Spectral embedding for dynamic networks with stability guarantees

Anonymous Author(s)
Affiliation
Address
email

Abstract

We consider the problem of embedding a dynamic network, to obtain time-evolving vector representations of each node, which can then be used to describe the changes in behaviour of a single node, one or more communities, or the entire graph. Given this open-ended remit, we wish to guarantee stability in the spatio-temporal positioning of the nodes: assigning the same position, up to noise, to nodes behaving similarly at a given time (cross-sectional stability) and a constant position, up to noise, to a single node behaving similarly across different times (longitudinal stability). These properties are defined formally within a generic dynamic latent position model. By showing how this model can be recast as a multilayer random dot product graph, we demonstrate that unfolded adjacency spectral embedding satisfies both stability conditions, allowing, for example, spatio-temporal clustering under the dynamic stochastic block model. We also show how alternative methods, such as omnibus, independent or time-averaged spectral embedding, lack one or the other form of stability.

1 Introduction

Consider a dynamic network in which nodes come and go, change abruptly or evolve, alone or in communities, and form connections accordingly. Our goal is to find a vector representation, $\hat{\mathbf{Y}}_i^{(t)} \in \mathbb{R}^d$, for every node i and time t, which could be used for a diversity of downstream analyses, such as clustering, time series analysis, classification, model selection and more. This problem is known as dynamic (or evolutionary) network (or graph) embedding, and a great number of scalable and empirically successful techniques have been put forward, with recent surveys by [40, 44]. We consider, among these, a subset about which it is reasonable to try to establish a certain statistical guarantee.

The novelty of this paper is *not* to propose a new procedure. Instead, it is to demonstrate that an existing procedure, unfolded adjacency spectral embedding (UASE) [14], has an important stability property. Given a sequence of symmetric adjacency matrices $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(T)} \in \{0,1\}^{n \times n}$, where $\mathbf{A}_{ij}^{(t)} = 1$ if nodes i and j form an edge at time t, UASE computes the rank d matrix factorisation of $\mathbf{A} := (\mathbf{A}^{(1)} | \cdots | \mathbf{A}^{(T)})$ to obtain $\mathbf{A} \approx \hat{\mathbf{X}}\hat{\mathbf{Y}}^\top$ using the singular value decomposition (precise details later). The matrix $\hat{\mathbf{Y}} \in \mathbb{R}^{nT \times d}$ contains, as rows, the desired representations $\hat{\mathbf{Y}}_1^{(1)}, \ldots, \hat{\mathbf{Y}}_n^{(1)}, \ldots, \hat{\mathbf{Y}}_1^{(T)}, \ldots, \hat{\mathbf{Y}}_n^{(T)}$.

The original motivation for UASE was to analyse multilayer (or multiplex) graphs under an appropriate extension of the random dot product graph model [14]. To evaluate UASE and other procedures on the task of dynamic network embedding, we instead consider the dynamic latent position model

$$\mathbf{A}_{ij}^{(t)} \overset{ind}{\sim} \text{Bernoulli}\left(f\left\{\mathbf{Z}_i^{(t)}, \mathbf{Z}_j^{(t)}\right\}\right), \tag{1}$$

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

for $1 \leq i < j \leq n$, $t \in [T]$, where $\mathbf{Z}_i^{(t)} \in \mathbb{R}^k$ represents the unknown position of node i at time t, and $f : \mathbb{R}^k \times \mathbb{R}^k \to [0, 1]$ is a symmetric function.

This model is well-established [34, 18, 12, 13, 28, 19, 7, 37, 9, 8], with recent reviews by [17] and [39], and inference usually proceeds on the basis of a parametric model for f (e.g. logistic in the latent position distance [34]) and for the dynamics of $\mathbf{Z}_i^{(t)}$ (e.g. a Markov process [34]). The model includes the dynamic standard, degree-corrected and mixed-membership stochastic block models as special cases, which were studied in [41, 45, 11, 24, 43, 42, 26, 2, 27, 16].

To make statistical sense of UASE under this latent position model, we must somehow connect its output $\hat{\mathbf{Y}}_i^{(t)}$ to $\mathbf{Z}_i^{(t)}$. To this end we construct a canonical representative of $\mathbf{Z}_i^{(t)}$, denoted $\mathbf{Y}_i^{(t)}$, that UASE can be seen to estimate. Although we make some regularity assumptions on f and the $\mathbf{Z}_i^{(t)}$, they are not modelled in an explicit, parametric way, and UASE can clearly be used in practice without having a specific model in mind. For example, we do not make a Markovian assumption on the evolution of $\mathbf{Z}_i^{(t)}$ and UASE can be used to uncover periodic behaviours.

The key purpose of imposing a dynamic latent position model is to allow us to put down certain embedding stability requirements. Using this framework, we can define precisely what we mean by two nodes, i and j, behaving "similarly" at times s and t respectively. In such cases, ideally we would have $\hat{\mathbf{Y}}_i^{(s)} \approx \hat{\mathbf{Y}}_j^{(t)}$. Two special cases, more easily seen to be desirable, are as follows: we assign the same position, up to noise, to nodes behaving similarly at a given time (cross-sectional stability) and a constant position, up to noise, to a single node behaving similarly across different times (longitudinal stability).

To achieve both cross-sectional and longitudinal stability is generally elusive. We show that two plausible alternatives, omnibus [21] and independent embedding of each $\mathbf{A}^{(t)}$, alternately exhibit one form of stability and not the other. More generally, we find existing procedures [4, 22, 35, 6, 47, 5, 23, 3, 2, 27, 32, 16] tend to trade one type of stability off against the other, e.g. via user-specified cost functions. As a side-note, it could be observed that omnibus embedding is not being evaluated on a task for which it was designed, since in the theory of [21] the graphs are identically distributed. Because the technique is so different from the others, we still feel it makes an interesting addition.

Our central contribution is to prove that UASE asymptotically provides both longitudinal and cross-sectional stability: for two nodes, i and j, behaving similarly at times s and t respectively, we have $\mathbf{Y}_i^{(s)} = \mathbf{Y}_j^{(t)}$ and, moreover, $\hat{\mathbf{Y}}_i^{(s)}$ and $\hat{\mathbf{Y}}_j^{(t)}$ have asymptotically equal error distribution. We emphasize that these properties hold without requiring any sort of global stability so, for example, if all nodes but one change behaviour, this last one will hold its position. In the asymptotic regime considered, we have $n \to \infty$, but T fixed so that, for example, our results are relevant to the case of two large graphs. The alternative regime where $T \to \infty$ grows but n is fixed is not easily handled by existing theory and to provide constant updates to UASE (or omnibus embedding) in a streaming context presents significant computational challenges.

The remainder of this article is structured as follows. Section 2 gives a pedagogical example demonstrating the cross-sectional and longitudinal stability of UASE in a two-step dynamic stochastic block model, while highlighting the instability of omnibus and independent spectral embedding. In Section 3, we prove a central limit theorem for UASE under a dynamic latent position model, and demonstrate that the distribution satisfies both stability conditions. In Section 4, we review the stability of other dynamic network embedding procedures. Section 5 presents an example of UASE applied to a dynamic network of social interactions in a French primary school, with a dynamic community detection example given in the main text and a further classification example provided in the Appendix. Section 6 concludes.

2 Motivating example

Suppose it is of interest to uncover dynamic community structure, for example, communities splitting or merging, new communities being born or communities dying. The dynamic stochastic block model [45, 43] provides a simple explicit model for this, in which two nodes connect at a certain point in time with probability only dependent on their current community membership. Suppose that at times

1 and 2, we have the following inter-community link probability matrices,

$$\mathbf{B}^{(1)} = \begin{pmatrix} 0.08 & 0.02 & 0.18 & 0.10 \\ 0.02 & 0.20 & 0.04 & 0.10 \\ 0.18 & 0.04 & 0.02 & 0.02 \\ 0.10 & 0.10 & 0.02 & 0.06 \end{pmatrix}, \quad \mathbf{B}^{(2)} = \begin{pmatrix} 0.16 & 0.16 & 0.04 & 0.10 \\ 0.16 & 0.16 & 0.04 & 0.10 \\ 0.04 & 0.04 & 0.09 & 0.02 \\ 0.10 & 0.10 & 0.02 & 0.06 \end{pmatrix}.$$

At time 1 there are four communities present, for example, a node of community 1 connects with a node of community 3 with probability 0.18. At time 2, this matrix changes so that communities 1 and 2 have merged, community 3 has moved, whereas community 4 is unchanged.

We simulate a dynamic network from this model over these two time steps, on $n = 1000$ nodes, equally divided among the four communities and investigate the results of three embedding techniques: UASE, omnibus, and independent spectral embedding, displayed in Figure 1. Given the dynamics described above, we contend that the following properties would be desirable:

1. *Cross-sectional stability*: The embeddings for communities 1 and 2 at time 2 are close.
2. *Longitudinal stability*: The embeddings for community 4 at times 1 and 2 are close.

Figure 1 illustrates that UASE has both properties: the blue and orange point clouds merge at time 2, and the red point cloud has the same location *and* shape over the two time points. On the other hand, omnibus embedding only has longitudinal stability (the blue and orange point clouds don't merge at time 2), whereas independent spectral embedding only has cross-sectional stability (the red point cloud is at different locations over the two time points).

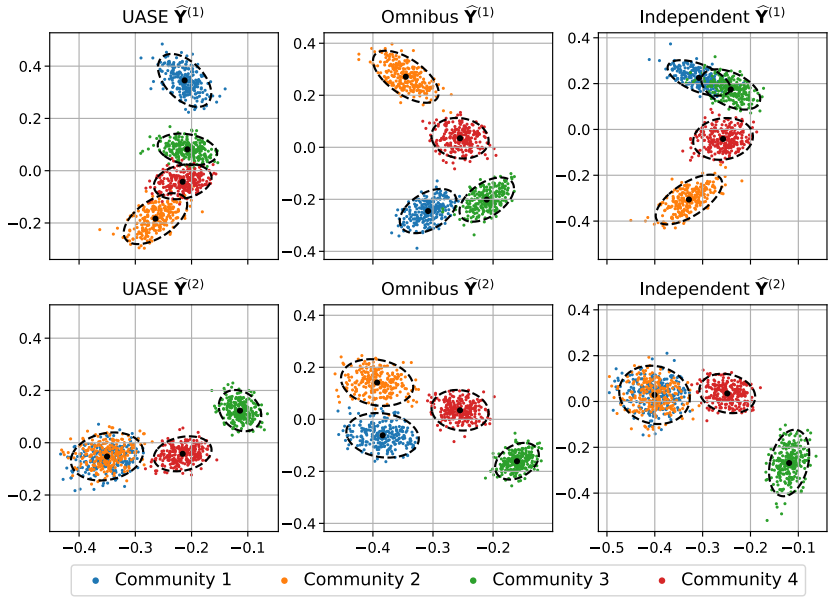

Figure 1: First two dimensions of the embeddings for the adjacency matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ using three different techniques: UASE, omnibus and separate embedding. The points are coloured according to true community membership, with the black dots showing the fitted community centroid and the ellipses a fitted 95% level Gaussian contour.

3 Setup

We consider a sequence of random graphs $G_1, \ldots, G_T$ distributed according to a dynamic latent position model as described above, in which the latent position sequences $\mathbf{Z}_i = (\mathbf{Z}_i^{(1)} | \cdots | \mathbf{Z}_i^{(T)})$ are independent of each other, and identically distributed according to a joint distribution $\mathcal{F}$ on $\mathcal{Z}^T$, for a bounded subset $\mathcal{Z}$ of $\mathbb{R}^k$. We can associate to the function f a compact, self-adjoint operator A on $L^2(\mathbb{R}^k)$ which has a—possibly infinite—sequence of non-zero eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$

with corresponding orthonormal eigenfunctions $u_1, u_2, \ldots \in L^2(\mathbb{R}^k)$, such that $Au_j = \lambda_j u_j$ (for a more detailed explanation of this construction, see for example [30]). We shall make the simplifying assumption that A has a *finite* number of non-zero eigenvalues, in which case f admits a canonical eigendecomposition

$$f(x, y) = \sum_{i=1}^{D} \lambda_i u_i(x) u_i(y), \tag{2}$$

assumed to hold everywhere, where after relabelling we may assume that $|\lambda_1| \geq \ldots \geq |\lambda_D|$. The individual graphs G_t defined as above can then be seen to be generated according to a GRDPG model [31]. Indeed, define a map $\phi : \mathcal{Z} \to \mathbb{R}^D$ by setting the ith coordinate of $\phi(\mathbf{z})$ to be $|\lambda_i|^{1/2} u_i(\mathbf{z})$, and let $\mathbf{X}_i^{(t)} = \phi(\mathbf{Z}_i^{(t)})$. Then $f(\mathbf{Z}_i^{(t)}, \mathbf{Z}_j^{(t)}) = g(\mathbf{X}_i^{(t)}, \mathbf{X}_j^{(t)})$, where $g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{I}_{p,q} \mathbf{y}$ and $\mathbf{I}_{p,q} = \mathrm{diag}(\mathrm{sgn}(\lambda_1), \ldots, \mathrm{sgn}(\lambda_D))$; precisely the definition of a GRDPG model. The following result shows that *collectively* the graphs are generated according to a *multilayer* RDPG model [14]:

Proposition 1. There exists a distribution $\mathcal{G}$ and matrices $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{\Lambda}$ such that $(\mathbf{A}, \mathbf{X}, \mathbf{Y}) \sim \mathrm{MRDPG}(\mathcal{G}, \mathbf{\Lambda})$, and linear maps $\mathcal{L} : \mathbb{R}^{TD} \to \mathbb{R}^d$ and $\mathcal{L}_t : \mathbb{R}^D \to \mathbb{R}^{d_t}$ (independent of n) such that $\mathbf{X}_i = \mathcal{L}\big((\mathbf{X}_i^{(1)} | \cdots | \mathbf{X}_i^{(T)})\big)$ and $\mathbf{Y}_i^{(t)} = \mathcal{L}_t(\mathbf{X}_i^{(t)})$, where $\mathbf{Y} = \mathrm{diag}(\mathbf{Y}^{(1)} | \cdots | \mathbf{Y}^{(T)})$.

The dimensions d_t and d are the asymptotic ranks of the Gram matrices $\mathbf{P}^{(t)} = \big(f(\mathbf{Z}_i^{(t)}, \mathbf{Z}_j^{(t)})\big)_{i,j \in [n]}$ and their concatenation $\mathbf{P} = (\mathbf{P}^{(1)} | \cdots | \mathbf{P}^{(T)})$, bounded above by D and TD, respectively.

We can extend our model to incorporate a range of sparsity regimes by scaling the function f by a sparsity factor ρ_n, which we assume is either constant and equal to 1, or else tends to zero as n grows, corresponding to dense and sparse regimes respectively. When this factor is present, the linear maps $\mathcal{L}$ and $\mathcal{L}_t$ (and consequently the points $\mathbf{X}_i$ and $\mathbf{Y}_i^{(t)}$) in Proposition 1 are scaled by a factor of $\rho_n^{1/2}$, and the distribution $\mathcal{G}$ is replaced by a similarly scaled distribution $\mathcal{G}_\rho$.

The UASE is constructed as follows. We form the concatenation $\mathbf{A} = (\mathbf{A}^{(1)} | \cdots | \mathbf{A}^{(T)})$, which admits a singular value decomposition of the form

$$\mathbf{A} = \mathbf{U}_\mathbf{A} \mathbf{\Sigma}_\mathbf{A} \mathbf{V}_\mathbf{A}^\top + \mathbf{U}_{\mathbf{A},\perp} \mathbf{\Sigma}_{\mathbf{A},\perp} \mathbf{V}_{\mathbf{A},\perp}^\top \tag{3}$$

where $\mathbf{\Sigma}_\mathbf{A}$ contains the d largest singular values of $\mathbf{A}$, and define the left and right embeddings

$$\hat{\mathbf{X}} = \mathbf{U}_\mathbf{A} \mathbf{\Sigma}_\mathbf{A}^{1/2} \in \mathbb{R}^{n \times d}, \quad \hat{\mathbf{Y}} = \mathbf{V}_\mathbf{A} \mathbf{\Sigma}_\mathbf{A}^{1/2} \in \mathbb{R}^{Tn \times d} \tag{4}$$

and further divide $\hat{\mathbf{Y}}$ into sub-embeddings $\hat{\mathbf{Y}}^{(t)} \in \mathbb{R}^{n \times d}$. Replacing $\mathbf{A}$ with $\mathbf{P}$ in this construction yields the *noise-free* embeddings $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}^{(t)}$, whose rows are known to be linear transformations of the rows of $\mathbf{X}$ and $\mathbf{Y}^{(t)}$ respectively (see Appendix for further details).

A desirable property of UASE is that since the matrices $\mathbf{A}^{(t)}$ follow a multilayer RDPG model there exist known asymptotic distributional results for the global embedding $\hat{\mathbf{X}}$ and—of particular interest to us—the graph-specific embeddings $\hat{\mathbf{Y}}^{(t)}$. In order to ensure the validity of these results, we will make the assumption that the sparsity factor ρ_n satisfies $\rho_n = \omega\left(\frac{\log^c(n)}{n^{1/2}}\right)$ for some universal constant $c > 1$.

The first of our distributional results states that after applying an orthogonal transformation—which leaves the structure of the resulting point cloud intact—the embedded points $\hat{\mathbf{Y}}_i^{(t)}$ converge in the Euclidean norm to the noise-free embedded points $\tilde{\mathbf{Y}}_i^{(t)}$ as the graph size grows:

Theorem 2. There exists a sequence of orthogonal matrices $\tilde{\mathbf{W}} \in \mathrm{O}(d)$ such that

$$\max_{i \in \{1, \ldots, n\}} \big\| \tilde{\mathbf{W}} \hat{\mathbf{Y}}_i^{(t)} - \tilde{\mathbf{Y}}_i^{(t)} \big\| = \mathrm{O}\left(\frac{\log^{1/2}(n)}{\rho_n^{1/2} n^{1/2}}\right) \tag{5}$$

almost surely for each t.

The second result (an extension of Theorem 2 in [14] incorporating ideas from [1]) states that after applying a second orthogonal transformation the above error converges in distribution to a fixed multivariate Gaussian distribution :

Theorem 3. For $\mathbf{z} \in \mathcal{Z}$, define

$$\mathbf{\Sigma}_t(\mathbf{z}) = \begin{cases} \mathbb{E}_\zeta\Big[f(\mathbf{z},\zeta_t)\big(1-f(\mathbf{z},\zeta_t)\big)\cdot\varphi(\zeta)\varphi(\zeta)^\top\Big] & \text{if } \rho_n = 1 \\ \mathbb{E}_\zeta\Big[f(\mathbf{z},\zeta_t)\cdot\varphi(\zeta)\varphi(\zeta)^\top\Big] & \text{if } \rho_n \to 0 \end{cases} \tag{6}$$

where $\zeta = (\zeta_1|\cdots|\zeta_T) \sim \mathcal{F}$ and $\varphi(\zeta) = \mathcal{L}\big((\phi(\zeta_1)|\cdots|\phi(\zeta_T))\big)$ for the linear map $\mathcal{L}$ appearing in Proposition 1. Then there exists a deterministic matrix $\mathbf{R}_* \in \mathbb{R}^{d\times d}$ and a sequence of orthogonal matrices $\mathbf{W} \in \mathrm{O}(d)$ such that, given $\mathbf{z} \in \mathcal{Z}$, for all $\mathbf{y} \in \mathbb{R}^d$ and for any fixed i,

$$\mathbb{P}\Big(n^{1/2}\mathbf{W}(\tilde{\mathbf{W}}\hat{\mathbf{Y}}_i^{(t)} - \tilde{\mathbf{Y}}_i^{(t)}) \le \mathbf{y} \mid \mathbf{Z}_i^{(t)} = \mathbf{z}\Big) \to \Phi\big(\mathbf{y}, \mathbf{R}_*\mathbf{\Sigma}_t(\mathbf{z})\mathbf{R}_*^\top\big) \tag{7}$$

for each t.

We can use these results to demonstrate one of the key advantages of using UASE over other embedding methods, namely that *UASE exhibits both cross-sectional and longitudinal stability*, in a sense that we shall now define. We say that two space-time positions $(\mathbf{z}, t)$ and $(\mathbf{z}', t')$ are *exchangeable* if $f(\mathbf{z}, \zeta_t) = f(\mathbf{z}', \zeta_{t'})$ with probability one, where $\zeta = (\zeta_1|\cdots|\zeta_T) \sim \mathcal{F}$, and that the positions are *exchangeable up to degree* if $f(\mathbf{z}, \zeta_t) = \alpha f(\mathbf{z}', \zeta_{t'})$ for some $\alpha > 0$.

Definition 4. A generic method for embedding a dynamic network, whose output is denoted $(\hat{\mathbf{Z}}_i^{(t)})_{i\in[n];t\in[T]}$, is said to exhibit *cross-sectional stability* if, for exchangeable pairs $(\mathbf{z}, t)$ and $(\mathbf{z}', t)$, the points $\hat{\mathbf{Z}}_i^{(t)}$ and $\hat{\mathbf{Z}}_j^{(t)}$ are asymptotically equal in distribution, conditional on $\mathbf{Z}_i^{(t)} = \mathbf{z}$ and $\mathbf{Z}_j^{(t)} = \mathbf{z}'$. Similarly, the method exhibits *longitudinal stability* if $\hat{\mathbf{Z}}_i^{(t)}$ and $\hat{\mathbf{Z}}_i^{(t')}$ are asymptotically equal in distribution, conditional on $\mathbf{Z}_i^{(t)} = \mathbf{Z}_i^{(t')} = \mathbf{z}$, for exchangeable pairs $(\mathbf{z}, t)$ and $(\mathbf{z}, t')$.

The following result—a consequence of Theorem 3—shows that UASE exhibits both types of stability:

Corollary 5. Conditional on $\mathbf{Z}_i^{(t)} = \mathbf{z}$ and $\mathbf{Z}_j^{(t')} = \mathbf{z}'$, the following properties hold:

1. If $(\mathbf{z}, t)$ and $(\mathbf{z}', t')$ are exchangeable then $\tilde{\mathbf{Y}}_i^{(t)} = \tilde{\mathbf{Y}}_j^{(t')}$ and $\mathbf{\Sigma}_t(\mathbf{z}) = \mathbf{\Sigma}_{t'}(\mathbf{z}')$, and so $\hat{\mathbf{Y}}_i^{(t)}$ and $\hat{\mathbf{Y}}_j^{(t')}$ are asymptotically equal in distribution.

2. If $(\mathbf{z}, t)$ and $(\mathbf{z}', t')$ are exchangeable up to degree then $\tilde{\mathbf{Y}}_i^{(t)} = \alpha\tilde{\mathbf{Y}}_j^{(t')}$, and, under a sparse regime, $\mathbf{\Sigma}_t(\mathbf{z}) = \alpha\mathbf{\Sigma}_{t'}(\mathbf{z}')$. In the asymptotic limit $\hat{\mathbf{Y}}_i^{(t)}$ and $\hat{\mathbf{Y}}_j^{(t')}$ follow distributions centred on a ray projecting from the origin in $\mathbb{R}^d$, with proportional covariance matrices under a sparse regime.

We can gain insight into our theoretical results by applying them in the context of common statistical models. For our motivating dynamic stochastic block model example detailed in Section 2, $f(\mathbf{Z}_i^{(t)}, \mathbf{Z}_j^{(t)}) = \mathbf{B}_{z_i,z_j}^{(t)}$, where $\mathbf{B}^{(t)}$ denotes the inter-community link probability matrices and z_i and z_j denote the community membership of the ith and jth nodes respectively. A pair $(\mathbf{z}, t)$ and $(\mathbf{z}', t')$ are exchangeable if and only if the corresponding rows of $\mathbf{B}^{(t)}$ and $\mathbf{B}^{(t')}$ are identical. Theorem 3 then predicts that the point cloud obtained via UASE should decompose into a finite number of *Gaussian* clusters, as we observe in Figure 1. Moreover, when combined with Corollary 5 the equality of the fourth rows of $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ implies that we should expect *identical* clusters (centre and shape) corresponding to the fourth community at both time points (indicating longitudinal stability) and similarly the equality of the first and second rows of $\mathbf{B}^{(2)}$ tells us that the communities should be indistinguishable at the second time point (indicating cross-sectional stability), behaviours we observe in Figure 1.

For the dynamic degree-corrected stochastic block model [15, 24], $f(\mathbf{Z}_i^{(t)}, \mathbf{Z}_j^{(t)}) = \omega_i^{(t)}\omega_j^{(t)}\mathbf{B}_{z_i,z_j}^{(t)}$, where the $\omega_i^{(t)} \in (0, 1]$ are node-specific parameters which describe the activity level of each node at a given time point. In this setting, a pair $(\mathbf{z}, t)$ and $(\mathbf{z}', t')$ are exchangeable *up to degree* if and only if the corresponding rows of $\mathbf{B}^{(t)}$ and $\mathbf{B}^{(t')}$ are identical. Consequently, Corollary 5 indicates that nodes belonging to the same community under such a model should have their corresponding UASE estimates distributed along a ray projecting from the origin, a behaviour that is exhibited in Figure 2.

4 Comparison

A method for embedding a dynamic network will generally admit one of the following descriptions:

1. A single, joint spectral decomposition, e.g. omnibus [21] or UASE [14];
2. A separate spectral decomposition obtained at each timepoint, from an individual or temporally smoothed graph [4, 35, 6, 2, 27, 32, 16];
3. A joint matrix factorisation, with smoothness constraints [22, 5, 47, 3, 23].

As mentioned in the introduction if, rather than obtaining a dynamic embedding, interest is solely on recovering dynamic community structure, several other approaches based on the dynamic stochastic block model are available. We also do not include in this comparison inference procedures which are based on an explicit parametric dynamic latent position network model [34, 12, 13, 18, 28, 19, 7, 37, 9, 8], reviews of which can be found in [17] and [39], typically Bayesian and/or variational, which are much harder to analyse and not usually regarded as embedding procedures *per se*.

We define the spectral embedding of an arbitrary symmetric matrix $\mathbf{M} \in \mathbb{R}^{N\times N}$ to be $\mathbf{X}_{\mathbf{M}} = \mathbf{U}|\mathbf{S}|^{1/2} \in \mathbb{R}^{N\times d}$ (removing the subscript if clear from context), where $\mathbf{S} \in \mathbb{R}^{d\times d}$ is a diagonal matrix containing the d largest eigenvalues of $\mathbf{M}$, by magnitude, and $\mathbf{U} \in \mathbb{R}^{N\times d}$ is a matrix containing corresponding orthonormal eigenvectors.

The omnibus method computes the spectral embedding of the matrix $\tilde{\mathbf{A}} \in \{0,1\}^{nT\times nT}$ whose k,ℓth $n\times n$ block is $(\mathbf{A}^{(k)} + \mathbf{A}^{(\ell)})/2$, giving a point cloud $\mathbf{X}_{\tilde{\mathbf{A}}} \in \mathbb{R}^{nT\times d}$ whose first n rows provide an embedding for the nodes at time 1, and so forth. In the second category, the sequence of spectral embeddings of the time-averaged matrices

$$\bar{\mathbf{A}}^{(t)} = \sum_k w_k \mathbf{A}^{(t-k)}, \tag{8}$$

is computed, where w_k are non-negative weights (e.g. sliding window average, exponential forgetting factors, and more), giving a point cloud $\mathbf{X}^{(t)}_{\bar{\mathbf{A}}} \in \mathbb{R}^{n\times d}$ for each time point. Alternatively, time-averaged versions of the normalised Laplacian or Katz matrix can be used. The description 'matrix factorisation', in the third category, is used to refer to a low-rank decomposition of a matrix, possibly under constraints (such as positivity, or smoothness over time), which is not based on a straightforward spectral decomposition. An example is the sequence of spectral embeddings computed from the matrices [23]

$$\text{argmin}_{\breve{\mathbf{A}}^{(1)},\ldots,\breve{\mathbf{A}}^{(T)}} \sum_{t=1}^{T} \|\mathbf{A}^{(t)} - \breve{\mathbf{A}}^{(t)}\|_F^2 + \alpha \sum_{t=1}^{T-1} \|\breve{\mathbf{A}}^{(t)} - \breve{\mathbf{A}}^{(t+1)}\|_F^2, \tag{9}$$

where $\alpha > 0$ (a high value favouring a smooth solution), subject to a low rank constraint on $\breve{\mathbf{A}}^{(1)},\ldots,\breve{\mathbf{A}}^{(T)}$, and again the normalised Laplacian (used by [23]) or Katz matrix could equally well be used.

4.1 Stability of other embeddings

To show how those methods fail to provide both cross-sectional and longitudinal stability, we establish some elementary facts about the spectral embedding $\mathbf{X}$ of a matrix $\mathbf{M}$. Although we consider the general case involving positive and negative eigenvalues, the issues we raise are very much present in the positive-definite case too. If $\mathbf{M}$ has signature (p,q) (p positive and q negative eigenvalues), its spectral embedding into $d = p+q$ dimensions satisfies $\mathbf{X} = \mathbf{M}\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{I}_{p,q}$ by the Moore-Penrose inverse, so that $\mathbf{X}_i = \mathbf{M}_i\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{I}_{p,q}$ where the subscript i picks out the ith row. Therefore,

(A1) $\mathbf{X}_i = \mathbf{X}_j$ if and only if $\mathbf{M}_i = \mathbf{M}_j$.

Given a subset of indices S, we denote $\mathbf{X}_S$ as the corresponding rows $(\mathbf{X})_{i\in S}$ and $\mathbf{M}_{S,S}$ as the matrix of corresponding rows and columns $(\mathbf{M}_{i,j})_{i,j\in S}$. Suppose we have two symmetric matrices $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$ of signature (p,q), with (non-equal) spectral embeddings $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ respectively. If there exists a subset S such that the principal minors $\mathbf{M}^{(1)}_{S,S}$ and $\mathbf{M}^{(2)}_{S,S}$ are equal with signature (p,q),

(A2) $\mathbf{X}_S^{(1)} = \mathbf{X}_S^{(2)}\mathbf{Q}$ for some $\mathbf{Q} \in \mathrm{O}(p, q)$.

Furthermore, in general, $\mathbf{X}_S^{(1)} \neq \mathbf{X}_S^{(2)}$, and the indefinite Procrustes matrix that best aligns $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$,

$$\hat{\mathbf{Q}} = \mathrm{argmin}_{\mathbf{Q} \in \mathrm{O}(p,q)} \|\mathbf{X}^{(1)}\mathbf{Q} - \mathbf{X}^{(2)}\|_F^2, \tag{10}$$

does not align the subsets of the spectral embeddings, $\mathbf{X}_S^{(1)} \neq \mathbf{X}_S^{(2)}\hat{\mathbf{Q}}$.

It will be considered sufficient to establish whether an embedding is stable when applied to the Gram matrices $\mathbf{P}^{(1)}, \ldots, \mathbf{P}^{(T)}$. A method found to be unstable in this noise-free condition is not expected to be stable when the matrices are replaced by their noisy observations $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(T)}$. For the case of omnibus embedding, we therefore consider the matrix $\tilde{\mathbf{P}} \in [0, 1]^{nT \times nT}$, the noise-free counterpart of $\tilde{\mathbf{A}}$. If $(\mathbf{z}, t)$ and $(\mathbf{z}, t')$ are exchangeable, then, conditional on $\mathbf{Z}_i^{(t)} = \mathbf{Z}_i^{(t')} = \mathbf{z}$, the rows $\tilde{\mathbf{P}}_{n(t-1)+i}$ and $\tilde{\mathbf{P}}_{n(t'-1)+i}$ are equal. But if $(\mathbf{z}, t)$ and $(\mathbf{z}', t)$ are exchangeable, then $\tilde{\mathbf{P}}_{n(t-1)+i} \neq \tilde{\mathbf{P}}_{n(t-1)+j}$, in general, when $\mathbf{Z}_i^{(t)} = \mathbf{Z}_j^{(t)} = \mathbf{z}$. By (A1), *omnibus embedding provides longitudinal but not cross-sectional stability*, a fact demonstrated in the central column of Figure 1.

Methods in the second and third category generally trade-off longitudinal against cross-sectional stability, seemingly never providing both. For example, if we pick $w_0 = 1, w_k = 0$ for $k \neq 0$ in (8) as an example of the second, or $\alpha = 0$ in (9) as an example of the third, the methods reduce to independent spectral embedding of the graphs, where cross-sectional stability is guaranteed. However, the embeddings are subject to transformations $\mathbf{Q} \in \mathrm{O}(p, q)$ [31], which will not be the same for all time steps, in general. By (A2) there exists a transformation which will realign the stable subset of nodes, but it requires knowing or estimating that stable subset, the latter difficult under noise. In this way, *independent spectral embedding provides cross-sectional but not longitudinal stability*. At the other extreme, where $w_k = 1/T$ in (8), or $\alpha \to \infty$ in (9), we evidently achieve longitudinal but not cross-sectional stability.

5 Real data

The Lyon primary school dataset shows the social interactions at a French primary school over two days in October 2009 [38]. The school consisted of 10 teachers and 241 students from five school years, each year divided into two classes. Face-to-face interactions were detected when radio-frequency identification devices worn by participants (10 teachers and 232 students gave consent to be included in the experiment) were in close proximity over an interval of 20 seconds and recorded as a pair of anonymous identifiers together with a timestamp. The data are available for download from the Network Repository website[1] [29].

A time series of networks was created by binning the data into hour-long windows over the two days, from 08:00 to 18:00 each day. If at least one interaction was observed between two people in a particular time window, an edge was created to connect the two nodes in the corresponding network. This results in a time series of graphs $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(20)}$ each with $n = 242$ nodes. Where a node is not active in a given time window, it is still included in the graph as an isolated node. This is compatible with the theory and method, and the node is embedded to the zero vector at that time point.

Given the unfolded adjacency matrix $\mathbf{A} = (\mathbf{A}^{(1)}|\cdots|\mathbf{A}^{(20)})$, an estimated embedding dimension $\hat{d} = 10$ was obtained using profile likelihood [48] and we construct the embeddings $\hat{\mathbf{Y}}^{(1)}, \ldots, \hat{\mathbf{Y}}^{(20)} \in \mathbb{R}^{n \times 10}$, taking approximately five seconds on a 2017 MacBook Pro. Figure 2 shows the first two dimensions of this embedding to visualise some of the structure in the data.

From this plot, we observe clustering of students in the same school class. For time windows corresponding to classroom time, for example, 09:00–10:00 and 15:00–16:00, the embedding forms rays of points in 10-dimensional space, with each ray broadly corresponding to a single school class. This is to be expected under a degree-corrected stochastic block model, and the distance along the ray is a measure of the node's activity level [20, 25, 33]. However, not all time windows exhibit this structure, for example, the different classes mix more during lunchtimes (time windows 12:00–13:00 and 13:00–14:00).

[1] https://networkrepository.com

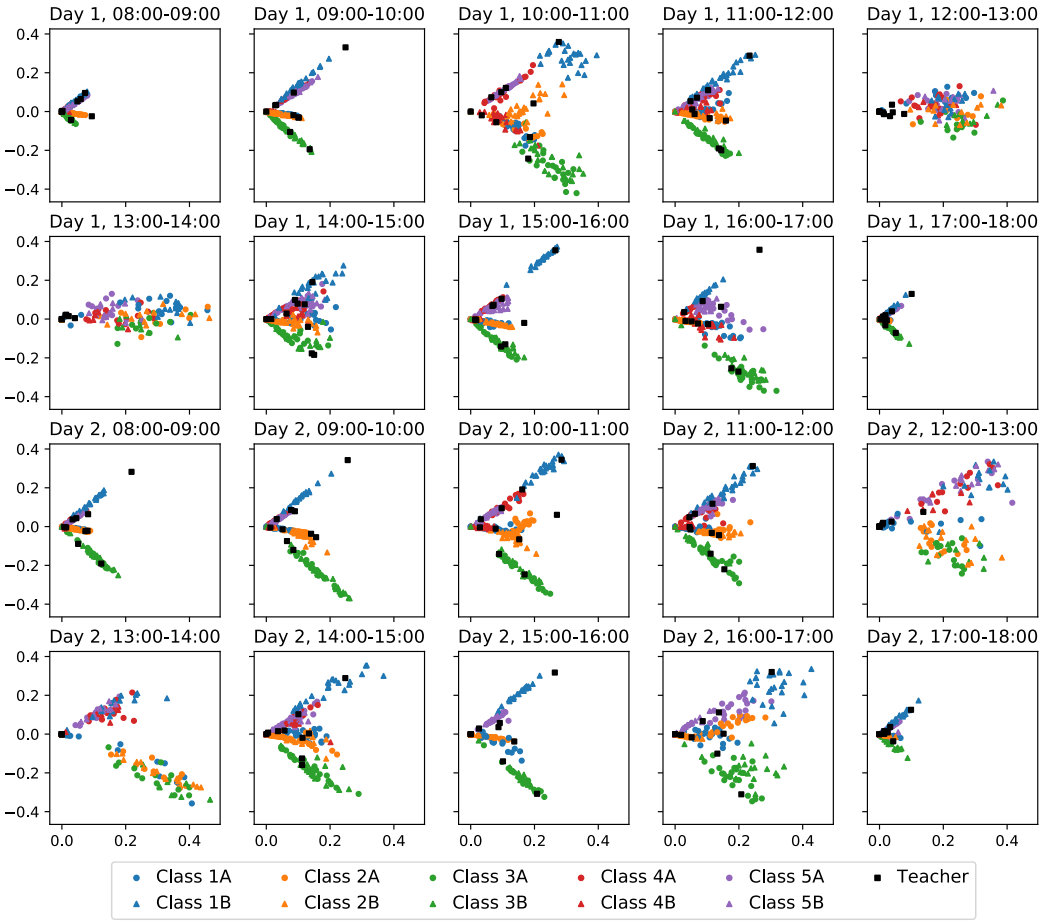

Figure 2: First two dimensions of the embeddings $\hat{\mathbf{Y}}^{(1)}, \ldots, \hat{\mathbf{Y}}^{(20)}$ of the unfolded adjacency matrix $\mathbf{A} = (\mathbf{A}^{(1)} | \cdots | \mathbf{A}^{(20)})$. The colours indicate different school years while the marker type distinguish the two school classes within each year.

5.1 Clustering

Following recommendations regarding community detection under a degree-corrected stochastic block model [33], we analyse UASE using spherical coordinates $\mathbf{\Theta}^{(t)} \in [0, 2\pi)^{n \times 9}$, for $t \in [T]$. Since UASE demonstrates cross-sectional and longitudinal stability, we can combine the embeddings into a single point cloud $\mathbf{\Theta} = (\mathbf{\Theta}^{(1)\top} | \cdots | \mathbf{\Theta}^{(T)\top})^\top \in \mathbb{R}^{nT \times 9}$ where each point represents a student or teacher in a particular time window. This allows us to detect people returning to a previous behaviour in the dynamic network. We fit a Gaussian mixture model with varying covariance matrices to the non-zero points in $\mathbf{\Theta}$ with 20–50 clusters increasing in increments of 5, with 50 random initialisations, taking approximately five minutes on a 2017 MacBook Pro. Using the Bayesian Information Criterion, we select the best fitting model (30 clusters) and assign the maximum a posteriori Gaussian cluster membership to each student in each time window.

Figure 3 shows how students in the ten classes move between these clusters over time. Each class has one or two clusters unique to it, for example, the majority of students in class 1A spend their classroom time (as opposed to break time) assigned to cluster 1 or cluster 25. This highlights the importance of longitudinal stability in UASE, as we are detecting points in the embedding returning to some part of latent space.

There are also instances of multiple school classes being assigned the same cluster at the same time period, for example, on the morning of day 1, classes 5A and 5B are mainly in cluster 28 suggesting they are having a joint lesson, and we see this behaviour again on day 2 with classes 3A and 3B in

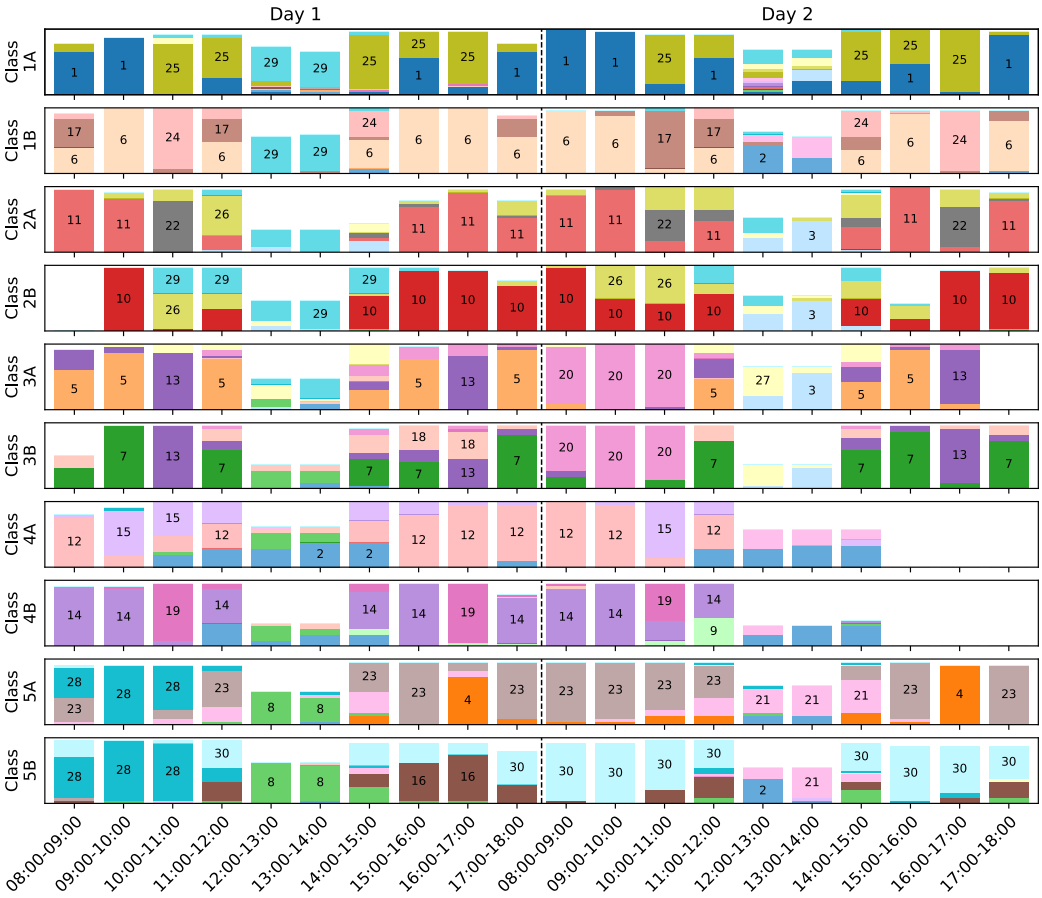

Figure 3: Bar chart showing the Gaussian cluster assignment of school classes over time. The height of each coloured bar represents the proportion of students, in that class and at that time, assigned to the corresponding Gaussian cluster, the total available height representing 100%. If the coloured bars do not sum to the full available height, the difference represents the proportion of inactive students. For legibility, only bars representing over 35% of the class are labelled with the cluster number.

cluster 20. In the lunchtime periods, particularly on day 1, the younger students (classes 1A–2B) mingle to form a larger cluster, as do the older students (classes 4A–5B), potentially explained by the cafeteria needing two sittings for lunch for space reasons [38]. This highlights the importance of cross-sectional stability in UASE, as it allows the grouping of nodes behaving similarly in a specific time window, irrespective of their potentially different past and future behaviours.

6 Conclusion

We prove that an existing procedure, UASE, allows dynamic network spectral embedding with longitudinal and cross-sectional stability guarantees. These properties make a range of subsequent spatio-temporal analyses possible using 'off-the-shelf' techniques for clustering, time series analysis, classification and more. While our goal is to facilitate exploratory data analysis across a range of scientific disciplines, we have an ethical concern relating to the possibility of matching a node's behaviour between two graphs, despite being labelled differently. This could be exploited to compromise someone's identity, for example, across online social networks. We hope to raise awareness of such opportunities, for instance, in the context of network data anonymisation.

References

[1] Agterberg, J., Tang, M., and Priebe, C. E. (2020). On two distinct sources of nonidentifiability in latent position random graph models. *arXiv preprint arXiv:2003.14250*.

[2] Bhattacharyya, S. and Chatterjee, S. (2018). Spectral clustering for multiple sparse networks: I. *arXiv preprint arXiv:1805.10594*.

[3] Chen, H. and Li, J. (2018). Exploiting structural and temporal evolution in dynamic link prediction. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 427–436.

[4] Chi, Y., Song, X., Zhou, D., Hino, K., and Tseng, B. L. (2007). Evolutionary spectral clustering by incorporating temporal smoothness. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 153–162.

[5] Deng, D., Shahabi, C., Demiryurek, U., Zhu, L., Yu, R., and Liu, Y. (2016). Latent space model for road networks to predict time-varying traffic. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1525–1534.

[6] Dunlavy, D. M., Kolda, T. G., and Acar, E. (2011). Temporal link prediction using matrix and tensor factorizations. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(2):1–27.

[7] Durante, D. and Dunson, D. B. (2014). Nonparametric Bayes dynamic modelling of relational data. *Biometrika*, 101(4):883–898.

[8] Durante, D., Mukherjee, N., and Steorts, R. C. (2017). Bayesian learning of dynamic multilayer networks. *The Journal of Machine Learning Research*, 18(1):1414–1442.

[9] Friel, N., Rastelli, R., Wyse, J., and Raftery, A. E. (2016). Interlocking directorates in Irish companies using a latent space model for bipartite networks. *Proceedings of the National Academy of Sciences*, 113(24):6629–6634.

[10] Gao, S., Denoyer, L., and Gallinari, P. (2011). Temporal link prediction by integrating content and structure information. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pages 1169–1174.

[11] Ho, Q., Song, L., and Xing, E. (2011). Evolving cluster mixed-membership blockmodel for time-evolving networks. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 342–350. JMLR Workshop and Conference Proceedings.

[12] Hoff, P. D. (2011). Hierarchical multilinear models for multiway data. *Computational Statistics & Data Analysis*, 55(1):530–543.

[13] Hoff, P. D. et al. (2011). Separable covariance arrays via the Tucker product, with applications to multivariate relational data. *Bayesian Analysis*, 6(2):179–196.

[14] Jones, A. and Rubin-Delanchy, P. (2021). The multilayer random dot product graph. *arXiv preprint arXiv:2007.10455*.

[15] Karrer, B. and Newman, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E (3)*, 83(1):016107, 10.

[16] Keriven, N. and Vaiter, S. (2020). Sparse and smooth: improved guarantees for spectral clustering in the dynamic stochastic block model. *arXiv preprint arXiv:2002.02892*.

[17] Kim, B., Lee, K. H., Xue, L., and Niu, X. (2018). A review of dynamic network models with latent variables. *Statistics surveys*, 12:105.

[18] Lee, N. H. and Priebe, C. E. (2011). A latent process model for time series of attributed random graphs. *Statistical inference for stochastic processes*, 14(3):231–253.

[19] Lee, N. H., Yoder, J., Tang, M., and Priebe, C. E. (2013). On latent position inference from doubly stochastic messaging activities. *Multiscale Modeling & Simulation*, 11(3):683–718.

[20] Lei, J. and Rinaldo, A. (2015). Consistency of spectral clustering in stochastic block models. *Ann. Statist.*, 43(1):215–237.

[21] Levin, K., Athreya, A., Tang, M., Lyzinski, V., Park, Y., and Priebe, C. E. (2017). A central limit theorem for an omnibus embedding of multiple random graphs and implications for multiscale network inference. *arXiv preprint arXiv:1705.09355*.

[22] Lin, Y.-R., Chi, Y., Zhu, S., Sundaram, H., and Tseng, B. L. (2008). Facetnet: a framework for analyzing communities and their evolutions in dynamic networks. In *Proceedings of the 17th international conference on World Wide Web*, pages 685–694.

[23] Liu, F., Choi, D., Xie, L., and Roeder, K. (2018). Global spectral clustering in dynamic networks. *Proceedings of the National Academy of Sciences*, 115(5):927–932.

[24] Liu, S., Wang, S., and Krishnan, R. (2014). Persistent community detection in dynamic social networks. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 78–89. Springer.

[25] Lyzinski, V., Sussman, D. L., Tang, M., Athreya, A., and Priebe, C. E. (2014). Perfect clustering for stochastic blockmodel graphs via adjacency spectral embedding. *Electron. J. Stat.*, 8(2):2905–2922.

[26] Matias, C. and Miele, V. (2015). Statistical clustering of temporal networks through a dynamic stochastic block model. *arXiv preprint arXiv:1506.07464*.

[27] Pensky, M., Zhang, T., et al. (2019). Spectral clustering in the dynamic stochastic block model. *Electronic Journal of Statistics*, 13(1):678–709.

[28] Robinson, L. F. and Priebe, C. E. (2012). Detecting time-dependent structure in network data via a new class of latent process models. *arXiv preprint arXiv:1212.3587*.

[29] Rossi, R. A. and Ahmed, N. K. (2015). The network data repository with interactive graph analytics and visualization. In *AAAI*.

[30] Rubin-Delanchy, P. (2020). Manifold structure in graph embeddings. In *Proceedings of the Thirty-fourth Conference on Neural Information Processing Systems*.

[31] Rubin-Delanchy, P., Cape, J., Priebe, C. E., and Tang, M. (2020). A statistical interpretation of spectral embedding: the generalised random dot product graph. *arXiv preprint arXiv:1709.05506v3*.

[32] Sanna Passino, F., Bertiger, A. S., Neil, J. C., and Heard, N. A. (2019). Link prediction in dynamic networks using random dot product graphs. *arXiv preprint arXiv:1912.10419*.

[33] Sanna Passino, F., Heard, N. A., and Rubin-Delanchy, P. (2020). Spectral clustering on spherical coordinates under the degree-corrected stochastic blockmodel. *arXiv e-prints*, pages arXiv–2011.

[34] Sarkar, P. and Moore, A. W. (2005). Dynamic social network analysis using latent space models. *ACM SigKDD Explorations Newsletter*, 7(2):31–40.

[35] Scheinerman, E. R. and Tucker, K. (2010). Modeling graphs using dot product representations. *Computational statistics*, 25(1):1–16.

[36] Seto, S., Zhang, W., and Zhou, Y. (2015). Multivariate time series classification using dynamic time warping template selection for human activity recognition. In *2015 IEEE Symposium Series on Computational Intelligence*, pages 1399–1406. IEEE.

[37] Sewell, D. K. and Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association*, 110(512):1646–1657.

[38] Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J.-F., Quaggiotto, M., Van den Broeck, W., Régis, C., Lina, B., et al. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. *PloS one*, 6(8):e23176.

[39] Turnbull, K. (2020). *Advancements in latent space network modelling*. PhD thesis, Lancaster University.

[40] Xie, Y., Li, C., Yu, B., Zhang, C., and Tang, Z. (2020). A survey on dynamic network embedding. *arXiv preprint arXiv:2006.08093*.

[41] Xing, E. P., Fu, W., Song, L., et al. (2010). A state-space mixed membership blockmodel for dynamic network tomography. *Annals of Applied Statistics*, 4(2):535–566.

[42] Xu, K. (2015). Stochastic block transition models for dynamic networks. In *Artificial Intelligence and Statistics*, pages 1079–1087. PMLR.

[43] Xu, K. S. and Hero, A. O. (2014). Dynamic stochastic blockmodels for time-evolving social networks. *IEEE Journal of Selected Topics in Signal Processing*, 8(4):552–562.

[44] Xue, G., Zhong, M., Li, J., Chen, J., Zhai, C., and Kong, R. (2021). Dynamic network embedding survey. *arXiv preprint arXiv:2103.15447*.

[45] Yang, T., Chi, Y., Zhu, S., Gong, Y., and Jin, R. (2011). Detecting communities and their evolutions in dynamic social networks—a Bayesian approach. *Machine learning*, 82(2):157–189.

[46] Zhao, B., Lu, H., Chen, S., Liu, J., and Wu, D. (2017). Convolutional neural networks for time series classification. *Journal of Systems Engineering and Electronics*, 28(1):162–169.

[47] Zhu, L., Guo, D., Yin, J., Ver Steeg, G., and Galstyan, A. (2016). Scalable temporal latent space inference for link prediction in dynamic social networks. *IEEE Transactions on Knowledge and Data Engineering*, 28(10):2765–2777.

[48] Zhu, M. and Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile likelihood. *Computational Statistics & Data Analysis*, 51(2):918–930.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] Scenarios where our technique for spectral embedding are not appropriate are discussion in Section 1.
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] See Section 6.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] See Section 3.
   (b) Did you include complete proofs of all theoretical results? [Yes] See Appendix.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] See supplemental material for data and Python Notebooks.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] Procedures for parameter choices, for example, embedding dimension and number of Gaussian mixtures, given in Section 5.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] Cross-validation error bars given for classification accuracy in Section 5.
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] Run times given for the exploratory data analysis in Section 5 for a 2017 MacBook Pro.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] Lyon primary school social interaction data [38].
   (b) Did you mention the license of the assets? [Yes] Data hosting website `https://networkrepository.com` [29].
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] Discussion of number of consenting individuals in the Lyon primary school social interaction experiment.
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes] Discussion of anonymisation in the Lyon primary school social interaction experiment.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]