DTU Library

Characterizing Polarization in Social Networks using the Signed Relational Latent Distance Model

Nakis, Nikolaos; Celikkanat, Abdulkadir; Boucherie, Louis; Djurhuus, Christian; Burmester, Felix; Holmelund, Daniel Mathias; Frolcová, Monika; Mørup, Morten

[Link back to DTU Orbit](#)

Characterizing Polarization in Social Networks using the Signed Relational Latent Distance Model

Nikolaos Nakis
Technical University
of Denmark

Abdulkadir Çelikkanat
Technical University
of Denmark

Louis Boucherie
Technical University
of Denmark

Christian Djurhuus
Technical University
of Denmark

Felix Burmester
Technical University
of Denmark

Daniel Mathias Holmelund
Technical University
of Denmark

Monika Frolcová
Technical University
of Denmark

Morten Mørup
Technical University
of Denmark

Abstract

Graph representation learning has become a prominent tool for the characterization and understanding of the structure of networks in general and social networks in particular. Typically, these representation learning approaches embed the networks into a low-dimensional space in which the role of each individual can be characterized in terms of their latent position. A major current concern in social networks is the emergence of polarization and filter bubbles promoting a mindset of "us-versus-them" that may be defined by extreme positions believed to ultimately lead to political violence and the erosion of democracy. Such polarized networks are typically characterized in terms of signed links reflecting likes and dislikes. We propose the Signed Latent Distance Model (SLDM) utilizing for the first time the Skellam distribution as a likelihood function for signed networks. We further extend the modeling to the characterization of distinct extreme positions by constraining the embedding space to polytopes, forming the Signed Latent relational dIstance Model (SLIM). On four real social signed networks of polarization, we demonstrate that the models extract low-dimensional characterizations that well predict friendships and animosity while SLIM provides interpretable visualizations defined by extreme positions when restricting the embedding space to polytopes.

1 INTRODUCTION

For several decades, the origin and influence of political polarization have been issues receiving considerable attention both within scholarly research and the public media (Hetherington, 2009). Several studies have demonstrated an increasing partisan polarization among the political elites, some of which rely on network science approaches, for instance, using co-voting similarity networks and modularity to model and explain the distinct aspects of the data (Moody and Mucha, 2013). Whereas polarization has been described in terms of communities and their boundary properties (Guerra et al., 2013), latent distance modeling has also been used to extract bipolar structures (Barberá et al., 2015).

Ideological polarization is the distance between policy preferences, typically of elites taking extreme stands on issues whereas the electoral behavior is denoted affective polarization. When these extremes are portrayed as existential in the media, they typically form an "us-versus-them"-mindset (Dagnes, 2019). From a social network perspective, the process of polarization has been described to occur when "homophily and influence become self-reinforcing when the attraction to those who are similar and differentiation from those who are dissimilar entail greater openness to influence. The result is network autocorrelation—the tendency for people to resemble their network neighbors" (DellaPosta et al., 2015).

To better capture ideological polarization, we turn to signed networks. Signed networks reflect complex social polarization better than unsigned networks because they capture positive, negative, and neutral relationships between entities. The study of signed networks goes back to the '50s and was motivated by friendly and hostile social relationships (Harary, 1953). Since then they have been used to study networks of Twitter users (Keuchenius et al., 2021)

and US Congress members (Thomas et al., 2006), two examples of polarized social networks (Garimella and Weber, 2017; Neal, 2020).

In this paper, we focus on polarization as extreme positions and argue that the multi-polarity of "us-versus-them" reinforced by homophily and influence can be characterized by a latent position model (i.e., the latent distance model (Hoff et al., 2002)) of networks confined to a constrained social space formed by a polytope, what we denote a sociotope. As such, the corners of the sociotope define distinct aspects (i.e., poles) formed by polarized networks' tendencies to self-reinforce homophily by positive ties driving those who are similar close as opposed to those that are negatively tied being repelled. This can be revealed in terms of the important multiple poles of social network defining corners of such sociotope. Within these corners, positive interactions between nodes place them in close proximity in space thereby accounting for homophily while negative interactions "push" nodes far apart (towards opposing poles) yielding the "us-versus-them" effect.

The conceptual idea of polytopes as formed by pure types can be traced back to Plato's forms, which characterize the physical world as a limited projection of the forms also referred to as ideal categories. Later, Carl Jung introduced the concept of universal archetypes, described as a collective unconscious, in which he related to Plato's forms by describing the forms as a Jungian version of the Platonian archetypes (Williamson, 1985). Employing the theoretical concept of archetypes to political and ideological polarization, the archetypes could be interpreted as genuine ideologies, while the ideological advocates can be expressed as a mixture of distinct ideologies.

Archetypal Analysis (AA) is a prominent framework for extracting polytopes in tabular data. AA was originally proposed by Cutler and Breiman (1994) as an unsupervised learning method that favors distinct aspects, archetypes, of the data in which observations are characterized by convex combinations (i.e., mixtures) of these archetypes as opposed to clustering procedures extracting prototypical observations (Mørup and Kai Hansen, 2010). AA has previously been used to model societal conflicts in Europe (Beugelsdijk et al., 2022). However, given that AA was proposed for tabular data, the applications are currently restricted to non-relational data. Thus, whereas the characterization of data in terms of distinct aspects and polytopes has a long history, such representation learning approaches have not previously been considered in the context of network analysis for the extraction of polarization by several extremes.

In the last years, representation learning of signed graphs has gathered substantial attention, with applications such as signed link prediction (Chiang et al., 2011), and community detection (Tzeng et al., 2020). Initial works extended the prominent random walks framework (Perozzi et al., 2014; Grover and Leskovec, 2016) to the analysis of signed networks. SIDE (Kim et al., 2018b) exploits truncated random walks on the signed graph with interaction signs for each node pair inferred based on balance theory (Cartwright and Harary, 1956). Balance theory is a socio-psychological theory admitting four rules: "The friend of my friend is my friend," "The friend of my enemy is my enemy," "The enemy of my friend is my enemy," and "The enemy of my enemy is my friend." POLE (Huang et al., 2022), also utilizes balance theory-based signed random walks to construct an auto-covariance similarity which is used to obtain the embedding space. Neural networks have also been adopted for the analysis of signed networks. Both SiNE (Wang et al., 2017) and SIGNET (Islam et al., 2018) combine balance theory and multi-layer neural networks to learn the network embeddings while SIGNet uses targeted node sampling to provide scalable inference. In addition, graph neural networks have also been studied in the context of signed graphs. More specifically, SiGAT (Huang et al., 2019) and SDGNN (Huang et al., 2021) combine balance and status theory with graph attention to learn signed network embeddings. The status theory is another important socio-psychological theory for directed relationships where for a source and a target node, a positive directed connection assumes a higher status of the target, i.e. {status(target) > status(source)}, while the inequality is opposite for a negative connection. Lastly, SLF (Xu et al., 2019) learns multiple latent factors of the signed network, modeling positive, negative, and neutral, as well as the absence of a relationship between node pairs.

A prominent approach for graph representation learning is the Latent Distance Model (Hoff et al., 2002) in which the tendency of nodes to connect is defined in terms of their proximity in latent space. Notably, the LDM can express the properties transitivity (*"a friend of a friend is a friend"*) and homophily (*"akin nodes tend to have links"*). Recently, it has been shown that LDMs can account for the structure of networks in ultra-low dimensions (Nakis et al., 2022, 2023; Çelikkanat et al., 2022). It has further been demonstrated that an LDM of one dimension can be used to extract bipolar network properties (Barberá et al., 2015).

For the modeling of signed networks for the characterization of polarization, we first present the Signed Latent Distance Model (SLDM). The model utilizes a likelihood function for weighted signed links based on the Skellam distribution (Skellam, 1946). The Skellam distribution is the discrete probability distribution of the difference between two independent Poisson random variables. It was introduced by John Gordon Skellam to model the dynamics of populations (Skellam, 1946). Since then it was used in medicine to model treatment measurements (Karlis and Ntzoufras, 2006), sports results (Karlis and Ntzoufras, 2008), as well as, econometric studies (Barndorff-Nielsen et al.,

2010). Furthermore, we introduce the Signed relational Latent dIstance Model (SLIM) being able to characterize the latent social space in terms of extreme positions forming polytopes inspired by archetypal analysis enabling archetypal analysis for relational data, i.e. relational AA (RAA). We apply SLDM and SLIM on four real signed networks believed to reflect polarization and demonstrate how SLIM uncovers prominent distinct positions (poles). To the best of our knowledge, this is the first work to model signed weighted networks using the Skellam distribution and the first time AA has been extended to relational data by leveraging latent position modeling approaches for the characterization of polytopes in social networks. **The implementation is available at:** *github.com/Nicknakis/SLIM_RAA*.

2 PROPOSED METHODOLOGY

Let $\mathcal{G} = (\mathcal{V}, \mathcal{Y})$ be a *signed graph* where $\mathcal{V} = \{1, \ldots, N\}$ denotes the set of vertices and $\mathcal{Y} : \mathcal{V}^2 \to \mathsf{X} \subseteq \mathbb{R}$ is a map indicating the weight of node pairs, such that there is an edge $(i, j) \in \mathcal{V}^2$ if the weight $\mathcal{Y}(i, j)$ is different from 0. In other words, $\mathcal{E} := \{(i, j) \in \mathcal{V}^2 : \mathcal{Y}(i, j) \neq 0\}$ indicates the set of edges of the network. Since many real networks consist of only integer-valued edges, in this paper, we set X to $\mathbb{Z}$, and we will call the graph *undirected* if the pairs (i, j) and (j, i) represent the same link. (The directed case is provided in the supplementary materials.) For simplicity, y_{ij} denotes each edge weight.

2.1 The Skellam Latent Distance Model (SLDM)

Our main purpose is to learn latent node representations $\{\mathbf{z}_i\}_{i \in \mathcal{V}} \in \mathbb{R}^K$ in a low dimensional space for a given signed network $\mathcal{G} = (\mathcal{V}, \mathcal{Y})$ ($K \ll |\mathcal{V}|$). Therefore, the edge weights can take any integer value to represent the positive or negative tendencies between the corresponding nodes. We model these signed interactions among the nodes using the Skellam distribution (Skellam, 1946), which can be formulated as the difference of two independent Poisson-distributed random variables ($y = N_1 - N_2 \in \mathbb{Z}$) with respect to the rates λ^+ and λ^-:

$$P(y|\lambda^+, \lambda^-) = e^{-(\lambda^+ + \lambda^-)} \left(\frac{\lambda^+}{\lambda^-}\right)^{y/2} \mathcal{I}_{|y|}\left(2\sqrt{\lambda^+ \lambda^-}\right),$$

where $N_1 \sim Pois(\lambda^+)$ and $N_2 \sim Pois(\lambda^-)$, and $\mathcal{I}_{|y|}$ is the modified Bessel function of the first kind and order $|y|$. To the best of our knowledge, the Skellam distribution has not been adapted before for modeling the network likelihood. More specifically, we propose a novel latent space model utilizing the Skellam distribution by adopting the latent distance model, which was proposed originally for undirected, and unsigned binary networks as a logistic regression model (Hoff et al., 2002). It was later extended to multiple generalized linear models (Hoff, 2005), including the Poisson regression model for integer-weighted net-

works. We can formulate the negative log-likelihood of a latent distance model under the Skellam distribution as:

$$\mathcal{L}(\mathcal{Y}) := \log p(y_{ij}|\lambda_{ij}^+, \lambda_{ij}^-)$$
$$= \sum_{i<j} (\lambda_{ij}^+ + \lambda_{ij}^-) - \frac{y_{ij}}{2} \log\left(\frac{\lambda_{ij}^+}{\lambda_{ij}^-}\right) - \log(I_{ij}^*),$$

where $I_{ij}^* := \mathcal{I}_{|y_{ij}|}\left(2\sqrt{\lambda_{ij}^+ \lambda_{ij}^-}\right)$. As it can be noticed, the Skellam distribution has two rate parameters, and we consider them to learn latent node representations $\{\mathbf{z}_i\}_{i \in \mathcal{V}}$ by defining them as follows:

$$\lambda_{ij}^+ = \exp\left(\gamma_i + \gamma_j - ||\mathbf{z}_i - \mathbf{z}_j||_2\right), \quad (1)$$
$$\lambda_{ij}^- = \exp\left(\delta_i + \delta_j + ||\mathbf{z}_i - \mathbf{z}_j||_2\right), \quad (2)$$

where the set $\{\gamma_i, \delta_i\}_{i \in \mathcal{V}}$ denote the node-specific random effect terms, and $||\cdot||_2$ is the Euclidean distance function. More specifically, γ_i, γ_j represent the "social" effects/reach of a node and the tendency to form (as a receiver and as a sender, respectively) positive interactions, expressing positive degree heterogeneity (indicated by $+$ as a superscript of λ). In contrast, δ_i, δ_j provide the "anti-social" effect/reach of a node to form negative connections, and thus models negative degree heterogeneity (indicated by $-$ as a superscript of λ).

By imposing standard normally distributed priors elementwise on all model parameters $\boldsymbol{\theta} = \{\boldsymbol{\gamma}, \boldsymbol{\delta}, \mathbf{Z}\}$, i.e., $\theta_i \sim \mathcal{N}(0, 1)$, We define a maximum a posteriori (MAP) estimation over the model parameters, via the loss function to be minimized (ignoring constant terms):

$$Loss = \sum_{i<j}\left(\lambda_{ij}^+ + \lambda_{ij}^- - \frac{y_{ij}}{2}\log\left(\frac{\lambda_{ij}^+}{\lambda_{ij}^-}\right)\right)$$
$$- \sum_{i<j} \log I_{|y_{ij}|}\left(2\sqrt{\lambda_{ij}^+ \lambda_{ij}^-}\right) \quad (3)$$
$$+ \frac{\rho}{2}\left(||\mathbf{Z}||_F^2 + ||\boldsymbol{\gamma}||_F^2 + ||\boldsymbol{\delta}||_F^2\right),$$

where $||\cdot||_F$ denotes the Frobenius norm. In addition, ρ is the regularization strength with $\rho = 1$ yielding the adopted normal prior with zero mean and unit variance. Importantly, by setting λ_{ij}^+ and λ_{ij}^- based on Eq. (11) and (2), the model effectively makes positive (weighted) links attract and negative (weighted links) deter nodes from being in proximity of each other.

2.2 Archetypal Analysis

Archetypal Analysis (AA) (Cutler and Breiman, 1994; Mørup and Kai Hansen, 2010) is an approach developed for the modeling of observational data in which the data is expressed in terms of convex combinations of characteristics (i.e. archetypes). The definition of the embedded data

points is given as follows:

$$\mathbf{X} \approx \mathbf{XCZ} \quad \text{s.t. } \boldsymbol{c}_d \in \Delta^N \text{ and } \mathbf{z}_j \in \Delta^K \tag{4}$$

where Δ^P denotes the standard simplex in $(P+1)$ dimensions such that $\mathbf{q} \in \Delta^P$ requires $q_i \geq 0$ and $\|\mathbf{q}\|_1 = 1$ (i.e. $\sum_i q_i = 1$). Notably, the archetypes given by the columns of $\mathbf{A} = \mathbf{XC}$ define the corners of the extracted polytope as convex combinations of the observations, whereas $\mathbf{Z}$ define how each observation is reconstructed as convex combinations of the extracted archetypes.

Whereas archetypal analysis constrains the representation to the convex hull of the data, other approaches to model pure/ideal forms have been Minimal Volume (MV) approaches defined by

$$\mathbf{X} \approx \mathbf{AZ} \quad \text{s.t. } vol(\mathbf{A}) = v \text{ and } \mathbf{z}_j \in \Delta^K, \tag{5}$$

in which $vol(\mathbf{A})$ defines the volume of $\mathbf{A}$. When $\boldsymbol{A}$ is a square matrix this can be defined by $vol(\mathbf{A}) = |det(\mathbf{A})|$, see also Hart et al. (2015); Zhuang et al. (2019) for a review on such end-member extraction procedures. A strength is that, as opposed to AA, the approach does not require the presence of pure observations, however, a drawback is a need for regularization tuning to define an adequate volume (Zhuang et al., 2019) whereas the exact computation of the volume of general polytopes requires the computation of determinants of the sum of all simplices defining the polytope (Büeler et al., 2000). Importantly, Archetypal Analysis and Minimal volume extraction procedures have been found to identify latent polytopes defining trade-offs in which vertices of the polytopes represent maximally enriched distinct aspects (archetypes), allowing identification of tasks or prominent roles the vertices of the polytope represent (Shoval et al., 2012; Hart et al., 2015). Due to the computational issues of regularizing high-dimensional volumes and the need for careful tuning of such regularization parameters, we presently focus on polytope extraction as defined through the AA formulation rather than the MV formulation.

2.3 A Generative Model of Polarization

Considering a latent space for the modeling of polarization, we presently extend the Skellam LDM and define polarization as extreme positions (pure forms/archetypes) that optimally represent the social dynamics observed in terms of the induced polytope - what we denote a sociotope, in which each observation is a convex combination of these extremes. In particular, we characterize polarization in terms of extreme positions in a latent space defined as a polytope akin to AA and MV.

In our generative model of polarization, we further suppose that the bias terms introduced in the definitions of the Poisson rates, $(\lambda_{ij}^+, \lambda_{ij}^-)$, are normally distributed. Since latent representations $\{\mathbf{z}_i\}_{i \in \mathcal{V}}$ according to AA and MV lie in the

standard simplex set Δ^K, we further assume that they follow a Dirichlet distribution. Formally, we can summarize the generative model as follows:

$$\gamma_i \sim \mathcal{N}(\mu_\gamma, \sigma_\gamma^2) \qquad\qquad \forall i \in \mathcal{V},$$
$$\delta_i \sim \mathcal{N}(\mu_\delta, \sigma_\delta^2) \qquad\qquad \forall i \in \mathcal{V},$$
$$\mathbf{a}_k \sim \mathcal{N}(\boldsymbol{\mu}_A, \sigma_A^2 \mathbf{I}) \qquad\qquad \forall k \in \{1, \ldots, K\},$$
$$\mathbf{z}_i \sim Dir(\boldsymbol{\alpha}) \qquad\qquad \forall i \in \mathcal{V},$$
$$\lambda_{ij}^+ = \exp\left(\gamma_i + \gamma_j - \|\mathbf{A}(\mathbf{z}_i - \mathbf{z}_j)\|_2\right),$$
$$\lambda_{ij}^- = \exp\left(\delta_i + \delta_j + \|\mathbf{A}(\mathbf{z}_i - \mathbf{z}_j)\|_2\right),$$
$$y_{ij} \sim Skellam(\lambda_{ij}^+, \lambda_{ij}^-) \qquad \forall(i,j) \in \mathcal{V}^2.$$

According to the above generative process, positive ($\boldsymbol{\gamma}$) and negative ($\boldsymbol{\delta}$) random effects for the nodes are first drawn, upon which the location of extreme positions $\mathbf{A}$ (i.e., corners of the polytope denoted archetypes) are generated. In addition, as the dimensionality of the latent space increases linearly with the number of archetypes, i.e. $\boldsymbol{A}$ is a square matrix, with probability zero archetypes will be placed in the interior of the convex hull of the other archetypes. Subsequently, the node-specific convex combinations $\mathbf{Z}$ of the generated archetypes are drawn, and finally, the weighted signed link is generated according to the node-specific biases and distances between dyads within the polytope utilizing the Skellam distribution.

2.4 The Signed Relational Latent Distance Model

For inference, we exploit how polytopes can be efficiently extracted using archetypal analysis. We, therefore, define the Signed Latent relational dIstance Model (SLIM) by defining a relational archetypal analysis approach endowing the generative model a parameterization akin to archetypal analysis in order to efficiently extract polytopes from relational data defined by signed weighted networks. Specifically, we formulate the relational AA in the context of the family of LDMs, as:

$$\lambda_{ij}^+ = \exp\left(\gamma_i + \gamma_j - \|\mathbf{A}(\mathbf{z}_i - \mathbf{z}_j)\|_2\right) \tag{6}$$
$$= \exp\left(\gamma_i + \gamma_j - \|\mathbf{RZC}(\mathbf{z}_i - \mathbf{z}_j)\|_2\right). \tag{7}$$
$$\lambda_{ij}^- = \exp\left(\delta_i + \delta_j + \|\mathbf{A}(\mathbf{z}_i - \mathbf{z}_j)\|_2\right) \tag{8}$$
$$= \exp\left(\delta_i + \delta_j + \|\mathbf{RZC}(\mathbf{z}_i - \mathbf{z}_j)\|_2\right). \tag{9}$$

Notably, in the AA formulation $\mathbf{X} = \mathbf{RZ}$ corresponds to observations formed by convex combinations $\mathbf{Z}$ of positions given by the columns of $\boldsymbol{R}^{K \times K}$. Furthermore, in order to ensure what is used to define archetypes $\mathbf{A} = \mathbf{XC} = \mathbf{RZC}$ corresponds to observations using these archetypes in their reconstruction $\mathbf{Z}$, we define $\boldsymbol{C} \in \boldsymbol{R}^{N \times K}$ as a gated version of $\mathbf{Z}$ normalized to the simplex such that $\boldsymbol{c}_d \in \Delta^N$ by defining

$$c_{nd} = \frac{(\mathbf{Z}^\top \circ [\sigma(\mathbf{G})]^\top)_{nd}}{\sum_{n'} (\mathbf{Z}^\top \circ [\sigma(\mathbf{G})]^\top)_{n'd}} \tag{10}$$

in which ○ denotes the elementwise (Hadamard) product and $\sigma(\mathbf{G})$ defines the logistic sigmoid elementwise applied to the matrix $\boldsymbol{G}$. As a result, the extracted archetypes are ensured to correspond to the nodes assigned the archetype, whereas the location of the archetypes can be flexibly placed in space as defined by $\mathbf{R}$. By defining $\mathbf{z}_i = \text{softmax}(\tilde{\mathbf{z}}_i)$ we further ensure $\mathbf{z}_i \in \Delta^K$.

Importantly, the loss function of Eq. (13) is adopted for the relational AA formulation forming the SLIM, with the prior regularization applied to the corners of the extracted polytope $\boldsymbol{A} = \mathbf{R}\mathbf{Z}\mathbf{C}$ instead of the latent embeddings $\boldsymbol{Z}$ imposing a standard elementwise normal distribution as prior $a_{k,k'} \sim \mathcal{N}(0,1)$. Furthermore, we impose a uniform Dirichlet prior on the columns of $\boldsymbol{Z}$, i.e. $(\boldsymbol{z}_i \sim Dir(\mathbf{1}_K)$, this only contributes constant terms to the joint distribution, and therefore the maximum a posteriori (MAP) optimization only constant terms. As a result, the loss function optimized is given by Eq. (13) replacing $\|\mathbf{Z}\|_F^2$ with $\|\boldsymbol{A}\|_F^2$.

Complexity analysis. With SLDM/SLIM being distance models, they scale prohibitively as $\mathcal{O}(N^2)$ since the node pairwise distance matrix needs to be computed. This does not allow the analysis of large-scale networks. For that, we adopt an unbiased estimation of the log-likelihood through random sampling. More specifically, gradient steps are based on the log-likelihood of the block formed by a sampled (per iteration and with replacement) set S of network nodes. This makes inference scalable defining an $\mathcal{O}(S^2)$ space and time complexity. More options for scalable inference of distance models have also been proposed in Nakis et al. (2022); Raftery et al. (2012).

3 RESULTS AND DISCUSSION

We extensively evaluate the performance of our proposed methods by comparing them to the prominent GRL approaches designed for signed networks. All experiments regarding SLDM/SLIM have been conducted on an 8 GB NVIDIA RTX 2070 Super GPU. In addition, we adopted the Adam optimizer Kingma and Ba (2017) with learning rate lr = 0.05 and for 5000 iterations. The sample size for the node set was chosen as approximately 3000 nodes for all networks. The initialization of the SLDM/SLIM frameworks is deterministic and based on the spectral decomposition of the normalized Laplacian (more details are provided in the supplementary).

Artificial networks. We first, introduce experiments on artificial networks, as generated by the generative process described in Section 2.3. We create two networks expressing different levels of polarization. Results are presented in Fig. 1. More specifically, sub-Figs 1a and 1e show the ground truth latent spaces generating the networks with adjacency matrices as shown by sub-Figs 1b and 1f, respectively. The inferred latent spaces of the two networks are provided in sub-Figs 1c and 1g where it is clear that the

Table 1: Network statistics; $|\mathcal{V}|$: # Nodes, $|\mathcal{Y}^+|$: # Positive links, $|\mathcal{Y}^-|$: # Negative links.

	$|\mathcal{V}|$	$|\mathcal{Y}^+|$	$|\mathcal{Y}^-|$	Density
Reddit	35,776	128,182	9,639	0.0001
Twitter	10,885	238,612	12,794	0.0021
wiki-Elec	7,117	81,277	21,909	0.0020
wiki-RfA	11,332	117,982	66,839	0.0014

model successfully distinguishes the difference in the level of polarization of the two networks. We also verify the generated networks based on the inferred parameters given by sub-Figs 1d and 1h. We observe that the model successfully generates sparse networks accounting for the positive and negative link imbalance.

Real networks. We employed four networks of varying sizes and structures. (**i**) *Reddit* is constructed based on hyperlinks representing the directed connections between two communities in a social platform (Kumar et al., 2018). (**ii**) *wikiRfA* and (**iii**) *wikiElec* are the election networks covering the different time intervals in which nodes indicate the users and the directed links show supporting, neutral, and opposing votes to be selected as an administrator on the Wikipedia platform (West et al., 2014; Leskovec et al., 2010). Finally, (**iv**) *Twitter* is an undirected social network built on the corpus of tweets concerning the highly polarized debate about the reform of the Italian Constitution (Ordozgoiti et al., 2020).

In our experiments, we consider the greatest connected component of the networks, and if the original network is temporal, we construct the static network by summing the weights of the links through time. For the experiments performed on undirected graphs, we similarly combine directed links to obtain the undirected version of the networks.

Baselines. We benchmark the performance of our proposed frameworks against five prominent graph representation learning methods, designed for the analysis of signed networks: (**i**) POLE (Huang et al., 2022) which learns the network embeddings by decomposing the signed random walks auto-covariance similarity matrix, (**ii**) SLF (Xu et al., 2019) learns embeddings that are the concatenation of two latent factors targeting positive and negative relations, (**iii**) SIGAT (Huang et al., 2019) is a graph neural network approach using graph attention to learn node embeddings, (**iv**) SIDE (Kim et al., 2018b) is another random walk based method for signed networks, (**v**) SIGNET (Islam et al., 2018) is a multi-layer neural network approach constructing a Hadamard product similarity to accommodate for signed proximity on the network pairwise relations.
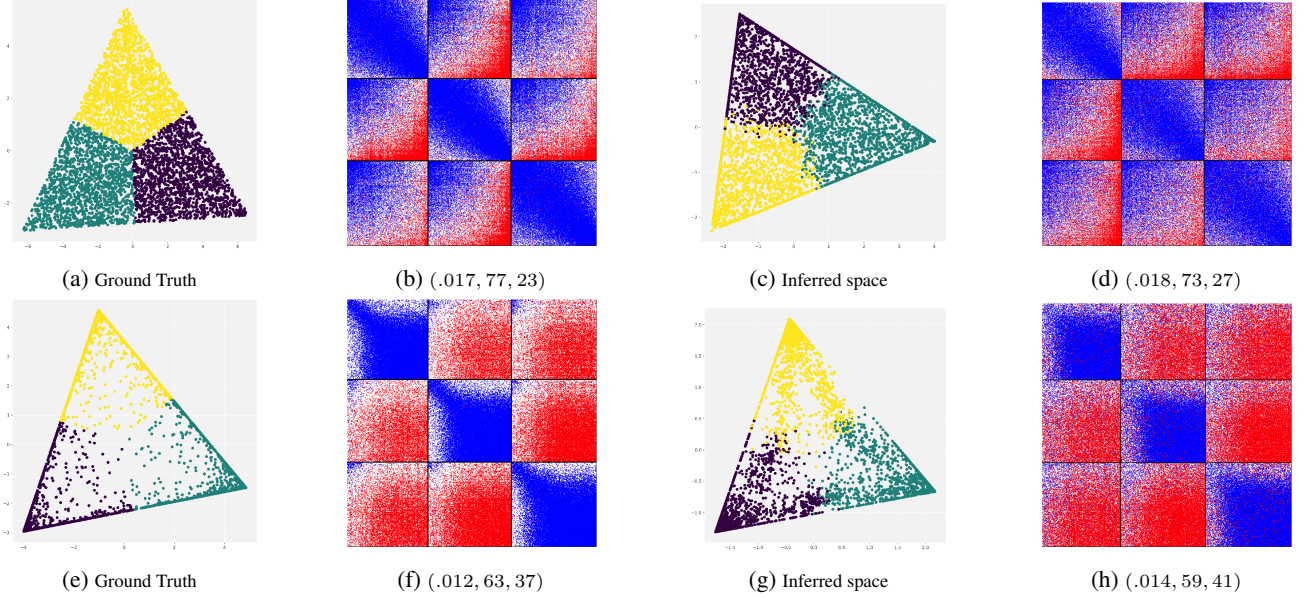
Figure 1: Two artificially generated networks with different levels of polarization $\{\mathbf{z}_i \sim Dir(\mathbf{1})$ (top row), and $\mathbf{z}_i \sim Dir(0.1 \cdot \mathbf{1})$ (bottom row)$\}$. Both size $N = 5000$ nodes and $K = 3$ archetypes. The first column shows the first two principal components of the original latent space $\tilde{Z} = AZ$, the second column the original adjacency matrix, while the parenthesis shows the network statistics as: (density,% of positive (blue) links,% of negative (red) links). The third column displays the first two principal components of the inferred latent space, and the fourth column is the SLIM generated network based on inferred parameters. All network adjacency matrices are ordered based on $\mathbf{z}_i$, in terms of maximum archetype membership and internally according to the magnitude of the corresponding archetype most used for their reconstruction.

Table 2: Area Under Curve (AUC-ROC) scores for representation size of $K = 8$.

	WikiElec			*WikiRfa*			*Twitter*			*Reddit*		
Task	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$
POLE	.809	.896	.853	.904	.921	.767	.965	.902	.922	x	x	x
SLF	**.888**	.954	**.952**	**.971**	.963	<u>.961</u>	.914	.877	.968	**.729**	**.955**	.968
SIGAT	.874	.775	.754	.944	.766	.792	**.998**	.875	.963	<u>.707</u>	.682	.712
SIDE	.728	.866	.895	.869	.861	.908	.799	.843	.910	.653	.830	.892
SIGNET	.841	.774	.635	.920	.736	.717	.968	.719	.891	.646	.547	.623
SLIM (OURS)	.862	<u>.965</u>	.935	.956	<u>.980</u>	.960	<u>.988</u>	**.963**	.972	.667	**.955**	**.978**
SLDM (OURS)	<u>.876</u>	**.969**	<u>.936</u>	<u>.963</u>	**.982**	**.963**	.986	<u>.962</u>	**.973**	.648	<u>.951</u>	<u>.975</u>

Table 3: Area Under Curve (AUC-PR) scores for representation size of $K = 8$.

	WikiElec			*WikiRfa*			*Twitter*			*Reddit*		
Task	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$
POLE	.929	.922	.544	.927	.937	.779	<u>.998</u>	.932	.668	x	x	x
SLF	**.964**	.926	**.787**	**.983**	.922	.881	.994	.870	.740	**.966**	<u>.956</u>	**.850**
SIGAT	<u>.960</u>	.724	.439	.969	.646	.497	**.999**	.861	.582	<u>.965</u>	.692	.232
SIDE	.907	.779	.608	.920	.806	.739	.974	.831	.469	.957	.820	.614
SIGNET	.944	.670	.298	.950	.572	.417	<u>.998</u>	.647	.248	.956	.510	.083
SLIM (OURS)	.953	<u>.956</u>	<u>.785</u>	.973	<u>.969</u>	**.907**	**.999**	<u>.962</u>	**.813**	.958	**.960**	**.850**
SLDM (OURS)	<u>.960</u>	**.963**	**.787**	<u>.977</u>	**.971**	<u>.912</u>	**.999**	**.963**	<u>.809</u>	.954	.955	<u>.846</u>

3.1 Link prediction

We evaluate performance considering the link prediction task considering the ability of our model to predict links of disconnected network pairs which should be connected, as well as, infer the sign of these links (positive or negative). For this, we remove/hide 20% of the total network links while preserving connectivity on the residual network. For the testing set, the removed edges are paired with a sample of the same number of node pairs that are not the edges of the original network to create zero instances. To learn the node embeddings, we make use of the residual network.

Predictions and evaluation metrics. For our methods we fit a logistic regression classifier on the concatenation of the corresponding Skellam rates and log-rates, as $\chi_{ij} = [\lambda_{ij}^{+}, \lambda_{ij}^{-}, \log \lambda_{ij}^{+}, \log \lambda_{ij}^{-}]$. Since our Skellam likelihood formulation relies both on the ratio and products of the rates, a concatenation can take advantage of a linear function of the rates, as well as, their ratio or product as allowed from the log transformation. For the baselines, we use five binary operators {average, weighted L1, weighted L2, concatenate, Hadamard product} to construct feature vectors. For each of these feature vectors, we fit a logistic regression model (except for the Hadamard product which is used directly for predictions). Since different operators provide different performances, for the baselines we choose the operator that returns the maximum performance per individual task. As a consequence of the class imbalances and the sparsity present in signed networks, we adopt robust evaluation metrics, such as area-under-curve of the receiver operating characteristic (AUC-ROC) and precision-recall (AUC-PR) curves. Lastly, we denote with "x" the performance of a baseline if it was unable to run due to high memory/runtime complexity.

Link sign prediction. In this setting, we utilize the link test set containing the negative/positive cases of removed connections. We then ask the models to predict the sign of the removed links. We denote the task of the link sign prediction task as $p@n$. In Table 2 we provide the AUC-ROC scores while in Table 3 the AUC-PR scores for the undirected case. Here we observe that our proposed models outperform the baselines in most networks while being competitive in the *Reddit* network against SLF. This specific baseline is the most competitive across networks showing high and consistent performance similar to SLIM and SLDM. Comparing now SLIM with SLDM we get mostly on-par results, verifying that constraining the model to a polytope still provides enough expressive capability as the unconstrained model while allowing for accurate extraction of "extreme" positions.

Signed link prediction. The second and more challenging task is to predict removed links against disconnected pairs of the network, as well as, infer the sign of each link correctly. For that, the test set is split into two sub-
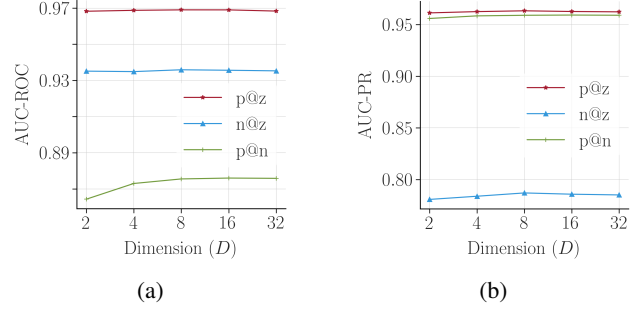

(a) (b)

Figure 2: *wikiElec*: Performance of SLIM across dimensions for different tasks, (a) Area-Under-Curve Receiver Operating Characteristic scores, (b) Area-Under-Curve Precision-Recall scores. Both AUC-ROC and AUC-PR scores are almost constant across different dimensions

sets positive/disconnected and negative/disconnected. We then evaluate the performance of each model on those subsets. The tasks of signed link prediction between positive and zero samples are denoted as $p@z$ while the negative against zero is $n@z$. We summarize our results by presenting AUC-ROC and AUC-PR scores in Table 2 and Table 3 respectively. Once more our models outperform the baselines in most networks and for both versions of signed link prediction. The SLF baseline is again the most competitive baseline yielding on-par results for *Reddit*.

Directed networks. Directed network results are provided in the supplementary. Since SLF has higher modeling capacity it outperforms the simple model formulation of SLDM and SLIM. For that, we explore and discuss formulations allowing for more capacity in the SLDM/SLIM model for the directed case (see supplementary).

Effect of dimensionality. In Figure 2, we provide the performance across dimensions for the different downstream task and for the *wikiElec* dataset. We observe that both AUC-ROC and AUC-PR scores are almost constant across different dimensions (note that as $R^{K \times K}$ dimensions for the SLIM is given by the number of archetypes), showcasing that increasing the models' capacity (in terms of dimensions) does not have a significant effect on the performance of these downstream tasks (similar results were observed for all networks and most of the baselines).

Visualizations. The RAA formulation facilitates the inference of a polytope describing the distinct aspects of networks. Here, we visualize the latent space across $K = 8$ dimensions for all of the corresponding networks. To facilitate visualizations we use Principal Component Analysis (PCA), and project the space based on the first two principal components of the final embedding matrix $\tilde{Z} = AZ$. In addition, we provide circular plots where each archetype of the polytope is mapped to a circle every $\text{rad}_k = \frac{2\pi}{K}$ radians, with K being the number of archetypes. Figure 3 contains three columns with the first denoting the PCA-

(a) *WikiElec*

(b) *WikiElec*

(c) *WikiElec*

(d) *WikiRfa*

(e) *WikiRfa*

(f) *WikiRfa*

(g) *Reddit*

(h) *Reddit*

(i) *Reddit*
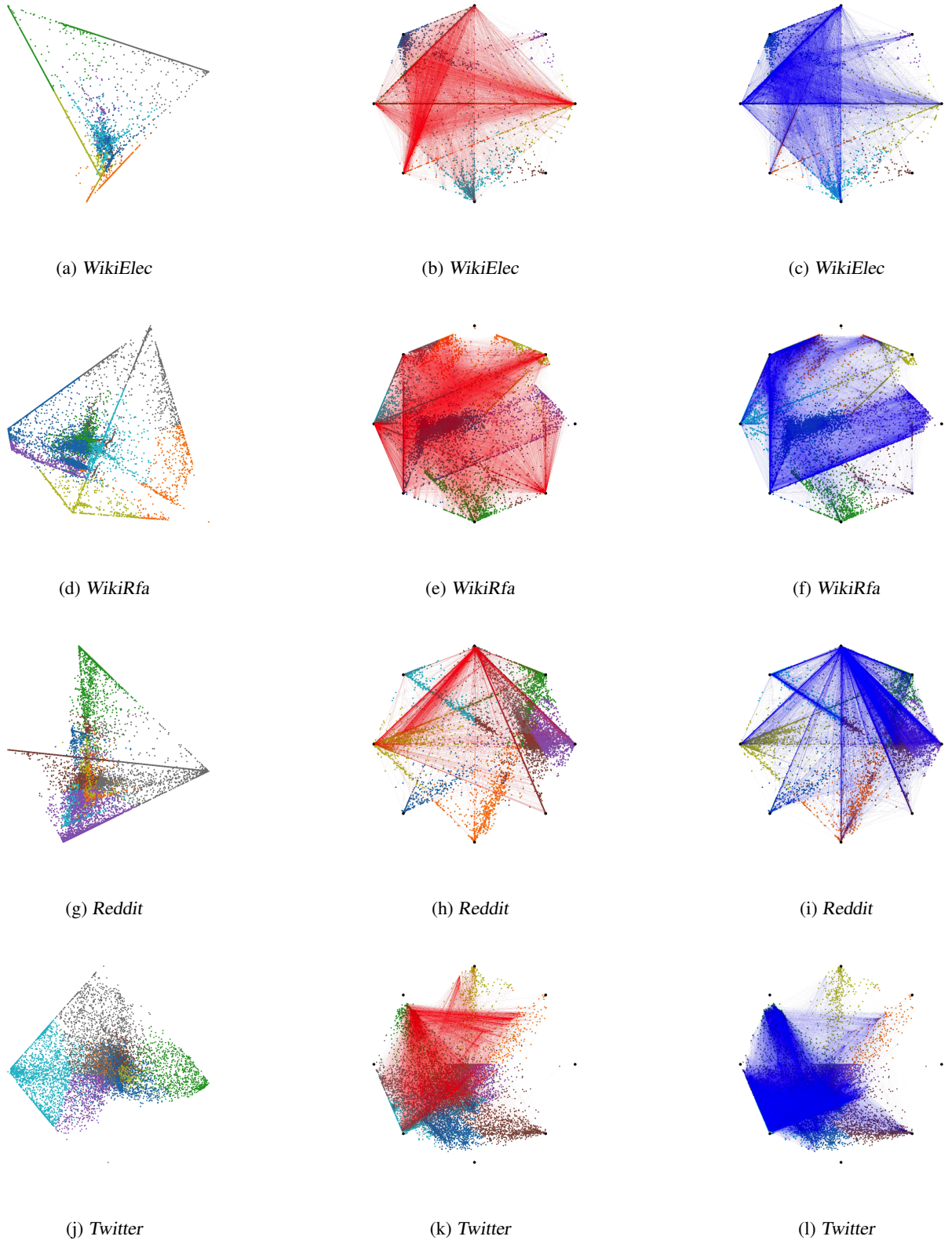
(j) *Twitter*

(k) *Twitter*

(l) *Twitter*

Figure 3: Inferred polytope visualizations for various networks. The first column showcases the $K = 8$ dimensional sociotope projected on the first two principal components (PCA) — second and third columns provide circular plots of the sociotope enriched with the negative (red) and positive (blue) links, respectively.

induced space while the second and third columns correspond to the circular plots enriched by the negative (red) and positive (blue) links, respectively. We observe how the polytope successfully uncovers extreme positional nodes. More specifically, all networks have at least one archetype which acts as a "dislike" hub and at least one as a "like" hub. Meaning that these archetypes contain high values of negative/positive interactions. For the *wiki-RfA* and *Twitter* networks we observe archetypes of very low degree, this is explained due to some only "disliked" nodes being pushed away from the main node population. These can be regarded as "outliers" of the sociotope. Nevertheless, such outliers are discovered since they provide high expressive power for the model.

Discussion. The Signed Relational Latent Distance Model has been presented for the undirected case setting, and we employed the Euclidean distance for both Skellam rates λ_{ij}^+, λ_{ij}^-. The capacity of the current formulation works well for undirected networks. Nevertheless, there are alternative model formulations, and keeping the distance identical for the positive and negative rates constrains the models' expressive capability, especially for the directed/bipartite signed network case. We therefore explore additional model formulations such as setting the Skellam rates as, $\lambda_{ij}^+ = \exp(\beta_i + \beta_j - ||\mathbf{z}_i - \mathbf{w}_j||_2)$ and $\lambda_{ij}^- = \exp(\gamma_i + \gamma_j - ||\boldsymbol{u}_i - \boldsymbol{w}_j||_2)$ in the supplementary material. Under this assumption, a positive directed relationship ($i \rightarrow j$) shows that node i "likes" node j and "dislikes" node j if it is negative. The latent embedding $\boldsymbol{w}_j$ is then the receiver position for the "likes" and "dislikes" with embeddings $\mathbf{z}_i$ and $\boldsymbol{u}_i$ being the sender positions for positive and negative relationships, respectively. In this case, we introduce three latent embeddings instead of the conventional two for the undirected case. The disparity of location $\mathbf{z}_i$ and $\boldsymbol{u}_i$ here can point out how polarity is formed between the two regions of the latent space (Please see the supplementary material for further discussion and results).

Another important design characteristic for the SLDM/SLIM frameworks is the choice of the prior/regularization of the different parameters. So far, we did not tune any regularization strength of the priors and simply adopted a normal distribution on the model parameters and non-informative uniform Dirichlet prior on **Z** in the case of SLIM. Potential tuning of the priors with cross-validation is expected to boost the performance and results.

A prominent characteristic of signed networks is the sparsity or, in other words, the excess of "zero" weights among node pairs. An intriguing direction to account for it might be the zero-inflated version of the Skellam distribution (Karlis and Ntzoufras, 2008). Here essentially, we can define a mixture model responsible for the imbalance between cases (sign-weighted links) and controls (neutral zero links) in the network. Such zero-inflated

SLDM/SLIM models can thereby define a generative process that can straightforwardly address different levels of network sparsity.

Whereas we consider the generalization of SLDM and SLIM to directed networks in the supplementary, a possible future direction should consider generalizations to bipartite networks in which we expect the directed generalizations to be applicable (Kim et al., 2018a; Nakis et al., 2022). Furthermore, networks of polarization typically evolve over time. Future work should thus investigate how the proposed modeling framework can be extended to characterize dynamic networks leveraging existing works by exploring dynamic extensions of latent space modeling approaches, including the diffusion model of (Sarkar and Moore, 2005) and approaches reviewed in Kim et al. (2018a).

4 CONCLUSION AND LIMITATIONS

The proposed Skellam Latent Distance Model (SLDM) and Signed Latent Relational Distance model (SLIM) provide easily interpretable network visualization with favorable performance in the link prediction tasks for weighted signed networks. In particular, endowing the model with a space constrained to polytopes (forming the SLIM) enabled us to characterize distinct aspects in terms of extreme positions in the social networks akin to conventional archetypal analysis but for graph-structured data. The Skellam distribution is considerably beneficial in modeling signed networks, whereas the relational extension of AA can be applied for other likelihood specifications, such as LDMs in general. This work thereby provides a foundation for using likelihoods accommodating weighted signed networks and representations akin to AA in general for analyzing networks.

The optimization for the SLDM/SLIM frameworks is a highly non-convex problem and thus relies on the quality of initialization in terms of convergence speed. In this regard, we use a deterministic initialization based on the normalized Laplacian. In addition, we observed that a maximum likelihood estimation of the model parameters became unstable when the network contained some nodes having only negative interactions. This is a direct consequence of the presence of the distance term $(\exp(+||\cdot||_2))$ for negative interactions, which can lead to overflow during inference. Nevertheless, the adopted MAP estimation was found to be stable across all networks. For real networks, the generative model created an "excess" of negative links increasing the overall network sparsity. For that, a modified *SLIM* excluding the regularization over the model parameters was introduced which achieved correct network sparsity (as shown in supplementary). Assuming priors over the model parameters created a bias over the generated network when compared to the ground truth network statistics.

Acknowledgements

References

F. Atay and H. Tunçel Gölpek. On the spectrum of the normalized laplacian for signed graphs: Interlacing, contraction, and replication. *Linear Algebra and its Applications*, 442:165–177, 02 2014. doi: 10.1016/j.laa.2013.08.022.

P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10):1531–1542, 2015.

O. Barndorff-Nielsen, D. Pollard, and N. Shephard. Integer-valued lévy processes and low latency financial econometrics. *Quantitative Finance*, 12, 10 2010.

S. Beugelsdijk, H. van Herk, and R. Maseland. The nature of societal conflict in europe; an archetypal analysis of the postmodern cosmopolitan, rural traditionalist and urban precariat. *JCMS*, n/a(n/a), 2022.

B. Büeler, A. Enge, and K. Fukuda. Exact volume computation for polytopes: a practical study. In *Polytopes—combinatorics and computation*, pages 131–154. Springer, 2000.

D. Cartwright and F. Harary. Structural balance: a generalization of heider's theory. *Psychological review*, 63 5: 277–93, 1956.

A. Çelikkanat, N. Nakis, and M. Mørup. Piecewise-velocity model for learning continuous-time dynamic node representations. In *The First Learning on Graphs Conference*, 2022.

K.-Y. Chiang, N. Natarajan, A. Tewari, and I. S. Dhillon. Exploiting longer cycles for link prediction in signed networks. In *CIKM*, page 1157–1162. Association for Computing Machinery, 2011.

A. Cutler and L. Breiman. Archetypal analysis. *Technometrics*, 36(4):338–347, 1994.

A. Dagnes. Us vs. them: Political polarization and the politicization of everything. In *Super Mad at Everything All the Time*, pages 119–165. Springer, 2019.

D. DellaPosta, Y. Shi, and M. Macy. Why do liberals drink lattes? *American Journal of Sociology*, 120(5):1473–1511, 2015.

V. R. K. Garimella and I. Weber. A long-term analysis of polarization on twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 11: 528–531, May 2017.

G. H. Golub and C. F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, USA, 1996. ISBN 0801854148.

A. Grover and J. Leskovec. Node2Vec: Scalable feature learning for networks. In *KDD*, pages 855–864, 2016.

P. Guerra, W. Meira Jr, C. Cardie, and R. Kleinberg. A measure of polarization on social media networks based on community boundaries. In *Proceedings of the international AAAI conference on web and social media*, volume 7, pages 215–224, 2013.

F. Harary. On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 2(2):143 – 146, 1953.

Y. Hart, H. Sheftel, J. Hausser, P. Szekely, N. B. Ben-Moshe, Y. Korem, A. Tendler, A. E. Mayo, and U. Alon. Inferring biological tasks using pareto analysis of high-dimensional data. *Nature methods*, 12(3): 233–235, 2015.

M. J. Hetherington. Review article: Putting polarization in perspective. *British Journal of Political Science*, 39 (2):413–448, 2009.

P. D. Hoff. Bilinear mixed-effects models for dyadic data. *JASA*, 100(469):286–295, 2005.

P. D. Hoff, A. E. Raftery, and M. S. Handcock. Latent space approaches to social network analysis. *JASA*, 97(460): 1090–1098, 2002.

J. Huang, H. Shen, L. Hou, and X. Cheng. Signed graph attention networks, 2019. URL https://arxiv.org/abs/1906.10958.

J. Huang, H. Shen, L. Hou, and X. Cheng. Sdgnn: Learning node representation for signed directed networks, 2021. URL https://arxiv.org/abs/2101.02390.

Z. Huang, A. Silva, and A. Singh. Pole: Polarized embedding for signed networks. *WSDM*, pages 390–400, 2022.

M. R. Islam, B. Aditya Prakash, and N. Ramakrishnan. SIGNet: Scalable embeddings for signed networks. In D. Phung, V. S. Tseng, G. I. Webb, B. Ho, M. Ganji, and L. Rashidi, editors, *Advances in Knowledge Discovery and Data Mining*, pages 157–169, Cham, 2018. Springer International Publishing.

D. Karlis and I. Ntzoufras. Bayesian analysis of the differences of count data. *Statistics in medicine*, 25:1885–905, 06 2006.

D. Karlis and I. Ntzoufras. Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference. *IMA Journal of Management Mathematics*, 20(2):133–145, 2008.

A. Keuchenius, P. Törnberg, and J. Uitermark. Why it is important to consider negative ties when studying polarized debates: A signed network analysis of a dutch cultural controversy on twitter. *PLOS ONE*, 2021.

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

J. Kim, H. Park, J.-E. Lee, and U. Kang. SIDE: Representation learning in signed directed networks. In *Proceedings of the 2018 World Wide Web Conference*, page 509–518. International World Wide Web Conferences Steering Committee, 2018b.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2017.

S. Kumar, W. L. Hamilton, J. Leskovec, and D. Jurafsky. Community interaction and conflict on the web. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, pages 933–943. International World Wide Web Conferences Steering Committee, 2018.

J. Leskovec, D. Huttenlocher, and J. Kleinberg. Predicting positive and negative links in online social networks. In *WWW*, page 641–650, 2010.

J. Moody and P. J. Mucha. Portrait of political party polarization. *Network Science*, 1(1):119–121, 2013. doi: 10.1017/nws.2012.3.

M. Mørup and L. Kai Hansen. Archetypal analysis for machine learning. In *2010 IEEE International Workshop on Machine Learning for Signal Processing*, pages 172–177, 2010. doi: 10.1109/MLSP.2010.5589222.

N. Nakis, A. Çelikkanat, S. L. Jørgensen, and M. Mørup. A hierarchical block distance model for ultra low-dimensional graph representations. 2022. URL https://arxiv.org/abs/2204.05885.

N. Nakis, A. Çelikkanat, and M. Mørup. HM-LDM: A hybrid-membership latent distance model. In *Complex Networks and Their Applications XI*, pages 350–363, Cham, 2023. Springer International Publishing. ISBN 978-3-031-21127-0.

Z. P. Neal. A sign of the times? weak and strong polarization in the u.s. congress, 1973–2016. *Social Networks*, 60:103–112, 2020.

B. Ordozgoiti, A. Matakos, and A. Gionis. Finding large balanced subgraphs in signed networks. In *Proceedings of The Web Conference 2020*, page 1378–1388, 2020.

B. Perozzi, R. Al-Rfou, and S. Skiena. DeepWalk: Online learning of social representations. *CoRR*, abs/1403.6652, 2014.

A. E. Raftery, X. Niu, P. D. Hoff, and K. Y. Yeung. Fast inference for the latent space network model using a case-control approximate likelihood. *Journal of Computational and Graphical Statistics*, 21(4):901–919, 2012.

P. Sarkar and A. Moore. Dynamic social network analysis using latent space models. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *NeurIPS*, volume 18, 2005.

O. Shoval, H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and U. Alon. Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science*, 336(6085):1157–1160, 2012.

J. G. Skellam. The frequency distribution of the difference between two poisson variates belonging to different populations. *Journal of the Royal Statistical Society. Series A (General)*, 109(Pt 3):296–296, 1946.

M. Thomas, B. Pang, and L. Lee. Get out the vote: Determining support or opposition from congressional floor-debate transcripts. *CoRR*, abs/cs/0607062, 2006.

R.-C. Tzeng, B. Ordozgoiti, and A. Gionis. Discovering conflicting groups in signed networks. In *NeurIPS*, 2020.

S. Wang, J. Tang, C. Aggarwal, Y. Chang, and H. Liu. *Signed Network Embedding in Social Media*, pages 327–335. 2017.

R. West, H. S. Paskov, J. Leskovec, and C. Potts. Exploiting social network structure for person-to-person sentiment analysis. *TACL*, 2:297–310, 2014.

E. Williamson. Plato's "eidos" and the archetypes of jung and frye. *Interpretations*, 16(1):94–104, 1985. ISSN 0196903X.

P. Xu, J. Wu, W. Hu, and B. Du. Link prediction with signed latent factors in signed social networks. *Proceedings of the Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, pages 1046–1054, 2019.

L. Zhuang, C.-H. Lin, M. A. Figueiredo, and J. M. Bioucas-Dias. Regularization parameter selection in minimum volume hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12):9858–9877, 2019.

A Directed Case Model Formulations

In this section, we describe how our proposed frameworks can be extended to the study of directed networks, and we further explore additional model formulations allowing for more capacity and expressive power.

A.1 The Skellam Latent Distance Model for the Directed Case (LDM)

Our main purpose here is to learn two latent node representations $\{\mathbf{z}_i\}_{i \in \mathcal{V}} \in \mathbb{R}^K$ and $\{\mathbf{w}_i\}_{i \in \mathcal{V}} \in \mathbb{R}^K$ in a low dimensional space for a given directed signed network $\mathcal{G} = (\mathcal{V}, \mathcal{Y})$ ($K \ll |\mathcal{V}|$). The two sets of the latent embeddings correspond to modeling directed relationships $i \to j$ of nodes, with $\mathbf{z}_i$ the source node and $\mathbf{w}_j$ the target node, and vice-versa for an oppositely directed relationship $i \leftarrow j$. Similar to the main paper, we can formulate the negative log-likelihood of a latent distance model under the Skellam distribution as:

$$
\begin{aligned}
\mathcal{L}(\mathcal{Y}) &:= \log p(y_{ij}|\lambda_{ij}^+, \lambda_{ij}^-) \\
&= \sum_{i,j} (\lambda_{ij}^+ + \lambda_{ij}^-) - \frac{y_{ij}}{2} \log \left(\frac{\lambda_{ij}^+}{\lambda_{ij}^-} \right) - \log \left(\mathcal{I}_{|y_{ij}|} \left(2\sqrt{\lambda_{ij}^+ \lambda_{ij}^-} \right) \right),
\end{aligned}
$$

For the directed case, the Skellam distribution has two rate parameters as well, and we consider them to learn latent node representations $\{\mathbf{z}_i\}_{i \in \mathcal{V}}$ and $\{\mathbf{w}_j\}_{j \in \mathcal{V}} \in \mathbb{R}^K$ by defining them as follows:

$$
\lambda_{ij}^+ = \exp\left(\beta_i + \gamma_j - ||\mathbf{z}_i - \mathbf{w}_j||_2\right), \tag{11}
$$

$$
\lambda_{ij}^- = \exp\left(\delta_i + \epsilon_j + ||\mathbf{z}_i - \mathbf{w}_j||_2\right), \tag{12}
$$

where the set $\{\beta_i, \gamma_i, \delta_i, \epsilon_i\}_{i \in \mathcal{V}}$ denote the node-specific random effect terms, and $||\cdot||_2$ is the Euclidean distance function. More specifically, the sender β_i and the receiver γ_j random effects represent the "social" reach of a node and the tendency to form positive interactions, expressing positive degree heterogeneity (indicated by $+$ as a superscript of λ). In contrast, δ_i and ϵ_j provide the "anti-social" sender and receiver effect of a node to form negative connections, and thus model negative degree heterogeneity (indicated by $-$ as a superscript of λ).

By imposing (as in the undirected case) standard normally distributed priors elementwise on all model parameters $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\epsilon}, \mathbf{Z}, \boldsymbol{W}\}$, i.e., $\theta_i \sim \mathcal{N}(0,1)$, We define a maximum a posteriori (MAP) estimation over the model parameters, via the loss function to be minimized (ignoring constant terms):

$$
\begin{aligned}
Loss &= \sum_{i,j} \left(\lambda_{ij}^+ + \lambda_{ij}^- - \frac{y_{ij}}{2} \log \left(\frac{\lambda_{ij}^+}{\lambda_{ij}^-} \right) \right) - \sum_{i,j} \log I_{|y_{ij}|} \left(2\sqrt{\lambda_{ij}^+ \lambda_{ij}^-} \right) \\
&+ \frac{\rho}{2} \left(||\mathbf{Z}||_F^2 + ||\mathbf{W}||_F^2 + ||\boldsymbol{\gamma}||_F^2 + ||\boldsymbol{\beta}||_F^2 + ||\boldsymbol{\delta}||_F^2 + ||\boldsymbol{\epsilon}||_F^2 \right),
\end{aligned}
\tag{13}
$$

where $||\cdot||_F$ denotes the Frobenius norm. In addition, ρ is the regularization strength with $\rho = 1$ yielding the adopted normal prior with zero mean and unit variance.

A.2 The Signed Relational Latent Distance Model for Directed Networks

We formulate the relational AA in the context of the family of LDMs and for directed networks, as:

$$
\lambda_{ij}^+ = \exp\left(\beta_i + \gamma_j - \|\mathbf{A}(\mathbf{z}_i - \mathbf{w}_j)\|_2\right) \tag{14}
$$

$$
= \exp\left(\beta_i + \gamma_j - \|\mathbf{R}[\mathbf{Z};\mathbf{W}]\mathbf{C}(\mathbf{z}_i - \mathbf{w}_j)\|_2\right). \tag{15}
$$

$$
\lambda_{ij}^- = \exp\left(\delta_i + \epsilon_j + \|\mathbf{A}(\mathbf{z}_i - \mathbf{w}_j)\|_2\right) \tag{16}
$$

$$
= \exp\left(\delta_i + \epsilon_j + \|\mathbf{R}[\mathbf{Z};\mathbf{W}]\mathbf{C}(\mathbf{z}_i - \mathbf{w}_j)\|_2\right). \tag{17}
$$

Notably, in the AA formulation $\mathbf{X} = \mathbf{R}[\mathbf{Z};\mathbf{W}]$ corresponds to observations formed by the concatenations of the convex combinations $\mathbf{Z}$ and $\boldsymbol{W}$ of positions given by the columns of $\boldsymbol{R}^{K \times K}$. Furthermore, in order to ensure what is used to define archetypes $\mathbf{A} = \mathbf{XC} = \mathbf{R}[\mathbf{Z};\mathbf{W}]\mathbf{C}$ corresponds to observations using these archetypes in their reconstruction $[\mathbf{Z};\mathbf{W}]$, we define $C \in \boldsymbol{R}^{2N \times K}$ as a gated version of $[\mathbf{Z};\mathbf{W}]$ normalized to the simplex such that $\boldsymbol{c}_d \in \Delta^{2N}$ by defining

$$
c_{nd} = \frac{([\mathbf{Z};\mathbf{W}]^\top \circ [\sigma(\mathbf{G})]^\top)_{nd}}{\sum_{n'} ([\mathbf{Z};\mathbf{W}]^\top \circ [\sigma(\mathbf{G})]^\top)_{n'd}} \tag{18}
$$

in which $\circ$ denotes the elementwise (Hadamard) product and $\sigma(\mathbf{G})$ defines the logistic sigmoid elementwise applied to the matrix $\mathbf{G}$. As a result, the extracted archetypes are ensured to correspond to the nodes assigned the archetype, whereas the location of the archetypes can be flexibly placed in space as defined by $\mathbf{R}$. By defining $\mathbf{z}_i = \text{softmax}(\tilde{\mathbf{z}}_i)$ and $\boldsymbol{w}_i = \text{softmax}(\tilde{\boldsymbol{w}}_i)$ we further ensure $\mathbf{z}_i, \boldsymbol{w}_i \in \Delta^K$.

As in the undirected case, the loss function of Eq. (13) is adopted for the relational AA formulation forming the SLIM, with the prior regularization applied to the corners of the extracted polytope $\mathbf{A} = \mathbf{R}[\mathbf{Z}; \mathbf{W}]\mathbf{C}$ instead of the latent embeddings $\mathbf{Z}, \mathbf{W}$ imposing a standard elementwise normal distribution as prior $a_{k,k'} \sim \mathcal{N}(0,1)$. Furthermore, we impose a uniform Dirichlet prior on the columns of $\mathbf{Z}, \mathbf{W}$, i.e. $(\mathbf{z}_i, \mathbf{w}_i \sim Dir(\mathbf{1}_K))$, this only contributes constant terms to the joint distribution. As a result, the loss function is given by Eq. (13) replacing $\|\mathbf{Z}\|_F^2$ and $\|\mathbf{W}\|_F^2$ with $\|\mathbf{A}\|_F^2$ for the maximum a posteriori (MAP) optimization.

A.3 Model Extensions for Additional Capacity

In the main paper, we briefly introduced an additional formulation for the rates of the Skellam distribution as adopted by our models. In this case (and for directed networks), the rates are:

$$\lambda_{ij}^+ = \exp(\beta_i + \gamma_j - ||\mathbf{z}_i - \mathbf{w}_j||_2) \text{ and } \lambda_{ij}^- = \exp(\delta_i + \epsilon_j - ||\mathbf{u}_i - \mathbf{w}_j||_2) \tag{19}$$

In this proposition, we have adopted three latent embeddings instead of the two previously described for the directed case. The disparity of location $\mathbf{z}_i$ and $\mathbf{u}_i$ here can point out how polarity is formed between the two regions of the latent space. This model specification introduces an additional regularization for the third embedding matrix $\mathbf{U}$ in the loss function of Equation (13). For the RAA case, we thereby define $\mathbf{X} = \mathbf{R}[\mathbf{Z}; \mathbf{U}; \mathbf{W}]$, i.e., as the concatenation of all three latent positions and with $\boldsymbol{C} \in \boldsymbol{R}^{3N \times K}$.

A.4 Directed case — Results and performance

In Table 4 and Table 5, we provide the results for the directed networks against various prominent baselines. Note that POLE is not defined for the directed case while SIDE failed to create embeddings for one-degree nodes. For the frameworks, we use two additional variations for SLDM and SLIM. The first ones are the SLDM REG=0.01 and SLIM REG=0.01, where we have used a regularization power $\rho = 0.01$ in Equation (13). This shows how performance is affected by less regularized parameters. In addition, we also provide results for SLDM-EXPR and SLIM-EXPR which denote the more expressive model as described in Subsection A.3. The results showcase our models' capability to outperform the baselines or provide competitive performance. Comparing now the SLDM and SLIM different variations we observe that performance is boosted by just using the vanilla methods. It seems that the most important trait is the regularization power of the model rather than the expressive capabilities that extra parameters provide to the model. Lastly, Figure 4 provides the same visualizations as in the main paper but for the directed networks.

Table 4: Area Under Curve (AUC-ROC) scores for varying representation sizes (Directed). The symbol '-' denotes that the corresponding model is not able to run on directed networks while 'x' that the model returned errors.

	WikiElec			WikiRfa			Reddit		
Task	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$
POLE	-	-	-	-	-	-	-	-	-
SLF	.938	.971	.980	.991	.980	.985	.823	.974	.984
SIGAT	.921	.750	.871	.988	.772	.927	.982	.713	.980
SIDE	x	x	x	x	x	x	x	x	x
SIGNET	.929	.907	.835	.991	.921	.873	.881	.757	.719
SLIM (OURS)	.910	.981	.963	.984	.989	.981	.713	.973	.982
SLDM (OURS)	.914	.977	.966	.983	.987	.978	.657	.937	.964
SLIM REG=0.01 (OURS)	.927	.989	.980	.992	.994	.990	.827	.982	.989
SLDM REG=0.01 (OURS)	.940	.989	.980	.984	.987	.976	.774	.982	.986
SLIM-EXPR (OURS)	.922	.984	.977	.987	.988	.982	.706	.930	.949
SLDM-EXPR (OURS)	.915	.987	.985	.989	.994	.992	.657	.965	.965

Table 5: Area Under Curve (AUC-PR) scores for varying representation sizes (Directed). The symbol '-' denotes that the corresponding model is not able to run on directed networks while 'x' that the model returned errors.

	WikiElec			WikiRfa			Reddit		
Task	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$	$p@n$	$p@z$	$n@z$
POLE	-	-	-	-	-	-	-	-	-
SLF	.981	.949	.890	.995	.954	.951	.978	.972	.919
SiGAT	.977	.689	.562	.993	.685	.714	.998	.727	.659
SIDE	x	x	x	x	x	x	x	x	x
SigNet	.979	.831	.577	.995	.840	.671	.988	.675	.233
SLIM (ours)	.971	.974	.852	.989	.981	.951	.962	.971	.874
SLDM (ours)	.972	.967	.862	.988	.978	.939	.952	.948	.861
SLIM reg=0.01 (ours)	.976	.983	.910	.995	.988	.973	.980	.982	.918
SLDM reg=0.01 (ours)	.981	.983	.912	.991	.976	.930	.972	.981	.911
SLIM-expr (ours)	.976	.978	.914	.992	.980	.953	.958	.938	.823
SLDM-expr (ours)	.973	.981	.936	.993	.987	.979	.949	.966	.871

B   Initialization

For the SLDM model, we used the Eigen-decomposition of the normalized Laplacian for singed networks (Atay and Tunçel Gölpek, 2014). Solving the eigenproblem for a few eigenvalues can be done efficiently through the Lanczos method (Golub and Van Loan, 1996), due to the high sparsity of real large-scale networks.

For SLIM, we would like to initialize matrix $\mathbf{A}$ based on the convex hull of the spectral decomposition of the normalized Laplacian. This is very costly since finding the convex hull has an exponential increase in complexity in terms of the dimensionality of the space. For that purpose, we use the furthest sum algorithm (Mørup and Kai Hansen, 2010) to discover guaranteed distinct aspects of the spectral space. Lastly, since we are unable to directly initialize $\mathbf{A}$, we use the furthest sum discovered points to initialize $\mathbf{R}$ while also tuning $\mathbf{G}$ for picking up the correct points in the latent space.

C   Bessel Function Approximation

We need to compute the modified Bessel function of the first kind and of order y for the implementation of our proposed approach, which is defined by

$$I_y(x) = (\frac{x}{2})^y \sum_{k=0}^{\infty} \frac{(\frac{x^2}{4})^k}{k!\,\Gamma(y+k+1)}$$

We approximate the actual value by only considering the first 50 terms of the infinite sum. Since we have small orders of y and small values of x, the series components converge to zero quickly. We observed that taking the first 50 components does not affect the performance/accuracy of the model.

D   Generating based on real networks

Here, we test how the model generates based on real networks. We use the *wikiElec* to train an $K = 8$ dimensional SLIM model and we then generate a network based on the inferred parameters. Results are shown in Fig. 5 where we observe that the generated network learns successfully the main structure of the network but generates more non-zero elements and more negative links thereby decreasing the sparsity and increasing the percentage of negative links when compared to the ground truth. Modifying *SLIM* to exclude the regularization over the model parameters achieves correct network sparsity as shown in Fig 6 with only a 2% increase in the inferred percentage of negative links when compared to the ground truth. Adding priors to the model creates a bias over the network generation. Lastly, the un-regularized SLIM boosted performance in the link prediction tasks ranging from 1% to 5% for the *wikiElec* network. Nevertheless, priors over the model parameters stabilize the inference when "extreme" negative nodes exist in the network (nodes with only negative links) that can also be considered outliers.

(a) *WikiElec*

(b) *WikiElec*

(c) *WikiElec*

(d) *WikiRfa*

(e) *WikiRfa*

(f) *WikiRfa*

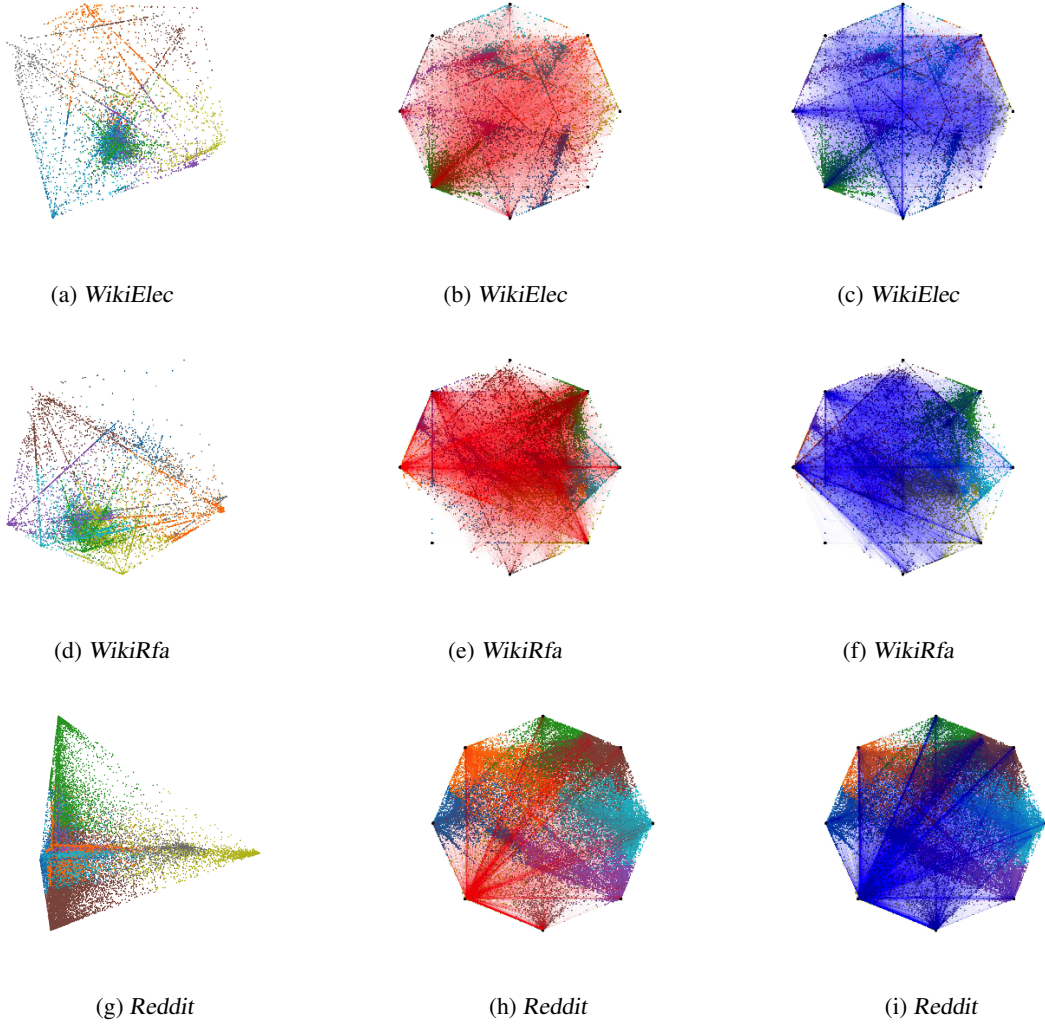(g) *Reddit*

(h) *Reddit*

(i) *Reddit*

Figure 4: Inferred polytope visualizations for various directed networks. The first column showcases the $K = 8$ dimensional sociotope projected on the first two principal components (PCA) of the combined embeddings (source, target) $[\mathbf{Z}; \mathbf{W}]$ — second and third columns provide circular plots of the sociotope enriched with the negative (red) and positive (blue) links, respectively.

E   Effect of Sampling Size

In the main paper, the sample size was set to the maximum number (~ 3000 nodes) that our 8GB GPU could fit in memory. Here, we provide a study on how different sample sizes affect the performance of the SLIM model. In Fig 7, we provide the performance across different tasks for the *wikiElec* dataset, considering sampling size of $\{10\%, 20\%, 30\%, 40\%, 50\%\}$. We observe with small differences almost constant performance across different sampling sizes. As we decrease the sampling size to 10% and 20% we observe some more significant decreases in the $p@n$ task performance. This is because we keep the total number of training iterations (it=5000) for all cases. Overall, smaller sampling sizes require additional iterations to converge to the performance of the model with larger sampling sizes.

F   Effect of Learning Rate

The learning rate for SLDM and SLIM was set to $lr = 0.05$. In Fig 8, we provide the performance across different tasks for the *wikiElec* dataset, considering three different learning rates $lr \in \{0.01, 0.05, 0.1\}$. We observe that the performance can be considered constant for the different learning rates, showing small sensitivity to the choice of this hyperparameter.

(a) Ground Truth: $(.003, 78\%, 22\%)$ (b) Generated: $(.006, 63\%, 37\%)$
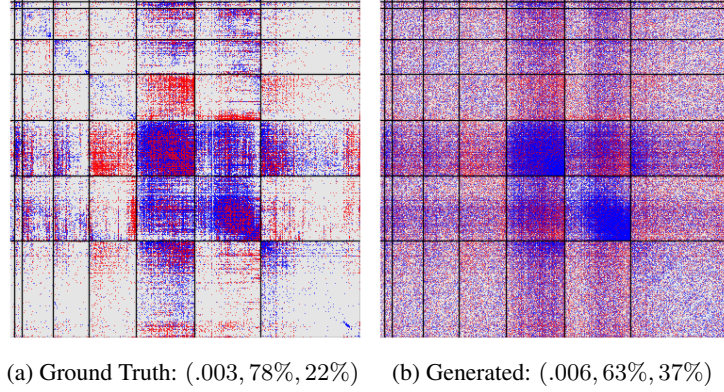
Figure 5: *wikiElec* ground truth (left) adjacency matrix and generated (right) adjacency matrix based on inferred parameters. The parenthesis shows the network statistics as: (density,% of positive (blue) links,% of negative (red) links). All network adjacency matrices are ordered based on z_i, in terms of maximum archetype membership and internally according to the magnitude of the corresponding archetype most used for their reconstruction.

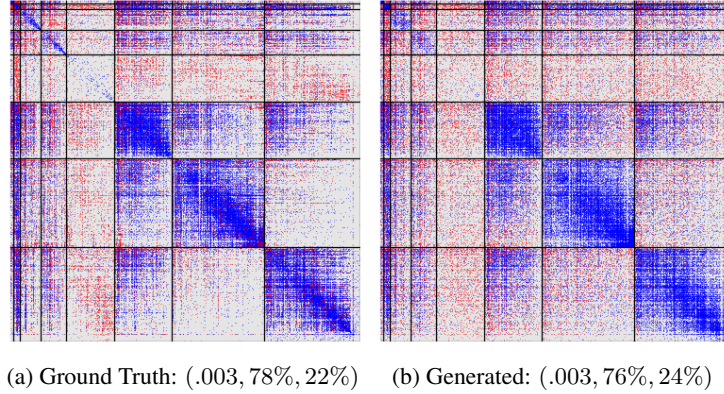(a) Ground Truth: $(.003, 78\%, 22\%)$ (b) Generated: $(.003, 76\%, 24\%)$

Figure 6: *wikiElec* ground truth (left) adjacency matrix and generated (right) adjacency matrix based on inferred parameters with a SLIM **without regularization priors** over the parameters. The parenthesis shows the network statistics as: (density,% of positive (blue) links,% of negative (red) links). All network adjacency matrices are ordered based on z_i, in terms of maximum archetype membership and internally according to the magnitude of the corresponding archetype most used for their reconstruction.
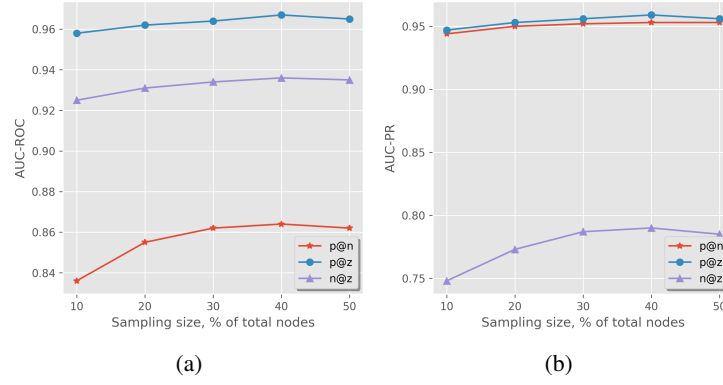
(a) (b)

Figure 7: *wikiElec*: Performance of SLIM across sample sizes for different tasks, (a) Area-Under-Curve Receiver Operating Characteristic scores, (b) Area-Under-Curve Precision-Recall scores. Both AUC-ROC and AUC-PR scores are almost constant across different dimensions
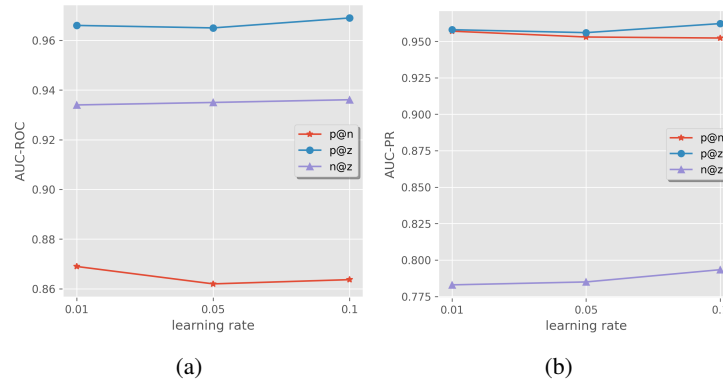

(a) (b)

Figure 8: *wikiElec*: Performance of SLIM across learning rates for different tasks, (a) Area-Under-Curve Receiver Operating Characteristic scores, (b) Area-Under-Curve Precision-Recall scores. Both AUC-ROC and AUC-PR scores are constant across different dimensions