CENTRE FOR **STOCHASTIC GEOMETRY**
AND ADVANCED **BIOIMAGING**

Ethan Anderes, Jesper Møller and Jakob G. Rasmussen

Isotropic covariance functions on graphs and their edges

No. 10, October 2017

Isotropic covariance functions on graphs and their edges

Ethan Anderes[1], Jesper Møller[2] and Jakob G. Rasmussen[2]

[1]University of California at Davis, anderes@ucdavis.edu
[2]Aalborg University, jm@math.aau.dk, jgr@math.aau.dk

Abstract

We develop parametric classes of covariance functions on linear networks and their extension to graphs with Euclidean edges, i.e., graphs with edges viewed as line segments or more general sets with a coordinate system allowing us to consider points on the graph which are vertices or points on an edge. Our covariance functions are defined on the vertices and edge points of these graphs and are isotropic in the sense that they depend only on the geodesic distance or on a new metric called the resistance metric (which extends the classical resistance metric developed in electrical network theory on the vertices of a graph to the continuum of edge points). We discuss the advantages of using the resistance metric in comparison with the geodesic metric as well as the restrictions these metrics impose on the investigated covariance functions. In particular, many of the commonly used isotropic covariance functions in the spatial statistics literature (the power exponential, Matérn, generalized Cauchy, and Dagum classes) are shown to be valid with respect to the resistance metric for any graph with Euclidean edges, whilst they are only valid with respect to the geodesic metric in more special cases.

Keywords: Geodesic metric, linear network, parametric classes of covariance functions, reproducing kernel Hilbert space, resistance metric, restricted covariance function properties.

1 Introduction

Linear networks are used to model a wide variety of non-Euclidean spaces occurring in applied statistical problems involving river networks, road networks, and dendrite networks, see e.g. Cressie et al. (2006), Ver Hoef et al. (2006), Ver Hoef and Peterson (2010), Okabe and Sugihara (2012), and Baddeley et al. (2015). However, the problem of developing valid random field models over networks is a decidedly difficult task. Compared to what is known for Euclidean spaces – where the results of Bochner and Schonberg allow practitioners to develop stationary Gaussian random fields using covariance functions, see e.g. Yaglom (1987) – the corresponding results for linear networks are few and far between. Even the fundamental notion of a stationary covariance function is, at best, ambiguous for linear networks. However, the notion of an isotropic covariance function can be made precise by requiring the function to depend only on a metric defined over the linear network. Often the easiest choice for such a metric is given by the length of the shortest path connecting two points, i.e., the geodesic metric. Still there are no general results which establish when a given function generates a valid isotropic covariance function with respect to this metric. Indeed, Baddeley et al. (2017) concluded that spatial point process models on a linear network with a pair correlation function which is only depending on shortest path distance "may be quite rare".

In this paper, we use Hilbert space embedding techniques to establish that many of the flexible isotropic covariance models used in spatial statistics are valid (i.e., positive semi-definite) over linear networks with respect the geodesic metric and a new metric introduced in Section 2.3. This new metric is called the resistance metric because it extends the classical resistance metric developed in electrical network theory. The validity of these covariance models do not hold, however, over the full parametric range available in Euclidean spaces. Moreover, we show the results for the geodesic metric apply to a much smaller class of linear networks and can not be extended to any graph which have three or more paths connecting two points on the linear network. This is in stark contrast to the resistance metric where we show there is no restriction on the type of linear network for which they apply.

We develop a generalization of a linear network which we call a graph with Euclidean edges. Essentially, this is a graph $(\mathcal{V}, \mathcal{E})$ where each edge $e \in \mathcal{E}$ is additionally associated to an abstract set in bijective correspondence with a line segment of $\mathbb{R}$. Treating the edges as abstract sets allows us to consider points on the graph that are either vertices or points on the edges, and the bijective assumption gives each edge set a (one-dimensional) Cartesian coordinate system for measuring distances between any two points on the edge (therefore the terminology Euclidean edges). The within-edge Cartesian coordinate system will be used to extend the geodesic and the resistance metric on the vertex set to the whole graph (including points on the edges). Our objective then is to construct parametric families of isotropic covariance functions over graphs with Euclidean edges with respect to the two metrics (in fact our covariance functions will be (strictly) positive definite). Thereby isotropic Gaussian random fields on the whole graph can be constructed; in turn, a rich class of isotropic point process models on the whole graph can be constructed via a log Gaussian Cox process (Møller et al., 1998; Møller and Waagepetersen, 2004). We leave this and other applications of our paper for future work.

1.1 Outline

The paper is organized as follows. In the remainder of this introduction we present the definition of a graph with Euclidean edges and finish with a summary of our main results. Then, in Section 2, the details of the geodesic and resistance metric on a graph with Euclidean edges are given. Section 3 establishes conditions for Hilbert space embeddings of our two metrics. This is the key result – used in Section 4 – for establishing that our parametric families of isotropic covariance functions are valid. Finally, Section 5 discusses restrictions of the parameters of these covariance functions. Technical details of our proofs are deferred to Appendix A.

1.2 Graphs with Euclidean edges

A linear network is typically defined as the union of a finite collection of line segments in $\mathbb{R}^2$ with distance between two points defined as the length of the shortest path connecting the points. This definition, although conceptually clear, does have limitations that restrict their application. For example, in the case of road networks:

- Bridges and tunnels can generate networks which do not have a planar representation as a union of line segments in $\mathbb{R}^2$.
- Varying speed limits or number of traffic lanes may require distances on line segments to be measured differently than their spatial extent.

A graph with Euclidean edges, defined below, is a generalization of linear networks that easily overcomes the above-mentioned limitations while still retaining the salient feature relevant to applications: that edges (or line segments) have a Cartesian coordinate system associated with them.

Definition 1.1. *A triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \{\varphi_e\}_{e\in\mathcal{E}})$ which satisfies the following conditions (a)–(d) is called a* graph with Euclidean edges*.*

(a) Graph structure*: $(\mathcal{V}, \mathcal{E})$ is a finite simple connected graph, meaning that the vertex set $\mathcal{V}$ is finite, the graph has no multiple edges or loops, and every pair of vertices is connected by a path.*

(b) Edge sets*: Each edge $e \in \mathcal{E}$ is associated with a unique abstract set, also denoted e, where the vertex set $\mathcal{V}$ and all the edge sets $e \in \mathcal{E}$ are mutually disjoint.*

(c) Edge coordinates*: For each edge $e \in \mathcal{E}$, if $u, v \in \mathcal{V}$ are the vertices connected by e, then φ_e is a bijection defined on $e \cup \{u, v\}$ (the union of the edge set e and the vertices $\{u, v\}$) such that φ_e maps e onto an open interval $(\underline{e}, \overline{e}) \subset \mathbb{R}$ and $\{u, v\}$ onto the endpoints $\{\underline{e}, \overline{e}\}$.*

(d) Distance consistency*: Let $d_\mathcal{G}(u, v)\colon \mathcal{V} \times \mathcal{V} \mapsto [0, \infty)$ denote the standard shortest-path weighted graph metric on the vertices of $(\mathcal{V}, \mathcal{E})$ with edge weights given by $\overline{e} - \underline{e}$ for every $e \in \mathcal{E}$. Then, for each $e \in \mathcal{E}$ connecting two vertices $u, v \in \mathcal{V}$, the following equality holds:*

$$d_\mathcal{G}(u, v) = \overline{e} - \underline{e}.$$

We write $u \in \mathcal{G}$ as a synonym for u in $\mathcal{V} \cup \bigcup_{e\in\mathcal{E}} e$, the whole graph given by the union of $\mathcal{V}$ and all edges $e \in \mathcal{E}$.

If we consider a linear network $\bigcup_{i\in\mathcal{I}} \ell_i$ consisting of closed line segments $\ell_i \subset \mathbb{R}^2$ which intersect only at their endpoints, we can easily construct a graph with Euclidean edges as follows. Let $\mathcal{V}$ be the set of endpoints of the line segments. Let the edge sets $e_i \in \mathcal{E}$ correspond to the relative interior of the corresponding line segment ℓ_i. Let the bijections φ_{e_i} be given by the inverse of path-length parameterization of ℓ_i. Then conditions (a)–(d) are easily seen to hold.

Any triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \{\varphi_e\}_{e\in\mathcal{E}})$ for which $(\mathcal{V}, \mathcal{E})$ forms a tree graph is automatically a graph with Euclidean edges given that conditions (b) and (c) hold. In this case, $\mathcal{G}$ is said to be a *Euclidean tree*. Figure 1 shows an example. If the graph $(\mathcal{V}, \mathcal{E})$, associated with a graph with Euclidean edges $\mathcal{G}$, forms a cycle, then $\mathcal{G}$ is said to be a *Euclidean cycle*. Conversely, if $(\mathcal{V}, \mathcal{E})$ forms a cycle graph with edge bijections $\{\varphi_e\}_{e\in\mathcal{E}}$, then the resulting triple $\mathcal{G} := (\mathcal{V}, \mathcal{E}, \{\varphi_e\}_{e\in\mathcal{E}})$ satisfies the conditions of Definition 1.1 whenever there are three or more vertices (to ensure there are no multiple edges) and for every $e_o \in \mathcal{E}$ the following inequality is satisfied:

$$\overline{e}_o - \underline{e}_o \leq \sum_{\substack{e\in\mathcal{E} \\ e\neq e_o}} (\overline{e} - \underline{e}). \tag{1.1}$$

The above condition guarantees that no edge spans more than *half* of the circumference of the cycle, implying that distance consistency holds for $\mathcal{G}$. Figure 2 illustrates examples of Euclidean cycles (the two first graphs) and an example of a graph violating both conditions (a) and (d) in Definition 1.1 (the last graph).

In all the above examples we have used spatial curves and line segments to represent the edges. It it worth pointing out that this is simply a visualization device. Indeed, the structure of a graph with Euclidean edges is completely invariant to the geometric shape of the visualized edges just so long as the path-length of each edge is preserved. This concept is important when considering the example given in Figure 3, where the edge represented by the diagonal line in the leftmost drawing represents a bridge or tunnel bypassing the other diagonal edge. Hence the lack of vertices at the intersection with that edge. Note that it is impossible to avoid this intersection when lengths of edges are fixed. This implies that this graph with Euclidean edges can not be represented as a linear network in $\mathbb{R}^2$.

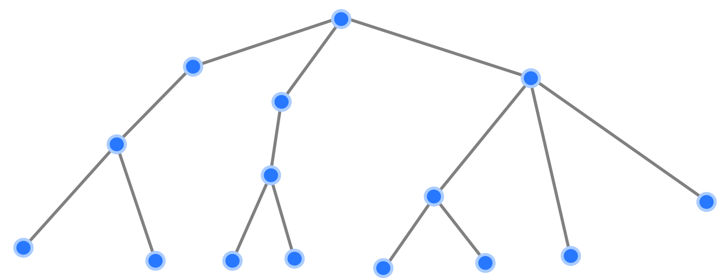


Figure 1: A Euclidean tree constructed from the linear network of grey lines. The blue dots represent the vertices.

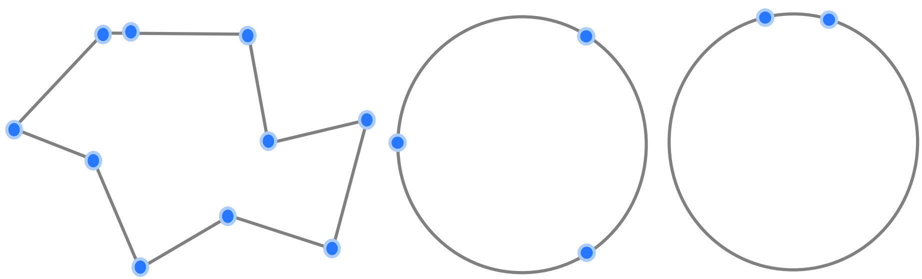


Figure 2: The two graphs on the left are Euclidean cycles. However, the right most graph is *not* a graph with Euclidean edges.

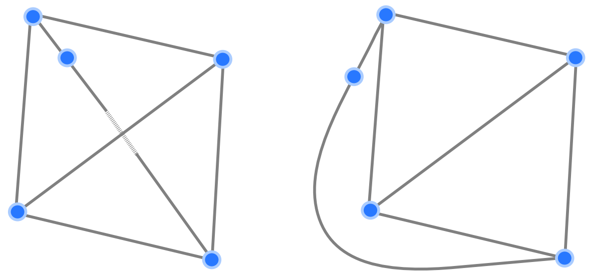


Figure 3: The two diagrams above show a graph with Euclidean edges $\mathcal{G}$ which can *not* be represented as a linear network in $\mathbb{R}^2$. The diagram on the left is drawn in a way that visually preserves edge length but forces an intersection that does not correspond to a vertex in $\mathcal{G}$ (the dashed segment indicates that one edge passes under the other). The diagram on the right is drawn without non-vertex intersections but requires curved segments that have length which do not correspond to the lengths determined by the edge bijections for a linear network.

1.3 Summary of main results

This section presents our main theorems explicitly, leaving the proofs and precise definitions for later sections.

Our first contribution is to establish sufficient conditions for a function $C : [0, \infty) \mapsto \mathbb{R}$ to generate a (strictly) positive definite function of the form $C(d(u, v))$ where $d(u, v)$ is a metric defined over the vertices and edge points of a graph with Euclidean edges $\mathcal{G}$; then we call $\mathcal{G} \times \mathcal{G} \ni (u, v) \mapsto C(d(u, v)) \in \mathbb{R}$ an *isotropic covariance function* and C its *radial profile*. We study two metrics, the *geodesic metric*, $d_{G,\mathcal{G}}$, as defined in Section 2.2, and a new *resistance metric*, $d_{R,\mathcal{G}}$, as developed in Section 2.3, which extends the resistance metric on the vertex set – from electrical network theory (Klein and Randić, 1993) – to the continuum of edge points on $\mathcal{G}$. As is apparent from the following two theorems, there are fundamental differences in terms of the generality of valid isotropic covariance functions when measuring distances under the two metrics.

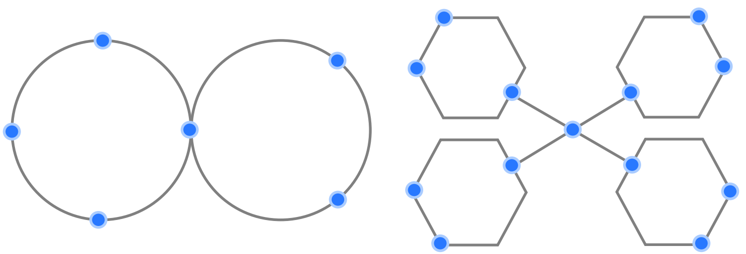


Figure 4: Examples of finite sequential 1-sums of cycles and trees. Left: A 1-sum of two Euclidean cycles. Right: A sequential 1-sum of four Euclidean cycles and one Euclidean tree.

In Theorem 1.2 below, we consider the *1-sum* of two graphs with Euclidean edges $\mathcal{G}_1$ and $\mathcal{G}_2$ having only a single point in common, $\mathcal{G}_1 \cap \mathcal{G}_2 = \{x_0\}$. This is defined explicitly in Section 3, but the concept is easy to visualize as the merging at x_0 and the concept easily extends to the case of three or more graphs with Euclidean edges;

Figure 4 gives two graphical illustrations.

Theorem 1.2. *For a graph with Euclidean edges $\mathcal{G}$ and any function $C : [0, \infty) \mapsto \mathbb{R}$ from the parametric classes in Table 1, we have the following:*

- $C(d_{R,\mathcal{G}}(u, v))$ *is (strictly) positive definite over* $(u, v) \in \mathcal{G} \times \mathcal{G}$*;*
- $C(d_{G,\mathcal{G}}(u, v))$ *is (strictly) positive definite over* $(u, v) \in \mathcal{G} \times \mathcal{G}$ *if* $\mathcal{G}$ *forms a finite sequential 1-sum of Euclidean cycles and trees.*

Table 1: Parametric classes of functions $C : [0, \infty) \mapsto \mathbb{R}$ which generate isotropic correlation functions $C(d_{R,\mathcal{G}}(\,\cdot\,,\,\cdot\,))$, i.e., when distance is measured by the resistance metric and $C(0) = 1$. Note: K_α denotes the modified Bessel function of order α.

Type	*Parametric form*	*Parameter range*
Power exponential	$C(t) = \exp(-\beta t^\alpha)$	$0 < \alpha \le 1,\ \beta > 0.$
Matérn	$C(t) = (\beta t)^\alpha K_\alpha(\beta t)$	$0 < \alpha \le \frac{1}{2},\ \beta > 0.$
Generalized Cauchy	$C(t) = (\beta t^\alpha + 1)^{-\xi/\alpha}$	$0 < \alpha \le 1,\ \beta, \xi > 0.$
Dagum	$C(t) = \left[1 - \left(\dfrac{\beta t^\alpha}{1 + \beta t^\alpha}\right)^{\xi/\alpha}\right]$	$0 < \alpha \le 1,\ 0 < \xi \le 1,\ \beta > 0.$

Theorem 1.2 shows that many of the commonly used isotropic covariance functions in the spatial statistics literature are valid with respect to the resistance metric for *any* graph with Euclidean edges $\mathcal{G}$. In contrast, the 1-sum restriction for the geodesic metric suggests a degeneracy when modeling isotropic covariance function in terms of the geodesic metric. The next result, Theorem 1.3, shows that for the geodesic metric, Theorem 1.2 can not be extended to the generality given for the resistance metric. Here, for $S = R$ or $S = G$, if there exists some $\beta > 0$ so that $e^{-\beta\, d_{S,\mathcal{G}}(u,v)}$ is not a positive semi-definite function over $(u, v) \in \mathcal{G} \times \mathcal{G}$, we say that $\mathcal{G}$ is a *forbidden graph (for the exponential class) with respect to the metric* $d_{S,\mathcal{G}}$. Figure 5 shows examples in case of the geodesic metric. Note that if a forbidden graph is present as a subgraph of $\mathcal{G}$, then $\mathcal{G}$ is forbidden as well.

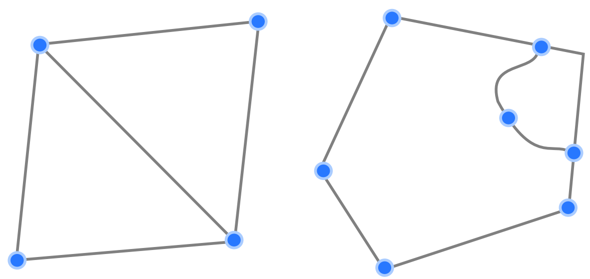


Figure 5: Examples of forbidden graphs for the exponential class with respect to the geodesic metric.

Theorem 1.3. *If $\mathcal{G}$ is a graph with Euclidean edges for which there exists three distinct paths connecting two points $u, v \in \mathcal{G}$, then $\mathcal{G}$ is a forbidden graph for the exponential class with respect to the geodesic metric.*

In Section 2.3, our resistance metric $d_{R,\mathcal{G}}$ is constructed to circumvent the apparent degeneracy seen in Theorem 1.3 for the geodesic metric. In particular, we connect $d_{R,\mathcal{G}}$ to the variogram of a canonical Gaussian random field. This connection provides the following key Hilbert space embedding result.

Theorem 1.4. *If $\mathcal{G}$ is a graph with Euclidean edges, there exists a Hilbert space H and an embedding $\varphi\colon \mathcal{G} \mapsto H$ such that*

$$\sqrt{d_{R,\mathcal{G}}(u,v)} = \left\|\varphi(u) - \varphi(v)\right\|_H \tag{1.2}$$

for all $u, v \in \mathcal{G}$ where $d_{R,\mathcal{G}}$ is the resistance metric developed in Section 2.3. If, in addition, $\mathcal{G}$ forms a sequential 1-sum of a finite number of Euclidean cycles and trees, then the above result also holds for the geodesic metric $d_{G,\mathcal{G}}$.

Theorem 1.4 is proved in Section 3 and forms the main tool used in Section 4 to establish Theorem 1.2 and Theorem 1.3. Finally, in Section 5 we establish constraints – depending on the graph structure of $\mathcal{G}$ – for features of any radial profile which generates an isotropic covariance function with respect to either metric, resistance or geodesic.

2 The geodesic and resistance metric on $\mathcal{G}$

In this section we develop the geodesic and resistance metric over graphs with Euclidean edges. The geodesic metric, developed in Section 2.2, is easily constructed once a concrete notion of a path is defined. The resistance metric, in contrast, requires decidedly more work and is developed in Section 2.3.

2.1 Notation and terminology

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \{\varphi_e\}_{e\in\mathcal{E}})$ be a graph with Euclidean edges. To stress the dependence on $\mathcal{G}$, write $\mathcal{V}(\mathcal{G}) \equiv \mathcal{V}$ and $\mathcal{E}(\mathcal{G}) \equiv \mathcal{E}$. If $u, v \in \mathcal{V}(\mathcal{G})$ are connected by an edge in $\mathcal{E}(\mathcal{G})$, we say they are *neighbours* and write $u \sim v$. If $u \in e \in \mathcal{E}(\mathcal{G})$, we let $\underline{u}, \overline{u}$ denote the neighbouring vertices which are connected by edge e and ordered so that $\underline{u}$ corresponds to $\underline{e}$ and $\overline{u}$ corresponds to $\overline{e}$. When $u \in \mathcal{V}(\mathcal{G})$, we define $\overline{u} = \underline{u} = u$. The distinction between $\underline{u}, \overline{u}$ and $\underline{e}, \overline{e}$ can be seen by noting that $\underline{e}, \overline{e} \in \mathbb{R}$ but $\underline{u}, \overline{u} \in \mathcal{V}(\mathcal{G})$.

Let $e \in \mathcal{E}(\mathcal{G})$ and $I \subseteq (\underline{e}, \overline{e})$ be a non-empty interval. Then $\varphi_e^{-1}(I)$ is called a *partial edge*, its two *boundaries* correspond to the two-point set $\varphi_e^{-1}(\overline{I} \setminus I^o)$, where $\overline{I}$ is the closure of I and I^o is the open interior of I, and its *length* is given by the Euclidean length of I. Thus the edge e is also a partial edge and its length is denoted $\mathrm{len}(e)$.

Two partial edges are called *incident* if they share a common boundary in $\mathcal{G}$. A *path* connecting two distinct points $u, v \in \mathcal{G}$ is denoted p_{uv} and given by an alternating sequence $u_1, e_1, u_2, e_2, \ldots, u_n, e_n, u_{n+1}$, where $u_1, \ldots, u_{n+1} \in \mathcal{G}$ are pairwise distinct, $u_1 = u$, $u_{n+1} = v$, and $e_1, e_2, \ldots, e_n$ are non-overlapping partial edges such that each e_i has boundary $\{u_i, u_{i+1}\}$. Moreover, the *length of* p_{uv} is denoted $\mathrm{len}(p_{uv})$ and defined as the sum of the lengths of $e_1, e_2, \ldots, e_n$.

2.2 Geodesic metric

For a graph with Euclidean edges $\mathcal{G}$, the *geodesic distance* is defined for all $u, v \in \mathcal{G}$ by

$$d_{G,\mathcal{G}}(u,v) = \inf\{\mathrm{len}(p_{uv})\} \tag{2.1}$$

where the infimum is over all paths connecting u and v. Using the consistency requirement given in Definition 1.1 (d), the following theorem is easily verified.

Theorem 2.1. *If $\mathcal{G}$ is a graph with Euclidean edges, then $d_{G,\mathcal{G}}(u,v)$ is a metric over $u, v \in \mathcal{G}$ satisfying the following.*

- *Restricting $d_{G,\mathcal{G}}$ to $\mathcal{V}(\mathcal{G})$ results in the standard weighted shortest-path graph metric with edge weights given by* $\mathrm{len}(e)$.
- *$d_{G,\mathcal{G}}$ is an extension of the Euclidean metric on each edge $e \in \mathcal{E}(\mathcal{G})$ induced by the bijection φ_e. That is, $d_{G,\mathcal{G}}(u,v) = |\varphi_e(u) - \varphi_e(v)|$ whenever $u, v \in e \in \mathcal{E}(\mathcal{G})$.*

2.3 Resistance metric

The resistance metric typically refers to a distance derived from electrical network theory on the vertices of a finite or countable graph with each edge representing a resistor with a given conductance, see e.g. Jorgensen and Pearse (2010) and the references therein. By definition, the resistance between two vertices u and v is the voltage drop when a current of one ampere flows from u to v. For a graph with Euclidean edges $\mathcal{G}$, there are two reasons why it is natural to consider an extension of the resistance metric, defined on just the vertices and edge conductance given by inverse geodesic length, to the continuum of edge points and vertices of $\mathcal{G}$. The first reason is purely mathematical: The resulting metric solves the degeneracy problem found in Theorem 1.3. Second, resistance may be a natural metric for applications associated with flow and travel time across street networks: For example, the total inverse resistance of resistors in parallel is equal to the sum of their individual inverse resistances; correspondingly, multiple pathways engender better flow.

In developing this extension, we take a somewhat non-standard approach and define a metric over $\mathcal{G}$ with the use of an auxiliary random field $Z_{\mathcal{G}}$ with index set $\mathcal{G}$. The resulting metric is then *defined* to be the variogram of $Z_{\mathcal{G}}$:

$$d_{R,\mathcal{G}}(u,v) := \mathrm{var}(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v)), \qquad u, v \in \mathcal{G}. \tag{2.2}$$

Theorem 2.2 and 2.3 below show that $d_{R,\mathcal{G}}$ does in fact give the natural extension of the electrical network resistance metric: $d_{R,\mathcal{G}}$ evaluated on any additional edge points will result in the same metric that would be obtained on the resulting discrete electrical network.

Before presenting the formal construction of $Z_{\mathcal{G}}$ and our results, we give a brief outline. The form of $Z_{\mathcal{G}}$ will be defined as a finite sum of independent zero-mean Gaussian random fields:

$$Z_{\mathcal{G}}(u) := Z_\mu(u) + \sum_{e\in\mathcal{E}(\mathcal{G})} Z_e(u), \qquad u \in \mathcal{G}. \tag{2.3}$$

The field Z_μ is characterized by the multivariate Gaussian vector $(Z_\mu(v); v \in \mathcal{V}(\mathcal{G}))$ whose covariance matrix is related to the so-called graph Laplacian in electrical network theory; this vector is linearly interpolated across the edges so that $Z_\mu(u)$ is defined for all points $u \in \mathcal{G}$. For each $e \in \mathcal{E}$, the random field Z_e is only defined to be non-zero on edge e and $Z_e(u) = B_e(\varphi_e(u))$ if $u \in e$ or u is a boundary point of e, where B_e is an independent Brownian bridge defined over $[\underline{e}, \overline{e}]$. Although the construction of $Z_{\mathcal{G}}$ appears ad-hoc, we will show that the variogram of the resulting random field $Z_{\mathcal{G}}$ results in the continuum extension of the resistance metric found in electrical network theory.

2.3.1 Construction of Z_μ

The random field Z_μ is constructed via analogy to electrical network theory and using the following ingredients. We view each edge in $\mathcal{G}$ as a resistor with conductance function $c : \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G}) \mapsto [0, \infty)$ given by

$$c(u, v) = \begin{cases} 1/d_{G,\mathcal{G}}(u, v) & \text{if } u \sim v, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Let $\mathbb{R}^{\mathcal{V}(\mathcal{G})}$ denote the vector space of real functions h defined on $\mathcal{V}(\mathcal{G})$; when convenient we view h as a vector indexed by $\mathcal{V}(\mathcal{G})$. Also let $u_o \in \mathcal{V}(\mathcal{G})$ be an arbitrarily chosen vertex called the origin; this is only introduced for technical reasons as explained below. Define $L : \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G}) \mapsto \mathbb{R}$ as the function/matrix with coordinates

$$L(u, v) = \begin{cases} 1 + c(u_o) & \text{if } u = v = u_o, \\ c(u) & \text{if } u = v \neq u_o, \\ -c(u, v) & \text{otherwise,} \end{cases} \tag{2.5}$$

where $c(u) := \sum_v c(u, v) = \sum_{v\sim u} c(u, v)$ corresponds to the sum of the conductances associated to the edges incident to vertex u. Obviously, L is symmetric and a simple calculation shows that for $z, w \in \mathbb{R}^{\mathcal{V}(\mathcal{G})}$,

$$z^T Lw = z(u_o)w(u_o) + \frac{1}{2}\sum_{u\sim v}(z(u) - z(v))c(u, v)(w(u) - w(v)), \tag{2.6}$$

so $z^T Lz = 0$ if and only if $z(u) = 0$ for all $u \in \mathcal{V}(\mathcal{G})$. Thus L is (strictly) positive definite with (strictly) positive definite matrix inverse L^{-1}. Notice that the matrix L is similar to what would be called the "Laplacian matrix" from electrical network theory, see e.g. Kigami (2003) and Jorgensen and Pearse (2010), except that L has the additional 1 added at u_o. The role of the origin u_o is to make L (strictly) positive definite, but the resistance metric will be shown to be invariant to this choice and have the correct form (see Theorem 2.2).

Now, the random field Z_μ is simply defined by linearly interpolating a collection of Gaussian random variables associated with the vertices $\mathcal{V}(\mathcal{G})$: Let $v_1, v_2, \ldots, v_n$ denote the vertices in $\mathcal{V}(\mathcal{G})$ and define Z_μ at these vertices by

$$(Z_\mu(v_1), \ldots, Z_\mu(v_n))^T \sim \mathcal{N}(0, L^{-1}). \tag{2.7}$$

To define the value of $Z_\mu(u)$ at any point $u \in \mathcal{G}$ we interpolate across each edge as follows

$$Z_\mu(u) = (1 - d(u))Z_\mu(\underline{u}) + d(u)Z_\mu(\overline{u}) \tag{2.8}$$

where $d(u)$ denotes the distance of u from $\underline{u}$ as a proportion of the length of the edge containing u, formally given by

$$d(u) = \begin{cases} d_{G,\mathcal{G}}(u, \underline{u})/d_{G,\mathcal{G}}(\underline{u}, \overline{u}) & \text{if } u \notin \mathcal{V}(\mathcal{G}), \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

Notice that the covariance function $R_\mu(u, v) := \operatorname{cov}(Z_\mu(u), Z_\mu(v))$ can be computed explicitly: For any $u, v \in \mathcal{G}$,

$$\begin{aligned} R_\mu(u, v) = d(u)d(v)L^{-1}(\overline{u}, \overline{v}) &+ [1 - d(u)][1 - d(v)]L^{-1}(\underline{u}, \underline{v}) \\ + d(u)[1 - d(v)]L^{-1}(\overline{u}, \underline{v}) &+ [1 - d(u)]d(v)L^{-1}(\underline{u}, \overline{v}). \end{aligned} \tag{2.10}$$

2.3.2 Construction of Z_e

The definition of Z_μ in the previous section used explicitly an analogy to electrical network theory. So it should come as no surprise that the variogram of Z_μ gives something related to the resistance metric. However, this will only be true at the vertices. What we want is the electrical network property to hold for all points of $\mathcal{G}$ without the necessity of re-computing the matrix L for additional edge points. By simply adding Brownian bridge fluctuations over each edge, this turns out to give the right amount of variability.

To formally define a Brownian bridge process over each edge $e \in \mathcal{E}(\mathcal{G})$, we use the edge bijection φ_e which identify points on e with points in the interval $(\underline{e}, \overline{e}) \subset \mathbb{R}$. For all $e \in \mathcal{E}(\mathcal{G})$, let B_e denote mutually independent Brownian bridges which are independent of Z_μ, where B_e is defined on $[\underline{e}, \overline{e}]$ so that $B_e(\underline{e}) = B_e(\overline{e}) = 0$. For any $u \in \mathcal{G}$, we define

$$Z_e(u) = \begin{cases} B_e(\varphi_e(u)) & \text{if } u \in e, \\ 0 & \text{otherwise.} \end{cases} \tag{2.11}$$

Letting $R_e(u, v) = \operatorname{cov}(Z_e(u), Z_e(v))$, we have for any $u, v \in \mathcal{G}$,

$$R_e(u, v) = \begin{cases} \big[d(u) \wedge d(v) - d(u)d(v)\big] d_{G,\mathcal{G}}(\underline{u}, \overline{u}) & \text{if } u, v \in e, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

Note that the covariance function $R_\mathcal{G}$ for the random field $Z_\mathcal{G}$, defined in (2.3), satisfies

$$R_\mathcal{G}(u, v) = R_\mu(u, v) + \sum_{e \in \mathcal{E}(\mathcal{G})} R_e(u, v), \qquad u, v \in \mathcal{G}. \tag{2.13}$$

2.3.3 Properties of $d_{G,\mathcal{G}}$ and $d_{R,\mathcal{G}}$

The following Theorem 2.2 shows that $d_{R,\mathcal{G}}$ is indeed the extension of the classical effective resistance on electrical networks and it is invariant to the choice of origin u_o (used in the construction of L in (2.5)). Further, Theorem 2.3 shows that $d_{R,\mathcal{G}}$ is invariant to the addition of vertices and removal of vertices with degree two. Finally, Theorem 2.4 characterizes $d_{R,\mathcal{G}}$ via an associated infinite dimensional reproducing kernel Hilbert space. The proofs of the theorems are given in Appendix A.2.

Theorem 2.2. *For a graph with Euclidean edges $\mathcal{G}$, $d_{R,\mathcal{G}}$ is indeed a metric, it is invariant to the choice of origin u_o, and it simplifies to the classic (effective) resistance metric over the vertices when $\mathcal{G}$ is considered to be an electrical network with nodes $\mathcal{V}(\mathcal{G})$, resistors given by the edges $e \in \mathcal{E}(\mathcal{G})$, and conductances given by* $1/\operatorname{len}(e)$ *for* $e \in \mathcal{E}(\mathcal{G})$.

An important property of the geodesic metric on graphs with Euclidean edges is that distances are, in some sense, invariant to the replacement of an edge by two new edges merging at a new degree 2 vertex. This is illustrated in Figure 2 where it is clear that geodesics are the same for the left-most graph and the middle graph (when the edge lengths are scaled so the circumferences are equal) regardless of the fact that the left-most graph has more vertices and edges.

Perhaps surprisingly, this important property also holds for $d_{R,\mathcal{G}}$. To state the result, we need to be precise about what it means to add a vertex on an edge and correspondingly remove a degree 2 vertex (merging the corresponding incident edges). The operations will be generically referred to as splitting and merging: For $u \in e \in \mathcal{E}(\mathcal{G})$, define the partial edges $\underline{u}u = \{\varphi_e^{-1}(t) : \underline{e} < t < \varphi_e(u)\}$ and $u\overline{u} = \{\varphi_e^{-1}(t) : \varphi_e(u) < t < \overline{e}\}$, and partition $e = \{\underline{u}u\} \cup \{u\} \cup \{u\overline{u}\}$. Then the operation of *splitting an edge $e \in \mathcal{E}(\mathcal{G})$ at $u \in e$* results in a new graph $\mathcal{G}_{\text{split}}$ with Euclidean edges which is obtained by adding u to $\mathcal{V}(\mathcal{G})$ and replacing $e \in \mathcal{E}(\mathcal{G})$ with new edges $\underline{u}u$ and $u\overline{u}$. The operation of *merging two edges $e_1, e_2 \in \mathcal{E}(\mathcal{G})$ which are incident to a degree two vertex $v \in \mathcal{V}(\mathcal{G})$* results in a new graph with Euclidean edges $\mathcal{G}_{\text{merge}}$ simply obtained by removing v from $\mathcal{V}(\mathcal{G})$ and replacing $e_1, e_2 \in \mathcal{E}(\mathcal{G})$ with the single merged edge given by $e_1 \cup \{v\} \cup e_2$.

Clearly, $\mathcal{G}$, $\mathcal{G}_{\text{merge}}$ and $\mathcal{G}_{\text{split}}$ are equal as point sets. It is also clear that the geodesic metric is *invariant to splitting edges and merging edges at degree two vertices* in the sense that

$$d_{G,\mathcal{G}}(u, v) = d_{G,\mathcal{G}'}(u, v)$$

for all $u, v \in \mathcal{G}$ whenever $\mathcal{G}'$ is obtained from $\mathcal{G}$ by a finite sequence of edge splitting operations and edge merging operations which meet at a degree two vertex. The following theorem shows this property also holds for the resistance metric.

Theorem 2.3. *For a graph with Euclidean edges $\mathcal{G}$, the resistance and geodesic metrics $d_{R,\mathcal{G}}$ and $d_{G,\mathcal{G}}$ are invariant to splitting edges and merging edges at degree two vertices.*

Theorem 2.2 and 2.3 show that $d_{R,\mathcal{G}}$ is the appropriate extension of the classic resistance metric over finite nodes of an electrical network to the continuum of edge points over a graph with Euclidean edges. Indeed, the following theorem is an extension of a well-known result from electrical network theory.

Theorem 2.4. *For a graph with Euclidean edges $\mathcal{G}$,*

$$d_{R,\mathcal{G}}(u,v) \le d_{G,\mathcal{G}}(u,v), \qquad u,v \in \mathcal{G},$$

with equality if and only if $\mathcal{G}$ is a Euclidean tree, in which case

$$R_{\mathcal{G}}(u,v) = \left\{ d_{G,\mathcal{G}}(u,u_o) + d_{G,\mathcal{G}}(v,u_o) - d_{G,\mathcal{G}}(u,v) \right\}/2 + 1, \qquad u,v \in \mathcal{G}. \tag{2.14}$$

The inequality in Theorem 2.4 is intuitively clear because multiple pathways engender better flow. Moreover, the explicit form of $R_{\mathcal{G}}$ in (2.14) not only gives an illustration of where u_o appears in $R_{\mathcal{G}}$ (and disappears in $d_{R,\mathcal{G}}$), but also illuminates the deep connection with the independent increments of a Brownian motion on $\mathbb{R}$ started at the origin 0. For example, if $\mathcal{G}$ has vertices 0 and 1 connected by a single edge $e = (0,1)$, φ_e is the identity and $u_0 = 0$, then (2.14) gives that $R_{\mathcal{G}}(u,v) = 1 + (|u| + |v| - |u-v|)/2$

Our final result on the resistance metric, although stated last and verified in Appendix A.1, gives the above three theorems as near corollaries and does so by characterizing the reproducing kernel Hilbert space of functions over $\mathcal{G}$ which is associated to the Gaussian random field $Z_{\mathcal{G}}$ (see e.g. Wahba (1990)). To state the result we need some notation for functions defined over $\mathcal{G}$. For $f : \mathcal{G} \mapsto \mathbb{R}$ and $e \in \mathcal{E}(\mathcal{G})$, we let $f_e : [\underline{e}, \overline{e}] \mapsto \mathbb{R}$ denote the restriction of f to e, interpreted as a function of the interval $[\underline{e}, \overline{e}]$. If f_e has a derivative Lebesgue almost everywhere, we denote this by f_e'; recall that the existence of f_e' is equivalent to absolute continuity of f_e.

Definition 2.5. *For a graph with Euclidean edges $\mathcal{G}$ and an arbitrarily chosen origin $u_o \in \mathcal{V}(\mathcal{G})$, let $\mathcal{F}$ be the class of functions $f : \mathcal{G} \mapsto \mathbb{R}$ which are continuous with respect to $d_{G,\mathcal{G}}$ and for all $e \in \mathcal{E}(\mathcal{G})$, f_e is absolutely continuous and $f_e' \in L^2([\underline{e}, \overline{e}])$. In addition, define the following quadratic form on $\mathcal{F}$:*

$$\langle f, g \rangle_{\mathcal{F}} := f(u_o)g(u_o) + \sum_{e \in \mathcal{E}(\mathcal{G})} \int_{\underline{e}}^{\overline{e}} f_e'(t) g_e'(t) \, \mathrm{d}t. \tag{2.15}$$

Theorem 2.6. *For a graph with Euclidean edges $\mathcal{G}$, the space $(\mathcal{F}, \langle \cdot\,, \cdot \rangle_{\mathcal{F}})$ is an infinite dimensional Hilbert space with reproducing kernel $R_{\mathcal{G}}(u,v)$ (given in (2.13)). Moreover, for $u, v \in \mathcal{G}$,*

$$\begin{aligned} d_{R,\mathcal{G}}(u,v) &= R_{\mathcal{G}}(u,u) + R_{\mathcal{G}}(v,v) - 2R_{\mathcal{G}}(u,v) \\ &= \sup_{f \in \mathcal{F}} \left\{ (f(u) - f(v))^2 : \|f\|_{\mathcal{F}} \le 1 \right\}. \end{aligned} \tag{2.16}$$

3 Hilbert space embedding of $d_{G,\mathcal{G}}$ and $d_{R,\mathcal{G}}$

This section proves the key Hilbert space embedding result given in Theorem 1.4. For this we first need to recall a theorem by Schoenberg (1935, 1938a) on relating Hilbert spaces and positive definite functions and establish a new theorem on embedding 1-sums of distance spaces. The exposition of both of these results are kept as general as possible, since they hold for arbitrary distance spaces.

Recall that (X, d) is called a *distance space* if $d(x, y)$ for $x, y \in X$ is a *distance* on X, i.e., d satisfies all the requirements of a metric with the possible exception of the triangle inequality. Let $\mathrm{Range}(X, d) = \{d(x, y) : x, y \in X\}$.

Definition 3.1. *Let (X, d) be a distance space and $g : \mathrm{Range}(X, d) \mapsto [0, \infty)$ a function. Then (X, d) is said to have a g-*embedding *into a* Hilbert space *$(H, \|\cdot\|_H)$, denoted $(X, d) \overset{g}{\hookrightarrow} H$, if there exists a map $\varphi\colon X \mapsto H$ which satisfies*

$$g(d(x, y)) = \|\varphi(x) - \varphi(y)\|_H$$

for all $x, y \in X$. The special case when g is the identity map is denoted simply $(X, d) \hookrightarrow H$.

The following fundamental theorem shows the connection between Hilbert space embeddings and positive semi-definite functions; it follows from Schoenberg (1935, 1938a). This turns out to be an extremely useful tool, both for constructing positive semi-definite functions and for proving the existence of Hilbert space embeddings.

Theorem 3.2. *Let (X, d) be a distance space and x_0 an arbitrary member of X. The following statements are equivalent.*

(I) *$(X, d) \hookrightarrow H$ for some Hilbert space H.*

(II) *$d(x, x_0)^2 + d(y, x_0)^2 - d(x, y)^2$ is positive semi-definite over $x, y \in X$.*

(III) *For every $\beta > 0$, the function $\exp(-\beta d(x, y)^2)$ is positive semi-definite over $x, y \in X$.*

(IV) *The inequality $\sum_{k,j=1}^{n} c_k c_j d(x_k, x_j)^2 \le 0$ holds for every $x_1, \ldots, x_n \in X$ and $c_1, \ldots, c_n \in \mathbb{R}$ for which $\sum_{k=1}^{n} c_k = 0$.*

It is common (in Wells and Williams (1975) for example) to call any distance space (X, d) which satisfies condition (IV) a *distance of negative type*. In the geo-statistical literature, however, if d satisfies (IV), then d^2 is said to be a *generalized covariance function of order* 0. In particular, for any random field Z, condition (IV) is a necessary property of the variogram $d(u, v)^2 = \mathrm{var}(Z(u) - Z(v))$.

The last concept needed to show Theorem 1.4 deals with the notion of the 1-sum of two distance spaces (Deza and Laurent, 1997). This operation allows one to construct new distance spaces (which are root-embeddable) by stitching multiple root-embeddable distance spaces together.

Definition 3.3. *Suppose (X_1, d_1) and (X_2, d_2) are two distance spaces such that $X_1 \cap X_2 = \{x_0\}$. Then the* 1-sum *of (X_1, d_1) and (X_2, d_2) is the distance space $(X_1 \cup X_2, d)$ defined by*

$$d(x, y) = \begin{cases} d_1(x, y) & \text{if } x, y \in X_1, \\ d_2(x, y) & \text{if } x, y \in X_2, \\ d_1(x, x_0) + d_2(x_0, y) & \text{if } x \in X_1 \text{ and } y \in X_2. \end{cases} \tag{3.1}$$

The key fact about 1-sums, for our purposes, is summarized in the following theorem, which is verified in Appendix A.3.

Theorem 3.4. *Suppose* (X_1, d_1) *and* (X_2, d_2) *are two distance spaces such that* $X_1 \cap X_2 = \{x_0\}$*. If* $(X_1, d_1) \overset{\sqrt{}}{\hookrightarrow} H_1$ *and* $(X_2, d_2) \overset{\sqrt{}}{\hookrightarrow} H_2$ *for two Hilbert spaces* H_1 *and* H_2*, then there exists a Hilbert space* H *such that* $(X_1 \cup X_2, d) \overset{\sqrt{}}{\hookrightarrow} H$ *where* $(X_1 \cup X_2, d)$ *is the* 1-sum *of* (X_1, d_1) *and* (X_2, d_2)*.*

We are now ready to prove Theorem 1.4, which is our key result for verifying Theorem 1.2 and 1.3.

Proof of Theorem 1.4. Suppose $\mathcal{G}$ is a graph with Euclidean edges. By (2.2), we trivially have

$$d(u, v)^2 = \mathrm{var}(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v))$$

where $d(u, v) := \sqrt{d_{R,\mathcal{G}}(u, v)}$. The fact that $d(u, v)^2$ is a variogram implies that Theorem 3.2 (IV) holds. Since (I) $\Leftrightarrow$ (IV) we have that $(\mathcal{G}, d) \hookrightarrow H$ for some Hilbert space H, and hence $(\mathcal{G}, d_{R,\mathcal{G}}) \overset{\sqrt{}}{\hookrightarrow} H$ as was to be shown.

For the geodesic metric, first assume $\mathcal{G}$ forms a tree graph. In this case, Theorem 2.4 implies $d_{G,\mathcal{G}} = d_{R,\mathcal{G}}$ and therefore $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\sqrt{}}{\hookrightarrow} H$ by the corresponding result for $d_{R,\mathcal{G}}$. Second, assume $\mathcal{G}$ forms a cycle graph (such as the left two graphs of Figure 3). For some constant $\lambda > 0$, there clearly exists a metric isometry between $(\mathcal{G}, \lambda d_{G,\mathcal{G}})$ and the unit circle $\mathbb{S}^1$ equipped with the great circle metric $d_{\mathbb{S}^1}$. Since $\exp(-\beta d_{\mathbb{S}^1}(x, y))$ is positive semi-definite over $\mathbb{S}^1 \times \mathbb{S}^1$ for all $\beta > 0$ (see Gneiting (2013), for example), the function $\exp(-\beta d_G(u, v))$ is positive semi-definite over $\mathcal{G} \times \mathcal{G}$ for all $\beta > 0$. Now, setting $d := \sqrt{d_{G,\mathcal{G}}}$, the equivalence (I) $\Leftrightarrow$ (III) in Theorem 3.2 implies $(\mathcal{G}, d) \hookrightarrow H$, hence $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\sqrt{}}{\hookrightarrow} H$ for some Hilbert space H. Finally, for the general result where $\mathcal{G}$ is a 1-sum of cycles and trees, we simply use Theorem 3.4 to conclude that $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\sqrt{}}{\hookrightarrow} H$ for some Hilbert space H. $\square$

4 Parametric families of isotropic covariance functions with respect to $d_{G,\mathcal{G}}$ and $d_{R,\mathcal{G}}$

In some sense Theorem 1.2 follows easily from the results presented in the previous section and the seminal work of Schoenberg and von Neumann (Schoenberg, 1938a,b; von Neumann and Schoenberg, 1941) connecting metric embeddings, Hilbert spaces and completely monotonic functions, but we review the necessary results for completeness. First, it will be convenient to adopt the following definitions.

Definition 4.1. *Suppose that* (X, d) *is a distance space. Then a function* $C\colon \mathrm{Range}(X, d) \mapsto \mathbb{R}$ *is said to be a* p.s.d. respective p.d. radial function on (X, d) *if* $C(d(x, y))\colon X \times X \mapsto \mathbb{R}$ *is positive semi-definite respective (strictly) positive definite.*

Definition 4.2. *A function* $f : [0, \infty) \mapsto \mathbb{R}$ *is said to be* completely monotonic on $[0, \infty)$ *if* f *is continuous on* $[0, \infty)$*, infinitely differentiable on* $(0, \infty)$ *and* $(-1)^j f^{(j)}(t) \geq 0$ *over* $(0, \infty)$ *for every integer* $j \geq 0$*, where* $f^{(j)}$ *denotes the* j*th derivative of* f *and* $f^{(0)} = f$*.*

There is a distinction, in the literature, between complete monotonicity on $[0, \infty)$ versus on $(0, \infty)$, the latter being fundamentally related to Bernstein functions and variograms (see Berg (2008); Wells and Williams (1975)). The following result shows that completely monotonic functions on $[0,\infty)$ are characterized as positive mixtures of scaled exponentials (see Theorem 3 and $3'$ in Schoenberg (1938b), for example).

Theorem 4.3. *The completely monotonic functions on $[0,\infty)$ are precisely those which admit a representation $f(t) = \int_0^\infty e^{-t\sigma}\,\mathrm{d}\mu(\sigma)$, where μ is a finite positive measure on $[0,\infty)$. Moreover, if f is completely monotonic on $[0,\infty)$ and $f'(t)$ is not a constant function of t on $(0,\infty)$, then $[0,\infty)\ni r\mapsto f(r^2)$ is a p.d. radial function over any Hilbert space.*

The following corollary is the main tool when verifying Theorem 1.2.

Corollary 4.4. *Let (X,d) be a distance space which satisfies*

$$(X,d)\overset{\surd}{\hookrightarrow} H$$

for some Hilbert space H. Then C is a p.s.d. radial function on (X,d) whenever C is completely monotonic on $[0,\infty)$. If, in addition, $C'(t)$ is a non-constant function of t over $(0,\infty)$ and for all $x,y\in X$, $d(x,y)=0$ if and only if $x=y$, then C is a p.d. radial function over (X,d).

Proof. Since $(X,d)\overset{\surd}{\hookrightarrow} H$ for some Hilbert space H, there exists a map $\varphi\colon X\mapsto H$ for which $d(x,y) = \|\varphi(x)-\varphi(y)\|_H^2$ for all $x,y\in X$. Assume C is completely monotonic on $[0,\infty)$. Then

$$C(d(x,y)) = \int_0^\infty \exp\big(-\sigma\|\varphi(x)-\varphi(y)\|_H^2\big)\,\mathrm{d}\mu(\sigma)$$

where μ is a finite (positive) measure on $(0,\infty)$. By a fundamental fact established in von Neumann and Schoenberg (1941), $\int_0^\infty \exp(-\sigma t^2)\,\mathrm{d}\mu(\sigma)$ is a p.s.d. radial function on any infinite dimensional Hilbert space (since μ is a finite positive measure on $(0,\infty)$). This clearly implies that $C(d(x,y))$ is positive semi-definite. To finish notice that if $d(x,y)=0\Leftrightarrow x=y$ and if $C'(t)$ is not a constant function of $t\in(0,\infty)$, then Theorem 4.3 implies C is a p.d. radial function. $\square$

Now, we turn to the proofs of Theorem 1.2 and 1.3.

Proof of Theorem 1.2. Suppose $\mathcal{G}$ is a graph with Euclidean edges and let C be any radial function from the parametric classes given in Table 1. Notice that $C'(t)<0$ for all $t>0$ and $d_{G,\mathcal{G}}(u,v)=0\Leftrightarrow u=v$ (similarly for $d_{R,\mathcal{G}}$). By Theorem 1.4 and Corollary 4.4, it will therefore be sufficient to show that C is completely monotonic.

Note that $t\mapsto f(\beta t^\alpha+\lambda)$ is completely monotonic on $[0,\infty)$ whenever f is completely monotonic on $[0,\infty)$, $\alpha\in(0,1]$, and $\beta,\lambda>0$ (see equation (1.6) in Miller and Samko (2001)). Therefore, $\exp(-\beta t^\alpha)$ and $(\beta t^\alpha+1)^{-\xi/\alpha}$ are completely monotonic on $[0,\infty)$ for $\beta,\xi>0$ and $\alpha\in(0,1]$, since both $\exp(-t)$ and $(t+1)^{-\xi/\alpha}$ are completely monotonic. This establishes the desired result for the power exponential class and the generalized Cauchy class in Table 1. For the Matérn class notice that equation 3.479 on page 365 of Gradshteyn and Ryzhik (2007) gives

$$\frac{\sqrt{\pi}}{2^{\frac{1}{2}+\nu}\Gamma(\nu)}\int_0^\infty \frac{\beta^{\nu-1}\exp\big(-t\sqrt{1+\beta}\big)}{\sqrt{1+\beta}}\,\mathrm{d}\beta = t^{\frac{1}{2}-\nu}K_{\frac{1}{2}-\nu}(t)$$

for all $\nu, t > 0$. This implies

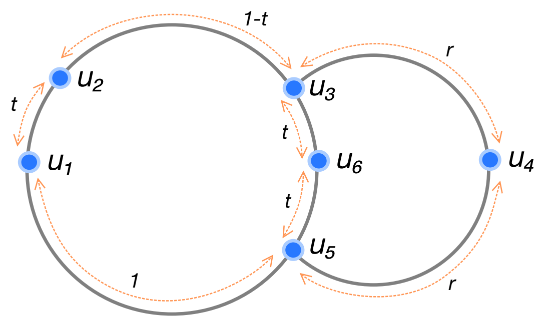
$$t^\alpha K_\alpha(t) = \frac{2^\alpha\sqrt{\pi}}{\Gamma\left(\frac{1}{2}-\alpha\right)} \int_1^\infty \left(\sigma^2-1\right)^{-\nu-1} \exp(-t\sigma)\,\mathrm{d}\sigma$$

which is completely monotonic on $[0,\infty)$ when $\alpha \in (0, \frac{1}{2})$, cf. Theorem 4.3. The boundary case $\alpha = \frac{1}{2}$ is resolved by noticing $\lim_{\alpha\uparrow\frac{1}{2}} t^\alpha K_\alpha(t) = \sigma e^{-\beta t}$ for some $\sigma, \beta > 0$. This establishes the desired result for the Matérn class in Table 1. Finally, Theorem 9 in Berg et al. (2008) establishes that $C(t) = 1 - \left(t^\beta/(1+t^\beta)\right)^\gamma$ is completely monotonic whenever $\beta\gamma \in (0,1]$ and $\beta \in (0,1]$, which establishes the desired result for the Dagum class in Table 1. □

Proof of Theorem 1.3. Let the situation be as in Theorem 1.3. By scaling $d_{G,\mathcal{G}}$ and possibly selecting new vertices on $\mathcal{G}$ by edge splitting operations (Theorem 2.3), we can obtain 6 vertices $\{u_1,\dots,u_6\}$ with edge lengths as given by the following geodesic pairwise distance matrix where $0 < t \le \frac{1}{2}$ and $0 < r \le 1$:

$$\{d_{G,\mathcal{G}}(u_i,u_j)\}_{i,j=1}^6 = \begin{pmatrix} 0 & t & 1 & r+1 & 1 & t+1 \\ t & 0 & 1-t & r-t+1 & t+1 & 1 \\ 1 & 1-t & 0 & r & 2t & t \\ r+1 & r-t+1 & r & 0 & r & r+t \\ 1 & t+1 & 2t & r & 0 & t \\ t+1 & 1 & t & r+t & t & 0 \end{pmatrix}$$

corresponding to the following subgraph:



Forming the matrix $\Sigma = \frac{1}{2}\{d_{G,\mathcal{G}}(u_i,u_1) + d_{G,\mathcal{G}}(u_1,u_j) - d_{G,\mathcal{G}}(u_i,u_j)\}_{i,j=2}^6$ gives

$$\Sigma = \begin{pmatrix} t & t & t & 0 & t \\ t & 1 & 1 & 1-t & 1 \\ t & 1 & r+1 & 1 & 1 \\ 0 & 1-t & 1 & 1 & 1 \\ t & 1 & 1 & 1 & t+1 \end{pmatrix}.$$

Setting $\boldsymbol{\xi} = (-1, -\xi, \xi, -1, 1)^T$, we have $\boldsymbol{\xi}^T\Sigma\boldsymbol{\xi} = \xi(r\xi - 2t)$, so $\boldsymbol{\xi}^T\Sigma\boldsymbol{\xi} < 0$ when $0 < \xi < 2t/r$, implying that $c(u_i,u_j) = \frac{1}{2}(d_{G,\mathcal{G}}(u_i,u_1) + d_{G,\mathcal{G}}(u_j,u_1) - d_{G,\mathcal{G}}(u_i,u_j))$ is not positive semi-definite over $\{u_1,\dots,u_6\}$. Then Theorem 3.2 gives the existence of a $\beta > 0$ such that $\exp(-\beta d_{G,\mathcal{G}}(u,v))$ is *not* positive semi-definite over $\{u_1,\dots,u_6\}$. □

5 Restricted covariance function properties

The restriction on the parameter α in Table 1 agrees with results for similar families of covariance functions for isotropic random fields on the d-dimensional sphere $\mathbb{S}^d$ (Gneiting, 2013). This may be no surprise, since a Euclidean cycle is similar to the circle $\mathbb{S}^1$. Below Corollary 5.2 shows that the restriction is in general also needed when considering a Euclidean tree $\mathcal{G}$, noting that if $\mathcal{G}$ has maximum degree $n < \infty$, then $\mathcal{G}$ has a star-shaped subgraph with $n+1$ vertices and n edges. Moreover, Corollary 5.4 shows that there are some quite severe limitations on the kind of covariance function that are valid for arbitrary Euclidean trees (and thus also arbitrary graphs with Euclidean edges).

In the following we only consider Euclidean trees. Then, by Theorem 2.4, $d_{G,\mathcal{G}} = d_{R,\mathcal{G}}$ and we use $d_{\cdot,\mathcal{G}}$ as a common notation for the two metrics.

Theorem 5.1. *If Z is a random field on a Euclidean tree $\mathcal{G}$ which contains a star-shaped tree subgraph $\mathcal{S}_n$ with $n+1$ vertices and n edges, and $\tilde{\alpha}, \tilde{\beta} > 0$ are numbers so that* $\mathrm{var}(Z_n(u) - Z_n(v)) = \tilde{\beta} d_{\cdot,\mathcal{G}}(u,v)^{\tilde{\alpha}} + o(d_{\cdot,\mathcal{G}}(u,v))$ *when $d_{\cdot,\mathcal{G}}(u,v) \to 0$, then $\tilde{\alpha} \leq \log(2n/(n-1))/\log(2)$.*

Proof. Let u_0 be the vertex in $\mathcal{S}_n$ with degree n and consider the variogram $d(u,v)^2 = \mathrm{var}(Z(u) - Z(v))$. By Theorem 3.2, $C(u,v) = d(u,u_0)^2 + d(v,u_0)^2 - d(u,v)^2$ is positive semi-definite over $\mathcal{S}_n \times \mathcal{S}_n$. For $i = 1,\dots,n$ and $\epsilon > 0$ sufficiently small, let $u_{i,\epsilon}$ be the point on the i^{th} edge which has $d_{\cdot,\mathcal{G}}$ distance ϵ from u_0. The assumption on $\mathrm{var}(Z(u) - Z(v))$ implies

$$\begin{aligned} d(u_{i,\epsilon}, u_0)^2 &= \tilde{\beta} d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_0)^{\tilde{\alpha}} + o\left(d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_0)\right) = \tilde{\beta}\epsilon^{\tilde{\alpha}} + o\left(\epsilon^{\tilde{\alpha}}\right) \\ d(u_{i,\epsilon}, u_{j,\epsilon})^2 &= \tilde{\beta} d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_{j,\epsilon})^{\tilde{\alpha}} + o\left(d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_{j,\epsilon})\right) = \tilde{\beta} 2^{\tilde{\alpha}}\epsilon^{\tilde{\alpha}} + o\left(\epsilon^{\tilde{\alpha}}\right) \end{aligned}$$

when $i \neq j$. Let Σ_ϵ be the $n \times n$ covariance matrix with $(i,j)^{\text{th}}$ entry

$$(\Sigma_\epsilon)_{i,j} = C(u_{i,\epsilon}, u_{j,\epsilon}) = \tilde{\beta}(2 - 2^{\tilde{\alpha}})\epsilon^{\tilde{\alpha}} + \tilde{\beta} 2^{\tilde{\alpha}} \epsilon^{\tilde{\alpha}} \delta_{ij} + o(\epsilon^{\tilde{\alpha}}),$$

then

$$\Sigma_\epsilon = \tilde{\beta}(2 - 2^{\tilde{\alpha}})\epsilon^{\tilde{\alpha}}\, 1_n\, 1_n^T + \tilde{\beta} 2^{\tilde{\alpha}} \epsilon^{\tilde{\alpha}} I_n + o(\epsilon^{\tilde{\alpha}}) A$$

where I_n is the $n \times n$ identity matrix, 1_n is the vector of length n with each coordinate equal to 1, and A is some $n \times n$ matrix not depending on ϵ. Now,

$$\begin{aligned} 0 \leq \det(\Sigma_\epsilon) &= \tilde{\beta}^n 2^{n\tilde{\alpha}} \epsilon^{n\tilde{\alpha}} \det\left((2^{1-\tilde{\alpha}} - 1)\, 1_n\, 1_n^T + I_n + o(1)A\right) \\ &= \tilde{\beta}^n 2^{n\tilde{\alpha}} \epsilon^{n\tilde{\alpha}} \left((2^{1-\tilde{\alpha}} - 1)n + 1\right) + o(\epsilon^{n\tilde{\alpha}}). \end{aligned}$$

Consequently $(2^{1-\tilde{\alpha}} - 1)n + 1 \geq 0$ as was to be shown. □

Corollary 5.2. *Let C be one of the functions given in Table 1 but with α outside the parameter range, i.e., $\alpha > \frac{1}{2}$ in case of the Matérn class and $\alpha > 1$ in case of the other three classes. Then there exists a Euclidean tree $\mathcal{G}$ so that $C(d_{\cdot,\mathcal{G}}(u,v))$ is not a covariance function.*

Proof. Suppose Z_n is a random field on a Euclidean tree $\mathcal{G}$ which contains a star-shaped graph $\mathcal{S}_n$ with $n+1$ vertices and n edges, with an isotropic covariance function with radial profile C and $\alpha > 0$.

If C is in the Matérn class, let

$$\tilde{\alpha} = \alpha + 1 - |\alpha - 1|, \qquad \tilde{\beta} = \frac{\beta^{\alpha+1-|\alpha-1|}\Gamma(|\alpha-1|)2^{|\alpha-1|-\alpha}}{\tilde{\alpha}\Gamma(\alpha)}.$$

By L'Hospital's Rule, equation 24.56 in Spiegel (1968) and equation 9.6.9 in Abramowitz and Stegun (1964),

$$\begin{aligned}
&\lim_{d_{\cdot,\mathcal{G}}(u,v)\to 0} \frac{\operatorname{var}(Z(u)-Z(v))}{d_{\cdot,\mathcal{G}}(u,v)^{\tilde{\alpha}}} \\
&\quad = \lim_{d_{\cdot,\mathcal{G}}(u,v)\to 0} \frac{2(1-\frac{1}{\Gamma(\alpha)2^{\alpha-1}}(\beta d_{\cdot,\mathcal{G}}(u,v))^{\alpha}K_\alpha(\beta d_{\cdot,\mathcal{G}}(u,v))}{d_{\cdot,\mathcal{G}}(u,v)^{\tilde{\alpha}}} \\
&\quad = \lim_{d_{\cdot,\mathcal{G}}(u,v)\to 0} \frac{\beta^{\alpha+1}}{\Gamma(\alpha)2^{\alpha-1}\tilde{\alpha}} d_{\cdot,\mathcal{G}}(u,v)^{\alpha-\tilde{\alpha}+1}K_{\alpha-1}(\beta d_{\cdot,\mathcal{G}}(u,v)) \\
&\quad = \lim_{d_{\cdot,\mathcal{G}}(u,v)\to 0} \frac{\beta^{\alpha+1-|\alpha-1|}\Gamma(|\alpha-1|)2^{|\alpha-1|-\alpha}}{\tilde{\alpha}\Gamma(\alpha)} d_{\cdot,\mathcal{G}}(u,v)^{\alpha-\tilde{\alpha}+1-|\alpha-1|} = \tilde{\beta}.
\end{aligned}$$

Hence Theorem 5.1 applies, and letting $n \to \infty$ we obtain $\tilde{\alpha} \le 1$ or equivalently $\alpha \le \frac{1}{2}$, thus proving the assertion.

If C is in one of the other three classes, it follows directly from L'Hospital's Rule that the requirement of the variogram $\operatorname{var}(Z_n(u)-Z_n(v))$ of Theorem 5.1 is satisfied for $\tilde{\alpha} = \alpha$ when $\tilde{\beta} = 2\beta$ in case of the power exponential class and $\tilde{\beta} = 2\beta\xi/\alpha$ in case of the generalized Cauchy or the Dagum class. Letting $n \to \infty$ in Theorem 5.1, we get that $\alpha \le 1$, thus completing the proof. □

Theorem 5.3. *Suppose $(u,v) \mapsto C(d_{\cdot,\mathcal{G}}(u,v))$ is a covariance function on a Euclidean tree $\mathcal{G}$ containing a star-shaped tree subgraph $\mathcal{S}_n$ with $n \ge 2$ edges of length larger than or equal to $t_0 > 0$. For all $t \in (0, t_0]$, we have*

$$-\frac{C(0)}{n-1} \le C(2t) \le C(0), \qquad \frac{nC(t)^2 - C(0)^2}{n-1} \le C(0)C(2t). \tag{5.1}$$

Proof. Denote $e_1, \ldots, e_n$ the edges of $\mathcal{S}_n$, and u_{n+1} their common vertex. Let $t \in (0, t_0)$ and $u_i \in e_i$ such that $d_{\cdot,\mathcal{G}}(u_{n+1}, u_i) = t$ for $i = 1, \ldots, n$. Note that $d_{\cdot,\mathcal{G}}(u_i, u_j) = 2t$ for $i, j = 1, \ldots, n$ and $i \ne j$. Let Σ denote the $(n+1) \times (n+1)$ matrix with the $(i,j)^{\text{th}}$ entry equal to $C(d_{\cdot,\mathcal{G}}(u_i, u_j))$, i.e.,

$$\Sigma_{i,j} = \begin{cases} C(0) & \text{if } i = j, \\ C(2t) & \text{if } i \ne j \text{ and } i, j < n+1, \\ C(t) & \text{otherwise.} \end{cases}$$

As Σ is a covariance matrix, its principal minors are non-negative determinants; these are of the form $\det(\Sigma_k)$ with $k \in \{1, \ldots, n\}$ or $\det(\Sigma'_k)$ with $k \in \{2, \ldots, n+1\}$; here, Σ_k denotes a $k \times k$ submatrix of Σ with the same rows and columns removed and where the $(n+1)^{\text{th}}$ row and column have been removed; and Σ'_k is defined in a similar way but where the $(n+1)^{\text{th}}$ row and column have not been removed. It is easily verified that

$$\det(\Sigma_k) = (C(0) - C(2t))^{k-1} \{(k-1)C(2t) + C(0)\}$$

for $k = 1, \ldots, n$, and hence either $0 \leq C(0) = C(2t)$ or both $C(0) > C(2t)$ and $(k-1)C(2t) + C(0) \geq 0$, implying the first inequality in (5.1), where we have let $k = n$ to obtain the highest lower bound. Moreover,

$$\det(\Sigma'_k) = \{C(0) - C(2t)\}^{k-2} \left\{C(0)^2 + (k-2)C(2t)C(0) - (k-1)C(t)^2\right\}$$

for $k = 2, \ldots, n+1$, and so either $|C(t)| \leq C(0) = C(2t)$ or both $C(0) > C(2t)$ and $C(0)^2 + (k-2)C(2t)C(0) - (k-1)C(t)^2 \geq 0$, implying the second inequality in (5.1), where we have used $k = n+1$ to get the highest lower bound. □

Corollary 5.4. *A function $(u, v) \mapsto C(d_{\cdot,\mathcal{G}}(u, v))$ which is a covariance function on all Euclidean trees has to be non-negative, and furthermore either have unbounded support or fulfill $C(t) = 0$ for all $t > 0$.*

Proof. Letting $n \to \infty$ and $t_0 \to \infty$, the first inequality in (5.1) implies non-negativity of C, and the second inequality in (5.1) implies $C(0)C(2t) \geq C(t)^2$ for all $t > 0$. Thus, if for some $t_1 > 0$ and all $t > t_1$ we have $C(2t) = 0$, then $C(t) = 0$ for all $t > t_1$, from which it follows by induction that $C(t) = 0$ for all $t > 0$. □

As an example of the severity of the restrictions, consider the "bounded linear model" for a variogram suggested on page 205 in Okabe and Sugihara (2012) in connection to linear networks. If there is a corresponding isotropic covariance function, its radial profile is given by

$$C(t) = \begin{cases} \beta_0 + \beta_1\beta_2 & \text{if } t = 0, \\ \beta_1(\beta_2 - t) & \text{if } 0 < t \leq \beta_2 \\ 0 & \text{if } t > \beta_2, \end{cases}$$

where $\beta_0, \beta_1, \beta_2 > 0$ are parameters. As this function has bounded support, by Corollary 5.4 this cannot be a valid covariance function on an arbitrary graph with Euclidean edges (or an arbitrary linear network). Of course this does not imply that the variogram for the bounded linear model is invalid, but numerical calculations show that the variogram is in general not conditionally negative semi-definite in the case of most star-shaped graphs. Therefore we believe it is not a proper variogram for an arbitrary graph with Euclidean edges.

A Proofs

A.1 Proof of Theorem 2.6

To verify Theorem 2.6, recall Definition 2.5. We use the notation I_A for an indicator function which is 1 on a set A and 0 otherwise. We need the following lemmas.

Lemma A.1. *$(\mathcal{F}, \langle \cdot\,, \cdot \rangle_{\mathcal{F}})$ is an inner product vector space, with metric $\|f\|_{\mathcal{F}} := \sqrt{\langle f, f\rangle_{\mathcal{F}}}$ given by*

$$\|f\|_{\mathcal{F}}^2 = f(u_o)^2 + \sum_{e\in\mathcal{E}(\mathcal{G})} \int_{\underline{e}}^{\overline{e}} f_e'(t)^2 \,\mathrm{d}t. \tag{A.1}$$

Proof. From (2.15) we obtain (A.1). Note that $\langle f, f\rangle_{\mathcal{F}} = 0$ implies both $f(u_o) = 0$ and for any $e \in \mathcal{E}(\mathcal{G})$, f_e is almost everywhere constant on e. The continuity requirement of $f \in \mathcal{F}$ then implies $\langle f, f\rangle_{\mathcal{F}} = 0 \Leftrightarrow f = 0$. Finally, $\langle \cdot\,, \cdot \rangle_{\mathcal{F}}$ is clearly symmetric, bilinear and positive semi-definite over $f \in \mathcal{F}$. □

For $f \in \mathcal{F}$, $u \in \mathcal{G}$ and $e \in \mathcal{E}(\mathcal{G})$, define

$$f_\mu(u) = (1 - d(u))f(\underline{u}) + d(u)f(\overline{u}), \qquad f_{e,r}(u) = \begin{cases} f(u) - f_\mu(u) & \text{if } u \in e, \\ 0 & \text{otherwise,} \end{cases}$$

where $d(u)$ is defined in (2.9). It will be convenient to denote the operations $f \mapsto f_\mu$ and $f \mapsto f_{e,r}$ with operator notation $\mathscr{P}_\mu : \mathcal{F} \mapsto \mathcal{F}$ and $\mathscr{P}_e : \mathcal{F} \mapsto \mathcal{F}$ given by $\mathscr{P}_\mu f = f_\mu$ and $\mathscr{P}_e f = f_{e,r}$. In addition, the inner product $\langle \cdot\,, \cdot \rangle_{\mathcal{F}}$ restricted to the function spaces $\mathscr{P}_\mu\mathcal{F}$ and $\mathscr{P}_e\mathcal{F}$ will be denoted $\langle \cdot\,, \cdot \rangle_\mu = \langle \cdot\,, \cdot \rangle_{\mathcal{F}}\big|_{\mathscr{P}_\mu\mathcal{F}\times\mathscr{P}_\mu\mathcal{F}}$ and $\langle \cdot\,, \cdot \rangle_{e,r} = \langle \cdot\,, \cdot \rangle_{\mathcal{F}}\big|_{\mathscr{P}_e\mathcal{F}\times\mathscr{P}_e\mathcal{F}}$.

Lemma A.2. *If $\mathcal{G}$ is a graph with Euclidean edges, then for all $e \in \mathcal{E}(\mathcal{G})$, $\mathscr{P}_\mu$ and $\mathscr{P}_e$ are mutually orthogonal projections and $\mathcal{F}$ is a direct sum:*

$$\mathcal{F} = \mathscr{P}_\mu\mathcal{F} \oplus \bigoplus_{e\in\mathcal{E}(\mathcal{G})} \mathscr{P}_e\mathcal{F}. \tag{A.2}$$

Proof. This is straightforwardly verified as soon as it is noted that $\mathscr{P}_\mu$ and $\mathscr{P}_e$ are self-adjoint operators, which follows from the fact that

$$\big[(f_\mu)_e\big]'(t) = \frac{f_e(\overline{e}) - f_e(\underline{e})}{\mathrm{len}(e)}, \qquad e \in \mathcal{E}(\mathcal{G}), t \in (\underline{e}, \overline{e}). \tag{A.3}$$

□

Lemma A.3. *Let $\mathcal{G}$ be a graph with Euclidean edges with vertices $\mathcal{V}$ and edges $\mathcal{E}$. Also let*

- *$(\mathbb{R}^{\mathcal{V}}, \langle \cdot\,, \cdot \rangle_L)$ denote the finite dimensional Hilbert space with inner product given by $\langle z, w\rangle_L = z^T L w$ as in* (2.6)*;*
- *H_e denote the infinite dimensional Hilbert space of absolutely continuous functions $f : [\underline{e}, \overline{e}] \mapsto \mathbb{R}$ such that $f' \in L^2([\underline{e}, \overline{e}])$ with boundary condition $f(\underline{e}) = f(\overline{e}) = 0$, and with inner product $\langle f, g\rangle_{H_e} := \int_{\underline{e}}^{\overline{e}} f'(t)g'(t)\,\mathrm{d}t$.*

Then we have the following.

(A) *$(\mathscr{P}_\mu\mathcal{F}, \langle \cdot\,, \cdot \rangle_\mu)$ is a finite dimensional Hilbert space which is isomorphic to $(\mathbb{R}^{\mathcal{V}}, \langle \cdot\,, \cdot \rangle_L)$ and has reproducing kernel R_μ (defined in* (2.10)*). Its inner product has simplified form for all $f, g \in \mathscr{P}_\mu\mathcal{F}$:*

$$\langle f, g\rangle_\mu = f(u_o)\,g(u_o) + \sum_{e\in\mathcal{E}} \frac{(f_e(\overline{e}) - f_e(\underline{e}))\,(g_e(\overline{e}) - g_e(\underline{e}))}{\mathrm{len}(e)}. \tag{A.4}$$

(B) *For each $e \in \mathcal{E}$, $(\mathscr{P}_e\mathcal{F}, \langle \cdot\,, \cdot \rangle_{e,r})$ is an infinite dimensional Hilbert space which is isomorphic to $(H_e, \langle \cdot, \cdot \rangle_{H_e})$ and has reproducing kernel R_e (given by (2.12)). Its inner product has simplified form for all $f, g \in \mathscr{P}_e\mathcal{F}$:*

$$\langle f, g \rangle_{e,r} = \int_{\underline{e}}^{\overline{e}} f_e'(t) g_e'(t)\, \mathrm{d}t. \tag{A.5}$$

(C) *$(\mathcal{F}, \langle \cdot\,, \cdot \rangle_{\mathcal{F}})$ is an infinite dimensional Hilbert space which is isomorphic to $\mathbb{R}^{\mathcal{V}} \otimes \bigotimes_{e\in\mathcal{E}} H_e$ and has reproducing kernel $R_{\mathcal{G}}$ (given by (2.13)). Its inner product has simplified form for all $f, g \in \mathcal{F}$:*

$$\langle f, g \rangle_{\mathcal{F}} = \langle f_\mu, g_\mu \rangle_\mu + \sum_{e\in\mathcal{E}} \langle f_{e,r}, g_{e,r} \rangle_{e,r}. \tag{A.6}$$

Proof. (A): There is a bijective correspondence between $z \in \mathbb{R}^{\mathcal{V}}$ and $f_\mu \in \mathscr{P}_\mu\mathcal{F}$ which simply corresponds to interpreting z as the values of f_μ on the vertices of $\mathcal{G}$. Then

$$\begin{aligned} f_\mu(u) &= (1 - d(u))z(\underline{u}) + d(u)z(\overline{u}), && \forall u \in \mathcal{G} \\ z(v) &= f_\mu(v), && \forall v \in \mathcal{V}. \end{aligned}$$

The bijection also preserves inner product because if $w \in \mathbb{R}^{\mathcal{V}}$ corresponds to $g_\mu \in \mathscr{P}_\mu\mathcal{F}$, then

$$\begin{aligned} \langle z, w \rangle_L &= z(u_o)w(u_o) + \frac{1}{2}\sum_{u\sim v} \frac{(z(u) - z(v))(w(u) - w(v))}{d_{G,\mathcal{G}}(u, v)} && \text{(by (2.6))} \\ &= f_\mu(u_o)\, g_\mu(u_o) + \frac{1}{2}\sum_{u\sim v} \frac{(f_\mu(u) - f_\mu(v))\,(g_\mu(u) - g_\mu(v))}{\mathrm{len}(e)} \\ &= \langle f_\mu, g_\mu \rangle_\mu \qquad \text{(by (A.3))}, \end{aligned}$$

where the above sums are over adjacent $u, v \in \mathcal{V}$. This establishes (A.4) and that $(\mathscr{P}_\mu\mathcal{F}, \langle \cdot\,, \cdot \rangle_\mu)$ is isomorphic to $(\mathbb{R}^{\mathcal{V}}, \langle \cdot\,, \cdot \rangle_L)$. It then remains to show that the reproducing kernel of $(\mathbb{R}^{\mathcal{V}}, \langle \cdot\,, \cdot \rangle_L)$, namely L^{-1}, is in bijective correspondence with R_μ. Indeed, for each $u \in \mathcal{G}$,

$$\begin{aligned} f_\mu(u) &= [\mathscr{P}_\mu f_\mu](u) \qquad (\text{since } \mathscr{P}_\mu^2 = \mathscr{P}_\mu) \\ &= (1 - d(u))f_\mu(\underline{u}) + d(u)f_\mu(\overline{u}) \\ &= (1 - d(u))z(\underline{u}) + d(u)z(\overline{u}) \\ &= (1 - d(u))\langle z, L^{-1}(\,\cdot\,, \underline{u})\rangle_L + d(u)\langle z, L^{-1}(\,\cdot\,, \overline{u})\rangle_L \\ &= \langle z, \underbrace{(1 - d(u))L^{-1}(\,\cdot\,, \underline{u}) + d(u)L^{-1}(\,\cdot\,, \overline{u})}_{:=R_L(\cdot,u)} \rangle_L \end{aligned} \tag{A.7}$$

where the function $R_L(\,\cdot\,, u)$ is a member of $\mathbb{R}^{\mathcal{V}}$. By (2.10), we can simply linear interpolate $R_L(\,\cdot\,, u)$ to find the corresponding member in $\mathscr{P}_\mu\mathcal{F}$ as follows

$$(1 - d(\,\cdot\,))R_L(\underline{\cdot}, u) + d(\,\cdot\,)R_L(\overline{\cdot}, u) = R_\mu(\,\cdot\,, u).$$

Therefore, by (A.7),

$$f_\mu(u) = \langle z, R_L(\,\cdot\,, u)\rangle_L = \langle f_\mu, R_\mu(\,\cdot\,, u)\rangle_\mu$$

where the second equality follows by the fact that inner products are preserved under the bijective correspondence. This completes the proof of (A).

(B): Let $e \in \mathcal{E}$. Note that H_e is equal to the constrained space $\{f \in \mathscr{H}_e : f(\overline{e}) = 0\}$ where $\mathscr{H}_e := \big\{f \in C([\underline{e}, \overline{e}]) : f' \in L^2([\underline{e}, \overline{e}]),\, f(\underline{e}) = 0\big\}$ corresponds to the *Cameron-Martin* Hilbert space (using inner product $\langle \cdot, \cdot \rangle_{H_e}$) with reproducing kernel $(s-\underline{e}) \wedge (t-\underline{e})$. Therefore, by Saitoh (1997) page 77, the subspace $(H_e, \langle \cdot, \cdot \rangle_{H_e})$ is also a Hilbert space with reproducing kernel given by

$$\tilde{R}_e(s,t) := (s-\underline{e}) \wedge (t-\underline{e}) - \frac{(s-\underline{e})(t-\underline{e})}{\overline{e}-\underline{e}} \qquad \underline{e} < s, t < \overline{e}. \tag{A.8}$$

Clearly, $f \in H_e$ and $f_{e,r} \in \mathscr{P}_e\mathcal{F}$ are in a bijective linear correspondence by the relation $f_{e,r}(u) = f(\varphi_e(u))I_e(u)$. By (2.12),

$$R_e(u,v) = \begin{cases} \tilde{R}_e(\varphi_e(u), \varphi_e(v)) & \text{if } u, v \in e, \\ 0 & \text{otherwise.} \end{cases} \tag{A.9}$$

Finally, for $f_{e,r}, g_{e,r} \in \mathscr{P}_e\mathcal{F}$ with corresponding $f, g \in H_e$, we obtain $\big\langle f_{e,r}, g_{e,r} \big\rangle_{e,r} = \big\langle f, g \big\rangle_{H_e}$, and so

$$\big\langle f_{e,r}, R_e(\,\cdot\,, v) \big\rangle_{e,r} = \big\langle f, \tilde{R}_e(\,\cdot\,, \varphi_e(v)) \big\rangle_{H_e} = f(\varphi_e(v)) = f_{e,r}(v).$$

Thereby (B) is verified.

(C): This follows immediately from Lemma A.2 and (A)–(B). □

Proof of Theorem 2.6. As $R_{\mathcal{G}}$ is the covariance function of $Z_{\mathcal{G}}$,

$$\operatorname{var}(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v)) = R_{\mathcal{G}}(u,u) + R_{\mathcal{G}}(v,v) - 2R_{\mathcal{G}}(u,v). \tag{A.10}$$

By Lemma A.3, $R_{\mathcal{G}}$ also serves as the reproducing kernel for $(\mathcal{F}, \langle \cdot, \cdot \rangle_{\mathcal{F}})$. Therefore, $(\mathcal{F}, \langle \cdot, \cdot \rangle_{\mathcal{F}})$ is *the* RKHS associated to $Z_{\mathcal{G}}$ (see Wahba (1990)). Now, to finish the proof we use standard Hilbert space arguments to show

$$\operatorname{var}(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v)) = \sup_{f \in \mathcal{F}} \big\{(f(u) - f(v))^2 : \|f\|_{\mathcal{F}} \le 1\big\}. \tag{A.11}$$

Let $u, v \in \mathcal{G}$ with $u \neq v$, and $f \in \mathcal{F}$ with $\|f\|_{\mathcal{F}} \le 1$. Use the reproducing property of $R_{\mathcal{G}}$ along with the Cauchy-Schwartz inequality to obtain

$$\begin{aligned}(f(u) - f(v))^2 &= \big\langle f,\, R_{\mathcal{G}}(\,\cdot\,, u) - R_{\mathcal{G}}(\,\cdot\,, v) \big\rangle_{\mathcal{F}}^2 \le \big\| R_{\mathcal{G}}(\,\cdot\,, u) - R_{\mathcal{G}}(\,\cdot\,, v) \big\|_{\mathcal{F}}^2 \\ &= \operatorname{var}(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v))\end{aligned}$$

by (A.10). Note that the function $f_o \in \mathcal{F}$ defined by

$$f_o(w) = (R_{\mathcal{G}}(w,u) - R_{\mathcal{G}}(w,v)) / \|R_{\mathcal{G}}(\,\cdot\,, u) - R_{\mathcal{G}}(\,\cdot\,, v)\|_{\mathcal{F}}$$

has norm $\|f_o\|_{\mathcal{F}} = 1$ and satisfies

$$(f_o(u) - f_o(v))^2 = \operatorname{var}(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v)).$$

This proves (A.11) as was to be shown. □

A.2 Proofs of Theorem 2.2, 2.3, and 2.4

We start by verifying Theorem 2.3 as it is used to prove Theorem 2.2.

Proof of Theorem 2.3. Since the operation of merging two edges at a degree two vertex v is the inverse of splitting the resulting edge at v, it will be sufficient to show that $\mathcal{G}'$ is isometric to $\mathcal{G}$ under the resistance metric when $\mathcal{G}'$ is obtained from $\mathcal{G}$ by splitting edge $e \in \mathcal{E}(\mathcal{G})$ at $u \in e$.

Let e_1 and e_2 denote the partial edges formed by splitting $e \in \mathcal{E}(\mathcal{G})$ at $u \in e$ such that $\underline{e}_1 = \underline{e}$. Since the sets $\mathcal{G}$ and $\mathcal{G}'$ are identical, their corresponding spaces of functions $\mathcal{F}$ and $\mathcal{F}'$ as given in Definition 2.5 are identical, however, $\mathcal{G}$ and $\mathcal{G}'$ induce different inner products on $\mathcal{F}$, denoted $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathcal{F}}$ and $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathcal{F},\text{split}}$, respectively. By Theorem 2.6, $d_{R,\mathcal{G}}$ and $d_{R,\mathcal{G}'}$ are completely determined by their inner products $\langle f, g\rangle_{\mathcal{F},\text{split}}$ and $\langle f, g\rangle_{\mathcal{F}}$. Therefore, we may suppose that both $\mathcal{G}$ and $\mathcal{G}'$ use the same origin u_o in their respective inner products. Then, for any $f, g \in \mathcal{F}$, the difference between the two inner products is

$$\begin{aligned}&\langle f, g\rangle_{\mathcal{F}} - \langle f, g\rangle_{\mathcal{F},\text{split}}\\ &\quad= \int_{\underline{e}}^{\overline{e}} f_e'(t)g_e'(t)\,\mathrm{d}t - \int_{\underline{e}_1}^{\overline{e}_1} f_{e_1}'(t)g_{e_1}'(t)\,\mathrm{d}t - \int_{\underline{e}_2}^{\overline{e}_2} f_{e_2}'(t)g_{e_2}'(t)\,\mathrm{d}t\end{aligned}$$

since the splitting operation on $e \in \mathcal{E}(\mathcal{G})$ at $u \in e$ only affects the term corresponding to e in (2.15). By Theorem 2.6, to show $d_{R,\mathcal{G}} = d_{R,\mathcal{G}'}$, it will be sufficient to show $\langle f, g\rangle_{\mathcal{F}} = \langle f, g\rangle_{\mathcal{F},\text{split}}$ for all $f, g \in \mathcal{F}$.

For any $f \in \mathcal{F}$ and $t \in [\underline{e}, \overline{e}]$, define

$$f_1(t) = f_e(t)\, I_{[\underline{e}_1,\overline{e}_1)}(t), \qquad f_2(t) = f_e(t)\, I_{[\underline{e}_2,\overline{e}_2]}(t).$$

Both f_1 and f_2 are almost everywhere differentiable and satisfy $f_1', f_2' \in L^2([\underline{e}, \overline{e}])$. Moreover, for any $f, g \in \mathcal{F}$, the fact that $f_1'(t)g_2'(t) \overset{\text{a.e.}}{=} 0$ and $f_2'(t)g_1'(t) \overset{\text{a.e.}}{=} 0$ implies that

$$\begin{aligned}\int_{\underline{e}}^{\overline{e}} f_e'(t)g_e'(t)\,\mathrm{d}t &= \int_{\underline{e}}^{\overline{e}} \left[f_1'(t) + f_2'(t)\right]\left[g_1'(t) + g_2'(t)\right]\mathrm{d}t\\ &= \int_{\underline{e}_1}^{\overline{e}_1} f_1'(t)g_1'(t)\,\mathrm{d}t + \int_{\underline{e}_2}^{\overline{e}_2} f_2'(t)g_2'(t)\,\mathrm{d}t.\end{aligned}$$

Note that for Lebesgue almost all numbers t, $f_{e_1}'(t) = f_1'(t)$ if $t \in [\underline{e}_1, \overline{e}_1]$ and $f_{e_2}'(t) = f_2'(t)$ if $t \in [\underline{e}_2, \overline{e}_2]$ (and similarly for g_{e_1}, g_{e_2}). Therefore,

$$\int_{\underline{e}}^{\overline{e}} f_e'(t)g_e'(t)\,\mathrm{d}t = \int_{\underline{e}_1}^{\overline{e}_1} f_{e_1}'(t)g_{e_1}'(t)\,\mathrm{d}t + \int_{\underline{e}_2}^{\overline{e}_2} f_{e_2}'(t)g_{e_2}'(t)\,\mathrm{d}t$$

which implies $\langle f, g\rangle_{\mathcal{F},\text{split}} = \langle f, g\rangle_{\mathcal{F}}$ as was to be shown. $\square$

Proof of Theorem 2.2. In the literature on resistance networks and metrics (see e.g. Kigami, 2003; Jorgensen and Pearse, 2010), given a conductance function c (i.e., a symmetric function associated to all pairs of adjacent vertices), the (effective) resistance distance between $u, v \in \mathcal{V}(\mathcal{G})$ is defined by

$$d_{\text{eff}}(u,v) = \sup_{z \in \mathbb{R}^{\mathcal{V}(\mathcal{G})}} \left\{ (z(u) - z(v))^2 : \frac{1}{2} \sum_{u_1 \sim u_2} c(u_1,u_2)(z(u_1) - z(u_2))^2 \leq 1 \right\} \quad \text{(A.12)}$$

(this is one of several equivalent definitions, cf. Theorem 2.3 in Jorgensen and Pearse (2010)). To relate d_{eff} and $d_{R,\mathcal{G}}$, recall (2.16) and that we have defined c by (2.4). Also, by Lemma A.3, each $f \in \mathcal{F}$ has an orthogonal decomposition $f = f_\mu + \sum_{e \in \mathcal{E}(\mathcal{G})} f_{e,r}$ where

$$\|f\|_{\mathcal{F}}^2 = \|f_\mu\|_\mu^2 + \sum_{e \in \mathcal{E}(\mathcal{G})} \|f_{e,r}\|_{e,r}^2$$

and $f_{e,r}(u) = f_{e,r}(u) = 0$ for all $u, v \in \mathcal{V}(\mathcal{G})$. Therefore, if $u, v \in \mathcal{V}(\mathcal{G})$, the term $(f(u) - f(v))^2$ in (2.16) simplifies to $(f_\mu(u) - f_\mu(v))^2$ and hence the supremum can be taken over $f \in \mathcal{F}$ such that $\|f_{e,r}\|_{e,r}^2 = 0$ for all $e \in \mathcal{E}(\mathcal{G})$. By Lemma A.3 (A), when $u, v \in \mathcal{V}(\mathcal{G})$, $d_{R,\mathcal{G}}(u,v)$ is equal to

$$\sup_{f \in \mathscr{P}_\mu \mathcal{F}} \left\{ (f_\mu(u) - f_\mu(v))^2 : f_\mu(u_o)^2 + \sum_{e \in \mathcal{E}(\mathcal{G})} \frac{\big([f_\mu]_e(\overline{e}) - [f_\mu]_e(\underline{e})\big)^2}{\text{len}(e)} \leq 1 \right\}.$$

Since the constant functions are all members of $\mathscr{P}_\mu \mathcal{F}$, we can subtract $f_\mu(u_o)$ from each $f_\mu \in \mathscr{P}_\mu \mathcal{F}$ and easily see that the supremum above can be taken over all $f \in \mathscr{P}_\mu \mathcal{F}$ which satisfy $f_\mu(u_o) = 0$. It is now easily seen that

$$d_{\text{eff}}(u,v) = d_{R,\mathcal{G}}(u,v), \qquad u, v \in \mathcal{V}(\mathcal{G}).$$

This also establishes that $d_{R,\mathcal{G}}$ does not depend on the choice of origin and that $d_{R,\mathcal{G}}$ is a metric on $\mathcal{V}(\mathcal{G})$ (see e.g. Jorgensen and Pearse, 2010, Lemma 2.6), and hence by the splitting operation on edges (Theorem 2.3) $d_{R,\mathcal{G}}$ is a metric on $\mathcal{G}$ as well. $\square$

Proof of Theorem 2.4. Theorem 2.2 and the theory on electrical networks imply

$$d_{R,\mathcal{G}}(u,v) \leq d_{G,\mathcal{G}}(u,v), \qquad u, v \in \mathcal{V}(\mathcal{G}), \quad \text{(A.13)}$$

with equality if and only if $\mathcal{G}$ is a tree graph (see e.g. Jorgensen and Pearse, 2010, Lemma 4.3). The fact that $d_{R,\mathcal{G}}$ and $d_{G,\mathcal{G}}$ are invariant to splitting edges (by Theorem 2.3) implies that (A.13) extends to any additional finite collection of edge points. Thereby, the first part of Theorem 2.4 is verified, where the *if and only if* follows since the tree property of $\mathcal{G}$ is also invariant to edge splitting.

Suppose $\mathcal{G}$ is a Euclidean tree with origin u_o. We show $R_{\mathcal{G}}(u,v) = \widetilde{R}(u,v)$ where

$$\widetilde{R}(u,v) := \tfrac{1}{2}\big[d_{G,\mathcal{G}}(u,u_o) + d_{G,\mathcal{G}}(v,u_o) - d_{G,\mathcal{G}}(u,v)\big] + 1, \qquad u, v \in \mathcal{G}.$$

To make it easier to distinguish fixed arguments of $\widetilde{R}$, let $\widetilde{R}_v(\,\cdot\,) := \widetilde{R}(\,\cdot\,, v)$ and $\widetilde{R}_{v,e}(\,\cdot\,) := \big[\widetilde{R}_v\big]_e(\,\cdot\,)$, where $\widetilde{R}_v : \mathcal{G} \mapsto \mathbb{R}$ and $\widetilde{R}_{v,e} : [\underline{e}, \overline{e}] \mapsto \mathbb{R}$. Clearly, $\widetilde{R}_v \in \mathcal{F}$, and so by uniqueness of reproducing kernels, to establish $\widetilde{R}(u,v) = R_{\mathcal{G}}(u,v)$, it suffices to show $\langle f, \widetilde{R}_v \rangle_{\mathcal{F}} = f(v)$ for every $f \in \mathcal{F}$ and $v \in \mathcal{G}$.

Suppose $v \in \mathcal{V}(\mathcal{G})$ and let p_{v,u_o} denote the unique geodesic path connecting v to u_o. For every $e \in \mathcal{E}(\mathcal{G})$,

- if $e \cap p_{v,u_o} = \varnothing$, then $\widetilde{R}'_{v,e} \equiv 0$;
- if $e \subset p_{v,u_o}$ and e is oriented toward u_o, then $\widetilde{R}'_{v,e} \equiv 1$;
- if $e \subset p_{v,u_o}$ and e is oriented toward v, then $\widetilde{R}'_{v,e} \equiv -1$.

This follows since as u is moving along e, the rate of change of $d_{G,\mathcal{G}}(u, u_o)$ and $-d_{G,\mathcal{G}}(u, v)$ cancel when $e \cap p_{v,u_o} = \varnothing$ but are both equal to 1 or -1 when $u \in p_{v,u_o}$ depending on the orientation of e. Since $\widetilde{R}_v(u_0) = 1$,

$$\langle f, R_v \rangle_{\mathcal{F}} = f(u_o) + \sum_{e \in \mathcal{E}(\mathcal{G})} \int_{\underline{e}}^{\overline{e}} f'_e(t) \widetilde{R}'_{v,e}(t) \, \mathrm{d}t$$

and hence by the items above and a telescoping sum cancellation,

$$\langle f, R_v \rangle_{\mathcal{F}} = f(u_o) + f(v) - f(u_o) = f(v).$$

A similar argument holds for the additional partial edge in p_{v,u_o} when v is not a vertex of $\mathcal{G}$. Hence the proof is completed.

□

A.3 Proof of Theorem 3.4

Theorem 3.4 follows from the proof given in Proposition 7.6.1 in Deza and Laurent (1997). We include a proof here for completeness and to resolve the fact that Deza and Laurent (1997) uses a slightly different nomenclature.

Proof of Theorem 3.4. We need to show $(X_1 \cup X_2, d) \xhookrightarrow{\sqrt{\ }} H$ or equivalently $(X_1 \cup X_2, \sqrt{d}\,) \hookrightarrow H$ for some Hilbert space H. Since $\sqrt{d}$ is a distance whenever d is, the equivalence of (I) and (IV) in Theorem 3.2 implies that it is sufficient to show that

$$\sum_{i,j \in \mathcal{I}_1} c_i c_j \big(\sqrt{d(u_i, u_j)}\,\big)^2 \leq 0 \tag{A.14}$$

when $u_1, \ldots, u_n \in X_1 \cup X_2$ and $c_1, \ldots, c_n \in \mathbb{R}$, with $\sum_{k=1}^n c_k = 0$. Define $\mathcal{I}_0 = \{i : u_i = x_0\}$, $\mathcal{I}_1 = \{i : u_i \in X_1\}$, $\mathcal{I}_2 = \{i : u_i \in X_2\}$, and

$$c_i^{(1)} = \begin{cases} c_i & \text{if } i \in \mathcal{I}_1 \setminus \mathcal{I}_0, \\ \dfrac{1}{|\mathcal{I}_0|} \displaystyle\sum_{j \in \mathcal{I}_2} c_j & \text{if } i \in \mathcal{I}_0, \end{cases} \qquad c_i^{(2)} = \begin{cases} c_i & \text{if } i \in \mathcal{I}_2 \setminus \mathcal{I}_0, \\ \dfrac{1}{|\mathcal{I}_0|} \displaystyle\sum_{j \in \mathcal{I}_1} c_j & \text{if } i \in \mathcal{I}_0, \end{cases}$$

where $|\mathcal{I}_0|$ denotes the number of indices in $\mathcal{I}_0$. Note that $\sum_{i \in \mathcal{I}_1} c_i^{(1)} = \sum_{i \in \mathcal{I}_2} c_i^{(2)} = 0$. By assumption and the equivalence of (I) and (IV) in Theorem 3.2,

$$\sum_{i,j \in \mathcal{I}_1} c_i^{(1)} c_j^{(1)} d_1(u_i, u_j) + \sum_{i,j \in \mathcal{I}_2} c_i^{(2)} c_j^{(2)} d_2(u_i, u_j) \leq 0. \tag{A.15}$$

Therefore,

$$\begin{aligned}
\sum_{i,j\in\mathcal{I}_1} c_i^{(1)}c_j^{(1)}d_1(u_i,u_j) &= \sum_{\substack{i\in\mathcal{I}_1\setminus\mathcal{I}_0\\ j\in\mathcal{I}_1}} c_i^{(1)}c_j^{(1)}d_1(u_i,u_j) + \sum_{\substack{i\in\mathcal{I}_0\\ j\in\mathcal{I}_1}} c_i^{(1)}c_j^{(1)}d_1(u_i,u_j)\\
&= \sum_{\substack{i\in\mathcal{I}_1\setminus\mathcal{I}_0\\ j\in\mathcal{I}_1}} c_i c_j^{(1)}d(u_i,u_j) + \frac{1}{|\mathcal{I}_0|}\sum_{\substack{i\in\mathcal{I}_0\\ j\in\mathcal{I}_1}}\sum_{k\in\mathcal{I}_2} c_k c_j^{(1)}d_1(x_0,u_j)\\
&= \sum_{\substack{i\in\mathcal{I}_1\setminus\mathcal{I}_0\\ j\in\mathcal{I}_1}} c_i c_j^{(1)}d(u_i,u_j) + \sum_{\substack{k\in\mathcal{I}_2\\ j\in\mathcal{I}_1}} c_k c_j^{(1)}[d(u_k,u_j) - d_2(x_0,u_k)]\\
&= \sum_{\substack{i\in\{1,\dots,n\}\\ j\in\mathcal{I}_1}} c_i c_j^{(1)}d(u_i,u_j) - \underbrace{\sum_{j\in\mathcal{I}_1} c_j^{(1)}}_{=0}\sum_{k\in\mathcal{I}_2} c_k d_2(x_0,u_k).
\end{aligned}$$

Applying a similar argument, first applied to the index j above, then to the second term below, gives

$$\sum_{i,j\in\mathcal{I}_1} c_i^{(1)}c_j^{(1)}d_1(u_i,u_j) + \sum_{i,j\in\mathcal{I}_2} c_i^{(2)}c_j^{(2)}d_2(u_i,u_j) = 2\sum_{i,j=1}^{n} c_i c_j d(u_i,u_j). \tag{A.16}$$

Now, (A.15) combined with (A.16) implies (A.14) as was to be shown. □

Acknowledgements

We thank Heidi Søgaard Christensen, James Sharpnack, Adrian Baddeley, and Gopalan Nair for useful comments and illuminating discussions.

Ethan Anderes is supported by a NSF CAREER grant (DMS-1252795) and the UC Davis Chancellor's Fellowship.

Jesper Møller and Jakob G. Rasmussen are supported by The Danish Council for Independent Research | Natural Sciences, grant DFF – 7014-00074 "Statistics for point processes in space and beyond", and by the "Centre for Stochastic Geometry and Advanced Bioimaging", funded by grant 8721 from the Villum Foundation.

Additionally Jakob G. Rasmussen is supported by the Australian Research Council, Discovery Grant DP130102322.

References

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, 1964.

A. Baddeley, G. Nair, S. Rakshit, and G. McSwiggan. "Stationary" point processes are uncommon on linear networks. *Stat*, 7:68–78, 2017.

A. J. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns: Methodology and Applications with R.* Interdisciplinary Statistics. Chapman & Hall/CRC, Boca Raton, 2015.

C. Berg. Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity. In J. Mateu and E. Porcu, editors, *Positive Definite Functions. From Schoenberg to Space-Time Challenges*, pages 15–45. University Jaume I, Castellon, 2008.

C. Berg, J. Mateu, and E. Porcu. The Dagum family of isotropic correlation functions. *Bernoulli*, 14:1134–1149, 2008.

N. Cressie, J. Frey, B. Harch, and M. Smith. Spatial prediction on a river network. *Journal of Agricultural, Biological, and Environmental Statistics*, 11:127–150, 2006.

M. Deza and M. Laurent. *Geometry of Cuts and Metrics.* Springer-Verlag, Berlin, 1997.

T. Gneiting. Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19: 1327–1349, 2013.

I. S. Gradshteyn and I.M Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, New York, 2007.

P. E. T. Jorgensen and E. P. J. Pearse. A Hilbert space approach to effective resistance metric. *Complex Analysis and Operator Theory*, 4:975–1013, 2010.

J. Kigami. Harmonic analysis for resistance forms. *Journal of Functional Analysis*, 204: 399–444, 2003.

D. J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12: 81–95, 1993.

K. S. Miller and S. G. Samko. Completely monotonic functions. *Integral Transforms and Special Functions*, 12:389–402, 2001.

J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, Boca Raton, 2004.

J. Møller, A. R. Syversveen, and R. P. Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25:451–482, 1998.

A. Okabe and K. Sugihara. *Spatial Analysis Along Networks: Statistical and Computational Methods.* Wiley, Chichester, 2012.

S. Saitoh. *Integral Transforms, Reproducing Kernels and Their Applications.* Research Notes in Mathematics Series, Volume 369, Taylor & Francis Ltd, Chapman & Hall/CRC, Harlow, 1997.

I. J. Schoenberg. Remarks to Maurice Frechet's article"Sur la definition axiomatique d'une classe d'espace distances vectoriellement applicable sur l'espace de Hilbert. *Annals of Mathematics*, 36:724–732, 1935.

I. J. Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44:522–536, 1938a.

I. J. Schoenberg. Metric spaces and completely monotone functions. *Annals of Mathematics*, 39:811–841, 1938b.

M. R. Spiegel. *Mathematical Handbook of Formulas and Tables.* McGraw/Hill, New York, 1968.

J. Ver Hoef and E. Peterson. A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*, 105:6–18, 2010.

J. Ver Hoef, E. Peterson, and D. Theobald. Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13:449–464, 2006.

J. von Neumann and I. J. Schoenberg. Fourier integrals and metric geometry. *Transactions of the American Mathematical Society*, 50:226–251, 1941.

G. Wahba. *Spline Models for Observational Data.* Society for Industrial and Applied Mathematics, Philadelphia, 1990. doi: 10.1137/1.9781611970128.

J. H. Wells and L. R. Williams. *Embeddings and Extensions in Analysis.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84, Springer-Verlag, Berlin, 1975.

A. M. Yaglom. *Correlation Theory of Stationary and Related Random Functions*, volume I. Springer-Verlag, New York, 1987.