The Annals of Statistics
2020, Vol. 48, No. 4, 2478–2503
https://doi.org/10.1214/19-AOS1896

ISOTROPIC COVARIANCE FUNCTIONS ON GRAPHS AND THEIR EDGES

BY ETHAN ANDERES[1], JESPER MØLLER[2,*] AND JAKOB G. RASMUSSEN[2,**]

[1]*Department of Statistics, University of California at Davis, anderes@ucdavis.edu*
[2]*Department of Mathematical Sciences, Aalborg University, *jm@math.aau.dk; **jgr@math.aau.dk*

We develop parametric classes of covariance functions on linear networks and their extension to graphs with Euclidean edges, that is, 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.

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 for example, Cressie et al. (2006), Cressie and Majure (1997), Gardner, Sullivan and Lembo (2003), Ver Hoef, Peterson and Theobald (2006), Ver Hoef and Peterson (2010), Okabe and Sugihara (2012) and Baddeley, Rubak and Turner (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 Schoenberg characterize the class of all stationary covariance functions; see for example, 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, that is, 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 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

Received June 2018; revised May 2019.
MSC2020 subject classifications. Primary 62M99; secondary 62M30, 60K99.
Key words and phrases. Geodesic metric, linear network, parametric classes of covariance functions, reproducing kernel Hilbert space, resistance metric, restricted covariance function properties.

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 cannot be extended to a graph that has 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 covariance functions over graphs with Euclidean edges which are isotropic with respect to the geodesic metric and the resistance metric developed below (in fact, our covariance functions will be (strictly) positive definite). Thereby, a rich class of isotropic Gaussian random fields on the whole graph can be constructed and inferred via likelihood methods. Finally, we remark that the validity of these isotropic covariance functions also allows the construction of isotropic point process models on the whole graph constructed via a log Gaussian Cox process (Møller, Syversveen and Waagepetersen (1998), Møller and Waagepetersen (2004)). We leave this and other applications of our paper for future work.

1.1. *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. 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 repeated edges or edge which joins a vertex to itself 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 $d_{\mathcal{G}}\colon \mathcal{V} \times \mathcal{V} \to [0, \infty)$ 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

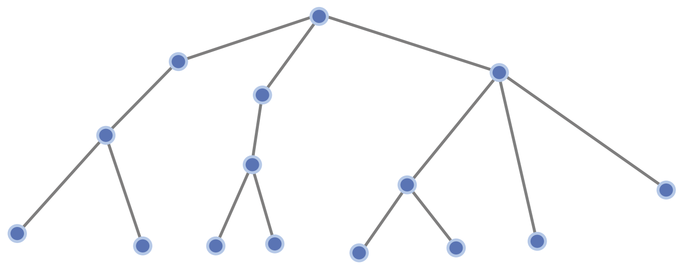


FIG. 1. *A Euclidean tree constructed from the linear network of grey lines. The dots represent the vertices.*

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 edge sets $e \in \mathcal{E}$.

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 each edge set $e_i \in \mathcal{E}$ correspond to the relative interior of the corresponding line segment ℓ_i. Let each bijection φ_{e_i} be given by the inverse of the 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 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:

$$\overline{e}_o - \underline{e}_o \leq \sum_{\substack{e \in \mathcal{E} \\ e \neq e_o}} (\overline{e} - \underline{e}). \tag{1}$$

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 (the last graph) when each φ_e is given by the inverse of path-length parametrization.

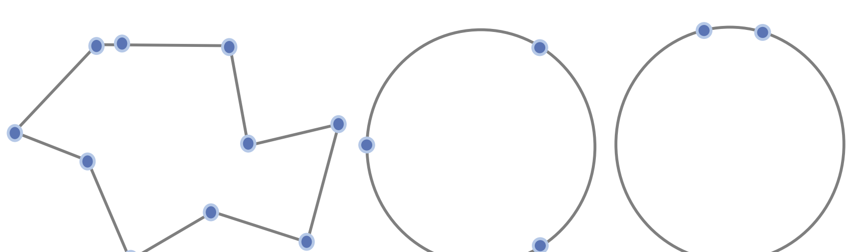


FIG. 2. *The two graphs on the left are Euclidean cycles. However, the right most graph is* not *a graph with Euclidean edges.*

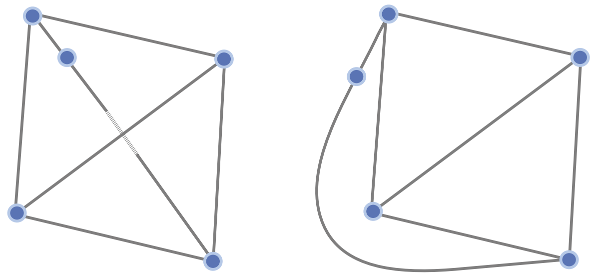


FIG. 3. *The two diagrams above show a graph with Euclidean edges $\mathcal{G}$ which can* not *be represented as a linear network in $\mathbb{R}^2$. The diagram on the left is drawn in a way that visually preserves edge length but forces an intersection that does not correspond to a vertex in $\mathcal{G}$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

In all of 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 and 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 cannot be represented as a linear network in $\mathbb{R}^2$.

1.2. *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 $C : [0, \infty) \to \mathbb{R}$ 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 $\mathcal{G} \times \mathcal{G} \ni (u, v) \to C(d(u, v)) \in \mathbb{R}$ 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.

In Theorem 1 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 of $\mathcal{G}_1$ and $\mathcal{G}_2$ 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. Further, we need to recall the following definition of a completely monotonic function, noting there is a distinction, in the literature, between complete monotonicity on $[0, \infty)$ vs. on $(0, \infty)$, the latter being fundamentally related to Bernstein functions and variograms (see Berg (2008), Wells and Williams (1975)).

DEFINITION 2. A function $f : [0, \infty) \to \mathbb{R}$ 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 jth derivative of f and $f^{(0)} = f$.

THEOREM 1. *Let $C : [0, \infty) \to \mathbb{R}$ be a completely monotone and nonconstant function*:

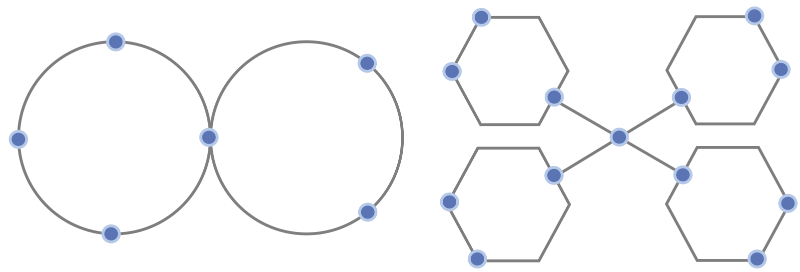


FIG. 4. *Examples of finite sequential* 1*-sums of cycles and trees. Left*: *A* 1*-sum of two Euclidean cycles. Right*: *A sequential* 1*-sum of four Euclidean cycles and one Euclidean tree.*

(i) *If $\mathcal{G}$ is a graph with Euclidean edges*, *then $C(d_{R,\mathcal{G}}(u,v))$ is* (*strictly*) *positive definite over $(u,v) \in \mathcal{G} \times \mathcal{G}$.*

(ii) *If $\mathcal{G}$ is a graph with Euclidean edges that forms a finite sequential* 1*-sum of Euclidean cycles and trees*, *then $C(d_{G,\mathcal{G}}(u,v))$ is* (*strictly*) *positive definite over $(u,v) \in \mathcal{G} \times \mathcal{G}$.*

A consequence of Theorem 1 is that many of the parametric classes of autocovariance functions used in spatial statistics are (strictly) positive definite with respect to $d_{R,\mathcal{G}}$ for general graphs and with respect to $d_{G,\mathcal{G}}$ for graphs which are 1-sums of Euclidean trees and cycles. Notice, however, this holds only after restricting the parametric range to ensure complete monotonicity on $[0,\infty)$, as in the radial functions given in Table 1. To see why the radial functions in Table 1 are completely monotonic, first note that $t \to f(\beta t^\alpha + \lambda)$ is completely monotonic if $\alpha \in (0,1]$, $\beta, \lambda > 0$ and f is completely monotonic (see equation (1.6) in Miller and Samko (2001)). Therefore, $\exp(-\beta t^\alpha)$ and $(\beta t^\alpha + 1)^{-\xi/\alpha}$ are completely monotonic 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. The complete monotonicity for the Matérn class in Table 1 was proved in Example 2 of Gneiting (2013). Finally, Theorem 9 in Berg, Mateu and Porcu (2008) establishes that $C(t) = 1 - (t^\beta/(1+t^\beta))^\gamma$ is completely monotonic, whenever $\beta\gamma \in (0,1]$ and $\beta \in (0,1]$, which proves the desired result for the Dagum class in Table 1.

In the special case where $\mathcal{G}$ is a Euclidean tree with geodesic metric $d_{G,\mathcal{G}}(u,v)$, the results of Theorem 1(ii) can be obtained from existing literature. Indeed, it is well known that the exponential covariance functions are positive definite (via ℓ_1 embedding, using Theorem 4.1 in Wells and Williams (1975) and Theorem 3.2.2 in Deza and Laurent (1997)) which implies that positive mixtures of exponential covariance functions are positive definite with respect to $d_{G,\mathcal{G}}(u,v)$. Now, the results of Schoenberg (outlined in Theorem 8 below) are sufficient to establish Theorem 1(ii) for this special case.

TABLE 1

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

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

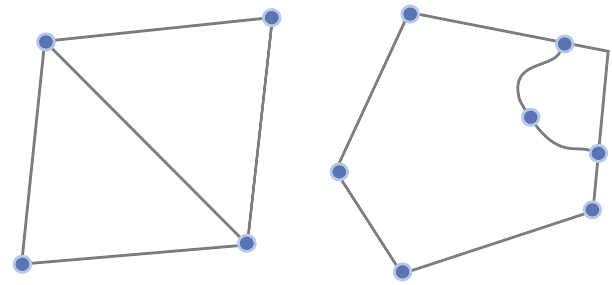


FIG. 5. *Examples of forbidden graphs for the exponential class with respect to the geodesic metric.*

The contrasting generality of the range of graphs $\mathcal{G}$ applicable in Theorem 1 for (i) versus (ii) hints at a degeneracy that occurs when modeling covariance functions which are isotropic with respect to the geodesic metric. The next result, Theorem 2, confirms this degeneracy by showing that for the geodesic metric, Theorem 1 cannot 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 semidefinite 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.

THEOREM 2. *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 we develop the new resistance metric $d_{R,\mathcal{G}}$ which is specifically designed to overcome the restrictions imposed by Theorem 2 for the geodesic metric on general graphs with Euclidean edges. We construct $d_{R,\mathcal{G}}$ as the variogram of a canonical Gaussian random field over $\mathcal{G}$ obtained by linearly interpolating a random vector on the vertices constructed from the graph Laplacian and then add independent Brownian bridges over each edge. While it is known that the (discrete) effective resistance metric can be expressed as the variogram of a random vector (see Lyons and Peres (2016) for an excellent exposition), it appears that the approach given here, namely, using the variogram of a canonical (continuous) random field to define a (continuum) resistance metric, is new. The advantage of this construction is that it gives the following key Hilbert space embedding result:

THEOREM 3. *If $\mathcal{G}$ is a graph with Euclidean edges, there exists a Hilbert space H and an embedding $\varphi\colon \mathcal{G} \to H$ such that*

$$\sqrt{d_{R,\mathcal{G}}(u, v)} = \|\varphi(u) - \varphi(v)\|_H \tag{2}$$

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 $d_{G,\mathcal{G}}$.

In some sense, the construction of $d_{R,\mathcal{G}}$ in Section 2.3 and the proof of Theorem 3 are the most important results of this paper. Once they are established, many of the results in this section follow almost immediately from well-known consequences of Schoenberg's work in the context of embeddings; see, for example, Wells and Williams (1975) or Jayasumana et al. (2013).

The next two theorems illustrate the scope of the results given above. In particular, Theorem 1(i) gives sufficient conditions for (strict) positive definiteness over *all* graphs with Euclidean edges $\mathcal{G}$. This does not preclude a less stringent sufficient condition that holds for a subcollection of graphs with Euclidean edges. For example, consider the case where $\mathcal{G}$ has a single edge connecting two vertices. Then both $d_{R,\mathcal{G}}$ and $d_{G,\mathcal{G}}$ are equivalent to the Euclidean metric on a compact interval $[0, c] \subset \mathbb{R}$ and, as such, contain a much richer collection of positive definite covariance functions than those established in Theorem 1. A less trivial example can be obtained by restricting $\mathcal{G}$ to be a Euclidean tree as in the next two theorems.

THEOREM 4. *Let $\mathcal{G}$ be a Euclidean tree with m leaves, where $m \geq 3$. Then $C(d_{G,\mathcal{G}}(u, v))$ and $C(d_{R,\mathcal{G}}(u, v))$ are positive semidefinite over $(u, v) \in \mathcal{G} \times \mathcal{G}$ whenever $C: [0, \infty) \to \mathbb{R}$ is given by*

$$C(t) = \int_0^\infty \omega_{\lceil m/2 \rceil}(\sigma t)\, \mathrm{d}\mu(\sigma), \tag{3}$$

where μ is a finite (positive) measure on $(0, \infty)$ and $\omega_n(t)$ is defined by

$$\omega_n(t) = \frac{\Gamma(n/2)}{\sqrt{\pi}\,\Gamma((n-1)/2)} \int_1^\infty \Omega_n(v^{1/2} t) v^{-n/2} (v-1)^{(n-3)/2}\, \mathrm{d}v$$

with $\Omega_n(t) = \Gamma(n/2)(2/t)^{(n-2)/2} J_{(n-2)/2}(t)$ and $J_\nu(t)$ denoting the Bessel function of the first kind and order ν.

The proof of the above result, given in Section 4, follows directly by the work of Cambanis, Keener and Simons (1983) once it is established that Euclidean trees with m leaves can be embedded in $\mathbb{R}^{\lceil m/2 \rceil}$ with ℓ_1 metric and $d_{R,\mathcal{G}} = d_{G,\mathcal{G}}$ on Euclidean trees (cf. Proposition 4 below). This ℓ_1 embedding result can also used in combination with Theorem 3.2 of Gneiting (1998a) to give a simplified criterion for the conclusion of Theorem 4.

THEOREM 5. *Let $\mathcal{G}$ be a Euclidean tree with m leaves, where $m \geq 3$. If $C: [0, \infty) \to \mathbb{R}$ is a continuous function such that $C^{(2\lceil m/2 \rceil - 2)}$ is convex and $\lim_{t\to\infty} C(t) = 0$, then both $C(d_{G,\mathcal{G}}(u, v))$ and $C(d_{R,\mathcal{G}}(u, v))$ are positive semidefinite over $(u, v) \in \mathcal{G} \times \mathcal{G}$.*

Finally, we notice that Theorem 4 shows that covariance functions on Euclidean trees may attain negative values, and, at the very end of Section 5, we give an example of a parametric family of covariance functions whose support can be made arbitrary small.

1.3. *Outline for the remainder of the paper.* Details of the geodesic metric and resistance metric over graphs with Euclidean edges, along with their theoretical properties used in subsequent sections, are given in Section 2. The resistance metric $d_{R,\mathcal{G}}$ is defined constructively as the variogram of a certain random field over $\mathcal{G}$, analogous to a Wiener process on $\mathbb{R}$. This construction has the advantage that it establishes the Hilbert space embedding result almost immediately (utilizing a theorem of Schoenberg). The difficulty, however, is in showing that $d_{R,\mathcal{G}}$ is indeed an extension of the classical effective resistance on any finite subgraph and is invariant to the graph operations of splitting and merging edges (cf. Proposition 3 below). The invariance result is important since it implies the resistance metric is, in some sense, intrinsic to the minimal graph structure of $\mathcal{G}$. The addition or removal of unnecessary vertices along an edge leaves $d_{R,\mathcal{G}}$ unchanged. Proofs of these theoretical properties rely heavily on Hilbert space methods and are deferred to the Appendix.

Sections 3 and 4 contain the proofs of all the results summarized in Section 1.2 and follow relatively easily given the results in Section 2. The Hilbert space embedding stated in

Theorem 3 is proved first in Section 3. The remaining four theorems summarized in Section 1.2 are proved in Section 4. 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 nonempty 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 $\operatorname{len}(e)$.

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 $u_1, e_1, u_2, e_2, \dots, u_n, e_n, u_{n+1}$, where $u_1, \dots, u_{n+1} \in \mathcal{G}$ are pairwise distinct, $u_1 = u$, $u_{n+1} = v$, and $e_1, e_2, \dots, e_n$ are nonoverlapping partial edges such that each e_i has boundary $\{u_i, u_{i+1}\}$. Moreover, the *length of p_{uv}* is denoted $\operatorname{len}(p_{uv})$ and defined as the sum of the lengths of $e_1, e_2, \dots, e_n$.

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

$$d_{G,\mathcal{G}}(u, v) = \inf\{\operatorname{len}(p_{uv})\}, \tag{4}$$

where the infimum is over all paths connecting u and v. Using the consistency requirement given in Definition 1(d), the following proposition is easily verified.

PROPOSITION 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* $\operatorname{len}(e)$.
- *$d_{G,\mathcal{G}}$ 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, for example, 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 edge 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 2. 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 nonstandard 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 $Z_{\mathcal{G}}$:

$$d_{R,\mathcal{G}}(u,v) := \operatorname{var}\big(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v)\big), \quad u, v \in \mathcal{G}. \tag{5}$$

Propositions 2 and 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:

$$Z_{\mathcal{G}}(u) := Z_\mu(u) + \sum_{e \in \mathcal{E}(\mathcal{G})} Z_e(u), \quad u \in \mathcal{G}. \tag{6}$$

The field Z_μ is characterized by a 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 $e \in \mathcal{E}(\mathcal{G})$, the random field Z_e is only defined to be nonzero 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 $c : \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G}) \to [0, \infty)$ given by

$$c(u,v) = \begin{cases} 1/d_{G,\mathcal{G}}(u,v) & \text{if } u \sim v, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

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 $L : \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G}) \to \mathbb{R}$ as the function/matrix with coordinates

$$L(u,v) = \begin{cases} 1 + c(u_o) & \text{if } u = v = u_o, \\ c(u) & \text{if } u = v \neq u_o, \\ -c(u,v) & \text{otherwise,} \end{cases} \tag{8}$$

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 $z, w \in \mathbb{R}^{\mathcal{V}(\mathcal{G})}$,

$$z^T L w = z(u_o)w(u_o) + \frac{1}{2} \sum_{u \sim v} \big(z(u) - z(v)\big) c(u,v) \big(w(u) - w(v)\big), \tag{9}$$

so $z^T L z = 0$ 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, for example, Kigami (2003) and Jorgensen and Pearse (2010), except that L has the additional 1 added at (u_o, 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 Proposition 2 below).

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 $v_1, v_2, \dots, v_n$ denote the vertices in $\mathcal{V}(\mathcal{G})$, and define Z_μ at these vertices by

$$\big(Z_\mu(v_1), \dots, Z_\mu(v_n)\big)^T \sim \mathcal{N}\big(0, L^{-1}\big). \tag{10}$$

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

$$Z_\mu(u) = \big(1 - d(u)\big) Z_\mu(\underline{u}) + d(u) Z_\mu(\overline{u}), \tag{11}$$

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

$$d(u) = \begin{cases} d_{G,\mathcal{G}}(u, \underline{u}) / d_{G,\mathcal{G}}(\underline{u}, \overline{u}) & \text{if } u \notin \mathcal{V}(\mathcal{G}), \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

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

$$\begin{aligned} R_\mu(u, v) = {} & d(u) d(v) L^{-1}(\overline{u}, \overline{v}) + \big[1 - d(u)\big]\big[1 - d(v)\big] L^{-1}(\underline{u}, \underline{v}) \\ & + d(u)\big[1 - d(v)\big] L^{-1}(\overline{u}, \underline{v}) + \big[1 - d(u)\big] d(v) L^{-1}(\underline{u}, \overline{v}). \end{aligned} \tag{13}$$

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 recomputing 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

$$Z_e(u) = \begin{cases} B_e\big(\varphi_e(u)\big) & \text{if } u \in e, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

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

$$R_e(u, v) = \begin{cases} \big[d(u) \wedge d(v) - d(u) d(v)\big] d_{G,\mathcal{G}}(\underline{u}, \overline{u}) & \text{if } u, v \in e, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Note that the covariance function $R_\mathcal{G}$ for the random field $Z_\mathcal{G}$, defined in (6), satisfies

$$R_\mathcal{G}(u, v) = R_\mu(u, v) + \sum_{e \in \mathcal{E}(\mathcal{G})} R_e(u, v), \quad u, v \in \mathcal{G}. \tag{16}$$

2.3.3. *Properties of $d_{G,\mathcal{G}}$ and $d_{R,\mathcal{G}}$.* The following Proposition 2 shows that $d_{R,\mathcal{G}}$ is indeed the extension of the classical effective resistance on electrical networks and is invariant to the choice of origin u_o (used in the construction of L in (8)). Further, Proposition 3 shows that $d_{R,\mathcal{G}}$ is invariant to the addition of vertices and removal of vertices with degree two. Finally, Proposition 4 characterizes $d_{R,\mathcal{G}}$ via an associated infinite dimensional reproducing kernel Hilbert space. The proofs of the propositions are given in Appendix A.2.

PROPOSITION 2. *For a graph with Euclidean edges $\mathcal{G}$, $d_{R,\mathcal{G}}$ is 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 two 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 two 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.

PROPOSITION 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 (so long as the resulting graph satisfies the conditions of Definition 1).*

Propositions 2 and 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. The next proposition, also analogous to results from electrical network theory, illustrates how multiple pathways between points of $\mathcal{G}$ lead to a reduction of $d_{R,\mathcal{G}}$ compared with $d_{G,\mathcal{G}}$.

PROPOSITION 4. *For any graph with Euclidean edges $\mathcal{G}$, we have*

$$d_{R,\mathcal{G}}(u, v) \le d_{G,\mathcal{G}}(u, v), \quad u, v \in \mathcal{G}, \tag{17}$$

with equality if and only if $\mathcal{G}$ is a Euclidean tree. If $\mathcal{G}$ is a Euclidean cycle with circumference $\omega = \sum_{e\in\mathcal{E}(\mathcal{G})} \operatorname{len}(e)$, then

$$d_{R,\mathcal{G}}(u,v) = d_{G,\mathcal{G}}(u,v) - \frac{d_{G,\mathcal{G}}(u,v)^2}{\omega}, \quad u,v\in\mathcal{G}. \tag{18}$$

The fact that $d_{R,\mathcal{G}}(u,v) = d_{G,\mathcal{G}}(u,v)$, if and only if $\mathcal{G}$ is a Euclidean tree, suggests that $d_{R,\mathcal{G}}$ can be viewed not only as an extension of the vertex (effective) resistance as established in Proposition 2 but also as an extension of $d_{G,\mathcal{G}}$ on trees to general graphs. Instead of extending the shortest path property of the geodesic metric, the resistance metric extends the validity of the covariance models given in Theorem 1, from $d_{G,\mathcal{G}}$ on trees to $d_{R,\mathcal{G}}$ on general graphs (noting that Theorem 2 implies both properties cannot be simultaneously extended to general graphs).

Notice that the quadratic term in (18), for the Euclidean cycles, explicitly quantifies how multiple paths leads to a reduction in (effective) resistance. Moreover, (18) can be rearranged, using the fact that $\omega = \operatorname{len}(p_{uv}) + \operatorname{len}(\widetilde{p}_{uv})$ where p_{uv} denotes the shortest path from u to v and $\widetilde{p}_{uv}$ denoting the longer path connecting u to v, to obtain

$$d_{R,\mathcal{G}}(u,v) = \big(\operatorname{len}(p_{uv})^{-1} + \operatorname{len}(\widetilde{p}_{uv})^{-1}\big)^{-1}.$$

In particular, $d_{R,\mathcal{G}}(u,v)$ is a function of both path lengths, strictly smaller than each, combined through what is called *parallel reduction* in electrical network theory.

We remark that (18) allows us to write any covariance model $C(d_{R,\mathcal{G}}(u,v))$ on a Euclidean cycle $\mathcal{G}$ in terms of the geodesic metric $d_{G,\mathcal{G}}$. Combined with Theorem 1, we conclude that $C(t - t^2/\omega)$ is strictly positive definite on the circle of radius $\omega/(2\pi)$ for every nonconstant completely monotonic function $C : [0,\infty) \to \mathbb{R}$. These results can be compared with the literature on isotropic autocovariance models on the circle. For example, in the case $C(t) = \exp(-\beta t)$ the positive definiteness of the autocovariance $\exp(-\beta(t - t^2/(2\pi)))$ with respect to the geodesic distance on $\mathbb{S}^1$ agrees with the conclusions of Pólya's theorem on the circle (see, e.g., Theorem 4 in Gneiting (1998b)).

Our final result on the resistance metric, although stated last and verified in Appendix A.1, gives the above three propositions 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 $f : \mathcal{G} \to \mathbb{R}$ and $e \in \mathcal{E}(\mathcal{G})$, we let $f_e : [\underline{e}, \overline{e}] \to \mathbb{R}$ denote the restriction of f to e and be 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 3. 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 $f : \mathcal{G} \to \mathbb{R}$ 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 $\mathcal{F}$:

$$\langle f, g\rangle_{\mathcal{F}} := f(u_o)g(u_o) + \sum_{e\in\mathcal{E}(\mathcal{G})} \int_{\underline{e}}^{\overline{e}} f_e'(t) g_e'(t)\,\mathrm{d}t. \tag{19}$$

PROPOSITION 5. *Let $\mathcal{G}$ be a graph with Euclidean edges with origin $u_o \in \mathcal{V}(\mathcal{G})$. Then, 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 (16), and resistance metric $d_{R,\mathcal{G}}(u,v)$, given in (5), satisfying*

$$d_{R,\mathcal{G}}(u,v) = \sup_{f\in\mathcal{F}} \big\{ \big(f(u) - f(v)\big)^2 : \|f\|_{\mathcal{F}} \le 1 \big\}, \tag{20}$$

$$R_{\mathcal{G}}(u,v) = 1 + \big\{ d_{R,\mathcal{G}}(u,u_o) + d_{R,\mathcal{G}}(v,u_o) - d_{R,\mathcal{G}}(u,v) \big\}/2, \tag{21}$$

for all $u, v \in \mathcal{G}$.

Notice that (21) illustrates how the additional 1, added to $c(u_o)$ in (8), translates to the dependence of $R_{\mathcal{G}}(u,v)$ on u_o. This was done to ensure the invertibility of L but the effect of which is canceled in the variogram of $Z_{\mathcal{G}}$. Moreover, (21) also illustrates the connection with classic Brownian motion on $\mathbb{R}$. 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 Propositions 4 and 5 show that $R_{\mathcal{G}}(u, v) = 1 + (|u| + |v| - |u - v|)/2$ which, up to an overall constant, is precisely the covariance function of Brownian motion.

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 3. 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. Since they hold for arbitrary distance spaces the exposition of both of these results are kept as general as possible.

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

DEFINITION 4. Let (X, d) be a distance space and $g : \operatorname{Range}(X, d) \to [0, \infty)$ 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 $\varphi\colon X \to H$ which satisfies

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

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

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

THEOREM 6. *Let (X, d) be a distance space and x_0 an arbitrary member of X. The following statements are equivalent*:

(I) $(X, d) \overset{\mathrm{id}}{\hookrightarrow} H$ *for some Hilbert space* H.
(II) $d(x, x_0)^2 + d(y, x_0)^2 - d(x, y)^2$ *is positive semidefinite over* $x, y \in X$.
(III) *For every* $\beta > 0$, *the function* $\exp(-\beta d(x, y)^2)$ *is positive semidefinite 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* $x_1, \dots, x_n \in X$ *and* $c_1, \dots, c_n \in \mathbb{R}$ *for which* $\sum_{k=1}^n c_k = 0$.

It is common (in Wells and Williams (1975), e.g.) to call any distance space (X, d), which satisfies condition (IV), a *distance of negative type*. In the geostatistical 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 $d(u, v)^2 = \operatorname{var}(Z(u) - Z(v))$.

The last concept needed to show Theorem 3 deals with the notion of the 1-sum of two distance spaces (Deza and Laurent (1997)). This operation allows us to construct new distance spaces (which are root embeddable) by stitching multiple root-embeddable distance spaces together.

DEFINITION 5. 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

$$d(x,y) = \begin{cases} d_1(x,y) & \text{if } x, y \in X_1, \\ d_2(x,y) & \text{if } x, y \in X_2, \\ d_1(x,x_0) + d_2(x_0,y) & \text{if } x \in X_1 \text{ and } y \in X_2. \end{cases} \tag{22}$$

The key fact about 1-sums, for our use, is summarized in the following theorem. We omit the proof and simply note that it can be found in Deza and Laurent (1997) (Proposition 7.6.1) with slightly different nomenclature.

THEOREM 7. *Suppose (X_1, d_1) and (X_2, d_2) are two distance spaces such that $X_1 \cap X_2 = \{x_0\}$. If $(X_1, d_1) \overset{\surd}{\hookrightarrow} H_1$ and $(X_2, d_2) \overset{\surd}{\hookrightarrow} H_2$ for two Hilbert spaces H_1 and H_2, then there exists a Hilbert space H such that $(X_1 \cup X_2, d) \overset{\surd}{\hookrightarrow} H$ 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 3, from Section 1.2, on the Hilbert space embedding of $d_{R,\mathcal{G}}$ and $d_{G,\mathcal{G}}$.

PROOF OF THEOREM 3. Suppose $\mathcal{G}$ is a graph with Euclidean edges. By (5) we trivially have

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

where $d(u,v) := \sqrt{d_{R,\mathcal{G}}(u,v)}$. The fact that $d(u,v)^2$ is a variogram implies that condition (IV) holds (from Schoenberg's result stated in Theorem 6). Since (I) $\iff$ (IV), we have that $(\mathcal{G}, d) \overset{\mathrm{id}}{\hookrightarrow} H$ for some Hilbert space H and, hence, $(\mathcal{G}, d_{R,\mathcal{G}}) \overset{\surd}{\hookrightarrow} H$, as was to be shown.

For the geodesic metric, first assume $\mathcal{G}$ forms a tree graph. In this case Proposition 4 implies $d_{G,\mathcal{G}} = d_{R,\mathcal{G}}$ and, therefore, $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\surd}{\hookrightarrow} H$ 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 2). 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 semidefinite over $\mathbb{S}^1 \times \mathbb{S}^1$ for all $\beta > 0$ (see Gneiting (2013), e.g.), the function $\exp(-\beta d_{G,\mathcal{G}}(u,v))$ is positive semidefinite over $\mathcal{G} \times \mathcal{G}$ for all $\beta > 0$. Now, setting $d := \sqrt{d_{G,\mathcal{G}}}$, the equivalence (I) $\iff$ (III) in Theorem 6 implies $(\mathcal{G}, d) \overset{\mathrm{id}}{\hookrightarrow} H$, hence $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\surd}{\hookrightarrow} H$ 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 7 to conclude that $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\surd}{\hookrightarrow} H$ for some Hilbert space H. $\square$

4. Isotropic covariance functions with respect to $d_{G,\mathcal{G}}$ and $d_{R,\mathcal{G}}$. In this section we prove all the results stated in Section 1.2, with the exception of Theorem 3 proved in the previous section. In some sense many of the proofs of these results follow easily from Theorem 3 and the seminal work of Schoenberg and von Neumann (Schoenberg (1938a, 1938b), von Neumann and Schoenberg (1941)) connecting metric embeddings, Hilbert spaces and completely monotonic functions (see also Gneiting (2013)), but we review the necessary results for completeness. The following result characterizes completely monotonic functions on $[0, \infty)$ as positive mixtures of scaled exponentials (see Theorems 2, 3 and 3′ in Schoenberg (1938b)).

THEOREM 8. *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* H *is a Hilbert space and* f *is a nonconstant completely monotonic function on* $[0,\infty)$ *then* $f(\|x-y\|_H^2)$ *is (strictly) positive definite over* $(x,y)\in H\times H$.

COROLLARY 1. *If* $C:[0,\infty)\to\mathbb{R}$ *is a nonconstant and completely monotonic function and* (X,d) *is a distance space which satisfies*

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

where H *is a Hilbert space, then* $C(d(x,y))$ *is positive semidefinite over* $(x,y)\in X\times X$. *If, in addition,* $d(x,y)=0\Leftrightarrow x=y$ *for all* $x,y\in X$, *then* $C(d(x,y))$ *is (strictly) positive definite over* $(x,y)\in X\times X$.

Corollary 1 follows easily from Theorem 8 since, if $(X,d)\overset{\surd}{\hookrightarrow} H$ for some Hilbert space H, then there exists a map $\varphi\colon X\to H$ for which $d(x,y)=\|\varphi(x)-\varphi(y)\|_H^2$ for all $x,y\in X$. Then, for a nonconstant and completely monotonic function C we have $C(d(x,y))=C(\|\varphi(x)-\varphi(y)\|_H^2)$ which is positive semidefinite, via Theorem 8. If, in addition, $d(x,y)=0\Leftrightarrow x=y$ for all $x,y\in X$, then φ maps one-to-one onto its range which implies that $C(d(x,y))=C(\|\varphi(x)-\varphi(y)\|_H^2)$ is strictly positive definite. This establishes Corollary 1.

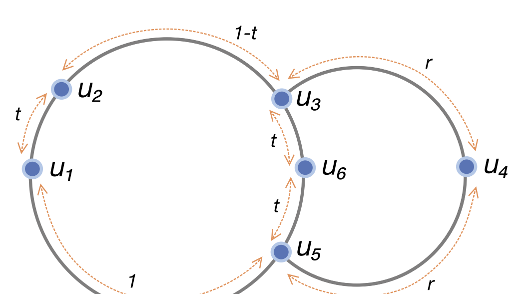
Now, we turn to the proofs of the remaining four theorems stated in Section 1.2: Theorems 1, 2, 4 and 5.

PROOF OF THEOREM 1. Suppose $\mathcal{G}$ is a graph with Euclidean edges, and let $C:[0,\infty)\to\mathbb{R}$ be a nonconstant and completely monotonic. The metric properties of both $d_{G,\mathcal{G}}$ and $d_{R,\mathcal{G}}$ imply that $d_{G,\mathcal{G}}(u,v)=0\Leftrightarrow u=v$ and $d_{R,\mathcal{G}}(u,v)=0\Leftrightarrow u=v$ for all $u,v\in\mathcal{G}$. Theorem 1 now follows immediately from Theorem 3 and Corollary 1. $\square$

PROOF OF THEOREM 2. By uniformly scaling $d_{G,\mathcal{G}}$ and possibly selecting new vertices on $\mathcal{G}$ by edge splitting operations (Proposition 3), one can obtain six vertices $u_1,\dots,u_6$ on $\mathcal{G}$ which have the following geodesic pairwise distance matrix where $0<t\le r\le 1$:

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

corresponding to the following subgraph:



The values $2t$, $2r$ and 2 represent the lengths of the three paths connecting vertices u_3 and u_5, ordered smallest to largest. Notice that in the case $t = r = 1$, one has $u_3 = u_5$ which implies the graph shown above will only have 5 distinct vertices. 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 semidefinite over $\{u_1, \ldots, u_6\}$. Then, Theorem 6 gives the existence of a $\beta > 0$ such that $\exp(-\beta d_{G,\mathcal{G}}(u, v))$ is *not* positive semidefinite over $\{u_1, \ldots, u_6\}$. $\square$

PROOF OF THEOREMS 4 AND 5. Let $\mathcal{G}$ be a Euclidean tree with m leaves where $m \geq 3$. Set $n = \lceil \frac{m}{2} \rceil$, and let $(\mathbb{R}^n, d_1)$ denote the usual ℓ_1 metric space so that $d_1(x, y) = \sum_{i=1}^n |x_i - y_i|$ for $x, y \in \mathbb{R}^n$.

First, we show that $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$ and $(\mathcal{G}, d_{R,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$. By well-known properties of tree graphs, $(\mathcal{V}(\mathcal{G}), d_{G,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$; see, for example, Proposition 11.1.4 in Deza and Laurent (1997). To extend this embedding from $\mathcal{V}(\mathcal{G})$ to all points in $\mathcal{G}$, it will be sufficient, by Theorem 3.2.2 in Deza and Laurent (1997), to show $(\mathcal{U}, d_{G,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$ for any finite subset $\mathcal{U} \subset \mathcal{G}$. Since $(\mathcal{U} \cup \mathcal{V}(\mathcal{G}), d_{G,\mathcal{G}})$ is also isometric to a tree graph with m leaves, we have that $(\mathcal{U} \cup \mathcal{V}(\mathcal{G}), d_{G,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$. This implies that $(\mathcal{U}, d_{G,\mathcal{G}})$ embeds into $(\mathbb{R}^n, d_1)$, via restriction, as was to be shown. Now, since $\mathcal{G}$ is a Euclidean tree, Proposition 4 implies $d_{G,\mathcal{G}} = d_{R,\mathcal{G}}$. Therefore, we also have $(\mathcal{G}, d_{R,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$.

Second, Theorem 3.1 of Cambanis, Keener and Simons (1983) implies $C(d_1(x, y))$ is positive semidefinite over $x, y \in \mathbb{R}^n$, thus proving Theorem 4 by $(\mathcal{G}, d_{G,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$ and $(\mathcal{G}, d_{R,\mathcal{G}}) \overset{\text{id}}{\hookrightarrow} (\mathbb{R}^n, d_1)$. Finally, Theorem 3.2 in Gneiting (1998a) establishes Theorem 5. $\square$

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 $\mathbb{S}^1$; in fact, all the covariance functions in Table 1 of Gneiting (2013) for the circle can be adapted when $\mathcal{G}$ is a Euclidean cycle. Below, Corollary 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 3 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 4 $d_{G,\mathcal{G}} = d_{R,\mathcal{G}}$, and we use $d_{\cdot,\mathcal{G}}$ as a common notation for the two metrics.

PROPOSITION 6. *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* $\operatorname{var}(Z_n(u) - Z_n(v)) = \tilde{\beta} d_{\cdot,\mathcal{G}}(u, v)^{\tilde{\alpha}} + o(d_{\cdot,\mathcal{G}}(u, v))$ *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 $d(u,v)^2 = \operatorname{var}(Z(u) - Z(v))$. By Theorem 6, $C(u,v) = d(u,u_0)^2 + d(v,u_0)^2 - d(u,v)^2$ is positive semidefinite over $\mathcal{S}_n \times \mathcal{S}_n$. For $i = 1, \ldots, n$ and $\epsilon > 0$ sufficiently small, let $u_{i,\epsilon}$ be the point on the ith edge which has $d_{\cdot,\mathcal{G}}$ distance ϵ from u_0. The assumption on $\operatorname{var}(Z(u) - Z(v))$ implies

$$d(u_{i,\epsilon}, u_0)^2 = \tilde{\beta} d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_0)^{\tilde{\alpha}} + o\big(d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_0)\big) = \tilde{\beta}\epsilon^{\tilde{\alpha}} + o(\epsilon^{\tilde{\alpha}}),$$
$$d(u_{i,\epsilon}, u_{j,\epsilon})^2 = \tilde{\beta} d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_{j,\epsilon})^{\tilde{\alpha}} + o\big(d_{\cdot,\mathcal{G}}(u_{i,\epsilon}, u_{j,\epsilon})\big) = \tilde{\beta} 2^{\tilde{\alpha}}\epsilon^{\tilde{\alpha}} + o(\epsilon^{\tilde{\alpha}})$$

when $i \neq j$. Let Σ_ϵ be the $n \times n$ covariance matrix with (i,j)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

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

where I_n is the $n \times n$ identity matrix, $\mathbf{1}_n$ 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,

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

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

COROLLARY 2. *Let C be one of the functions given in Table 1 but with α outside the parameter range, that is, $\alpha > \frac{1}{2}$ 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 (1969),

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

Hence, Proposition 6 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 $\mathrm{var}(Z_n(u) - Z_n(v))$ of Theorem 6 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 Proposition 6, we get that $\alpha \le 1$, thus completing the proof. □

PROPOSITION 7. *Suppose $(u, v) \to C(d_{\cdot,\mathcal{G}}(u, v))$ 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*

$$-\frac{C(0)}{n-1} \le C(2t) \le C(0), \qquad \frac{nC(t)^2 - C(0)^2}{n-1} \le C(0)C(2t). \tag{23}$$

PROOF. Denote $e_1, \dots, e_n$ 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 $i = 1, \dots, n$. Note that $d_{\cdot,\mathcal{G}}(u_i, u_j) = 2t$ for $i, j = 1, \dots, n$ and $i \ne j$. Let Σ denote the $(n+1) \times (n+1)$ matrix with the (i, j)th entry equal to $C(d_{\cdot,\mathcal{G}}(u_i, u_j))$, that is,

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

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

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

for $k = 1, \dots, n$, and hence either $0 \le C(0) = C(2t)$ or both $C(0) > C(2t)$ and $(k-1)C(2t) + C(0) \ge 0$, implying the first inequality in (23), where we have let $k = n$ to obtain the highest lower bound. Moreover,

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

for $k = 2, \dots, n+1$, and so either $|C(t)| \le C(0) = C(2t)$ or both $C(0) > C(2t)$ and $C(0)^2 + (k-2)C(2t)C(0) - (k-1)C(t)^2 \ge 0$, implying the second inequality in (23), where we have used $k = n+1$ to get the highest lower bound. □

COROLLARY 3. *A function $(u, v) \to C(d_{\cdot,\mathcal{G}}(u, v))$ which is a covariance function on all Euclidean trees has to be nonnegative 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 (23) implies nonnegativity of C, and the second inequality in (23) implies $C(0)C(2t) \ge C(t)^2$ 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 $t > 0$. □

To illustrate the scope of the covariance function restrictions given above, it is instructive to consider the 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

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

where $\beta_0, \beta_1, \beta_2 > 0$ are parameters. As this function has bounded support, by Corollary 3 this cannot be a valid covariance function on an arbitrary graph with Euclidean edges (or an arbitrary linear network). However, as remarked in the paragraph proceeding Theorem 4, this does not preclude the positive definiteness of $C(t)$ on a particular *fixed* tree graph. Indeed, Theorem 5 can be invoked to imply that when $\mathcal{G}$ is a Euclidean tree with $m \ge 3$ leaves, then $C(t)^\alpha$ is positive semidefinite (with respect to $d_{R,\mathcal{G}} = d_{G,\mathcal{G}}$) for any $\alpha \ge 2\lceil m/2 \rceil - 1$.

APPENDIX: PROOFS

A.1. Proof of Proposition 5. To verify Proposition 5, recall Definition 3. 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 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*

$$\|f\|_{\mathcal{F}}^2 = f(u_o)^2 + \sum_{e\in\mathcal{E}(\mathcal{G})} \int_{\underline{e}}^{\overline{e}} f_e'(t)^2 \,\mathrm{d}t. \tag{24}$$

PROOF. From (19) we obtain (24). 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 semidefinite over $f \in \mathcal{F}$. □

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

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

where $d(u)$ is defined in (12). It will be convenient to denote the operations $f \to f_\mu$ and $f \to f_{e,r}$ with operator notation $\mathscr{P}_\mu : \mathcal{F} \to \mathcal{F}$ and $\mathscr{P}_e : \mathcal{F} \to \mathcal{F}$ 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 $\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_{\mathcal{F}}|_{\mathscr{P}_\mu\mathcal{F}\times\mathscr{P}_\mu\mathcal{F}}$ and $\langle\cdot,\cdot\rangle_{e,r} = \langle\cdot,\cdot\rangle_{\mathcal{F}}|_{\mathscr{P}_e\mathcal{F}\times\mathscr{P}_e\mathcal{F}}$.

LEMMA 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*:

$$\mathcal{F} = \mathscr{P}_\mu\mathcal{F} \oplus \bigoplus_{e\in\mathcal{E}(\mathcal{G})} \mathscr{P}_e\mathcal{F}. \tag{25}$$

PROOF. This is straightforwardly verified as soon as it is noted that $\mathscr{P}_\mu$ and $\mathscr{P}_e$ are selfadjoint operators which follows from the fact that

$$\big[(f_\mu)_e\big]'(t) = \frac{f_e(\overline{e}) - f_e(\underline{e})}{\mathrm{len}(e)}, \quad e \in \mathcal{E}(\mathcal{G}), t \in (\underline{e}, \overline{e}). \tag{26}$$

□

LEMMA 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* (9);
- H_e *denote the infinite dimensional Hilbert space of absolutely continuous functions* $f : [\underline{e}, \overline{e}] \to \mathbb{R}$ *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* (13)). *Its inner product has simplified form for all* $f, g \in \mathscr{P}_\mu\mathcal{F}$:

$$\langle f, g\rangle_\mu = f(u_o)g(u_o) + \sum_{e\in\mathcal{E}} \frac{(f_e(\overline{e}) - f_e(\underline{e}))(g_e(\overline{e}) - g_e(\underline{e}))}{\mathrm{len}(e)}. \tag{27}$$

(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* (15)). *Its inner product has simplified form for all* $f, g \in \mathscr{P}_e\mathcal{F}$:

$$\langle f, g\rangle_{e,r} = \int_{\underline{e}}^{\overline{e}} f_e'(t)g_e'(t)\,\mathrm{d}t. \tag{28}$$

(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* (16)). *Its inner product has simplified form for all* $f, g \in \mathcal{F}$:

$$\langle f, g\rangle_{\mathcal{F}} = \langle f_\mu, g_\mu\rangle_\mu + \sum_{e\in\mathcal{E}} \langle f_{e,r}, g_{e,r}\rangle_{e,r}. \tag{29}$$

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,

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

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

$$\begin{aligned} \langle z, w\rangle_L &= z(u_o)w(u_o) + \frac{1}{2}\sum_{u\sim v} \frac{(z(u) - z(v))(w(u) - w(v))}{d_{G,\mathcal{G}}(u, v)} \quad \text{(by (9))} \\ &= f_\mu(u_o)g_\mu(u_o) + \frac{1}{2}\sum_{u\sim v} \frac{(f_\mu(u) - f_\mu(v))(g_\mu(u) - g_\mu(v))}{\mathrm{len}(e)} \\ &= \langle f_\mu, g_\mu\rangle_\mu \quad \text{(by (26))}, \end{aligned}$$

where the above sums are over adjacent $u, v \in \mathcal{V}$. This establishes (27) 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 $u \in \mathcal{G}$,

$$
\begin{aligned}
f_\mu(u) &= [\mathscr{P}_\mu f_\mu](u) \quad \bigl(\text{since } \mathscr{P}_\mu^2 = \mathscr{P}_\mu\bigr) \\
&= \bigl(1-d(u)\bigr) f_\mu(\underline{u}) + d(u) f_\mu(\overline{u}) \\
&= \bigl(1-d(u)\bigr) z(\underline{u}) + d(u) z(\overline{u}) \\
&= \bigl(1-d(u)\bigr)\bigl\langle z, L^{-1}(\cdot,\underline{u})\bigr\rangle_L + d(u)\bigl\langle z, L^{-1}(\cdot,\overline{u})\bigr\rangle_L \\
&= \bigl\langle z, \underbrace{\bigl(1-d(u)\bigr)L^{-1}(\cdot,\underline{u}) + d(u)L^{-1}(\cdot,\overline{u})}_{:=R_L(\cdot,u)}\bigr\rangle_L,
\end{aligned}
\tag{30}
$$

where the function $R_L(\cdot,u)$ is a member of $\mathbb{R}^{\mathcal{V}}$. By (13) we can simply linear interpolate $R_L(\cdot,u)$ to find the corresponding member in $\mathscr{P}_\mu\mathcal{F}$ as follows:

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

Therefore, by (30)

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

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 $\mathscr{H}_e := \{f \in C([\underline{e},\overline{e}]) : f' \in L^2([\underline{e},\overline{e}]), f(\underline{e}) = 0\}$ 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

$$\tilde{R}_e(s,t) := (s-\underline{e}) \wedge (t-\underline{e}) - \frac{(s-\underline{e})(t-\underline{e})}{\overline{e}-\underline{e}}, \quad \underline{e} < s,t < \overline{e}. \tag{31}$$

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 (15)

$$R_e(u,v) = \begin{cases} \tilde{R}_e\bigl(\varphi_e(u),\varphi_e(v)\bigr) & \text{if } u,v \in e, \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

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

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

Thereby, (B) is verified.

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

PROOF OF PROPOSITION 5. Lemma 3 establishes that $R_{\mathcal{G}}$ is the reproducing kernel for $(\mathcal{F}, \langle\cdot,\cdot\rangle_{\mathcal{F}})$. Since $R_{\mathcal{G}}$ is also the covariance function of $Z_{\mathcal{G}}$, we have that $(\mathcal{F}, \langle\cdot,\cdot\rangle_{\mathcal{F}})$ is *the* RKHS associated to $Z_{\mathcal{G}}$ (see Wahba (1990)). To prove (20), we use standard Hilbert space arguments to show

$$\operatorname{var}\bigl(Z_{\mathcal{G}}(u) - Z_{\mathcal{G}}(v)\bigr) = \sup_{f\in\mathcal{F}}\bigl\{\bigl(f(u)-f(v)\bigr)^2 : \|f\|_{\mathcal{F}} \le 1\bigr\} \tag{33}$$

the left-hand side being the definition of $d_{R,\mathcal{G}}(u,v)$. 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

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

Note that the function $f_o \in \mathcal{F}$ defined by

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

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

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

This proves (33) and, hence, (20).

To show (21), notice that the definition of inner product on $\mathcal{F}$, in (19), implies that $\langle f, \mathbf{1}\rangle_{\mathcal{F}} = f(u_o)$ where $\mathbf{1}$ denotes the member of $\mathcal{F}$ which has constant value 1 over all points on $\mathcal{G}$. Now, the reproducing kernel property of $R_{\mathcal{G}}$ implies $\langle f, \mathbf{1}\rangle_{\mathcal{F}} = \langle f, R_{\mathcal{G}}(\cdot,u_o)\rangle_{\mathcal{F}}$ for all $f \in \mathcal{F}$. Therefore, $R_{\mathcal{G}}(\cdot,u_o) = \mathbf{1}$ and hence $R_{\mathcal{G}}(u,u_o) = 1$ for all $u \in \mathcal{G}$. Finally, notice

$$d_{R,\mathcal{G}}(u,u_o) = R_{\mathcal{G}}(u,u)+R_{\mathcal{G}}(u_o,u_o)-2R_{\mathcal{G}}(u,u_o) = R_{\mathcal{G}}(u,u)-1,$$

which gives

$$\begin{aligned}d_{R,\mathcal{G}}(u,v) &= R_{\mathcal{G}}(u,u)+R_{\mathcal{G}}(v,v)-2R_{\mathcal{G}}(u,v)\\ &= 2+d_{R,\mathcal{G}}(u,u_o)+d_{R,\mathcal{G}}(v,u_o)-2R_{\mathcal{G}}(u,v).\end{aligned}$$

This proves (21), as was to be shown. □

A.2. Proofs of Propositions 2, 3 and 4. We start by verifying Proposition 3, as it is used to prove Proposition 2.

PROOF OF PROPOSITION 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 3 are identical; however, $\mathcal{G}$ and $\mathcal{G}'$ induce different inner products on $\mathcal{F}$, denoted $\langle\cdot,\cdot\rangle_{\mathcal{F}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{F},\mathrm{split}}$, respectively. By Proposition 5, $d_{R,\mathcal{G}}$ and $d_{R,\mathcal{G}'}$ are completely determined by their inner products $\langle f,g\rangle_{\mathcal{F},\mathrm{split}}$ 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

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

since the splitting operation on $e \in \mathcal{E}(\mathcal{G})$ at $u \in e$ only affects the term corresponding to e in (19). By Proposition 5, to show $d_{R,\mathcal{G}} = d_{R,\mathcal{G}'}$, it will be sufficient to show $\langle f,g\rangle_{\mathcal{F}} = \langle f,g\rangle_{\mathcal{F},\mathrm{split}}$ 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

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

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. □

PROOF OF PROPOSITION 2. In the literature on resistance networks and metrics (see, e.g., Jorgensen and Pearse (2010), Kigami (2003)), 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

$$\begin{aligned}&d_{\text{eff}}(u, v) \\ &\quad = \sup_{z \in \mathbb{R}^{\mathcal{V}(\mathcal{G})}} \left\{ \big(z(u) - z(v)\big)^2 : \frac{1}{2} \sum_{u_1 \sim u_2} c(u_1, u_2)\big(z(u_1) - z(u_2)\big)^2 \le 1 \right\}\end{aligned} \tag{34}$$

(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 (20) and that we have defined c by (7). Also, by Lemma 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 (20) 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 3(A), when $u, v \in \mathcal{V}(\mathcal{G})$, $d_{R,\mathcal{G}}(u, v)$ is equal to

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

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 (Proposition 3) $d_{R,\mathcal{G}}$ is a metric on $\mathcal{G}$ as well. □

PROOF OF PROPOSITION 4. Proposition 2 and the theory of electrical networks imply

$$d_{R,\mathcal{G}}(u,v) \le d_{G,\mathcal{G}}(u,v), \quad u,v \in \mathcal{V}(\mathcal{G}), \tag{35}$$

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 Proposition 3) implies that (35) extends to any additional finite collection of edge points. Thereby, (17) of Proposition 4 is verified, where the *if and only if* follows, since the tree property of $\mathcal{G}$ is also invariant to edge splitting.

To show (18), suppose $\mathcal{G}$ is a Euclidean cycle with circumference ω. Let $u, v \in \mathcal{G}$ be arbitrary and $s \in \mathcal{G}$ be the polar opposite of the midpoint of the geodesic path connecting u to v. By a sequence of edge splits and merges, we may construct a new graph $\mathcal{G}'$, equaling $\mathcal{G}$ as a point set, but with vertices $\{u, v, s\}$ and edges connecting $u \sim v$, $v \sim s$ and $u \sim s$ with corresponding edge lengths $d_{G,\mathcal{G}}(u,v)$, $d_{G,\mathcal{G}}(v,s)$ and $d_{G,\mathcal{G}}(u,s)$, respectively. Notice that the L matrix for $\mathcal{G}'$, constructed via (8), has a particularly simple inverse given by

$$\begin{aligned} L^{-1} &= \begin{pmatrix} c_1+c_2 & -c_1 & -c_2 \\ -c_1 & c_1+c_3 & -c_3 \\ -c_2 & -c_3 & c_2+c_3+1 \end{pmatrix}^{-1} \\ &= \frac{1}{b}\begin{pmatrix} b+c_1+c_3 & b+c_1 & b \\ b+c_1 & b+c_1+c_2 & b \\ b & b & b \end{pmatrix}, \end{aligned}$$

where $c_1 := 1/d_{G,\mathcal{G}}(u,v)$, $c_2 := 1/d_{G,\mathcal{G}}(u,s)$ and $c_3 := 1/d_{G,\mathcal{G}}(v,s)$ are the edge conductances of $\mathcal{G}'$ and $b := c_1c_2 + c_1c_3 + c_2c_3$. Now, since $u, v \in \mathcal{V}(\mathcal{G}')$,

$$\begin{aligned} d_{R,\mathcal{G}'}(u,v) &= L^{-1}(u,u) + L^{-1}(v,v) - 2L^{-1}(u,v) \\ &= \frac{b+c_1+c_3}{b} + \frac{b+c_1+c_2}{b} - 2\frac{b+c_1}{b} \\ &= d_{G,\mathcal{G}}(u,v) - \frac{d_{G,\mathcal{G}}(u,v)^2}{\omega}, \end{aligned}$$

which, along with the fact that $d_{R,\mathcal{G}}(u,v) = d_{R,\mathcal{G}'}(u,v)$ by Proposition 3, completes the proof of (18). □

Acknowledgments. We thank Heidi Søgaard Christensen, James Sharpnack, Adrian Baddeley and Gopalan Nair for useful comments and illuminating discussions. We are very grateful to an anonymous referee who pointed out that the work of Cambanis, Keener and Simons (1983), Berens and Xu (1997) and Gneiting (1998a) can be used to prove Theorems 4 and 5.

The first author was supported in part by NSF Grants DMS-1252795, DMS-1812199 and a UC Davis Chancellor's Fellowship.

The second and third authors were 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.

The third author was supported by the Australian Research Council, Discovery Grant DP130102322.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

BADDELEY, A. J., RUBAK, E. and TURNER, R. (2015). *Spatial Point Patterns: Methodology and Applications with R. Interdisciplinary Statistics*. CRC Press/CRC, Boca Raton, FL.

BADDELEY, A., NAIR, G., RAKSHIT, S. and MCSWIGGAN, G. (2017). "Stationary" point processes are uncommon on linear networks. *Stat* **6** 68–78. MR3613182 https://doi.org/10.1002/sta4.135

BERENS, H. and XU, Y. (1997). l-1 summability of multiple Fourier integrals and positivity. *Math. Proc. Cambridge Philos. Soc.* **122** 149–172. MR1443593 https://doi.org/10.1017/S0305004196001521

BERG, C. (2008). Stieltjes–Pick–Bernstein–Schoenberg and their connection to complete monotonicity. In *Positive Definite Functions. From Schoenberg to Space–Time Challenges* (J. Mateu and E. Porcu, eds.) 15–45. Univ. Jaume I, Castellon.

BERG, C., MATEU, J. and PORCU, E. (2008). The Dagum family of isotropic correlation functions. *Bernoulli* **14** 1134–1149. MR2543589 https://doi.org/10.3150/08-BEJ139

CAMBANIS, S., KEENER, R. and SIMONS, G. (1983). On α-symmetric multivariate distributions. *J. Multivariate Anal.* **13** 213–233. MR0705548 https://doi.org/10.1016/0047-259X(83)90023-4

CRESSIE, N. and MAJURE, J. J. (1997). Spatio-temporal statistical modeling of livestock waste in streams. *J. Agric. Biol. Environ. Stat.* **2** 24–47. MR1812249 https://doi.org/10.2307/1400639

CRESSIE, N., FREY, J., HARCH, B. and SMITH, M. (2006). Spatial prediction on a river network. *J. Agric. Biol. Environ. Stat.* **11** 127–150.

DEZA, M. M. and LAURENT, M. (1997). *Geometry of Cuts and Metrics. Algorithms and Combinatorics* **15**. Springer, Berlin. MR1460488 https://doi.org/10.1007/978-3-642-04295-9

GARDNER, B., SULLIVAN, P. and LEMBO, A. JR (2003). Predicting stream temperatures: Geostatistical model comparison using alternative distance metrics. *Canadian Journal of Fisheries and Aquatic Sciences* **60** 344–351.

GNEITING, T. (1998a). On α-symmetric multivariate characteristic functions. *J. Multivariate Anal.* **64** 131–147. MR1621926 https://doi.org/10.1006/jmva.1997.1713

GNEITING, T. (1998b). Simple tests for the validity of correlation function models on the circle. *Statist. Probab. Lett.* **39** 119–122. MR1652540 https://doi.org/10.1016/S0167-7152(98)00042-X

GNEITING, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327–1349. MR3102554 https://doi.org/10.3150/12-BEJSP06

JAYASUMANA, S., HARTLEY, R., SALZMANN, M., LI, H. and HARANDI, M. (2013). Kernel methods on the Riemannian manifold of symmetric positive definite matrices. In *Computer Vision and Pattern Recognition* (*CVPR*), 2013 *IEEE Conference on* 73–80. IEEE Press, New York.

JORGENSEN, P. E. T. and PEARSE, E. P. J. (2010). A Hilbert space approach to effective resistance metric. *Complex Anal. Oper. Theory* **4** 975–1013. MR2735315 https://doi.org/10.1007/s11785-009-0041-1

KIGAMI, J. (2003). Harmonic analysis for resistance forms. *J. Funct. Anal.* **204** 399–444. MR2017320 https://doi.org/10.1016/S0022-1236(02)00149-0

KLEIN, D. J. and RANDIĆ, M. (1993). Resistance distance. *J. Math. Chem.* **12** 81–95. MR1219566 https://doi.org/10.1007/BF01164627

LYONS, R. and PERES, Y. (2016). *Probability on Trees and Networks. Cambridge Series in Statistical and Probabilistic Mathematics* **42**. Cambridge Univ. Press, New York. MR3616205 https://doi.org/10.1017/9781316672815

MILLER, K. S. and SAMKO, S. G. (2001). Completely monotonic functions. *Integral Transforms Spec. Funct.* **12** 389–402. MR1872377 https://doi.org/10.1080/10652460108819360

MØLLER, J., SYVERSVEEN, A. R. and WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25** 451–482. MR1650019 https://doi.org/10.1111/1467-9469.00115

MØLLER, J. and WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes. Monographs on Statistics and Applied Probability* **100**. CRC Press/CRC, Boca Raton, FL. MR2004226

OKABE, A. and SUGIHARA, K. (2012). *Spatial Analysis Along Networks: Statistical and Computational Methods*. Wiley, Chichester.

SAITOH, S. (1997). *Integral Transforms, Reproducing Kernels and Their Applications. Pitman Research Notes in Mathematics Series* **369**. Longman, Harlow. MR1478165

SCHOENBERG, I. J. (1935). Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert" [MR1503246]. *Ann. of Math.* (2) **36** 724–732. MR1503248 https://doi.org/10.2307/1968654

SCHOENBERG, I. J. (1938a). Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* **44** 522–536. MR1501980 https://doi.org/10.2307/1989894

SCHOENBERG, I. J. (1938b). Metric spaces and completely monotone functions. *Ann. of Math.* (2) **39** 811–841. MR1503439 https://doi.org/10.2307/1968466

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

VER HOEF, J. M. and PETERSON, E. E. (2010). A moving average approach for spatial statistical models of stream networks. *J. Amer. Statist. Assoc.* **105** 6–18. MR2757185 https://doi.org/10.1198/jasa.2009.ap08248

VER HOEF, J. M., PETERSON, E. and THEOBALD, D. (2006). Spatial statistical models that use flow and stream distance. *Environ. Ecol. Stat.* **13** 449–464. MR2297373 https://doi.org/10.1007/s10651-006-0022-8

VON NEUMANN, J. and SCHOENBERG, I. J. (1941). Fourier integrals and metric geometry. *Trans. Amer. Math. Soc.* **50** 226–251. MR0004644 https://doi.org/10.2307/1989957

WAHBA, G. (1990). *Spline Models for Observational Data*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. SIAM, Philadelphia, PA. MR1045442 https://doi.org/10.1137/1.9781611970128

WELLS, J. H. and WILLIAMS, L. R. (1975). *Embeddings and Extensions in Analysis*. Springer, New York. MR0461107

YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions*. *Vol. II*. *Springer Series in Statistics*. Springer, New York. MR0915557