| Statistica Sinica Preprint No: SS-2024-0345 | |
|---|---|
| **Title** | Temporally-evolving Generalised Networks and Their Reproducing Kernels |
| **Manuscript ID** | SS-2024-0345 |
| **URL** | http://www.stat.sinica.edu.tw/statistica/ |
| **DOI** | 10.5705/ss.202024.0345 |
| **Complete List of Authors** | Tobia Filosi,<br>Claudio Agostinelli and<br>Emilio Porcu |
| **Corresponding Authors** | Emilio Porcu |
| **E-mails** | georgepolya01@gmail.com |

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



# Temporally-Evolving Generalised Networks and their Kernels

Tobia Filosi, Claudio Agostinelli and Emilio Porcu

*University of Trento, University of Trento, Khalifa University and ADIA Lab*

*Abstract:* This paper considers time-evolving generalised network, intended as networks where (i) the edges connecting the nodes are nonlinear, (ii) stochastic processes are continuously indexed over both vertices and edges and (iii) the topology is allowed to change over time, that is: vertices and edges can disappear at subsequent time instants and edges may change in shape and length. Topological structures satisfying (i) and (ii) are usually represented through special classes of graphs, termed graphs with Euclidean edges. We build a rigorous mathematical framework for time-evolving networks. We consider both cases of linear and circular time, where, for the latter, the generalised network exhibits a periodic structure. Our findings allow to illustrate pros and cons of each setting. Our approach allows to build proper semi-distances for the temporally-evolving topological structures of the networks. Generalised networks become semi-distance spaces whenever equipped with semi-distances. Our final effort is then devoted to guiding the reader through the appropriate choice of classes of functions that allow to build random fields on the time-evolving networks, via their kernels, that are composed with the temporally-evolving semi-distances topological structure.

*Key words and phrases:* Generalised networks, reproducing kernels, semi-distance spaces, time-evolving graphs.

## 1. Introduction

Data indexed on networks have increasingly attracted the interest of the statistical and machine learning communities. The driver for such an interest is the need to represent complex systems where quantifying the interactions between components covers a fundamental importance. Networks provide an, albeit abstract, flexible framework that

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

may describe and sometimes reveal relationships and dependencies among variables stemming from diverse applications, ranging from transportation infrastructures and communication systems to biological structures and climate-related processes. Within the realm of space-time statistics, networks naturally arise in several applications, since they allow to extend the classic modelling stochastic processes machinery over irregular, yet structured, and often non-Euclidean spaces (Borovitskiy et al., 2023; Bolin et al., 2024).

Traditionally, the statistical literature has focused on data defined over static networks, where the topology is assumed fixed throughout the analysis (Tsonis and Roebber, 2004). Such an assumption is deemed unrealistic for many applications where the underlying system evolves dynamically over time. In several examples, the connectivity structure changes over time: edges may appear or vanish; nodes may emerge or disappear. Further, the geometry governing the relationships between entities might change as a function of time. As a consequence, the temporally-evolving nature of many networks poses challenges to modelling, inference and prediction.

There has been some work related to purely spatial stochastic processes continuously defined over networks or for space-time processes where the network topology does not change as a function of time (Anderes et al., 2020; Bolin et al., 2024; Porcu et al., 2023).

The attention has been specially devoted to Gaussian processes. One group of authors (Anderes et al., 2020; Porcu et al., 2023, 2022) directly models the covariance function as a function of some *distance* that is properly defined over a network, while another group of authors (Bolin and Lindgren, 2011; Bolin et al., 2024) obtains valid covariance functions on the basis of a given class of SPDEs. Both approaches have positive features. The SPDEs group works with more general topologies that allow, for instance, for multiple edges connecting any two vertices of the network; however, this point is not very relevant as the results in Anderes et al. (2020) imply that multiple edges can be achieved

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

at the expense of adding arbitrary degree-two vertices over the network, without altering its intrinsic topology. Further, Bolin et al. (2025) achieve a class of Gaussian random fields having a covariance function that is arbitrarily differentiable at the origin, with the advantages implied by such property in terms of kriging prediction. The approach by Anderes et al. (2020) can instead count on generality, as a wealth of parametric examples is available thanks to this approach. However, none of the covariance functions proposed in Anderes et al. (2020) is differentiable at the origin. Another advantage of working with covariance functions is that Porcu et al. (2023) provide a method based on direct construction that allows to create several parametric classes of nonseparable space-time covariance functions.

However, none of these approaches considers topologies that can evolve over time. Several real-life applications motivate this research field. For instance, in neuroscience, functional brain connectivity evolves throughout development, learning, and disease progression (Medaglia et al., 2015; Uddin and Karlsgodt, 2018); in environmental science, river networks change due to erosion, sediment transport, or climatic shifts (James and Roulet, 2007; Masselink et al., 2017); in engineering, transportation and energy infrastructures adapt through expansion, failure, or technological advances (Newman et al., 2011).

A recent contribution (Porcu et al., 2025) classifies the role of networks in climate data into three paradigms: networks *of* data, networks *over* which data are defined, and networks *for* data. Networks *of* data are constructed using similarity (synchronous) measures among observations. Networks *for* data pertain to the use of network-based models in machine learning, such as graph neural networks. Finally, networks *over* data—the focus of this paper, called geophysical networks therein—describe domains where the network topology is pre-defined, such as river systems or transportation grids. Within the third paradigm, the temporal evolution of the network remains a critical and

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

unexplored challenge in statistical modelling.

This paper works under the paradigm of covariance-based constructions for Gaussian processes defined on temporally-evolving networks. There are several intricacies inherent to this challenge: the notion of distance —a cornerstone in defining covariance structures— is itself dynamic, as paths between points may lengthen, shorten, or be redefined over time. Second, the integration of temporal dynamics into spatial structures is a non-trivial task. Separable covariance models treating time and space independently are definitely inadequate to describe scenarios where spatial distances depend explicitly on time horizon. Third, covariance functions must be positive definite, and such a task entails a fair amount of mathematical work.

We provide a comprehensive framework for modelling stochastic processes over temporally-evolving networks, while extending the existing theory of graphs with Euclidean edges (Anderes et al., 2020) to temporal dynamics. Our constructions accounts for either linear and periodic temporal evolutions, reflecting systems that evolve arbitrarily or cyclically over time. Our first argument to advocate in favour of a periodic construction is that it suits perfectly to several real-world phenomena, where both linear-time evolution (*e.g.* long-term trends) and cyclic oscillations (*e.g.* seasonal components) might happen. To make an example, consider temperatures in a given geographical area: apparently there might be strong correlations between: (i) contiguous spatial points at a given time, which are represented by means of spatial edges; (ii) the same points considered at contiguous times, which are represented by means of temporal edges between temporal layers and (iii) the same points considered at the same periods of the year, which are considered in the model as they are exactly the same point in the temporally evolving graph. In addition, in many real-world applications, the network underlying a system is only partially observable. As a consequence, it could be hard or impossible to specify the whole time-evolving (not periodic) network in cases of long

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

time series. The periodic assumption, when all in all reasonable, may be a great help in this circumstance as well.

Our method embeds temporal dynamics through the addition of temporal edges into the network structure, and allows for semi-metrics that capture both spatial and temporal proximities in a unified manner. Clearly, the semi-metrics cover a fundamental part to provide the new covariance structures. Our contribution is not a mere technical generalisation, but responds to concrete statistical needs: in hydrology, the dynamic connectivity of catchments affects the propagation of pollutants or the prediction of flood events (James and Roulet, 2007); in neuroscience, evolving functional networks influence the interpretation of longitudinal brain imaging studies (Medaglia et al., 2015); in engineering, the resilience of infrastructures under dynamic conditions requires models that incorporate changing connectivity patterns Newman et al. (2011).

To illustrate how our framework may apply to a concrete scientific problem, consider longitudinal brain connectivity studies using fMRI data (Damaraju et al., 2014). Within this framework, dynamic functional connectivity (dFC) is associated with evolving patterns of interaction between brain regions through the so-called time-resolved correlation networks (Hutchison et al., 2013; Preti et al., 2017). Existing models make use of sliding-window correlations, or of vector autoregressive approaches applied to static graphs (Lindquist et al., 2014), which fail to account for continuous topological changes such as the emergence or dissolution of edges, or shifts in network topology.

By contrast, our framework models the data via a temporally-evolving graph: each layer represents the functional network at a given time and temporal edges can link corresponding nodes (or regions) across successive layers with weights that reflect continuity or similarity. We can then define a Gaussian process over this evolving network using resistance-based semi-distances (see Section 4), allowing us to model neural activity as a continuous stochastic process indexed by a time-varying domain.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

For statisticians, this approach may pave the way for inference tasks such as predicting connectivity or activity on unobserved regions or times, quantifying uncertainty across layers and detecting anomalous structural changes, which are key for understanding developmental trajectories or disease progression (Medaglia et al., 2015).

Another example may come from modelling vascular networks in tumour imaging. Tumour angiogenesis — where new vasculature forms in response to tumour growth — can be seen as a dynamic stochastic process, with vessel topology and perfusion patterns evolving over time (Folkman, 2002; Kim et al., 2016). By modelling perfusion as a Gaussian process over the evolving vascular graph, our approach may allow for interpolation at unsampled sites, identification of emergent or collapsing vascular pathways, and quantification of structural and functional uncertainty. Hence, the new framework proposed here provides potentially valuable insights for prognosis or therapy planning in oncology.

As a final contextualisation of our method, urban transportation systems are inherently dynamic: the topology of routes, road availabilities, and patterns associated with passenger flows change over time: construction, accidents, policy interventions (*e.g.*, road closures or congestion pricing), or daily temporal rhythms such as rush hours (Lilleborge et al., 2025). In metropolitan areas, for instance, public transit networks — buses, trains, and subways — have a temporally dynamical connectivity with nodes (stations, stops) and edges (routes or paths) are active only at specific time intervals or vary in travel time and reliability throughout the day (Zheng et al., 2014; Rico et al., 2021).

According to our framework, the transportation system may be modelled as a time-indexed sequence of graphs, where each layer corresponds to the transportation network at a given time (*e.g.*, every 15 minutes). Edges across temporal layers translate into continuity or transitions in accessibility, *e.g.*, based on predicted travel times or real-time availability of connections. Our dynamical covariance functions can model quantities

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

such as commuter flow, vehicle congestion, or service reliability and aligns with the realities of smart-city infrastructure and urban analytics.

The contributions of this work are organised as follows. Section 2 recalls the main mathematical objects that will be used. Section 3 builds the skeleton of our construction, *i.e.* time-evolving graphs, which are exploited in Sections 4 and 5, where time-evolving graphs with Euclidean edges are defined for the linear time and circular time cases, respectively. Section 6 illustrates how it is possible to build kernels on such a structure and present some examples, whilst Section 7 presents some numerical experiments. Finally, Section 8 provides conclusions and final remarks. In addition, in the Supplementary Material Section S1 we recall some standard mathematical definitions used throughout. Section S2 contains some additional insight about the definition of the processes. Section S3 provides complete details on the construction of the RKHS introduced in Section 6. In Section S4, we present a marginal result about semi-distances defined via variograms. Furthermore, we defer proofs to Section S5 for a neater exposition of the main text. Finally, Section S6 presents some additional plots.

## 2. Mathematical background

This material is largely expository and provides the mathematical background and notation used in the next sections. For the unfamiliar reader, Section S1 of the Supplementary Material provides basic definitions and concepts used in network theory.

### 2.1 Gaussian random fields over semi-distance spaces

Let us begin with a brief introduction about Gaussian random fields (Stein, 1999). Let $X$ be a non-empty set and let $k : X \times X \to \mathbb{R}$ be a function. Then $k$ is a *positive semi-definite* function (or a *kernel*, or a *covariance function*) if and only if it is symmetric and, for all $n \in \mathbb{N}^+$, $x_1, \ldots, x_n \in X$ and $a_1, \ldots, a_n \in \mathbb{R}$, $\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j k(x_i, x_j) \geq 0$. If, in

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



addition, whenever the above relation is an equality, then necessarily $a_1 = \cdots = a_n = 0$, $k$ is *(strictly) positive definite*.

We denote $Z$ a real-valued random field over $X$, *videlicet*: for each $x \in X$, $Z(x)$ is a real-valued random variable. Then $Z$ is called *Gaussian* if, for all $n \in \mathbb{N}^+$ and $x_1, \ldots, x_n \in X$, the random vector $\boldsymbol{Z} := (Z(x_1), \ldots, Z(x_n))^\top$, with $\top$ denoting the transpose operator, follows a $n$-variate Gaussian distribution.

A Gaussian random field $Z$ on $X$ is completely determined by its first two moments: the mean function $\mu_Z : X \to \mathbb{R}$, $x \mapsto \mathbb{E}\{Z(x)\}$, with $\mathbb{E}$ denoting the expectation operator, and the covariance function (kernel) $k_Z : X \times X \to \mathbb{R}$, $(x_1, x_2) \mapsto \mathbb{C}\mathrm{ov}\,(Z(x_1), Z(x_2))$. A necessary and sufficient condition for a function $k_Z$ to be a covariance function (a kernel) of some random field $Z$ is to be symmetric and positive semi-definite.

A mapping $d : X \times X \to \mathbb{R}$ is called *semi-distance* on $X$ (or, equivalently, $(X, d)$ is a *semi-distance space*) if the following conditions hold for each $x, y \in X$: $d(x, y) \geq 0$, $d(x, y) = 0 \iff x = y$ and $d(x, y) = d(y, x)$. In addition, $d$ is called a *distance* on $X$ (or, equivalently, $(X, d)$ is a *distance space*) if it is a semi-distance and the triangle inequality holds, namely, for all $x, y, z \in X$: $d(x, y) + d(y, z) \geq d(x, z)$. The covariance function $k_Z$ is called isotropic for the semi-distance space $(X, d)$ if there exists a mapping $\psi : D_X^d \to \mathbb{R}$ such that $k_Z(x, y) = \psi\{d(x, y)\}$, for $x, y \in X$. Here, $D_X^d := \{d(x_1, x_2) : x_1, x_2 \in X\}$ is the diameter of $X$. Theorem 1 in Section 6 provides a useful tool to construct isotropic kernels on arbitrary domains. For a Gaussian random field $Z$ on $X$, we define its *variogram* $\gamma_Z : X \times X \to \mathbb{R}$ through

$$\gamma_Z(x_1, x_2) := \mathbb{V}\mathrm{ar}\,\{Z(x_1) - Z(x_2)\}, \qquad x_1, x_2 \in X, \tag{2.1}$$

with $\mathbb{V}\mathrm{ar}$ denoting the variance operator. Schoenberg (1942) proves that $\gamma_Z$ is a variogram if and only if the mapping $\exp\{-\gamma_Z(\cdot, \cdot)\}$ is positive definite on $X \times X$.

Let $(X_1, d_1)$ and $(X_2, d_2)$ be two semi-distance spaces. Then, the triple $(X_1 \times X_2, d_1, d_2)$ is called a *product semi-distance space*. Menegatto et al. (2020) define *isotropy* over a

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



product semi-distance space through continuous functions $\psi : D_{X_1}^{d_1} \times D_{X_2}^{d_2} \to \mathbb{R}$ such that, for $(x_1, x_2), (x_1', x_2') \in X_1 \times X_2$,

$$((x_1, x_2), (x_1', x_2')) \mapsto \psi\{d_1(x_1, x_1'), d_2(x_2, x_2')\}, \tag{2.2}$$

is positive definite. The above definition naturally arises from spatio-temporal settings: suppose we have a *static* semi-distance space $(X, d)$ that represents some spatial structure and $(T, d_T)$ representing time, where $T \subseteq \mathbb{R}$ and, usually, $d_T(t_1, t_2) = |t_1 - t_2|$. In such a case, Equation (2.2) can be re-adapted to define kernels. This is the setting adopted by Porcu et al. (2023) and by Tang and Zimmerman (2024).

Later on, we deviate from the literature and we instead consider a distance space $X_t$ that evolves over time, $t \in T$. Hence, our domain is written as $\{(x_t, t) : x_t \in X_t, t \in T\}$, where $t$ describes *time*, and the graph coordinate $x_t$ is constrained on the space $X_t$. Such a framework entails a way more sophisticated construction to equip such a space with a proper distance.

## 2.2 Graphs with Euclidean edges

We start with a formal definition of graphs with Euclidean edges. We slightly deviate from the definition provided by Anderes et al. (2020), for the reasons that will be clarified subsequently. For a definition of graph, see Definition 1 in the Supplementary Material.

**Definition 1** (Graph with Euclidean edges)**.** Consider a simple, connected and weighted graph $G = (V, E, w)$, where $w : E \to \mathbb{R}^+$ represents the weight mapping. Then, $G$ is called a *graph with Euclidean edges* provided that the following conditions hold.

1. Edge sets. Each edge $e \in E$ is associated to the compact segment, also denoted by $e$, $[0, \ell(e)]$, where $\ell(e) := w(e)^{-1}$ may be interpreted as the *length* of the edge $e$.

2. Linear edge coordinates. Each point on an edge $e \in E$, that may be formally seen as a point $u \in [0, \ell(e)]$, where $e = (\underline{u}, \overline{u})$ is uniquely determined by the

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



endpoints $\underline{u}$ and $\overline{u}$ of $e$ and its relative distance $\delta_e(u) := \frac{u}{\ell(e)} = u\, w(e)$ from $\underline{u}$, that is $u = (\underline{u}, \overline{u}, \delta_e(u))$, so that $\underline{u} = (\underline{u}, \overline{u}, 0) = (\overline{u}, \underline{u}, 1)$ and $\overline{u} = (\underline{u}, \overline{u}, 1) = (\overline{u}, \underline{u}, 0)$.

Henceforth, we shall assume the existence of a total order relation on the set of vertices $V$ and that every edge is represented through the ordered pair $(v_1, v_2)$, where $v_1 < v_2$. In particular, for each $u \in e$, the endpoints of $e$, $\underline{u}$ and $\overline{u}$ satisfy the relation $\underline{u} < \overline{u}$. Although this assumption is not essential for the following, we think that in practice it simplifies the exposition, as each point on a given edge $e$ is written in exactly one way and, therefore, there is only one quantity $\delta_e(u)$. See Figure 1 (right) for an illustration.

Notice that our Definition 1 does not require any *distance consistency* opposed to Anderes et al. (2020, Definition 1, (d)). The reason is that our setting does not need a bridge between geodesic and resistance metrics. In addition, we have restricted the space of possible bijections from each edge onto closed intervals to be linear: the main reason is that the focus of this paper is not to explore isometric embeddings, but to provide suitable topological structures evolving over time, and stochastic processes attached to them. We stress that the framework introduced through Definition 1 is more general than linear networks, for at least two reasons: (i) in our framework, the weights of each edge be chosen independently from the others, and (ii) our framework needs no restriction on the network structure, *e.g.*, as shown in Figure 1 (right), edges may cross without sharing the crossing point.

## 2.3 Graph laplacian and resistance metric

The resistance metric has been widely used in graph analysis, as it is more natural than the shortest-path metric when, for instance, considering flows or transport networks, where multiple roads between two given points may share the total flow (Jorgensen and Pearse, 2010). Below we introduce the effective resistance distance for an undirected

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



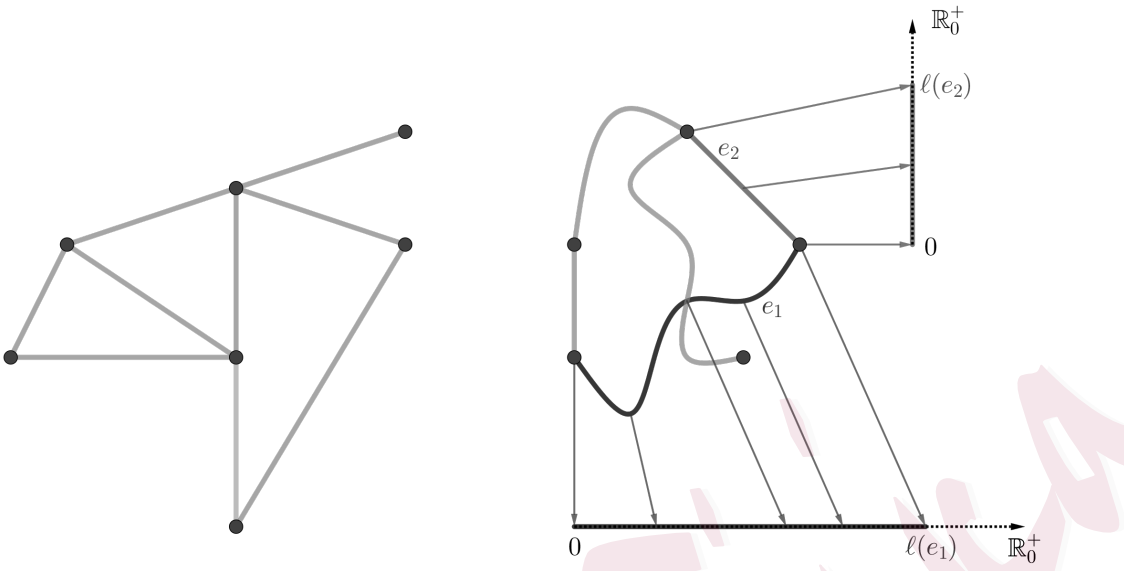



Figure 1: Left: a linear network. Right: a graph with Euclidean edges, where the bijections between the edges $e_1$ and $e_2$ and their respective real segments $[0, \ell(e_1)]$ and $[0, \ell(e_2)]$ are stressed.

and connected graph, we briefly report its definition and its mathematical construction.

Let $G = (V, E, w)$ be a simple, weighted and connected graph and let $W$ its adjacency matrix, that is: $W(v_1, v_2) = w\{(v_1, v_2)\}$, where we set $w\{(v_1, v_2)\} = 0$ whenever $v_1 \not\sim v_2$ (*i.e.*, $v_1$ is not adjacent to $v_2$). In addition, for each node $v \in V$, we define its *degree* as the sum of the weights of the edges adjacent to it. Let $D$ be the degree matrix of $G$, *i.e.* the diagonal matrix where each diagonal element is the degree of the corresponding vertex. Then, the *laplacian matrix* (or simply *laplacian*) of $G$ is the matrix $L := D - W$.

Laplacian matrices enjoy several properties (Devriendt, 2022): they are symmetric, diagonally dominant (thus positive semi-definite) and singular with exactly one null eigenvalue, corresponding to the eigenvector $\mathbf{1}_n$. Furthermore, they have non-positive off-diagonal entries and positive main-diagonal entries.

A graph $G = (V, E)$ is called a *resistor graph* if the edges $e \in E$ represent electrical resistors and the nodes represent contact points. Given a resistor graph, the *effective resistance distance* $R$ between two vertices is defined as the voltage drop between them when injecting one Ampere of current in one and withdrawing one Ampere from the

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

other. Several mathematical formulations of this concept have been provided, and the reader is reminded, among others, to Jorgensen and Pearse (2010, Subsection 2.1). Throughout, we follow Ghosh et al. (2008). Let $G = (V, E)$ be a resistor graph. For each $v_1 \sim v_2 \in V$, let $r(v_1, v_2) \in \mathbb{R}^+$ denote the resistance of the resistor that connects $v_1$ and $v_2$. In addition, for each $v_1, v_2 \in V$, define the weight (which plays the role of the physical conduttance) $w\{(v_1, v_2)\} := 1/r\{(v_1, v_2)\}$ if $v_1 \sim v_2$ and $w\{(v_1, v_2)\} := 0$ if $v_1 \not\sim v_2$. Let $L$ be the laplacian matrix of $G$ with the above-defined weights, $L^+$ its Moore-Penrose generalised inverse (see Definition 2 in the Supplementary Material). Finally let $e_{v_i}$ denote the vector with all zeroes, except a one at position $v_i$. Then the effective resistance distance $R$ between two nodes $v_1$ and $v_2$ is given by $R(v_1, v_2) = (e_{v_1} - e_{v_2})^\top L^+ (e_{v_1} - e_{v_2})$.

## 3. Time-evolving graphs with Euclidean edges

In this Section we provide the construction of time-evolving graphs with Euclidean edges. While providing full specifications below, we start with some description and a graphical representation to give an intuition of how these graphs work. The *rationale* behind this construction is shown in Figure 3: we build a new graph containing all the graphs a different time instants, with additional edges connecting the same nodes at different times. However, the mathematical description of such an object is quite involved as it requires several steps and substantial notation. To guide the reader in following, we provide a sketch of our procedure in Box 2. **Step 1.** is completely general and does not require any topological structure on every *marginal* graph $G_t$, for a given time $t$. Yet, having graphs with Euclidean edges that evolve over time requires some more work, and this fact justifies **Step 2.**, which allows for connectivity, being one of the properties *sine qua non* of a graph with Euclidean edges. **Step 4.** is not mathematically necessary to guarantee the validity of the structure, but it is justified by computational and intuitive reasons as explained throughout.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

**Step 1.** Define a time-evolving graph as a properly defined sequence of graphs indexed by discrete time instants.

**Step 2.** Define *connected equivalent simple* time-evolving graphs by completing a time evolving graph through a set of edges that connect the same nodes at different time instants.

**Step 3.** Over the connected equivalent simple graph, we can now define, for every time $t$, a graph with Euclidean edges, $G_t$.

**Step 4.** Define a time-evolving graph of order 1 to exploit computational advantages.

Box 2: A sketch of our construction.

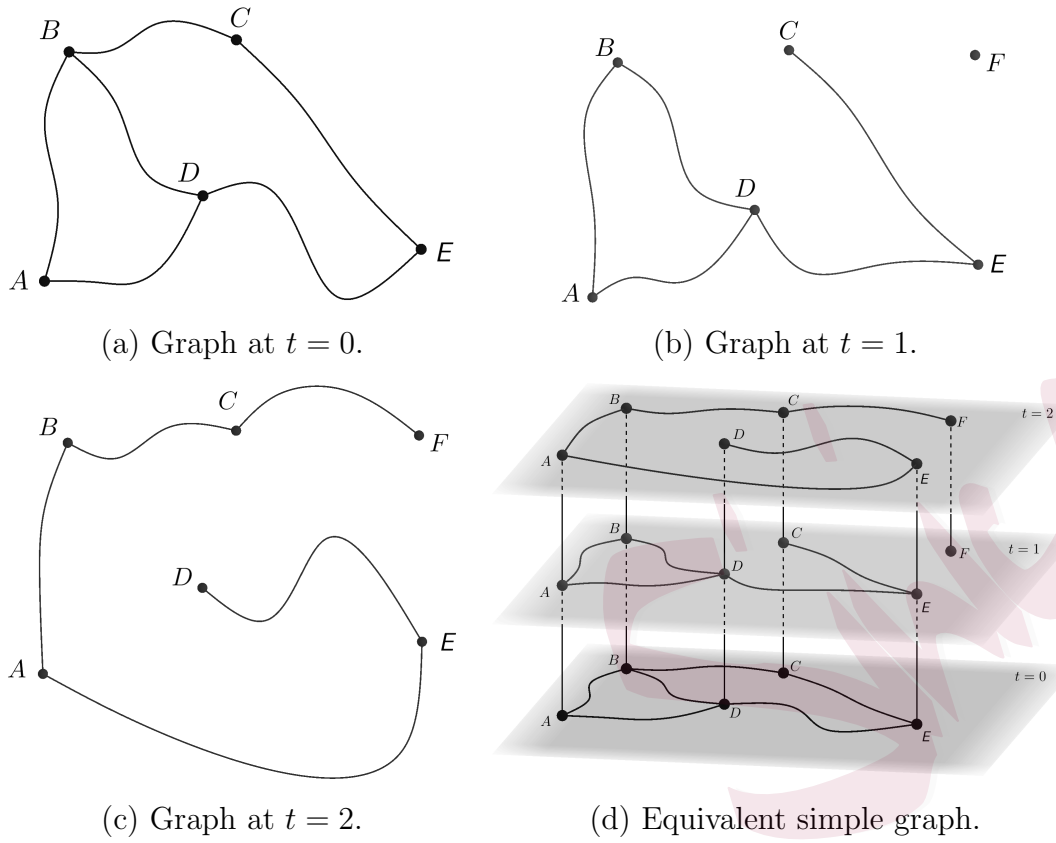
**Step 1.** starts with the definition of *equivalent simple graph*. The intuitive idea behind the construction formalised next is depicted in Figure 3: consider $m$ layers, each representing a different temporal instant (namely a graph $G_t$), and connect them by means of additional intra-time edges, which account for the time-dependency of the graphs.

**Definition 2** (Time-evolving graph)**.** Let $T = \{0, ..., m-1\}$ be a (finite) collection of time instants. To every time instant $t \in T$ we associate a simple undirected and weighted graph $G_t = (V_t, E_t, w_t)$, with $V_t \cap V_{t'} = \emptyset$ whenever $t \neq t'$. For an edge $e_t \in E_t$, the corresponding weight is denoted $w(e_t) := w_t(e_t)$. We use $n_t := |V_t|$ for the number of vertices at time $t$.

Let $G = \{G_0, ..., G_{m-1}\}$ be the associate finite collection of these graphs. Call $V := \bigcup_t V_t$ the set of vertices, $n := |V|$ the total number of vertices, and $E_S := \bigcup_t E_t$ the set of *spatial* edges. Finally, if $v \in V$, whenever convenient we shall write $t(v)$ for the unique value $t$ such that $v \in V_t$.

Let $s : V \to S$ be a mapping from $V$, where $S$ is a set of labels, such that $s(v_1) \neq s(v_2)$ whenever $v_1$ and $v_2$ are two distinct vertices belonging to the same graph $G_t$, $t \in T$. Two vertices $v_1 \neq v_2 \in V$ are considered the same vertex at different times if $s(v_1) = s(v_2)$. We call the triple $\boldsymbol{G} = (T, G, s)$ a *time-evolving graph*. We assume $\boldsymbol{G}$ to be *connected*. If this is not the case, we henceforth consider each connected component independently

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345




(a) Graph at $t = 0$.

(b) Graph at $t = 1$.

(c) Graph at $t = 2$.

(d) Equivalent simple graph.

Figure 3: An example of an equivalent simple graph (bottom-right), with $m = 3$, $S = \{A, B, C, D, E, F\}$, $n_1 = 5$, $n_2 = n_3 = 6$. The coloured edges belong to $E_S$, whilst the black ones belong to $E_T$. The temporal *slices* at time instants $t = 0, 1, 2$ are reported, respectively, at quadrants (a), (b) and (c).

from the others.

While Definition 2 provides a flexible framework to manage graphs that evolve over time, we are going to merge its underlying idea with the one of graph with Euclidean edges presented in Subsection 2.2. **Step 2.** of our procedure intends to *complete* the time-evolving graph as in Definition 2 so to ensure spatio-temporal connectivity.

Let $\boldsymbol{G}$ be a time-evolving graph. We define its *equivalent simple* time-evolving graph, $\widetilde{\boldsymbol{G}} = (V, \widetilde{E})$ as the graph with edges $\widetilde{E} := E_S \cup E_T$, with $E_T$ a set of additional edges (called *temporal* edges throughout) that connect the same nodes at different time instants. More precisely, $E_T$ is a subset of $\{(v_1, v_2) \in V \times V \,:\, s(v_1) = s(v_2),\, t(v_1) \neq t(v_2)\}$. To each new edge $e = (v_1, v_2) \in E_T$ a weight $w(e) > 0$ is assigned, while all the other weights remain unchanged. It is reasonable to choose a weight that depends only on the

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

temporal distance $|t(v_1) - t(v_2)|$. One possibility is to choose $w(e) := \alpha\,|t(v_1) - t(v_2)|^{-r}$, with $\alpha, r > 0$ given scale factors. Although we assume this particular expression with $r := 1$ in all the following examples, we stress that any choice leads to a valid model as long as $w(e) > 0$. Notice that an equivalent simple time-evolving graph is a graph with Euclidean edges.

**Step 2.** ensures that it becomes feasible to assign, to each temporal label $t \in T$, a graph with Euclidean edges to the equivalent simple graph associated with a given time-evolving graph (**Step 3.**). We note that the choice of the temporal edges needs care. Indeed, the set of possible temporal edges for a fixed label $s \in S$ may grow quadratically in the number of considered temporal instants $m$. An interesting class of time-evolving graph that allows to control the structure of temporal edges is the time-evolving graph of order $p$, formalised below.

**Definition 3** (Time-evolving graph of order $p$)**.** Let $p \in \mathbb{N}^+$. A *time-evolving graph of order* $p$ is a time-evolving graph $\boldsymbol{G} = (V, E)$, where

$$E_T \subseteq \{(v_1, v_2) \in V \times V \,:\, s(v_1) = s(v_2),\, 1 \leq |t(v_1) - t(v_2)| \leq p\}\,. \tag{3.1}$$

This class of models restricts the sets of temporal edges so that they cannot link directly nodes more than $p$ time steps away. We stress that mathematical results following hold in the general case, even for the case of non adjacent layers. Yet, being an order $p$ graph simplifies both interpretation and computational aspects. In fact, it allows for a plain representation of an equivalent simple graph. Furthermore, there are non-negligible computational reasons. Indeed, for large networks, order $p$ graphs have a sparse Laplacian matrices (block $(2p+1)$-diagonal) of the associated equivalent simple graph. This entails huge computational savings in both terms of storage and computation. Finally, allowing edges between non-adjacent layers could lead to a huge number of weights $\alpha$'s. In particular, under the assumption of temporally homogeneous weights (the weight of a temporal edges depends only on the temporal distance between

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

the layers it connects), we would have $m-1$ possible weights if order $m-1$ graph is assumed. For a large collection of time instants, this can become computationally unfeasible.

We call a time-evolving graph of order $p$ $\boldsymbol{G}$ *temporally complete* when the set $E_T$ is identically equal to the set in the right hand side of (3.1). Albeit such a property is not required to prove our theoretical results, it is operationally useful as it allows to avoid removing or adding temporal edges.

## 4. Resistance metrics for linear time

We start this section by noting that defining the classic resistance metrics between nodes of the temporally evolving graph is not an issue. Yet, we are dealing with a graph where (semi-)distances should be computed between any pair of points lying continuously over the edges.

For the case of static graphs with Euclidean edges, Anderes et al. (2020) provide an ingenious construction that allows for a suitable continuously-defined metric on the basis of Brownian bridges and their variograms.

The idea is to follow a similar path, by defining a Gaussian process that is continuously indexed over the edges of an equivalent simple connected graph associated with a given time-evolving graph.

Before going into technical details, we present a brief outline. Following Anderes et al. (2020), we are going to define a semi-distance on all the points of the graph, namely its vertices and the points on its edges. To this aim, we define a Gaussian process $Z$ on every point of the time-evolving equivalent simple graph and then *define* the semi-distance between two points as the variogram of such process, *i.e.*, for each $u_1, u_2 \in \widetilde{\boldsymbol{G}}$:

$$d(u_1, u_2) := \gamma_Z(u_1, u_2), \tag{4.1}$$

with $\gamma_Z$ as being defined through (2.1). In such a way, we can directly apply Theorem 1

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



stated in Section 6 to obtain kernels. Here, $Z := Z_V + Z_E$ is the sum of two independent Gaussian processes defined on the equivalent simple graph $\widetilde{\boldsymbol{G}}$. The process $Z_V$ accounts for the structure of the graph (namely its vertices and the weights of its edges) and plays the role of major source of variability, whilst $Z_E$ adds some variability on the edges and accounts for the temporal relationship between the same edge at different times. To have a better intuition about the specific needs for and roles of both $Z_V$ and $Z_E$, the reader is deferred to Remark 1 in the Supplementary Material.

### 4.1 Formal construction of $Z_V$ and $Z_E$

We start by defining the process $Z_V$ at the vertices $V$ as a zero-mean multivariate normal random variable: $Z_V\big|_V \sim \mathcal{N}\left(0, (L^\star)^{-1}\right)$, where $L^\star := L + \boldsymbol{x}\boldsymbol{x}^\top$ with L the laplacian matrix associated to the graph $\widetilde{\boldsymbol{G}}$ and $\boldsymbol{x}$ being a vector such that $\mathbf{1}_n^\top \boldsymbol{x} \neq 0$. The role of the inverse laplacian matrix is explained in Remark 1 in the Supplementary Material, whilst the reason underlying the addition of $\boldsymbol{x}\boldsymbol{x}^\top$ is that $L$ is singular, whilst $L^\star$ is strictly positive definite. This will turn out to be necessary in the construction of the reproducing kernel Hilbert space, outlined in Proposition 7 and delved into in the Supplementary Material. However, we stress that the semi-distance and the resulting covariance functions remain the same for $\boldsymbol{x} = 0$ (thus $L^\star = L$), as shown in Proposition 1 below. Notice that Anderes et al. (2020) chose $\boldsymbol{x}$ as a vector of the canonical base of $\mathbb{R}^n$. The process $Z_V\big|_V$ is then extended to the whole equivalent simple graph via a sheer linear interpolation:

$$Z_V(u) := (1 - \delta_e(u))\, Z_V(\underline{u}) + \delta_e(u) Z_V(\overline{u}), \tag{4.2}$$

where $u = (\underline{u}, \overline{u}, \delta_e(u))$, with $e = (\underline{u}, \overline{u})$ and $\delta_e(u)$ as in Definition 1. The choice is motivated by the fact that the resulting semi-distance is an extension of the classic resistance distance.

**Proposition 1.** *The variogram of $Z_V$, $(u_1, u_2) \mapsto \gamma_{Z_V}(u_1, u_2)$, does not depend on the*

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



*choice of* $\boldsymbol{x}$.

The construction of $Z_E$ is a bit more complex, as $Z_E$ is piecewise defined on a suitable partition of $E$. For each $e = (v_1, v_2) \in E_S$, we define the *lifespan* of $e$, written $\mathrm{ls}(e)$, as the maximal connected set of time instants, $t$, for which the edge $e$ exists. More formally, $\mathrm{ls}(e)$ is defined as the maximal (with respect to the inclusion partial order) subset of $T$ such that: (i) $t(v_1) \in \mathrm{ls}(e)$; (ii) $\mathrm{ls}(e)$ is connected, that is $\forall t_1 < t_2 \in \mathrm{ls}(e)$, $\{t_1, \ldots, t_2\} \subseteq \mathrm{ls}(e)$; and (iii) $\forall t \in \mathrm{ls}(e)$, there exists $(v_1', v_2') \in E_s$ such that $s(v_1') = s(v_1)$, $s(v_2') = s(v_2)$ and $t(v_1') = t(v_2') = t$.

Figure 3 allows to visualise the situation. The lifespan of the edge $(A, B)$ at time $t = 0$ is $\{0, 1, 2\}$; the lifespan of $(C, E)$ at time $t = 1$ is $\{0, 1\}$ and the one of $(B, C)$ at time $t = 2$ is $\{2\}$.

We now define the *life* of $e$ (denoted $\mathrm{lf}(e)$) as the set of edges that represent $e$ at different times and have the same lifespan $\mathrm{ls}(e)$, that is:

$$\mathrm{lf}(e) := \{(v_1', v_2') \in E_S : s(v_1') = s(v_1),\ s(v_2') = s(v_2),\ t(v_1') = t(v_2') \in \mathrm{ls}(e)\}\,.$$

For convenience, we define the life for temporal edges as well: if $e \in E_T$, $\mathrm{lf}(e) := \{e\}$. The set $\{\mathrm{lf}(e) : e \in E_S\}$ forms a partition of all the spatial edges $E_S$, and $\{\mathrm{lf}(e) : e \in E_T\}$ is a partition of $E_T$. For convenience, we set $\Lambda := \left\{\mathrm{lf}(e) : e \in \widetilde{E}\right\}$ the set of all lives and we indicate with $\lambda$ a generic element of $\Lambda$. Finally, we write $u \in \lambda$ whenever the edge containing $u$ belongs to $\lambda$. The main idea is to consider the life of each spatial and temporal edge and define a suitable process on it, being independent from the others. Let us consider a spatial edge $e \in E_S$ and its lifespan $\mathrm{ls}(e)$. Consider now the set $\mathrm{ls}(e) \times [0, 1]$ and define on it a zero-mean Gaussian process $B(t, \delta)$ whose covariance function is given by $k_B\{(t_1, \delta_1), (t_2, \delta_2)\} := k_T(|t_1 - t_2|)\, k_{BB}(\delta_1, \delta_2)$, with $t_1, t_2 \in \mathrm{ls}(e)$ and $\delta_1, \delta_2 \in [0, 1]$. Here $k_T$ is a temporal kernel defined on $\mathbb{N}$ such that $k_T(0) = 1$ and $k_{BB}(\delta_1, \delta_2) := \min(\delta_1, \delta_2) - \delta_1\delta_2$ is the kernel of the standard Brownian bridge on $[0, 1]$. Notice that the spatial marginals of the process $B$ are standard Brownian bridges. We

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



stress that the process $B(t,\delta)$ is only needed for the definition of the process $Z_E$ on $\mathrm{lf}(e)$, denoted $Z_E\big|_{\mathrm{lf}(e)}$, as follows: given an edge $e' = (\underline{u}, \overline{u}) \in \mathrm{lf}(e)$ and given a point $u = (\underline{u}, \overline{u}, \delta)$ on it,

$$Z_E\big|_{\mathrm{lf}(e)}(u) := \sqrt{\ell(e')}\, B(t(\underline{u}), \delta). \tag{4.3}$$

Finally, for each temporal edge $e = [0, \ell(e)] \in E_T$, we define the process $Z_E$ on it as an independent (from both $Z_V$ and $Z_E$ on $E_S$) Brownian bridge on $[0, \ell(e)]$, having covariance function given by $\mathbb{C}\mathrm{ov}\left(Z_E\big|_e(\delta_1), Z_E\big|_e(\delta_2)\right) = \ell(e)\left(\min(\delta_1, \delta_2) - \delta_1\delta_2\right)$ (see an example of realisation in Figure 1). This concludes the construction of the process on the whole set of the edges.

## 4.2 Mathematical properties of the construction

We remind the reader that the process $Z$ is Gaussian, being the sum of two independent Gaussian processes. Hence, the finite dimensional distributions of $Z$ are completely specified through the second order properties, namely the covariance function. The following result provides an analytical expression for the covariance function associated with $Z$.

**Proposition 2.** *Let $u_1, u_2 \in \widetilde{\boldsymbol{G}}$, with $u_i = (\underline{u}_i, \overline{u}_i, \delta_i)$, $i = 1, 2$. Let $Z = Z_V + Z_E$, with $Z_V$ as defined through (4.2) and $Z_E$ as defined in the previous subsection (see 4.3). Then,*

$$\begin{aligned} k_Z(u_1, u_2) = {} & \boldsymbol{\delta}_1^\top \left(L^\star\right)^{-1} \boldsymbol{\delta}_2 \\ & + \mathbb{1}_{\mathrm{lf}(e_1)=\mathrm{lf}(e_2)} \sqrt{\ell(e_1)\ell(e_2)}\, k_T\left\{\left|t(\underline{u}_1) - t(\underline{u}_2)\right|\right\}\left\{\min\left(\delta_1, \delta_2\right) - \delta_1\delta_2\right\}, \end{aligned} \tag{4.4}$$

*where $\boldsymbol{\delta}_i$ ($i \in \{1, 2\}$) are the $n-$dimensional vectors whose entries are given by*

$$\boldsymbol{\delta}_j := \begin{cases} 1 - \delta_i & \text{if } j = \underline{u}_i, \\ \delta_i & \text{if } j = \overline{u}_i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

While noting that this construction is completely general, we also point out that conditional independence properties, whenever aimed, can be achieved through a proper choice of the temporal kernel $k_T$. A reasonable choice for $k_T$ is the correlation function

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



of an autoregressive process of order one: $k_T(h) = \phi^{|h|}$, where $\phi \in (-1, 1)$ is a free parameter and $h \in \mathbb{Z}$ is the lag. Notice that the special case $\phi = 0$, for which $k_T(h) = \mathbb{1}_{h=0}$, corresponds to the static resistance metric provided by Anderes et al. (2020). Figure 1 in the Supllementary Materials depicts some realisations for the process $Z_E$ over an edge for different values of the parameter $\phi$. The parameter $\phi$ for the edges plays a similar role of the parameter $\alpha$ for the nodes: their are both closely related to the inter-dependency of the process at different times. Indeed, $\phi$ measures how much the process $Z_E$ is correlated between two times $t_1, t_2 \in \mathrm{ls}(e)$. Analogously, the value $\alpha$ as weight of the temporal edge $e \in E_T$, is related to the partial correlation of the endpoints of $e$ given everything else. As a consequence, it is natural to choose a high value of $\phi$ for high values of $\alpha$ and vice-versa. We notice that it is natural to choose non-negative values for $\phi$, as we usually expect a non-negative correlation between the values of $Z_E$ for close times. Using equation (4.1), we get the following expression for the semi-distance between any to points $u_1, u_2 \in \widetilde{\boldsymbol{G}}$.

$$d(u_1, u_2) = k_Z(u_1, u_1) + k_Z(u_2, u_2) - 2k_Z(u_1, u_2). \tag{4.6}$$

The formal statement below provides a complete description of the space $(\widetilde{\boldsymbol{G}}, d)$, that is, the time-evolving graph $\widetilde{\boldsymbol{G}}$ equipped with the semi-distance $d$.

**Proposition 3.** *Let $d$ be the mapping defined at (4.6). Then, the pair $(\widetilde{\boldsymbol{G}}, d)$ is a semi-distance space.*

One might ask whether a stronger assertion holds for the pair $(\widetilde{\boldsymbol{G}}, d)$ as defined above. The next statement provides a negative answer. A counterexample is given by the graph depicted in Figure 1 in the Supplementary Material (see the proof of Proposition 4 therein for more details).

**Proposition 4.** *Let $d$ be the mapping defined at (4.6). Then, the pair $(\widetilde{\boldsymbol{G}}, d)$ is not, in general, a metric space.*

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



**Remark 1.** Although in general our extension to the classic resistance distance is not a metric, it retains some of its properties: (i) $(V, d\big|_V)$ is a metric space and (ii) for all $t$, $(G_t, d\big|_{G_t})$ coincides with the restriction on $G_t$ of the resistance metric of Anderes et al. (2020) computed on the whole graph $\widetilde{\boldsymbol{G}}$, hence it is a metric and it is invariant to splitting edges and merging edges at degree 2 vertices (Anderes et al., 2020, Propositions 2 and 3).

### 4.3 Time-evolving linear networks

Anderes et al. (2020) defined graphs with Euclidean edges as a generalisation of linear networks, and Euclidean trees with a given number of leaves. For both cases, edges are linear. This case is not especially interesting for the framework proposed here. The reason is that a simple isometric embedding arguments as in Tang and Zimmerman (2024) proves that one can embed a time-evolving linear network in $\mathbb{R} \times \mathbb{R}^2 = \mathbb{R}^3$, where the first component indicates time. As a consequence, it is immediate to build a vast class of covariance functions on a time-evolving linear network by a sheer restriction of a given covariance function defined on $\mathbb{R}^3$. However, such a method does not take into account the *structure* of the graph: two points that are close in $\mathbb{R}^3$ but far in the time-evolving graph could have a high correlation. Section 6 illustrates how to build kernels over the special topologies proposed in this paper. Apparently, the choices are more restrictive than the ones available for the case of linear networks, but they ensure that the spatio-temporal structure is taken into account.

## 5. Circular time and periodic graphs

Perhaps the main drawback of using the resistance semi-distance in the layer graphs that express the spatio-temporal variability is that, when adding one or more new layers, the semi-distances between the points of the previous layers may change (decrease). Indeed,

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

whenever new paths between a couple of points are added, the effective semi-distance between such points decreases, as the current meets less resistance. This presents a critical interpretation problem: for a given time series, let new data be added on a daily basis. Then, inference routines may provide different results when compared to the results of same inference techniques applied to the updated time series. Indeed, as the semi-distances may vary, the covariances between the same space-time points may vary as well.

Here, we consider time-evolving *periodic* networks, *i.e.* time-evolving networks whose evolution repeats after a fixed amount of time instants (number of layers). Not only does this construction solve the above-mentioned issue, but it also suits many phenomena whose evolution present both linear and periodic components.

**Definition 4** (Time-evolving periodic graph)**.** Let $\boldsymbol{G} = \{G_0, G_1, \ldots\}$ be a countable sequence of graphs. Then, $\boldsymbol{G}$ is a *time-evolving periodic graph* if there exists a natural number $m \geq 3$ such that, for all $t \in \mathbb{N}$, $G_t = G_{t+m}$.

Its equivalent simple periodic graph $\widetilde{\boldsymbol{G}}$ is built by connecting $G_0, \ldots, G_{m-1}$ via a proper set of temporal edges $E_T$. A special case is when the set of edges is

$$E_T = \{(v_1, v_2) \in V \times V \,:\, s(v_1) = s(v_2),\ |t(v_1) - t(v_2)| \equiv \pm 1 \pmod{m}\},$$

which provides a temporally-complete periodic graph of order 1.

Each point in the resulting space-time is denoted by its true time $t \in \mathbb{R}_0^+$, by the endpoints of the edge $e$ it lies on and by the relative distance $\delta_e(u)$ from the first one: we write $u = (t, \underline{u}, \overline{u}, \delta_e(u))$, where $e = (\underline{u}, \overline{u})$. Notice that $t \in \mathbb{N}$ whenever $u$ belongs to a temporal layer, while $t$ is not integer if $u$ belongs to the inner part of a temporal edge $e \in E_T$. Given a point $u \in V \cup \bigcup E_S$, we sometimes write $\tau(u)$ as the unique layer $\tau \in \{0, \ldots, m-1\}$ that contains $u$, *i.e.* $\tau(u) \equiv t(u) \pmod{m}$.

We start by noting that the previously-mentioned issue about linear time-evolving graphs is overcome by this construction. Indeed, once the full periodic structure has been

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

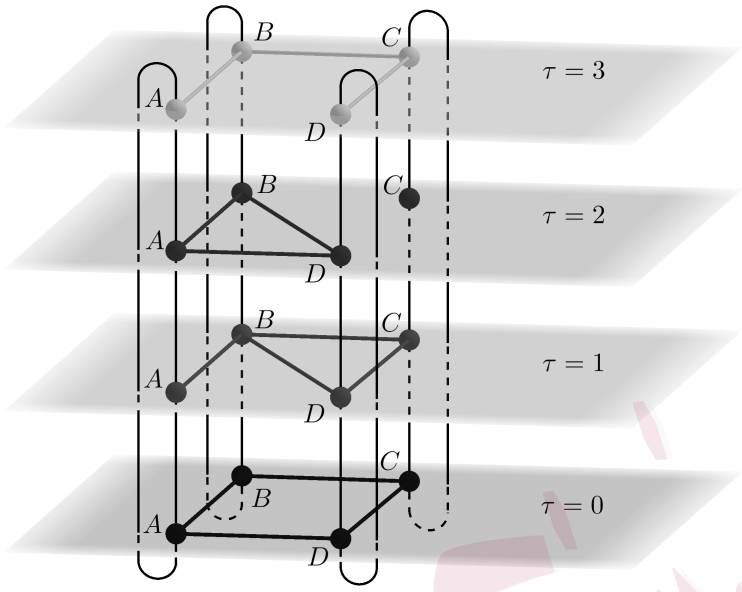



Figure 4: An example of an equivalent simple graph for a periodic time-evolving graph with $m = 4$ and $S = \{A, B, C, D\}$. The horizontal edges belong to $E_S$, whilst the vertical ones belong to $E_T$.

established, the Laplacian matrix needs be computed only once, regardless of how many new time points are added. A second remark comes from the metric construction, which necessarily needs to be adapted to a periodic process. Otherwise, some counter-intuitive properties can arise. Suppose the semi-distance $d(u_1, u_2)$ is defined as in (4.1). Then, for any couple of points $u_1 = (t_1, \underline{u}_1, \overline{u}_1, \delta_e(u_1))$ and $u_2 = (t_2, \underline{u}_2, \overline{u}_2, \delta_e(u_2))$ with $\underline{u}_1 = \underline{u}_2$, $\overline{u}_1 = \overline{u}_2$ and $\delta_e(u_1) = \delta_e(u_2)$, even when $t_1 \neq t_2$, the semi-distance would be identically equal to zero. Hence, a different definition for the process $Z$ is necessary.

For a point $u = (t, \underline{u}, \overline{u}, \delta_e(u)) \in \widetilde{\boldsymbol{G}}$, we define the process $Z$ for the periodic graph as $Z(u) := Z_V(u) + Z_E(u) + \beta W(t)$, where $\beta > 0$ is a given parameter, $W$ is a standard Wiener process, $Z_V$ is the same process as in the linear-time case, while $Z_E$, albeit similar, presents some difference with respect to the construction given in Subsection 4.1, aimed to capture the time structure of the periodic graph.

Let $e = (v_1, v_2) \in E_S$: we define the *lifespan* of $e$ as the maximal subset of $\{0, \ldots, m-1\}$ such that: (i) $\tau(v_1) \in \mathrm{ls}(e)$; (ii) $\mathrm{ls}(e)$ is connected, *i.e.* if $\tau_1 < \tau_2 \in \mathrm{ls}(e)$, then $\{\tau_1, \ldots, \tau_2\} \subseteq \mathrm{ls}(e)$ or $\{\tau_2, \ldots, m-1, 0, \ldots, \tau_1\} \subseteq \mathrm{ls}(e)$; and (iii) $\forall \tau \in$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



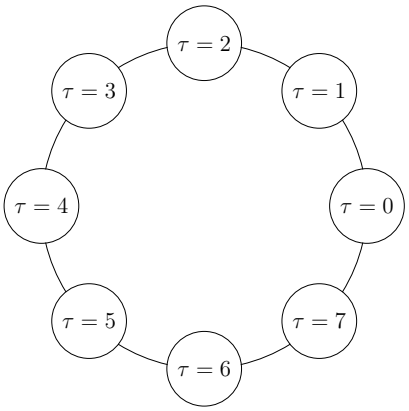



Figure 5: Conditional dependence structure of the process $Z_T$ for $m = 8$.

$\mathrm{ls}(e)$, $\exists (v_1', v_2') \in E_S$ such that $s(v_1') = s(v_1)$, $s(v_2') = s(v_2)$ and $\tau(v_1') = \tau(v_2') = \tau$. Figure 4 depicts this situation. Here, the lifespan of the edge $(A, B)$ at time $\tau = 0$ is $\{0, 1, 2, 3\}$; the lifespan of $(A, D)$ at time $\tau = 2$ is $\{2\}$; the lifespan of $(C, D)$ at time $\tau = 3$ is $\{0, 1, 3\}$. The definition of life of any edge $e \in E$ remains unchanged: if $e \in E_S$,

$$\mathrm{lf}(e) := \{(v_1', v_2') \in E_S : s(v_1') = s(v_1),\ s(v_2') = s(v_2),\ \tau(v_1') = \tau(v_2') \in \mathrm{ls}(e)\}$$

while, if $e \in E_T$, $\mathrm{lf}(e) := \{e\}$. The definition of $Z_E$ is now identical to the one of the linear-time graph, except the choice of the temporal kernel $k_T$. Indeed, we ought to consider that the time is now cyclic in the dependence structure of the temporal layers $\tau \in \{0, ..., m-1\}$. It is reasonable to model the process underlying the temporal kernel $k_T$ by means of a graphical model, as it embodies the idea of conditional independence. For a given spatial edge $e \in E_S$, we distinguish two cases: whether the lifespan of $e$ is the whole temporal set $T = \{0, ..., m-1\}$ or not.

### 5.1 The lifespan coincides with $T$

In this case, we define the covariance matrix of a zero-mean Gaussian random vector $Z_T : \{0, ..., m-1\} \to \mathbb{R}$ via its precision matrix. More precisely, let $G_T$ be the circulant graph with $m$ nodes (labelled by $\tau \in \{0, ..., m-1\}$) and $m$ edges between adjacent nodes, as shown in Figure 5. We associate each edge with a given weight $\rho \in \left[0, \frac{1}{2}\right)$, which represents the partial correlation between subsequent times. As a consequence,

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



the precision matrix $\Theta_{Z_T}$ is the circulant matrix

$$\Theta_{Z_T} = \kappa \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 & -\rho \\ -\rho & 1 & -\rho & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\rho \\ -\rho & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}.$$

Here, $\kappa > 0$ is a normalising constant whose role is to make the covariance matrix $\Sigma_{Z_T} := (\Theta_{Z_T})^{-1}$ a correlation matrix (namely the variances of every entry of $Z_T$ should be 1). Notice that the matrix $\Sigma_{Z_T}$ is a symmetric circulant matrix: as a consequence, it is possible to store only its first column, which will be denoted by $\sigma_{Z_T} \in \mathbb{R}^m$. In Figure 2 in the Supplementary Material, the values of the vector $\sigma_{Z_T}$ are plotted for some values of $m$ and $\rho$.

### 5.2 The lifespan does not coincide with $T$

In this case, the life of the edge $e$ is interrupted. Thus, it is reasonable to consider the different parts of the life of $e$ as independent. To this aim, we consider the subgraph of $G_T$ that represents the evolution of the edge $e$. More precisely, we remove from $G_T$ all the nodes $\tau$ for which the edge $e$ does not exist and we remove from $G_T$ all the edges whose at least one endpoint has been eliminated. Next, we consider all the connected components of the so-obtained graph and define an autoregressive model on each of them, independently from the others (similarly to the linear case of Section 4.2). More precisely, we define the covariance matrix of the process $Z_T$ as a block-diagonal matrix whose diagonal blocks are of the form

$$\begin{bmatrix} 1 & \phi & \dots & \phi^{j-1} \\ \phi & 1 & \dots & \phi^{j-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{j-1} & \phi^{j-2} & \dots & 1 \end{bmatrix},$$

being $\phi \in (-1, 1)$ the lag-1 correlation and $j$ the number of times $\tau$ that belong to $\mathrm{ls}(e)$, *i.e.* $j := |\mathrm{ls}(e)|$.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



### 5.3 Second-order properties of $Z$ in the circular case

The following result illustrates the analytic expression for the covariance function associated with $Z$ in the construction of the distance associated with $\widetilde{\boldsymbol{G}}$ in the periodic case.

**Proposition 5.** *Let $u_1 = (t_1, \underline{u}_1, \overline{u}_1, \delta_e(u_1)) \in \widetilde{\boldsymbol{G}}$ and $u_2 = (t_2, \underline{u}_2, \overline{u}_2, \delta_e(u_2)) \in \widetilde{\boldsymbol{G}}$. Then the kernel of the process $Z$ defined on $\widetilde{\boldsymbol{G}}$ enjoys the following representation:*

$$k_Z(u_1, u_2) = \boldsymbol{\delta}_1^\top \left(L^\star\right)^{-1} \boldsymbol{\delta}_2 + \beta^2 \min(t_1, t_2) + \mathbb{1}_{\mathrm{lf}(e_1)=\mathrm{lf}(e_2)} \sqrt{\ell(e_1)\ell(e_2)}\, k_T\left\{\tau(\underline{u}_1), \tau(\underline{u}_2)\right\} \left\{\min\left(\delta_1, \delta_2\right) - \delta_1\delta_2\right\} \tag{5.1}$$

*where $\boldsymbol{\delta}_i$ ($i \in \{1, 2\}$) are defined in Equation (4.5).*

Notice that the unique differences with the expression (4.4) are the different choices for the temporal kernel $k_T$ and the additional addend $\beta^2 \min(t_1, t_2)$. The latter ensures that the same points at different times have a strictly positive semi-distance, as, combining equations (4.6) and (5.1) for such points $u_1$ and $u_2$, we get $d(u_1, u_2) = \beta^2 |t_1 - t_2|$. We conclude this section with a formal assertion regarding the mapping $d$ as being introduced for the case of a periodic graph $\widetilde{\boldsymbol{G}}$.

**Proposition 6.** *$(\widetilde{\boldsymbol{G}}, d)$ is a semi-distance space.*

## 6. Kernels and covariance functions

The next result introduces and characterise the Reproducing Kernel Hilbert Space (RKHS) for equivalent simple graphs.

**Proposition 7.** *Let $\widetilde{\boldsymbol{G}}$ be an equivalent simple graph of either a time evolving graph or a periodic time evolving graph. Denote $\mathcal{F}$ the set of functions $f : \widetilde{\boldsymbol{G}} \to \mathbb{R}$ that are continuous w.r.t. the shortest-path metric and that, for each edge $e \in \widetilde{E}$ the restriction $f_e$ of $f$ on $e$ is absolutely continuous and $f_e'$ belongs to $L^2(e)$. Further, define the quadratic*

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

*form* $\langle \cdot, \cdot \rangle_{\mathcal{H}} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ *as follows:*

$$\langle f, g \rangle_{\mathcal{H}} = \langle \mathcal{P}_V f, \mathcal{P}_V g \rangle_V + \sum_{\lambda \in \Lambda} \langle \mathcal{P}_\lambda f, \mathcal{P}_\lambda g \rangle_\lambda, \tag{6.1}$$

$\langle \cdot, \cdot \rangle_V$, $\langle \cdot, \cdot \rangle_\lambda$, $\mathcal{P}_V$ *and* $\mathcal{P}_\lambda$ *are defined in Section S3 in the Supplementary Material. Then, the space* $(\mathcal{F}, \langle \cdot, \cdot \rangle_{\mathcal{F}})$ *is an infinite dimensional Hilbert space with reproducing kernel given by (4.4). The construction of the RKHS of the kernel (5.1) is obtained by the direct sum of* $\mathcal{H}$ *with the RKHS of a (scaled) Wiener process, as explained in Step 5 in the Supplementary Material (S3).*

Next, we explore the possibility of defining a large class of covariance functions on the equivalent simple graph. Indeed, variograms can be composed with certain classes of functions to create valid covariance functions associated with semi-distance spaces. A function $\psi : [0, +\infty) \to \mathbb{R}$ is called *completely monotone* if it is continuous on $[0, +\infty)$, infinitely differentiable on $(0, +\infty)$ and for each $i \in \mathbb{N}$ it holds $(-1)^i \psi^{(i)}(x) \geq 0$, where $\psi^{(i)}$ denotes the $i^{\text{th}}$ derivative of $\psi$ and $\psi^{(0)} := \psi$. By the Bernstein's theorem (Bernstein, 1929), completely monotone functions are the Laplace transforms of positive and bounded measures. Some example of parametric families of completely monotone functions are presented in Table 1.

| Type | $\psi(x)$ | Parameter range |
|---|---|---|
| Power exponential | $e^{-\beta x^\alpha}$ | $0 < \alpha \leq 1$, $\beta > 0$ |
| Matérn | $\frac{2^{1-\alpha}}{\Gamma(\alpha)}(\beta x)^\alpha K_r(\beta x)$ | $0 < \alpha \leq \frac{1}{2}$, $\beta > 0$ |
| Generalised Cauchy | $(\beta x^\alpha + 1)^{-\xi/\alpha}$ | $0 < \alpha \leq 1$, $\beta > 0$, $\xi > 0$ |
| Dagum | $1 - \left(\frac{\beta x^\alpha}{1+\beta x^\alpha}\right)^{\xi/\alpha}$ | $0 < \alpha, \xi \leq 1$, $\beta > 0$ |
| | $\frac{\log(1+\beta x)}{\beta x}$ | $\beta > 0$ |
| | $\{1 + \alpha \log(1+\beta x)\}^{-\xi}$ | $\alpha, \beta, \xi > 0$ |
| | $e^{-\alpha}(1+\beta x)^{\frac{\alpha}{\beta x}}$ | $\alpha, \beta > 0$ |
| | $\frac{1}{(\alpha+\xi)^\mu}\left(\alpha + \frac{\xi}{1+\beta x}\right)^\mu$ | $\alpha, \beta, \xi, \mu > 0$ |

Table 1: Examples of completely monotone functions $0 < x \mapsto \psi(x)$ such that $\psi(0^+) = 1$. Here, $K_r$ denotes the modified Bessel function of the second kind. The first four rows are from Anderes et al. (2020, Table 1), whilst the others have been taken from Miller and Samko (2001).

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

By Theorem 6 and Corollary 1 in Anderes et al. (2020) it is immediate to show the next result, which allows to define covariance functions on arbitrary domains. As a sheer consequence, we obtain in Proposition 8 a positive definite kernel on the time-evolving graph.

**Theorem 1.** *Let $Z$ be a stochastic process defined on a set $X$ such that $\mathbb{E}\{Z^2(x)\} < +\infty$ for all $x \in X$. Define $d : X \times X \to \mathbb{R}_0^+$ as*

$$d(x_1, x_2) := \gamma_Z(x_1, x_2) = \mathbb{V}\mathrm{ar}\{Z(x_1) - Z(x_2)\}. \tag{6.2}$$

*In addition, let $\psi$ be a non-constant completely monotone function on $[0, +\infty)$. Then the following holds:*

1. *$(X, d) \overset{\sqrt{\cdot}}{\hookrightarrow} H$ for some Hilbert space $H$, that is: there exist a Hilbert space $H$ and a function $\xi : X \to H$ such that, given $x_1, x_2 \in X$: $\sqrt{d(x_1, x_2)} = \|\xi(x_1) - \xi(x_2)\|_H$;*
2. *the function $(x_1, x_2) \mapsto \psi\{d(x_1, x_2)\}$ is positive semi-definite;*
3. *if, in addition, $d$ is a semi-distance on $X$ (see Proposition 1 in the Supplementary Material for a useful characterisation), then $(x_1, x_2) \mapsto \psi\{d(x_1, x_2)\}$ is strictly positive definite.*

The first point comes from the fact that $d$, being a variogram, is a generalised covariance function of order 0 and by the equivalence (IV) $\iff$ (I) in Anderes et al. (2020, Theorem 6); whilst the other two statements follow from Anderes et al. (2020, Corollary 1).

**Proposition 8.** *Let $\psi : [0, \infty) \to \mathbb{R}$ be a continuous, completely monotone function on the positive real line, with $\psi(0) < \infty$. Let $d : \widetilde{\boldsymbol{G}} \times \widetilde{\boldsymbol{G}} \to \mathbb{R}$ be the mapping defined at (4.6). Then, the function $k_Z(u_1, u_2) = \psi\{d(u_1, u_2)\}$, where $u_1, u_2 \in \widetilde{\boldsymbol{G}}$, is a strictly positive definite function.*

Proposition 8 provides a very easy recipe to build kernels over time evolving graphs, whatever the temporal structure (linear or periodic). Any element from the Table 1 is a

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



good candidate for such a composition. We do not report the corresponding algebraic forms, instead we illustrate how these covariance works through two practical examples. We believe that the free parameters and the large number of analytically-tractable completely monotone functions provide a wide range of models that could fit several real-world frameworks.

### 6.1 Linear time

We start by considering the graph in Figure 6. Here, we have $m = 3$ time instants. Further, $V = \{A_0, B_0, C_0, D_0, A_1, B_1, C_1, D_1, A_2, C_2, D_2\}$. We focus on the semi-distances as well as the covariances between the points $A_0 = (A_0, B_0, \delta(A_0) = 0)$, $P := (C_0, D_0, \delta(P) = 0.8)$ and $Q := (C_2, D_2, \delta(Q) = 0.5)$. All the spatial edges $E_S$ have weight 1, whilst the temporal edges have weight $\alpha > 0$. Finally, we use $k_T(h) = \phi^{|h|}$ as defined in Subsection 4.2, with $\phi \in (-1, 1)$ free parameter. Figure 6 clearly shows the effect the temporal edge parameter $\alpha$ plays on the semi-distances: while it has a considerable impact on the semi-distances $d(A_0, Q)$ and $d(P, Q)$, it shows a negligible effect on $d(A_0, P)$. This is reasonable given the graph structure: $A_0$ and $P$ belong to the same layer ($t = 0$) and they are connected from both the paths $A_0, D_0, P$ and $A_0, B_0, C_0, P$, which completely lie on $t = 0$ (and therefore they do not change with $\alpha$). On the other hand, $Q$ can be connected to $A_0$ and $P$ only via paths that include temporal edges. As a consequence, if $\alpha \to 0^+$, both $d(A_0, Q)$ and $d(P, Q)$ will go to infinity.

The plot on the right of Figure 6 shows the effect of the correlation parameter $\phi$ as well. Whilst it does not influence the semi-distances concerning $A_0$ (since it is a vertex), as it increase, it reduces the semi-distances between $P$ and $Q$. Clearly, the effect is more significant for large values of the parameter $\alpha$. Indeed, when $\alpha$ is small, the semi-distances between nodes at different time instants are large. As a consequence,

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



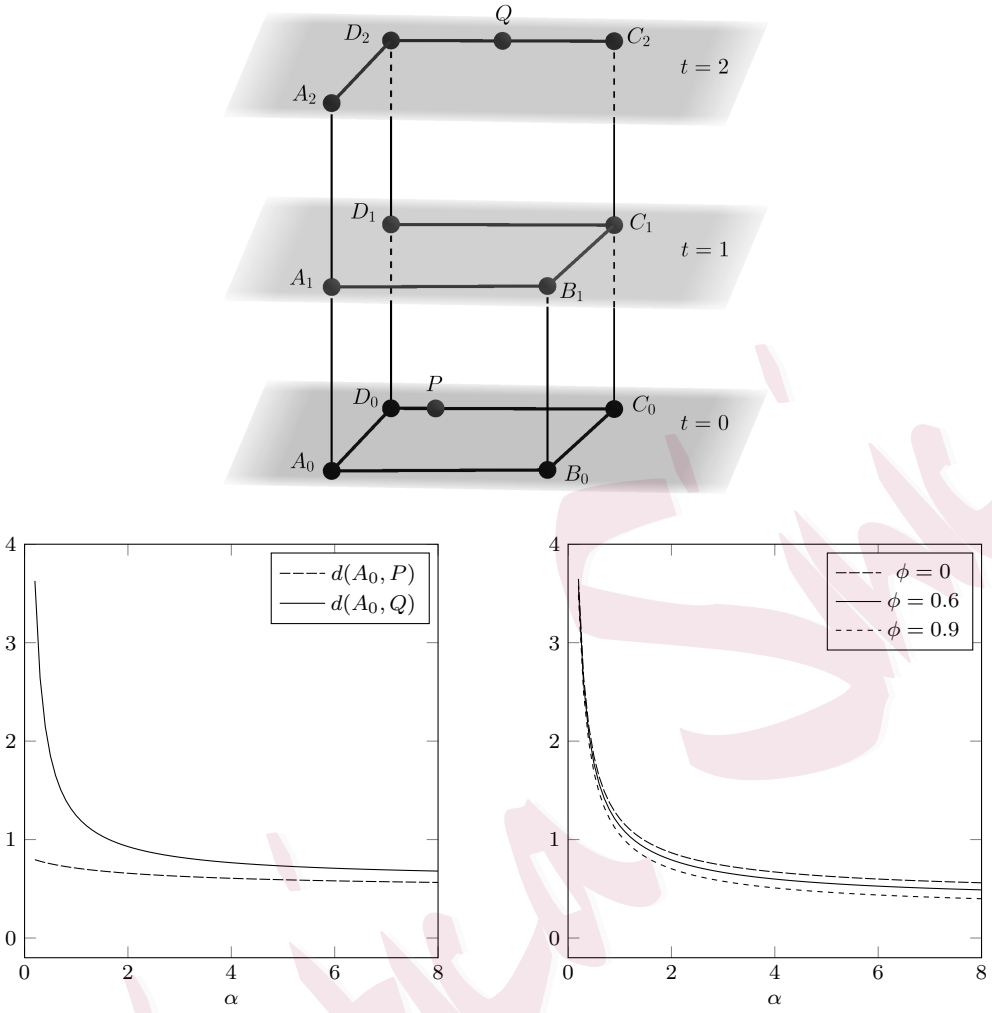



Figure 6: Top: equivalent simple graph taken as an example for the linear-time case. Bottom: semi-distances between the points $A_0$, $P$ and $Q$. Notice that while $d(A_0, P)$ and $d(A_0, Q)$ (left) do not depend on $\phi$, the semi-distance $d(P, Q)$ (right) decreases as $\phi$ increases.

the second line of equation (4.4) becomes negligible when compared to the first one. Figure 3 in the Supplementary Material shows the resulting effect of the parameter $\alpha$ on the correlations between $A_0$, $P$ and $Q$ generated by the composition of two completely monotone functions taken from Table 1 and the semi-distances shown in Figure 6.

## 6.2 Circular time

We are going to analyse the time-evolving periodic graph depicted in Figure 7. In this case, we have $m = 8$ and $V = \{A_0, B_0, A_1, B_1, \ldots, A_7, B_7\}$. We will compare the

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



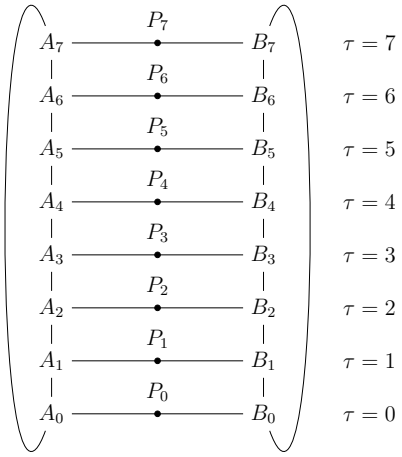



Figure 7: Equivalent simple graph taken as an example for the circular-time case.

semi-distances and the covariances between the points $P_0 := (0, A_0, B_0, \delta(P_0) = 0.5)$ and $P_t := (t, A_\tau, B_\tau, \delta(P_t) = 0.5)$, where $t \in \mathbb{N}$ and $\tau \equiv t \pmod m$. Here, all the spatial edges have weight 1, whilst the temporal ones have weight $\alpha > 0$. We use the temporal kernel $k_T$ as described in Subsection 5.1, with $\rho$ and $\beta$ free parameters.

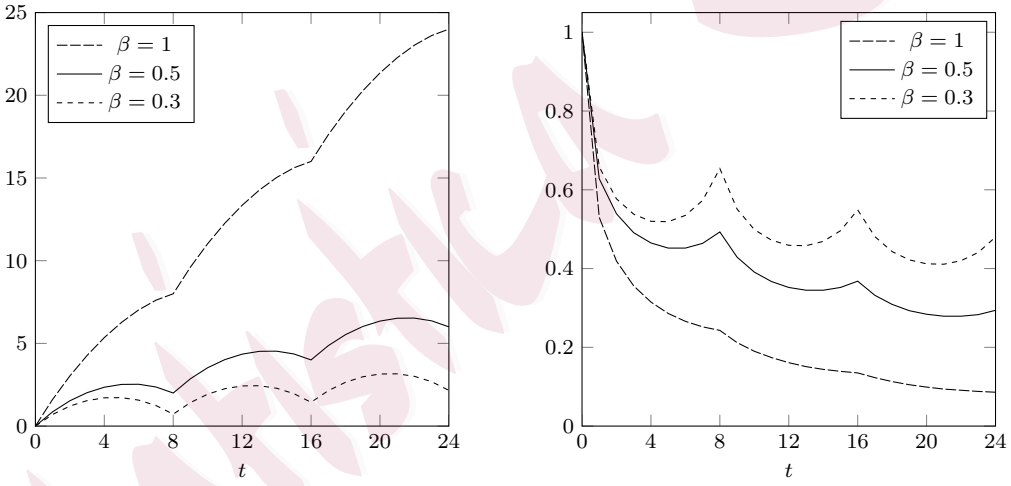



Figure 8: Semi-distances (left) and covariances (right) for the graph in 7 between the points $P_0$ and $P_t$ for $\rho = 0.45$ and $\alpha = 1$. Covariances have been generated via the exponential kernel with parameters $\alpha = 0.5$ and $\beta = 0.5$ (see Table 1).

The semi-distances and covariances in Figure 8 show the effect of the parameter $\beta$ on our construction. It adds a linear component $\beta^2 t$, which allows to calibrate the semi-distance (and, as a result, the covariance) between the points at different time instants. In such a way, the effect of the periodicity is increased by setting a low $\beta$ and becomes negligible when $\beta$ grows. Furthermore, notice the spikes the covariance functions show: they perfectly embody the periodic setting of a process, as introduced

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



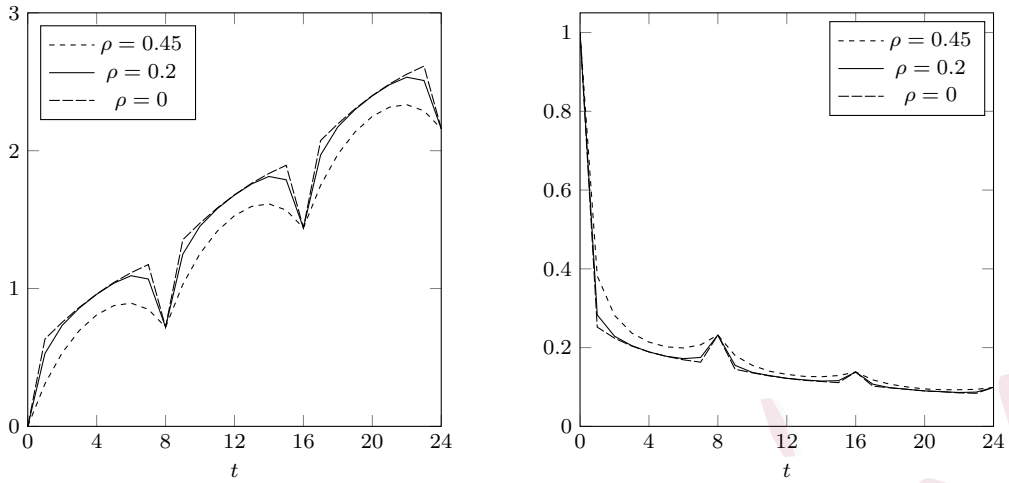



Figure 9: Semi-distances (left) and covariances (right) for the graph in 7 between the points $P_0$ and $P_t$ for $\alpha = 10$ and $\beta = 0.3$. Covariances have been generated via the Dagum kernel with parameters $\alpha = 1$, $\beta = 2$ and $\xi = 0.5$ (see Table 1).

in Section 5. We notice that, although isotropic covariances are decreasing functions of the spatial (semi-)distances, Figures 8 and 9 show valid covariance functions, as the semi-distance of our setting is completely different from the Euclidean distance on $\mathbb{R}^n$ as it takes into account the spatio-temporal structure of the time-evolving graph.

In Figure 9, it is possible to visualise the effect of the partial correlation parameter $\rho \in [0, \frac{1}{2})$. First, notice that its role is particularly significant when the weight $\alpha$ is high. Indeed, for low $\alpha$'s, the covariance structure of the vertices given by the inverse laplacian matrix is dominant. Yet, when $\alpha$ is high, the nodes at different time instants are considered close to each other and the resulting (semi-)distances given by the sheer first line of (5.1) are low. Thus, the parameter $\rho$ (which enters in the second line of (5.1)) has a greater influence. Clearly, the greater is $\rho$, the lower is the semi-distance, as it correlates different Brownian bridge realisations.

## 6.3 A note on continuous time

There may be several approaches to extend this work to continuous-time settings. The first, and perhaps most intuitive, is to perform a limit operation, increasing the number of time instants for a fixed time interval (either linear or periodic). However, this would

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

increase the number of vertices and, as a consequence, the dimension of the Laplacian matrix $L$ to be inverted, leading to extremely high computational burdens. Another approach could be to use the values of the process $Z$ on temporal edges $E_T$ connecting time $t \in T$ to $t+1 \in T$ to define the process and, thus, the (semi-)metric on the whole interval $[t, t+1]$ for each couple of points belonging to both $G_t$ and $G_{t+1}$.

## 7. Numerical experiments

In order to illustrate our approach, we performed a numerical experiment to show how it can handle a time-evolving graph. More specifically, we considered the periodic time-evolving graph $\widetilde{\boldsymbol{G}}_1$ depicted in Figure 10: at any $\tau \in \{0, 1, 2\}$, it has two vertices, say $A$ and $B$; yet, while at both $\tau = 0$ and $\tau = 2$ they are connected by an edge of weight 1, at $\tau = 1$ they are not connected. We then set parameters $\alpha := 1$, $\beta := 0.5$ and

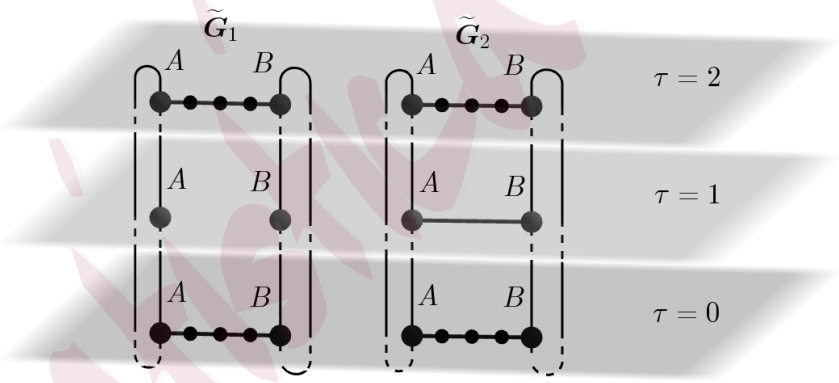



Figure 10: The two periodic time-evolving graphs used for the numerical experiment. Coloured points are vertices, while black ones are points on the spatial edges taken as regular grid.

$\gamma := 0.6$. Notice that in this graph the parameter $\rho$ has no effect, since there is no edge whose lifespan coincides with $\{0, 1, 2\}$. We compared $\widetilde{\boldsymbol{G}}_1$ with $\widetilde{\boldsymbol{G}}_2$, that is a "regularised" version of $\widetilde{\boldsymbol{G}}_1$: we assume that at time $\tau = 1$ the vertices $A$ and $B$ are indeed connected by the edge with weight 1. We considered a regular grid of points on $\widetilde{\boldsymbol{G}}_1$, including all the vertices and 3 equally-spaced points on each spatial edge and an evolution of 2 complete periods ($t \in \{0, \ldots, 5\}$), resulting in 24 spatio-temporal points: 6 vertices and

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

6 points on the edges, times 2 periods. We considered the same grid on $\widetilde{\boldsymbol{G}}_2$ as well. Next, we computed their $24 \times 24$ semi-distance matrix $D$ and covariances matrix $\Sigma = \psi(D)$, where we used the completely monotone function $\psi(x) := e^{-x}$. We simulated $N := 10^4$ iid samples from the multivariate normal distribution $\mathcal{N}_{24}(0, \Sigma)$. Next, we performed a maximum likelihood estimation of the parameters $\alpha, \beta, \gamma, \rho$ under both $\widetilde{\boldsymbol{G}}_1$ and $\widetilde{\boldsymbol{G}}_2$ and compared the results (see Table 2). Clearly, the maximum likelihood estimates under

| | $\theta_{\widetilde{\boldsymbol{G}}_1}$ | $\hat{\theta}_{\widetilde{\boldsymbol{G}}_1}$ | $\hat{\theta}_{\widetilde{\boldsymbol{G}}_2}$ |
|---|---|---|---|
| $\alpha$ | 1.0000 | 0.9980 | 0.7825 |
| $\beta$ | 0.5000 | 0.5006 | 0.5076 |
| $\gamma$ | 0.6000 | 0.6059 | – |
| $\rho$ | – | – | 0.2882 |
| $\log L$ | -16934.6 | -16934.2 | -18991.3 |
| $\mathbb{E}\lvert\Delta D\rvert$ | 0.0000 | 0.0016 | 0.0857 |
| $\max\lvert\Delta D\rvert$ | 0.0000 | 0.0037 | 0.8600 |
| $\mathbb{E}\lvert\Delta\Sigma\rvert$ | 0.0000 | 0.0004 | 0.0232 |
| $\max\lvert\Delta\Sigma\rvert$ | 0.0000 | 0.0009 | 0.3042 |

Table 2: Comparison between the maximum likelihood estimates under $\widetilde{\boldsymbol{G}}_1$ and $\widetilde{\boldsymbol{G}}_2$. Blank spaces denote that the parameter is not influent in the model.

the true underlying graph structure $\widetilde{\boldsymbol{G}}_1$ are extremely close to the true ones, whilst the "regularised" graph $\widetilde{\boldsymbol{G}}_2$ clearly is not able to capture the actual data structure. This behaviour can also be deduced from the log-likelihood and the summary statistics that report the mean and the maximum discrepancies of the matrices $\Delta D := \hat{D} - D$ and $\Delta\Sigma := \hat{\Sigma} - \Sigma$. Such matrices are also represented in Figures 4 and 5 in the Supplementary Material. We point out that the effect of the introduction of edge connecting $A$ and $B$ at time $\tau = 1$ in the graph $\widetilde{\boldsymbol{G}}_2$ is particularly significant in the semi-distances and, thus, in the covariances between such points. However, it also influences the general spatio-temporal structure of the graph and, as a consequence, of the process.

## 8. Conclusion: statistical implications and broader impact

We have contributed with a first attempt to define distances and kernels for a class of graphs whose topology can evolve over time, a setting of growing importance across a range of scientific and engineering domains. This work generalises Anderes et al. (2020),

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

who contributed through static graphs and provides a new tool to space-time analysis over networks by introducing dynamics into metrics and consequently on the covariance functions, to provide a more faithful representation of certain types of networks.

The examples provided in the introduction explain the relevance of having adaptive topologies, and consequently adaptive covariance functions, in a wealth of practical real-life situations (among them, urban mobility systems, active river networks, or biological connectivity structures) and hence a rigorous and valid support for statistical modelling over temporally-adaptive networks.

The methodology we proposed might be outperformed in some real-world cases by *ad hoc* models, since these are specifically built to capture the intrinsic dynamics of the relative phenomena. Nevertheless, to our best knowledge, this is the first approach that provides a unified framework for time-evolving graphs. As such, it opens many aspects for future researches. The next step will be to inspect the finite sample properties of our covariance functions on real-life datasets, which is beyond our scope for this paper. A second step will be concerned about how to efficiently simulate Gaussian processes over dynamical networks, adapting, for instance, the methods in Alegría et al. (2026) to the time-evolving setting.

A fundamental aspect for future work will be to understand statistical inference aspects, maximum likelihood estimation, and kriging misspecification, under fixed domain asymptotics (Bevilacqua et al., 2019; Bolin and Wallin, 2024) as well as Bayesian inference and its computational aspects.

Through examples from neuroscience, hydrology, and transportation science, we have illustrated several fields that could benefit from this contribution. For instance, modelling dynamic functional connectivity in fMRI studies, evolving discharge patterns in stream networks, or real-time passenger flow in urban transit systems all benefit from our ability to define stochastic processes over changing topologies.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345

Finally, our approach might be useful to integrate climate networks - called Tsonis networks in Porcu et al. (2025) - typically constructed through measures of statistical similarity, with geophysical networks - such as river basins - to improve the statistical understanding of one or the other. Merging networks having different topologies and different temporal evolution is a key of success to understand certain climate phenomena (Tsonis and Roebber, 2004).

We believe the methodological advances proposed in this work will be useful in a growing number of real-world problems as it offers a new statistical lens on temporally-evolving systems and provides both theoretical rigour and practical tools to model, infer, and understand data in nonstationary and structurally dynamic environments.

## Supplementary Material

The online Supplementary Material contains 6 sections. Section S1 contains additional mathematical background; Section S2 contains some additional remarks about the motivations behind the definition on $Z_V$ and $Z_E$; Section S3 explores the reproducing kernel Hilbert space properties of our construction; Section S4 we state some general results that may be useful to the definition of kernels on arbitrary sets; Section S5 presents the proofs of the results in the main document; while Section S6 contains some additional plots.

## Acknowledgements

The authors are grateful to Havard Rue for insightful discussions during the preparation of this manuscript. This project has also been sustained by the prompt help of Valeria Simoncini and Valter Moretti.

Claudio Agostinelli was partially funded by BaC INF-ACT S4 - BEHAVE-MOD PE00000007 PNRR M4C2 Inv. 1.3 - NextGenerationEU, CUP: I83C22001810007 and

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



by the PRIN funding scheme of the Italian Ministry of University and Research (Grant No. P2022N5ZNP).

## References

Alegría, A., X. Emery, T. Filosi, and E. Porcu (2026). Computationally efficient algorithms for simulating isotropic gaussian random fields on graphs with euclidean edges. *Journal of Computational and Graphical Statistics 0*(0), 1–14.

Anderes, E., J. Møller, and J. G. Rasmussen (2020). Isotropic covariance functions on graphs and their edges. *Annals of Statistics 48*(4), 2478–2503.

Bernstein, S. (1929). Sur les fonctions absolument monotones. *Acta Mathematica 52*, 1–66. doi: 10.1007/BF02592679.

Bevilacqua, M., T. Faouzi, R. Furrer, and E. Porcu (2019). Estimation and prediction using generalized wendland covariance functions under fixed domain asymptotics. *The Annals of Statistics 47*(2), 828–856.

Bolin, D. and F. Lindgren (2011). Spatial Models Generated by Nested Stochastic Partial Differential Equations, with an Application to Global Ozone Mapping. *The Annals of Applied Statistics 5*(1), 523–550.

Bolin, D., L. Riera-Segura, and A. B. Simas (2025). A new class of non-stationary Gaussian fields with general smoothness on metric graphs. 10.48550/arXiv.2501.11738.

Bolin, D., A. Simas, and J. Wallin (2024). Gaussian Whittle–Matérn fields on metric graphs. *Bernoulli 30*, 1611–1639. doi: 10.3150/23-BEJ1647.

Bolin, D. and J. Wallin (2024). Spatial confounding under infill asymptotics. doi: 10.48550/arXiv.2403.18961.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



Borovitskiy, V., M. R. Karimi, V. R. Somnath, and A. Krause (2023). Isotropic gaussian processes on finite spaces of graphs. In F. Ruiz, J. Dy, and J.-W. van de Meent (Eds.), *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, Volume 206 of *Proceedings of Machine Learning Research*, pp. 4556–4574. PMLR.

Damaraju, E., A. Caprihan, J. R. Lowe, E. A. Allen, V. D. Calhoun, and J. P. Phillips (2014). Functional connectivity in the developing brain: a longitudinal study from 4 to 9 months of age. *Neuroimage 84*, 169–180.

Devriendt, K. (2022). Effective resistance is more than distance: Laplacians, simplices and the Schur complement. *Linear Algebra and its Applications 639*, 24–49. doi: 10.1016/j.laa.2022.01.002.

Folkman, J. (2002). Role of angiogenesis in tumor growth and metastasis. *Seminars in Oncology 29*(6 Suppl 16), 15–18.

Ghosh, A., S. Boyd, and A. Saberi (2008). Minimizing effective resistance of a graph. *SIAM Review 50*(1), 37–66. doi: 10.1137/050645452.

Hutchison, R. M., T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. D. Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage 80*, 360–378.

James, A. L. and N. T. Roulet (2007). Investigating hydrologic connectivity and its association with threshold change in runoff response in a temperate forested watershed. *Hydrological Processes: An International Journal 21*(25), 3391–3408.

Jorgensen, P. E. T. and E. P. J. Pearse (2010). A Hilbert space approach to effective

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



resistance metric. *Complex Analysis and Operator Theory 4*(4), 975–1013. doi: 10.1007/s11785-009-0041-1.

Kim, S. H., H. S. Lee, B. J. Kang, B. J. Song, H.-B. Kim, A. Lee, and M.-S. Jin (2016). Dynamic contrast-enhanced mri perfusion parameters as imaging biomarkers of angiogenesis in invasive ductal carcinoma. *PLOS ONE 11*(12), e0168632.

Lilleborge, K., S. Martino, G.-A. Fuglstad, F. Lindgren, and R. Ingebrigtsen (2025). Joint modelling of line and point data on metric graphs. doi: 10.48550/arXiv.2505.01175.

Lindquist, M. A., Y. Xu, M. B. Nebel, and B. S. Caffo (2014). Evaluating dynamic bivariate correlations in resting-state fmri: a comparison study and a new approach. *NeuroImage 101*, 531–546.

Masselink, R. J., T. Heckmann, A. J. Temme, N. S. Anders, H. P. Gooren, and S. D. Keesstra (2017). A network theory approach for a better understanding of overland flow connectivity. *Hydrological Processes 31*(1), 207–220.

Medaglia, J. D., M.-E. Lynall, and D. S. Bassett (2015). Cognitive network neuroscience. *Journal of cognitive neuroscience 27*(8), 1471–1491.

Menegatto, V., C. Oliveira, and E. Porcu (2020). Gneiting class, semi-metric spaces and isometric embeddings. *Constructive Mathematical Analysis 3*(2), 85–95.

Miller, K. and S. Samko (2001). Completely monotonic functions. *Integral Transforms and Special Functions 12*(4), 389–402. doi: 10.1080/10652460108819360.

Newman, M., A.-L. Barabási, and D. J. Watts (2011). *The structure and dynamics of networks*. Princeton University Press.

Porcu, E., X. Emery, and A. P. Peron (2022). Nested covariance functions on graphs with euclidean edges cross time. *Electronic Journal of Statistics 16*(2), 4222–4246. 10.1214/22-EJS2039.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



Porcu, E., T. Filosi, and H. Simon (2025). A triad of networks and a triad of fusions for the other climate crisis. doi: 10.48550/arXiv.2510.09728.

Porcu, E., P. A. White, and M. G. Genton (2023). Stationary nonseparable space-time covariance functions on networks. *Journal of the Royal Statistical Society Series B: Statistical Methodology 85*(5), 1417–1440. 10.1093/jrsssb/qkad082.

Preti, M. G., T. A. W. Bolton, and D. Van De Ville (2017). The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage 160*, 41–54.

Rico, J., J. Barateiro, and A. Oliveira (2021). Graph neural networks for traffic forecasting. doi: 10.48550/arXiv.2104.13096.

Schoenberg, I. J. (1942). Positive Definite Functions on Spheres. *Duke Math. Journal 9*, 96–108.

Stein, M. L. (1999). *Statistical Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York.

Tang, J. and D. L. Zimmerman (2024). Space-time covariance models on networks. *Electronic Journal of Statistics 18*(1), 490–514.

Tsonis, A. A. and P. J. Roebber (2004). Architecture and dynamics of a climate network. *Physica A 333*, 497–504.

Uddin, L. Q. and K. H. Karlsgodt (2018). Future directions for examination of brain networks in neurodevelopmental disorders. *Journal of Clinical Child & Adolescent Psychology 47*(3), 483–497.

Zheng, Y., L. Capra, O. Wolfson, and H. Yang (2014). Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST) 5*(3), 1–55.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0345



Department of Mathematics, University of Trento, Trento, I–38123, Italy

E-mail: (tobia.filosi@unitn.it, claudio.agostinelli@unitn.it)

Department of Mathematics, Khalifa University, Abu Dhabi, UAE, ADIA Lab, Abu Dhabi, UAE and School of Computer Science and Statistics, Trinity College, Dublin, Ireland

E-mail: (emilio.porcu@ku.ac.ae )

Statistica Sinica