n} (D\mu)_i \leqslant \beta \leqslant \max_{1 \leqslant i \leqslant n} (D\mu)_i.$$

We can plug in a specific measure ν to obtain both upper and lower bounds on β. Pick two vertices $a, b \in V$ at maximal distance $d(a, b) = \operatorname{diam}(G)$ and set

$$\nu = \frac{1}{2}(\delta_a + \delta_b).$$

The triangle inequality shows that for any other vertex v

$$\frac{\operatorname{diam}(G)}{2} = \frac{d(a, b)}{2} \leqslant \frac{d(a, v) + d(v, b)}{2} = (D\mu)(v).$$

Thus,

$$\beta \geqslant \min_{1 \leqslant i \leqslant n} (D\mu)_i \geqslant \frac{\mathrm{diam}(G)}{2}.$$

□

In particular, in our application, it shows that a suitable rescaling of the energy always grows at least by $\mathrm{diam}(G)$ in each step since

$$\begin{aligned}(m+1)m \cdot E_{m+1} &= m(m-1) \cdot E_m + 2\sum_{i=1}^{m} d(x_{m+1}, x_i) \\ &= m(m-1) \cdot E_m + \max_{v \in V} 2\sum_{i=1}^{m} d(v, x_i) \\ &\geqslant m(m-1) \cdot E_m + m \cdot \mathrm{diam}(G).\end{aligned}$$

Abbreviating $F_m = m(m-1)E_m$, we have $F_{m+1} \geqslant F_m + m \cdot \mathrm{diam}(G)$ and thus

$$F_m \geqslant \mathrm{diam}(G) \sum_{k=1}^{m-1} k = (1 + o(1)) \cdot \mathrm{diam}(G) \cdot \frac{m^2}{2},$$

which implies the desired statement.

Step 3. Recalling $\alpha = \lim_{m \to \infty} E_m$, we can use the inequality

$$\inf_{k \geqslant m} E_k \geqslant \left(1 - \frac{16}{m}\right) E_m$$

in combination with $\alpha \geqslant \inf_{k \geqslant m} E_k$ to conclude, for m sufficiently large,

$$E_m \leqslant \left(1 + \frac{20}{m}\right) \alpha.$$

These two facts combined suggest that we would except E_m to approach its limit from below (though this will not be used in subsequent arguments). Let now $\varepsilon > 0$ be arbitrary and let m be so large that

$$\forall\, k \geqslant m : \qquad E_k \geqslant \alpha - \varepsilon.$$

We introduce a sequence of (transport cost) functions $T_m : V \to \mathbb{R}$ via

$$T_m(v) = \sum_{w \in V} d(v, w) \mu_m(w),$$

where μ_m is the empirical probability distribution of the first m elements of the sequence

$$\mu_m = \frac{1}{m} \sum_{\ell=1}^{m} \delta_{x_\ell}.$$

We observe that the μ_m-weighted average of T_m is explicit since

$$\sum_{v\in V} \mu_m(v) T_m(v) = \sum_{v,w\in V} \mu_m(v) d(v,w) \mu_m(w)$$
$$= \frac{1}{m^2} \sum_{i,j=1}^{m} d(x_i, x_j) = \frac{m-1}{m} E_m,$$

which, for m very large, is very close to E_m which, in turn, is at least $\alpha - \varepsilon$. We will now show that the maximal value of $T_m(v)$ is also close to $\alpha - \varepsilon$.

Step 4. The basic ingredient is a continuity estimate: if $\|T_m\|_{\ell^\infty}$ was much larger than α, then we would expect this to also be true for T_{m+1} since μ_{m+1} is fairly similar to μ_m when m is large (the difference is a single additional vertex). Indeed, we have

$$\mu_{m+1} = \frac{1}{m+1} \sum_{k=1}^{m+1} \delta_{x_k} = \frac{m}{m+1} \mu_m + \frac{1}{m+1} \delta_{x_{m+1}},$$

and therefore, since $T_m \geqslant 0$ and $T_m(v) \leqslant \mathrm{diam}(G)$,

$$\|T_{m+1} - T_m\|_{\ell^\infty} = \max_{v\in V} \left| T_{m+1}(v) - T_m(v) \right|$$
$$\leqslant \max_{v\in V} \left| \frac{m}{m+1} T_m(v) + \frac{d(x_{k+1}, v)}{m+1} - T_m(v) \right| \leqslant \frac{\mathrm{diam}(G)}{m+1}.$$

Suppose now there exists $v \in V$ such that

$$T_m(v) \geqslant \alpha + \delta$$

for some $\delta > 0$. Then the value of $T_{m+1}, T_{m+2}, \ldots$ at the same vertex is still large at least for the next few iterations of the process and we have for all $\ell \in \mathbb{N}$

$$\|T_{m+\ell}\|_{\ell_\infty} \geqslant T_{m+\ell}(v) \geqslant \alpha + \delta - \frac{\ell}{m+1} \mathrm{diam}(G),$$

and thus, for

$$1 \leqslant \ell \leqslant \frac{m\delta}{4\,\mathrm{diam}(G)},$$

we have that

$$T_{m+\ell}(v) \geqslant \alpha + \delta/2.$$

Step 5. Let $\varepsilon > 0$ fixed and all m be so large that

$$\forall\, k \geqslant m \qquad \alpha - \varepsilon \leqslant E_k \leqslant \left(1 + \frac{20}{k}\right)\alpha.$$

This will now be shown to imply that

$$\|T_m\|_{\ell^\infty} \leqslant \alpha + 3\sqrt{\mathrm{diam}(G)}\sqrt{\varepsilon}.$$

Let us suppose that, for some $\delta > 0$, we have $T_m(v) \geqslant \alpha + \delta$. We use Step 4 with $\ell = m\delta/(4\,\mathrm{diam}(G))$ (note that this may not be an integer and one may want to use $\lfloor \ell \rfloor$; instead, the subsequent arguments are not sensitive to this change and we omit them in the interest of simplicity of exposition). We argue that since in each step, we add at least $\alpha + \delta/2$,

$$\begin{aligned}(m+\ell)(m+\ell-1)E_{m+\ell} &\geqslant m(m-1)E_m + 2\ell m\left(\alpha + \frac{\delta}{2}\right) \\ &\geqslant m(m-1)(\alpha-\varepsilon) + 2\ell m\left(\alpha + \frac{\delta}{2}\right).\end{aligned}$$

We will now bound the first term from above and its lower bound from below. For the bound from above, we first recall that, for m sufficiently large,

$$E_{m+\ell} \leqslant \left(1 + \frac{20}{m}\right)\alpha,$$

and therefore, using $\delta \leqslant \mathrm{diam}(G)$,

$$\begin{aligned}(m+\ell)(m+\ell-1)E_{m+\ell} &\leqslant \left(1 + \frac{20}{m}\right)\alpha(m+\ell)^2 \\ &= \left(1 + \frac{20}{m}\right)\alpha m^2\left(1 + \frac{\delta}{4\,\mathrm{diam}(G)}\right)^2 \\ &\leqslant \alpha m^2\left(1 + \frac{\delta}{4\,\mathrm{diam}(G)}\right)^2 + \frac{20}{m}\alpha m^2\left(1 + \frac{\delta}{4\,\mathrm{diam}(G)}\right)^2 \\ &\leqslant 40\alpha m + \alpha m^2\left(1 + \frac{\delta}{4\,\mathrm{diam}(G)}\right)^2.\end{aligned}$$

Using $\alpha \leqslant \mathrm{diam}(G)$, this leads to

$$(m+\ell)(m+\ell-1)E_{m+\ell} \leqslant 40\alpha m + \alpha m^2 + \frac{\alpha\delta m^2}{2\,\mathrm{diam}(G)} + \frac{m^2\delta^2}{16\,\mathrm{diam}(G)}.$$

For the lower bound, we argue that by plugging in the definition of ℓ, we have

$$m(m-1)(\alpha-\varepsilon) + 2\ell m\left(\alpha + \frac{\delta}{2}\right) \geqslant (m-1)^2(\alpha-\varepsilon) + \frac{m^2\delta\alpha}{2\,\mathrm{diam}(G)} + \frac{m^2\delta^2}{4\,\mathrm{diam}(G)}.$$

We see that this leads to a contradiction as soon as

$$(m-1)^2\varepsilon + 40\alpha m \leqslant \frac{m^2\delta^2}{8\,\mathrm{diam}(G)},$$

and thus, for m sufficiently large,

$$\delta \leqslant 3\sqrt{\operatorname{diam}(G)}\sqrt{\varepsilon}.$$

This shows that

$$\limsup_{m\to\infty} \|T_m\|_{\ell^\infty} \leqslant \alpha.$$

Step 6. We will now refine this last statement and show that lim sup can be replaced by a lim on supp μ_m. Combining the estimates from the previous sections

$$\begin{aligned}\frac{m-1}{m}E_m &= \frac{1}{m^2}\sum_{i,j=1}^{m} d(x_i,x_j)\\ &= \sum_{v,w\in V}\mu_m(v)d(v,w)\mu_m(w)\\ &= \sum_{v\in V} T_m(v)\mu_m(v).\end{aligned}$$

Recalling that μ_m is a probability measure, that

$$\lim_{m\to\infty} E_m = \alpha$$

as well as

$$\lim_{m\to\infty} \|T_m\|_{\ell^\infty} = \alpha,$$

we conclude that T_m has to be close to maximal in each vertex carrying a nonvanishing portion of the probability mass in the sense of (3) of the statement. □

3.2 | Proof of Theorem 2

Proof. Let $G = (V, E)$ be arbitrary and let μ be a balanced measure supported on m vertices $w_1, \dots, w_m \in V$. We define the map $\phi : V \to \mathbb{R}^m_{\geqslant 0}$ via

$$\phi(v) = (\mu(w_1)d(v,w_1), \dots, \mu(w_m)d(v,w_m)).$$

ϕ maps V into the positive orthant $\mathbb{R}^m_{\geqslant 0}$. Moreover, we observe that since

$$T(v) = \sum_{w\in V} d(w,v)\mu(v) \qquad \text{assumes its maximum in the support}$$

and is constant on $\operatorname{supp}\mu$, we have, for each $1 \leqslant i \leqslant m$,

$$\|\phi(v)\|_{\ell^1} = \sum_{j=1}^{m} d(v, w_j)\mu(w_j) \leqslant \sum_{j=1}^{m} d(w_i, w_j)\mu(w_j).$$

Summing over i

$$\|\phi(v)\|_{\ell^1} \leqslant \sum_{i=1}^{m} \mu(w_i) \sum_{j=1}^{m} d(w_i, w_j)\mu(w_j).$$

Note that these inequalities are equations whenever v is in the support of μ. At this point, we recall the definition of a balanced measure: the total transport cost is maximized when transporting to vertex in the support of the measure. Therefore,

$$\sum_{i=1}^{m} \mu(w_i) \sum_{j=1}^{m} d(w_i, w_j)\mu(w_j) = \max_{v \in V} T(v).$$

We trivially have

$$\begin{aligned}\|\phi(v) - \phi(w)\|_{\ell^1} &\leqslant \sum_{j=1}^{m} \mu(w_j)|d(v, w_j) - d(w, w_j)| \\ &\leqslant \sum_{j=1}^{m} \mu(w_j) d(v, w) = d(v, w),\end{aligned}$$

which shows that the embedding is 1-Lipschitz with respect to $\ell^1(\mathbb{R}^m)$. Let us now fix two arbitrary vertices $v_1, v_2 \in \operatorname{supp}\mu$. The entry of $\phi(v_1) - \phi(v_2)$ in the v_1-th position is $-d(v_1, v_2)\mu(v_2)$. This implies

$$\|\phi(v_1) - \phi(v_2)\|_{\ell^\infty(\mathbb{R}^m)} \geqslant d(v_1, v_2)\mu(v_2).$$

This inequality is trivially correct also when $v_1 = v_2$. Thus, summing v_2 over all vertices in the support, we get

$$\frac{1}{m} \sum_{v_2 \in \operatorname{supp}(\mu)} \|\phi(v_1) - \phi(v_2)\|_{\ell^\infty(\mathbb{R}^m)} \geqslant \frac{1}{m} \sum_{v_2 \in \operatorname{supp}(\mu)} d(v_1, v_2)\mu(v_2).$$

By definition of a balanced measure and Lemma 3,

$$\frac{1}{m} \sum_{v_2 \in \operatorname{supp}(\mu)} d(v_1, v_2)\mu(v_2) = \max_{v} T(v) \geqslant \frac{\operatorname{diam}(G)}{2m}.$$

□

3.3 | Proof of Proposition 1

Proof. We fix $D \in \mathbb{R}^{n\times n}$ to be the distance matrix. Throughout the proof, we will identify probability measures with probability vectors. We first show that critical points of the functional are balanced. Suppose now that μ is a critical point and ν is an arbitrary signed measure with total weight $\nu(V) = 0$ such that

$$\operatorname{supp}\nu \subseteq \operatorname{supp}\mu = \{v \in V : \mu(v) > 0\}.$$

Then, for some $\varepsilon_0 > 0$ (depending only on μ and ν) and all $|\varepsilon| \leqslant \varepsilon_0$, we have that $\mu + \varepsilon\nu$ is also a probability measure, and thus, since μ is a critical point,

$$\langle(\mu + \varepsilon\nu), D(\mu + \varepsilon\nu)\rangle = \langle\mu, D\mu\rangle + o(\varepsilon).$$

Since D is symmetric, the linear term of the left-hand side is given by

$$\langle\nu, D\mu\rangle + \langle D\nu, \mu\rangle = 2\langle D\mu, \nu\rangle = 0.$$

This, in turn, implies that $D\mu$ restricted to $\operatorname{supp}\mu$ has to be constant since otherwise we could construct a signed measure ν for which the equation is not satisfied, which shows that μ could not have been a critical point. Suppose now that this constant c, the restriction of $D\mu$ onto $\operatorname{supp}\mu$, is different from $\|D\mu\|_{\ell^\infty}$ (in which case it has to satisfy $c < \|D\mu\|_{\ell^\infty}$). Then there exists a vertex $w \in V$ with $w \notin \operatorname{supp}\mu$ and $(D\mu)(w) > c$. Then, by taking any $w_2 \in \operatorname{supp}(\mu)$, we can construct the measure $\nu = \delta_w - \delta_{w_2}$ and see that, for ε sufficiently small, $\mu + \varepsilon\nu$ is still a probability measure with larger energy that again contradicts the fact that μ was a critical point. This shows that $D\mu$ equals $\|D\mu\|_{\ell^\infty}$ when restricted to $\operatorname{supp}\mu$ and is therefore a balanced measure. Let us now suppose that μ is a balanced measure. Let ν be an arbitrary signed measure such that $\mu + \varepsilon\nu$ is a probability measure for all ε sufficiently small (this means that the negative entries of ν have to be contained in $\operatorname{supp}\mu$). Then

$$\langle(\mu + \varepsilon\nu), D(\mu + \varepsilon\nu)\rangle = \langle\mu, D\mu\rangle + 2\varepsilon\langle D\mu, \nu\rangle + o(\varepsilon).$$

Since the negative weights of ν are contained in $\operatorname{supp}\mu$ and since $D\mu$ is maximal in $\operatorname{supp}\mu$, we have that $\langle D\mu, \nu\rangle \leqslant 0$ with equality if and only if all the positive weight of ν is also contained points where $D\mu$ assumes its maximum. This implies the statement. □

3.4 | Proof of Proposition 2

Proof. Let $G = (V, E)$ be fixed, let $x_1, \dots, x_k \in V$ be given, and suppose that $v \in V$ is *not* a boundary vertex. We will then prove that

$$\sum_{i=1}^{k} d(v, x_i) < \max_{w\in V} \sum_{i=1}^{k} d(w, x_i),$$

which establishes the desired result. Since v is not a boundary vertex, this means that for each vertex x_i,

$$\frac{1}{\deg(v)} \sum_{(v,w)\in E} d(w, x_i) \geqslant d(v, x_i).$$

Summing over all $1 \leqslant i \leqslant k$, we get that

$$f(v) = \sum_{i=1}^{k} d(v, x_i) \leqslant \frac{1}{\deg(v)} \sum_{(v,w)\in E} \sum_{i=1}^{k} d(w, x_i) = \frac{1}{\deg(v)} \sum_{(v,w)\in E} f(w).$$

This shows that the value of f in a vertex $v \notin \partial G$ can be bounded from above by the average value in an adjacent vertex. Therefore, if f assumes its maximum in any nonboundary vertex, it also assumes the maximum in all adjacent vertices. It is then possible, recalling that we are working with a connected graph, to hop from vertex to vertex until one ends up being adjacent to a boundary vertex, and thus, the maximum is also assumed at the boundary. □

ACKNOWLEDGMENTS

S.S. is supported by the NSF (DMS-2123224) and the Alfred P. Sloan Foundation.

JOURNAL INFORMATION

Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society's research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. R. Alexander, *Generalized sums of distances*, Pacific J. Math. **56** (1975), no. 2, 297–304.
2. R. Alexander, *Two notes on metric geometry*, Proc. Amer. Math. Soc. **64** (1977), no. 2, 317–320.
3. R. Alexander and K. Stolarsky, *Extremal problems of distance geometry related to energy integrals*, Trans. Amer. Math. Soc. **193** (1974), 1–31.
4. M. Aouchiche and P. Hansen, *Distance spectra of graphs: a survey*, Linear Algebra Appl. **458** (2014), 301–386.
5. L. Brown and S. Steinerberger, *Positive-definite functions, exponential sums and the greedy algorithm: a curious phenomenon*, J. Complexity **61** (2020), 101485.
6. L. Brown, *Sequences of well-distributed vertices on graphs and spectral bounds on optimal transport*, J. Fourier Anal. Appl. **27**, article 36 (2021).
7. G. Björck, *Distributions of positive mass, which maximize a certain generalized energy integral*, Ark. Mat. **3** (1958), 255–269.
8. D. Carando, D. Galicer, and D. Pinasco, *Energy integrals and metric embedding theory*, Int. Math. Res. Not. IMRN **2015**, no. 16, 7417–7435.
9. J. Cleary, S. Morris, and D. Yost, *Numerical geometry–numbers for shapes*, Amer. Math. Monthly **93** (1986), no. 4, 260–275.
10. A. Cloninger and H. Mhaskar, *A low discrepancy sequence on graphs*, J. Fourier Anal. Appl. **27**, article 76 (2021).
11. O. Gross, *The rendezvous value of metric space*, Advances in game theory, Princeton University Press, Princeton, NJ, 1964, pp. 49–53.
12. A. Hinrichs, P. Nickolas, and R. Wolf, *A note on the metric geometry of the unit ball*, Math. Z. **268** (2011), no. 3–4, 887–896.

13. J. Kleinberg, A. Slivkins, and T. Wexler, *Triangulation and embedding using small sets of beacons*, J. ACM (JACM) **56** (2009), 1–37.
14. R. Kritzinger, *Uniformly distributed sequences generated by a greedy minimization of the L_2 discrepancy*, arXiv:2109.06298.
15. G. Larcher, W. Schmid, and R. Wolf, *On the approximation of certain mass distributions appearing in distance geometry*, Acta Math. Hungar. **87** (2000), no. 4, 295–316.
16. G. C. Linderman and S. Steinerberger, *Numerical Integration on Graphs: where to sample and how to weigh*, Math. Comp. **89** (2020), 1933–1952.
17. J. v. Neumann, *Zur Theorie der Gesellschaftsspiele*, Math. Ann. **100** (1928), no. 1, 295–320.
18. P. Nickolas and R. Wolf, *Finite quasihypermetric spaces*, Acta Math. Hungar. **124** (2009), no. 3, 243–262.
19. P. Nickolas and R. Wolf, *Distance geometry in quasihypermetric spaces*, I. Bull. Aust. Math. Soc. **80** (2009), no. 1, 1–25.
20. P. Nickolas and R. Wolf, *Distance geometry in quasihypermetric spaces. II*, Math. Nachr. **284** (2011), no. 2–3, 332–341.
21. P. Nickolas and R. Wolf, *Distance geometry in quasihypermetric spaces. III*, Math. Nachr. **284** (2011), no. 5–6, 747–760.
22. E. Ng and H. Zhang, *Predicting internet network distance with coordinates-based approaches*, IEEE INFOCOM, 2002.
23. S. Steinerberger, *A nonlocal functional promoting low-discrepancy point sets*, J. Complexity **54** (2019), 101410.
24. S. Steinerberger, *Dynamically defined sequences with small discrepancy*, Monatshefte Math. **191** (2020), 639–655.
25. S. Steinerberger, *Polynomials with Roots on the Unit Circle: regularity of Leja sequences*, Mathematika **67** (2021), 553–268.
26. S. Steinerberger, *On combinatorial properties of greedy Wasserstein minimization*, arXiv: 2207.08043.
27. S. Steinerberger, *The boundary of a graph and its isoperimetric inequality*, arXiv:2201.03489.
28. S. Steinerberger, *Curvature on graphs via equilibrium measures*, J. Graph Theory, to appear.
29. S. Steinerberger, *The first eigenvector of a distance matrix is nearly constant*, Discrete Math. **346** (2023), 113291.
30. C. Thomassen, *The rendezvous number of a symmetric matrix and a compact connected metric space*, Amer. Math. Monthly **107** (2000), no. 2, 163–166.
31. R. Wolf, *On the average distance property and certain energy integrals*, Ark. Mat. **35** (1997), no. 2, 387–400.
32. R. Wolf, *Averaging distances in real quasihypermetric Banach spaces of finite dimension*, Israel J. Math. **110** (1999), 125–151.
33. T. Zamfirescu, *On longest paths and circuits in graphs*, Math. Scand. **38** (1976), 211–239.