DIRECTED SCATTERING FOR KNOWLEDGE GRAPH-BASED CELLULAR SIGNALING ANALYSIS

Aarthi Venkat[1,], Joyce Chew[2,*], Ferran Cardoso Rodriguez[3], Christopher J. Tape[3],*
Michael Perlmutter[4,§], Smita Krishnaswamy[1,5,§]

[1]Program for Computational Biology and Bioinformatics, Yale University
[2]Department of Mathematics, University of California, Los Angeles [3]Department of Oncology, University College London Cancer Institute
[4]Department of Mathematics, Boise State University [5]Departments of Genetics and Computer Science, Yale University
[*] These authors contributed equally. [§] These authors jointly supervised the work.

ABSTRACT

Directed graphs are a natural model for many phenomena, in particular scientific knowledge graphs such as molecular interaction or chemical reaction networks that define cellular signaling relationships. In these situations, source nodes typically have distinct biophysical properties from sinks. Due to their ordered and unidirectional relationships, many such networks also have hierarchical and multiscale structure. However, the majority of methods performing node- and edge-level tasks in machine learning do not take these properties into account, and thus have not been leveraged effectively for scientific tasks such as cellular signaling network inference. We propose a new framework called Directed Scattering Autoencoder (DSAE) which uses a directed version of a geometric scattering transform, combined with the non-linear dimensionality reduction properties of an autoencoder and the geometric properties of the hyperbolic space to learn latent hierarchies. We show this method outperforms numerous others on tasks such as embedding directed graphs and learning cellular signaling networks.

Index Terms— geometric scattering, graph signal processing, hyperbolic geometry, directed graphs, biological networks

1. INTRODUCTION

Graph representation learning methods, or methods which learn low-dimensional feature representations for each node in a graph, have shown strong performance and widespread utility for preserving structural information for downstream tasks, including link prediction, community detection, and node classification [1]. However, most existing approaches are designed for undirected graphs. This is despite the fact that many relations of interest can be best described as directed; in fact, common benchmarks for graph representation learning, including citation networks, are directed and symmetrized for evaluation. Such real-world relationships also often exhibit hierarchical structure, motivating approaches that leverage hyperbolic geometry [2, 3, 4]. Notably, many scientific networks are directed but are not regularly benchmarked or represented with direction or latent hierarchies [5]. This includes graphs for cellular signaling, consisting of molecular interactions occurring in a highly ordered series resulting in directed and tree-like relationships [6].

Recently, efforts to incorporate the geometric scattering transform, a generalization of the Euclidean scattering transform to the graph domain, have proven effective in capturing multiscale graph structure and overcoming oversmoothing for undirected graph neural networks [7]. While there exist directed graph neural network-based approaches for supervised tasks [8], we hypothesized geometric scattering would better preserve the graph geometry and improve representations for unsupervised learning. Here, we introduce our framework, Directed Scattering Autoencoder (DSAE), which incorporates a directed version of the geometric scattering transform based on the eigendecomposition of the magnetic Laplacian. Additionally, as the resulting feature representation is high-dimensional,

we employ an autoencoder architecture with either Euclidean or hyperbolic geometry to learn a meaningful low-dimensional representation, where the hyperbolic version accounts for additional hierarchical structure. We evaluate DSAE on two previously benchmarked WebKB graphs [8] and three cellular signaling graphs [9, 10, 11], comparing against 11 undirected, directed, knowledge graph-based, and hyperbolic-based graph representation learning methods. DSAE performs best for link prediction on four of five graphs due to the ability to account for directedness, higher-order, and hierarchical structure. Second, we describe how to adapt our pipeline by adding additional regularizations to use DSAE for data-driven cellular signaling network inference, demonstrating the efficacy of this approach on a mouse visual cortex dataset with matched spatial data.

2. BACKGROUND

2.1. Magnetic Laplacian

Many methods in (undirected) graph signal processing rely on the eigendecomposition of the graph Laplacian, $L = D - A$. The eigenvalues are viewed as generalized frequencies and the eigenvectors are viewed as generalized Fourier modes. However, it is not straightforward to extend these methods to directed graphs since the naive definition of the graph Laplacian will in general not be symmetric or diagonalizable. While a number of solutions have been proposed [12, 13, 14], here, we shall use a complex Hermitian matrix $L^{(q)}$, known as the magnetic Laplacian (see, e.g, [15, 16]).

To construct the magnetic Laplacian, we let A be the (asymmetric) adjacency matrix of a directed graph $G = (V, E)$, $N = |V|$, let $A^{(\text{sym})} = \frac{1}{2}(A + A^T)$ be its symmetrized counterpart, and let $D^{(\text{sym})}$ be the diagonal degree matrix corresponding to $A^{(\text{sym})}$, i.e. $D_{j,j}^{(\text{sym})} = \sum_{k=0}^{N-1} A_{j,k}^{(\text{sym})}$, and $D_{j,k}^{(\text{sym})} = 0$ if $j \neq k$. Then, we capture directional information via the phase matrix $\Theta^{(q)}$, where for $q \geq 0$, we define $\Theta^{(q)} = 2\pi q(A - A^T)$. Letting $i = \sqrt{-1}$, this allows us to define the complex Hermitian adjacency matrix by

$$H^{(q)} = A^{(\text{sym})} \odot \exp(i\Theta^{(q)}),$$

where $\odot$ denotes the Hadamard product and the exponentiation is performed term-by-term. We then define the unnormalized and normalized magnetic Laplacians by

$$L_U^{(q)} = D^{(\text{sym})} - H^{(q)} \quad \text{and} \quad L_N^{(q)} = D^{(\text{sym})-1/2} L_U^{(q)} D^{(\text{sym})-1/2}$$

respectively. Since $\Theta^{(q)}$ is skew-symmetric, it is clear that both $L_U^{(q)}$ and $L_N^{(q)}$ are Hermitian, and Theorem 1 of [8] shows that they are positive semi-definite. Therefore, they both admit orthonormal bases of eigenvectors with non-negative eigenvalues $\mathbf{u}_k$, $0 \leq k \leq N - 1$, $L^{(q)}\mathbf{u}_k = \lambda_k\mathbf{u}_k$, $0 \leq \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{N-1}$ (where $L^{(q)}$ is either

$L_U^{(q)}$ or $L_N^{(q)}$).

2.2. The Geometric Scattering Transform

The geometric scattering transform [17, 18, 19] is a recently proposed multiscale, multi-order signal transform that applies an alternating sequence of wavelet transforms and non-linear activations to an input signal $\mathbf{x} : V \to \mathbb{R}$, building off of analogous work for functions defined on $\mathbb{R}^n$ [20]. Given a wavelet frame, $\mathcal{W}_J = \{W_j\}_{j=0}^J \cup \{A_J\}$ (where W_j and A_J are $n \times n$ matrices) and an input signal $\mathbf{x} : V \to \mathbb{R}$, which we identify with the vector $\mathbf{x}[k] = \mathbf{x}(v_k)$, we define first- and second-order scattering coefficients of $\mathbf{x}$, for $0 \le j_1 \le j_2 \le J$, by

$$S[j_1]\mathbf{x} = A_J M W_1 j \mathbf{x}, \quad S[j_1, j_2]\mathbf{x} = A_J M W_{j_2} M W_{j_1} \mathbf{x},$$

where M denotes the entry-wise modulus (absolute value). The zeroth-order coefficient is defined simply by $A_J\mathbf{x}$.

2.3. Hyperbolic geometry

Hyperbolic geometry is a non-Euclidean geometry that embeds tree-structured data with arbitrarily low distortion [21]. Hyperbolic spaces have constant negative curvature, where curvature measures how a geometric object deviates from a flat plane. Here, we use the d-dimensional Poincaré ball model with constant negative curvature $-c$ ($c > 0$), defining smooth manifold $(\mathbb{D}^{d,c}, g_x)$, where $\mathbb{D}^{d,c} = \{x \in \mathbb{R}^d : \|x\|^2 < \frac{1}{c}\}$ and g_x is the Poincaré metric tensor [2]. The exponential map $\exp_x^c : \mathcal{T}_x^c \mapsto \mathbb{D}^{d,c}$ and the logarithmic map $\log_x^c : \mathbb{D}^{d,c} \mapsto \mathcal{T}_x^c$ map between the tangent space at x ($\mathcal{T}_x^c$) and $\mathbb{D}^{d,c}$ [22]. The closed-form for these maps at the origin are:

$$\exp_0^c(v) = \tanh(\sqrt{c}\|v\|) \frac{v}{\sqrt{c}\|v\|}, \log_0^c(y) = \tanh^{-1}(\sqrt{c}\|y\|) \frac{v}{\sqrt{y}\|y\|}.$$

2.4. Related Work

Learning meaningful representation of nodes for directed graphs has relevant applications in works related to directed graph neural networks [8] and knowledge graph embeddings [23], where knowledge graphs are defined as directed, heterogeneous multigraphs. Additionally, [24] recently explored a directed-graph version of the graph scattering transform on synthetic data. Concurrently, there have been developments in hyperbolic geometry in machine learning for preserving latent hierarchies [2, 22, 3, 25, 26, 27]. While some preliminary work incorporates hyperbolic geometry into directed graph learning [4], no work has accounted for this geometry for biological networks. Most embedding approaches do not evaluate on biological networks, and benchmarks for these networks often test on undirected (or symmetrized) versions [5]. This is despite the fact that much biological analysis considers the directionality of the relationship (e.g., between two interacting molecules [6, 28]).

3. METHODS

Our goal is to learn a representation of the nodes on a directed graph that preserves both directional and hierarchical information (Fig. 1a). To accomplish this, we introduce a method which first applies a directed-graph version of the scattering transform, then applies an autoencoder restructure the scattering coefficients into a compressed representation for improved performance on downstream tasks.

First, following the lead of [24], we use the (normalized or unnormalized) magnetic Laplacian $L^{(q)}$ to construct a directed-graph wavelet frame $\mathcal{W}_J = \{W_j\}_{j=0}^J \cup \{A_J\}$. As in Section 2.1, the q parameter associated to the magnetic Laplacian will determine the way in which $L^{(q)}$ encodes directional information. As in traditional wavelet constructions, the j parameter denotes the scale of the wavelet W_j with J being the maximal scale.

More specifically, for $t \ge 0$, we define the graph heat-kernel by $H_t = \sum_{k=0}^{N-1} e^{-\lambda_k t} \mathbf{u}_k \mathbf{u}_k^*$, where $\star$ denotes conjugate transpose, and define a wavelet frame for a fixed $J \ge 0$, by $\mathcal{W}_J = \{W_j\}_{j=0}^J \cup \{A_J\}$ where $W_0 = \mathrm{Id} - H_1$, $A_J = H_{2^J}$, and $W_j = H_{2^{j-1}} - H_{2^j}$ for $1 \le j \le J$.
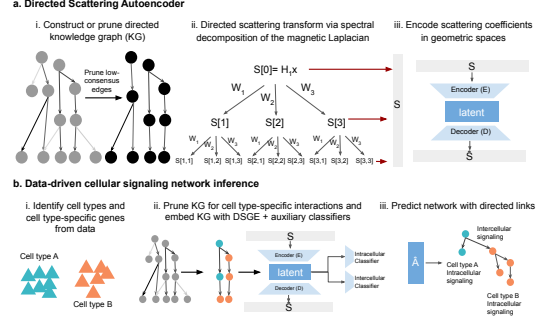

Fig. 1. DSAE and cellular signaling network inference pipeline.

We will assume that we are given C signals $\mathbf{x}_1, \ldots, \mathbf{x}_C$ and will compute the zeroth-, first-, and second-order scattering coefficients of each $\mathbf{x}_i$ using the formulas from Section 2.2. However, inspired by [7], we use H_1 instead of A_J in the final filtering (e.g., we compute second-order scattering coefficients by $H_1 M W_{j_2} M W_{j_1} \mathbf{x}$). We then let $S[v]$ denote the concatenation of all zeroth-, first-, and second-order scattering coefficients evaluated at the vertex v and let $S(G) = \{S[v] : v \in V\}$.

The node representations $S[v]$ are typically redundant and unnecessarily high-dimensional. Therefore, in order to reduce the dimension of our preliminary scattering representation, we apply an autoencoder $D \circ E$ so that

$$S(G) \approx \hat{S}(G) = D(E(S(G))).$$

In particular, while training our autoencoder, we aim to minimize the mean squared error (MSE) reconstruction loss

$$\mathcal{L}_{\mathrm{recon}} = \|S(G) - \hat{S}(G)\|_2^2 = \sum_{v \in V} \|S[v] - \hat{S}[v]\|_2^2.$$

We consider two versions of the feed-forward encoder and decoder layers to learn a d-dimensional embedding, where $d \ll$ the dimensionality of $S[v]$. First, we consider matrix operations carried out via Euclidean geometry (DSAE-Euc). Second, since each $S[v]$ has inherent tree-like structure (Fig. 1), we consider an analog to Euclidean operations for hyperbolic space to preserve hierarchical structure: Möbius addition $\oplus_c$ and Möbius matrix-vector multiplication $\otimes_c$ [22] (DSAE-Hyp) . For the encoder in the latter case, we map $S[v]$ to $\mathbb{D}^{d,c}$, the d-dimensional Poincaré ball of radius c, via the exponential map $h_0^v = \exp_0^c(S[v])$ (see [22, 3]), then employ hyperbolic feed-forward layers, i.e.

$$h_\ell^v \leftarrow \sigma((W_\ell \otimes_c h_{\ell-1}^v) \oplus_c \exp_0^c(b_\ell))$$

for hidden representation h_ℓ at layer ℓ, $1 \le \ell \le L$, where L is the number of layers. W_ℓ is the weight matrix (which is shared across v), b_ℓ is the bias, and σ is the activation. For the decoder, we use the same construction and map to the Euclidean space for reconstruction via $\log_0^c(h_L^v)$. We view the encoder output $E(S[v])$ as the hidden representation of each vertex v. In Section 4, we show this representation may be used for tasks such as link prediction.

When applying our method, we may also use an additional network to regularize our autoencoder. When doing so, we assume that there is some property of interest $p(v)$ which is important for down-

9762

stream tasks. Therefore, we train an additional network F (or possibly multiple additional networks) so that $F(E(S[v])) \approx p(v)$ and penalize this network by

$$\mathcal{L}_{\text{prop}} = \sum_{v \in V} \text{dist}(p(v), F(E(S[v]))),$$

where the distance function will vary depending on the task of interest. This regularizing network helps our autoencoder learn a compressed representation of the scattering coefficients which preserves information relevant for a given application but discards less relevant information. In Section 5, we will provide an example of how such regularizing networks may be applied to cellular signaling networks.

4. LINK PREDICTION IN DIRECTED GRAPHS

We evaluate the performance of our method on the task of link direction prediction on two WebKB graphs modeling links between websites at different universities [8], and three cellular signaling graphs [9, 10, 11]. In the latter case, the nodes are genes and the edges encode signaling relationships, including transcriptional regulation (within cells), or protein-protein binding after secretion (between cells). While constructing the biological graphs, we prune the network for interaction confidence, restrict the graph to the largest connected component, and give all remaining edges unit weight so as not to overemphasize well-studied genes.

We note that all graphs are directed and hierarchical; less than 20% of edges are bidirectional for the website graphs, and less than 5% of the edges are bidirectional for the biological graphs. Further, the Krachkardt hierarchy score (Khs) [29] is over 0.75 (between 0 and 1, where 1 is most hierarchical and 0 least hierarchical) for all graphs, and the Ollivier-Ricci curvature [30] (averaged for all edges) is negative. The k-NN graphs of the directed scattering coefficients also have negative average Ollivier-Ricci curvature (data not shown). Further details on the dataset are available in Table 1.

	Type	# Nodes	# Edges	Rec.	Khs	ORC
Texas	WebKB	183	325	18%	0.91	-0.14
Cornell	WebKB	183	298	12%	0.96	-0.30
OmniPath	Post-translational and transcriptional	8909	58919	4%	0.76	-0.51
SIGNOR_ppi	Post-translational	4743	11109	2%	0.92	-0.56
iPTMnet	Post-translational	1642	3650	0%	0.98	-0.44

Table 1. Statistics for each directed network. Rec: Reciprocity, Khs: Krackhardt hierarchy score. ORC: Ollivier-Ricci curvature.

On each of these graphs, we are given a partially complete version of the network (i.e., the original graph, but with some edges removed) and the task is to predict whether a given edge, (u, v) or its reciprocal edge (v, u) belongs to the original network. As baseline methods, we considered shallow approaches – node2vec [31] and Poincaré map (PM) [2]; graph neural network approaches – graph autoencoder (GAE) [32], hyperbolic graph convolutional network (HGCN) [3], and directed graph neural network MagNet [8]; a knowledge graph-based approach – TransE [23]; and the undirected graph scattering transform (UDS) [17]. When applicable, we use a single input signal chosen to be a standard Gaussian random vector.

Table 2 summarizes comparisons of DSAE-Hyp and DSAE-Euc to the other approaches. Both versions of DSAE perform quite well. DSAE-Hyp is the top performing method on Texas, Cornell, and SIGNOR and is the second best on OmniPath. DSAE-Euc is the top method on OmniPath and second best on Texas and iPTMnet.

Method	Texas	Cornell	OmniPath	SIGNOR	iPTMnet
DSAE-Hyp	**0.915**	**0.936**	<u>0.846</u>	**0.905**	0.983
DSAE-Euc	<u>0.913</u>	0.809	**0.870**	0.893	<u>0.988</u>
node2vec	0.860	0.714	0.608	0.703	0.804
PM	0.591	0.574	0.545	0.579	0.530
PM-D	0.617	0.678	0.581	0.564	0.602
GAE	0.807	0.584	0.716	0.780	0.873
GAE-D	0.806	0.695	0.712	0.793	0.881
HGCN	0.729	0.639	0.586	0.711	0.816
HGCN-D	0.799	0.754	0.714	0.733	0.910
MagNet	0.888	0.930	0.831	<u>0.898</u>	**0.989**
TransE	nan	nan	0.793	nan	nan
TransE-E	0.801	<u>0.933</u>	0.777	<u>0.898</u>	<u>0.988</u>
UDS-AE	0.845	0.893	0.668	0.783	0.976

Table 2. Link Direction Prediction mean AUROC. -D refers to implementation with asymmetric matrices -E refers to implementation with edge attributes. nan refers to undefined runs (no edge attributes). Best performance bolded, second best underlined.

4.1. Ablation experiment

To test the importance of the autoencoder and the robustness to q (the charge parameter) and J (the scale parameter), we performed a series of ablations (shown in Table 3) with DSAE-Hyp, which performed best on average. Ablations showed q=0.0 (which does not take into account the edge direction) performed worse than $q = 0.1$ and $q = 0.2$ at all scales (with all other optimal hyperparameters), showing that edge direction was useful for link prediction. Directed scattering without the autoencoder (no AE) performed worse than the full DSAE method on average (setting J and q to be the same as the optimal performing version of DSAE).

Method	Texas	Cornell	OmniPath	SIGNOR	iPTMnet
no AE	0.903	0.929	0.843	0.903	0.981
q=0.0 J=5	0.750	0.641	0.519	0.549	0.922
q=0.0 J=10	0.747	0.642	0.518	0.561	0.921
q=0.0 J=15	0.751	0.641	0.518	0.564	0.921
q=0.1 J=5	0.906	0.930	**0.846**	**0.905**	0.980
q=0.1 J=10	**0.915**	0.932	**0.846**	**0.905**	0.980
q=0.1 J=15	0.911	0.932	0.845	0.904	0.979
q=0.2 J=5	0.900	0.935	0.833	0.904	**0.983**
q=0.2 J=10	0.899	**0.936**	**0.846**	0.904	0.982
q=0.2 J=15	0.899	0.935	0.845	0.903	0.982

Table 3. Ablation Link Direction Prediction mean AUROC. no AE refers to scattering without autoencoder. Best performance bolded.

4.2. Training details

We trained with an Adam / RiemmanianAdam optimizer for 50 epochs with a patience of 10 epochs with latent dimension $d = 128$. The train/val/test split was 85/5/10, where the training subgraph was first built from a minimum spanning tree with edges added until split proportion met. We optimized for the following: walk length and number of walks (node2vec); learning rate, bias, dropout, number of layers, activation, and weight decay (remaining methods); attention versus no attention (GAE); q (MagNet, DSAE-Hyp, DSAE-Euc); c (HGCN, Poincaré map, DSAE-Hyp); and J (DSAE-Hyp, DSAE-Euc). Results are reported as mean AUROC of five runs[1].

---

[1]github.com/KrishnaswamyLab/Directed-Hierarchical-Gene-Networks

9763

5. INFERRING A DIRECTED GRAPH FROM DATA

Here, we show that we can use DSAE to infer signaling connections fusing data and knowledge graph embeddings (Fig. 1b). Given single-cell RNA sequencing data (i.e., high-dimensional point-cloud data where observations are cells and the features are gene expression), we want to understand what genes are signaling to each other within and between cell types based on their expression and known gene interactions. That is, we aim to infer a directed gene-gene graph defined by two within-cell type gene networks connected by between-cell type interactions. While most prior work identifies correlated gene pairs [6, 33], this is, to our knowledge, the first approach to infer a directed gene network between cell types.

5.1. Constructing the cellular signaling network

We let $G = (V, E)$ be a directed graph where each vertex v represents a gene and edges represent genes which interact within or between cells (such as OmniPath or SIGNOR). We assume each gene corresponds to a certain cell type and let V_i denote the set of all genes corresponding to cell type i. Given two cell types i and j, we let $G_{i,j} = (V_{i,j}, E_{i,j})$ be the largest connected component of the induced subgraph with vertices $V_i \cup V_j$.

5.2. Applying DSAE to Cellular Signaling Networks

In order to structure the latent space for cellular signaling inference, when applying DSAE to the graph $G_{i,j}$ constructed above, we regularize our autoencoder with two additional classifiers that predict cell labels from hidden representation $E(S[v])$. In particular, we have an intracellular classifier, which predicts the cell type, and an intercellular classifier, which predicts whether or not the node is capable of intercellular communication. Each of these classifiers are penalized by supervised contrastive losses $\mathcal{L}_{\text{intracellular}}$ and $\mathcal{L}_{\text{intercellular}}$ [34]. The intracellular classifier ensures that nodes from the same cell type remain close. The intercellular classifier, on the other hand, ensures that nodes which are capable of intercellular communication and are connected to nodes from the opposite cell type remain close. These additional classifiers allow us to jointly learn a shared representation for both unsupervised reconstruction and supervised intracellular and intercellular classification. The total loss is then
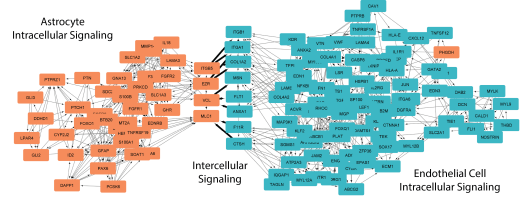
$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_{\text{recon}} + \beta \mathcal{L}_{\text{intracellular}} + \gamma \mathcal{L}_{\text{intercellular}}. \quad (1)$$

After training DSAE, we then replace the initial $E_{i,j}$ with a new edge set. First, we add an undirected edge between vertices v_i and v_j if v_i is amongst the k-th nearest (in the embedded space) neighbors of v_j or vice-versa. Then, we assign a direction via a linear (logistic regression) classifier trained on the edges from the original $E_{i,j}$.

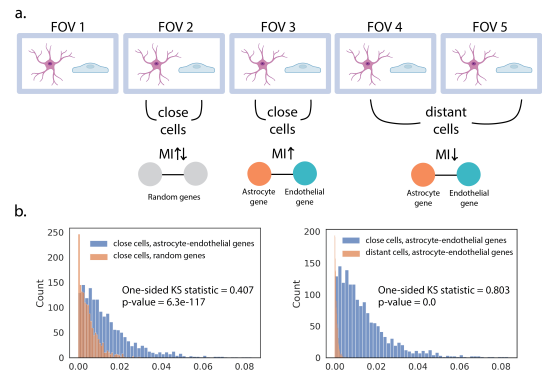### 5.3. Case study: Mouse visual cortex scRNA-seq data

We take two previously annotated cell types (astrocytes and endothelial cells) from an scRNA-seq dataset of the mouse visual cortex [35] and identify cell type-specific genes as genes with log2 fold change > 2 and p-value < 0.05 in a given cell type from a Wilcoxon rank sum test [36]. We prune the OmniPath network (version 1.0.5) to only these genes and label nodes based on cell type and intercellular connection, then apply DSAE-Euc with $(\alpha, \beta, \gamma) = (10, 5, 1)$ in (1), resulting in an embedding informed by intracellular and intercellular connections in addition the original graph G. We then build a directed network with a k-NN graph (k=5) and visualize the network with Cytoscape [37] (Fig. 2). We see that four astrocyte genes and eight endothelial cell genes are involved in intercellular signaling.

Following [33], we validate our network with matched mouse visual cortex data with spatial resolution [38]. We reasoned cells in the same spatial field-of-view (FOV) are close and are more likely to signal to each other than cells in different FOVs. True signaling pathway genes should have higher mutual information (MI) of spatial expression in close cells than distant cells, whereas random

Fig. 2. Learned cellular signaling network between cell types.

gene pairs may have high or low MI between close cells (Fig. 3a). We confirm that genes in our network have a higher MI than random genes between close astrocytes and endothelial cells (one-sided Kolmogorov-Smirnov (KS) test [39] statistic = 0.407, p=6.3e-117). Genes in our network also have a higher MI between close astrocytes and endothelial cells than between distant astrocytes and endothelial cells (one-sided KS test statistic = 0.803, p=0.0) (Fig. 3b).

Fig. 3. Evaluation of learned network with spatial data.

6. CONCLUSION

We have introduced DSAE, a novel method for node embeddings which trains a (regularized) autoencoder on top of a directed-graph variation of the scattering transform. The autoencoder learns a compressed representation and generates improved embeddings, in particular with the hyperbolic variation that better represents latent tree-like motifs common to real-world graphs such as biological networks. We show that DSAE is effective at representing five directed graphs due to the ability to preserve directed, multiscale, and hierarchical graph structure. Moreover, when properly regularized, DSAE preserves relevant information (for a given task), including learning active and novel gene-gene interactions for cellular signaling analysis. We emphasize this is the first known application of directed scattering to a real-world setting, providing a key contribution to both graph signal processing and biomedical research.

7. REFERENCES

[1] William L. Hamilton, Rex Ying, and Jure Leskovec, "Representation learning on graphs: Methods and applications," *IEEE Data Eng. Bull.*, vol. 40, no. 3, pp. 52–74, 2017.

[2] Maximillian Nickel and Douwe Kiela, "Poincaré embeddings for learning hierarchical representations," in *Advances in Neural Information Processing Systems*, 2017, vol. 30.

[3] Ines Chami, Rex Ying, Christopher Ré, and Jure Leskovec, "Hyperbolic graph convolutional neural networks," Oct. 2019.

[4] Ines Chami, Adva Wolf, Da-Cheng Juan, Frederic Sala, Sujith Ravi, and Christopher Ré, "Low-dimensional hyperbolic knowledge graph embeddings," May 2020.

[5] Aditya Pratapa, Amogh P. Jalihal, Jeffrey N. Law, Aditya Bharadwaj, and T. M. Murali, "Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data," *Nature Methods*, vol. 17, no. 2, pp. 147–154, Feb. 2020.

[6] Erick Armingol, Adam Officer, Olivier Harismendy, and Nathan E Lewis, "Deciphering cell-cell interactions and communication from gene expression," *Nat. Rev. Genet.*, vol. 22, no. 2, pp. 71–88, Feb. 2021.

[7] Frederik Wenkel, Yimeng Min, Matthew Hirn, Michael Perlmutter, and Guy Wolf, "Overcoming oversmoothness in graph convolutional networks via hybrid scattering networks," *arXiv:2201.08932*, 2022.

[8] Xitong Zhang, Yixuan He, Nathan Brugnone, Michael Perlmutter, and Matthew Hirn, "MagNet: A neural network for directed graphs," in *Advances in Neural Information Processing Systems*, 2021, pp. 27003–27015.

[9] Dénes Türei, Alberto Valdeolivas, Lejla Gul, Nicolàs Palacio-Escat, Michal Klein, Olga Ivanova, Márton Ölbei, Attila Gábor, Fabian Theis, Dezső Módos, Tamás Korcsmáros, and Julio Saez-Rodriguez, "Integrated intra- and intercellular signaling knowledge for multicellular omics analysis," *Mol. Syst. Biol.*, vol. 17, no. 3, pp. e9923, Mar. 2021.

[10] Prisca Lo Surdo, Marta Iannuccelli, Silvia Contino, Luisa Castagnoli, Luana Licata, Gianni Cesareni, and Livia Perfetto, "SIGNOR 3.0, the SIGnaling network open resource 3.0: 2022 update," *Nucleic Acids Res.*, vol. 51, no. D1, pp. D631–D637, Jan. 2023.

[11] Hongzhan Huang, Cecilia N Arighi, Karen E Ross, Jia Ren, Gang Li, Sheng-Chih Chen, Qinghua Wang, Julie Cowart, K Vijay-Shanker, and Cathy H Wu, "iPTMnet: an integrated resource for protein post-translational modification network discovery," *Nucleic Acids Res.*, vol. 46, no. D1, pp. D542–D550, Jan. 2018.

[12] Joya A Deri and Jose M F Moura, "Spectral projector-based graph Fourier transforms," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 6, pp. 785–795, Sept. 2017.

[13] Stefania Sardellitti, Sergio Barbarossa, and Paolo Di Lorenzo, "On the graph Fourier transform for directed graphs," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 6, pp. 796–811, Sept. 2017.

[14] Fan Chung, "Laplacians and the Cheeger inequality for directed graphs," *Ann. Comb.*, vol. 9, no. 1, pp. 1–19, Apr. 2005.

[15] Satoshi Furutani, Toshiki Shibahara, Mitsuaki Akiyama, Kunio Hato, and Masaki Aida, "Graph signal processing for directed graphs based on the Hermitian Laplacian," in *Machine Learning and Knowledge Discovery in Databases*. 2020.

[16] Michaël Fanuel, Carlos M. Alaíz, Ángela Fernández, and Johan A.K. Suykens, "Magnetic eigenmaps for the visualization of directed networks," *Applied and Computational Harmonic Analysis*, vol. 44, pp. 189–199, 2018.

[17] Feng Gao, Guy Wolf, and Matthew Hirn, "Geometric scattering for graph data analysis," in *International Conference on Machine Learning*. PMLR, 2019, pp. 2122–2131.

[18] Fernando Gama, Alejandro Ribeiro, and Joan Bruna, "Diffusion scattering transforms on graphs," in *7th International Conference on Learning Representations, ICLR 2019*, 2019.

[19] Dongmian Zou and Gilad Lerman, "Graph convolutional neural networks via scattering," *Applied and Computational Harmonic Analysis*, 2020.

[20] Stéphane Mallat, "Group invariant scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[21] Rik Sarkar, "Low distortion delaunay embedding of trees in hyperbolic plane," in *Graph Drawing*, pp. 355–366. Springer Berlin Heidelberg, 2012.

[22] Octavian-Eugen Ganea, Gary Bécigneul, and Thomas Hofmann, "Hyperbolic neural networks," May 2018.

[23] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in neural information processing systems*, vol. 26, 2013.

[24] Joyce Chew, Matthew Hirn, Smita Krishnaswamy, Deanna Needell, Michael Perlmutter, Holly Steach, Siddharth Viswanath, and Hau-Tieng Wu, "Geometric scattering on measure spaces," *arXiv:2208.08561*, 2022.

[25] Ines Chami, Albert Gu, Dat P Nguyen, and Christopher Ré, "Horopca: Hyperbolic dimensionality reduction via horospherical projections," in *International Conference on Machine Learning*. PMLR, 2021, pp. 1419–1429.

[26] Panagiotis Kyriakis, Iordanis Fostiropoulos, and Paul Bogdan, "Learning hyperbolic representations of topological features," *arXiv preprint arXiv:2103.09273*, 2021.

[27] Ruochen Yang, Frederic Sala, and Paul Bogdan, "Efficient representation learning for higher-order data with simplicial complexes," in *Learning on Graphs Conference*. PMLR, 2022, pp. 13–1.

[28] Mirjana Efremova, Miquel Vento-Tormo, Sarah A Teichmann, and Roser Vento-Tormo, "CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes," *Nat. Protoc.*, vol. 15, no. 4, pp. 1484–1506, Apr. 2020.

[29] David Krackhardt, "Graph theoretical dimensions of informal organizations," *Computational organization theory, pages 107–130. Psychology Press*.

[30] Yann Ollivier, "Ricci curvature of Markov chains on metric spaces," *Journal of Functional Analysis*, vol. 256, no. 3, pp. 810–864, 2009.

[31] Aditya Grover and Jure Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 2016, pp. 855–864.

[32] Thomas N. Kipf and Max Welling, "Variational graph auto-encoders," *CoRR*, vol. abs/1611.07308, 2016.

[33] Yuxuan Hu, Tao Peng, Lin Gao, and Kai Tan, "CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data," *Sci. Adv.*, vol. 7, no. 16, Apr. 2021.

[34] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan, "Supervised contrastive learning," 2020.

[35] Tasic et al., "Adult mouse cortical cell taxonomy revealed by single cell transcriptomics," *Nature Neuroscience*, vol. 19, no. 2, pp. 335–346, Jan. 2016.

[36] Winston Haynes, *Wilcoxon Rank Sum Test*, p. 2354–2355, Springer New York, 2013.

[37] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Res.*, vol. 13, no. 11, pp. 2498–2504, Nov. 2003.

[38] Linus et al., "Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+," *Nature*, vol. 568, no. 7751, pp. 235–239, Mar. 2019.

[39] F. Massey, "The Kolmogorov-Smirnov test for goodness of fit," *J. Am. Stat. Assoc.*, vol. 46, no. 253, 1951.