



HAL
open science

Deep Latent Position Block Model for the clustering of nodes in Mutli-graphs

Seydina Ousmane Niang, Charles Bouveyron, Marco Corneli, Pierre Latouche,
Rémi Boutin

► **To cite this version:**

Seydina Ousmane Niang, Charles Bouveyron, Marco Corneli, Pierre Latouche, Rémi Boutin. Deep Latent Position Block Model for the clustering of nodes in Mutli-graphs. 7th international AI conference, SophI.A Summit - SOPHIA SUMMIT 2024, Nov 2024, Valbonne (06560), France. hal-04909806

HAL Id: hal-04909806

<https://hal.science/hal-04909806v1>

Submitted on 24 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DEEP LATENT POSITION BLOCK MODEL FOR THE CLUSTERING OF NODES IN MUTLI-GRAPHS

S. Niang¹, M. Corneli^{1,2}, P. Latouche^{3,4}, C. Bouveyron¹, R. Boutin⁵

¹ Université Côte d'Azur, Inria, CNRS, Laboratoire J.A. Dieudonné, Maasai team, Nice, France

² Université Côte d'Azur, Laboratoire CEPAM, Nice, France

³ Université Clermont Auvergne, CNRS Laboratoire, Aubière, France

⁴ Institut Universitaire de France (IUF), Paris, France

⁵ Sorbonne Université, Université Paris Cité, CNRS, UMR 8001

1. Introduction

Goal: Cluster nodes in a multidimensional social network.

To this end, we introduce:

- An extension of the [deep latent position block model \(Deep-LPBM\)](#) decoder to multidimensional networks.
- A specific encoder based on graph convolutional networks for each network and a multi-layer perceptron to aggregate them into a unified continuous embedding Z representing [partial membership](#) probabilities to the cluster.
- An evaluation of the performance of the approach and an application to real-world social networks.

2. Deep-LPBM

Data structure: The observed network is identified by a given $N \times N$ adjacency matrix $A = (A_{ij})_{i,j \in 1, \dots, N}$ whose entry (i, j) is one if the individual i is in relation with the individual j and zero otherwise. A is supposed symmetric.

The DDPM model assumes:

- Nodes are spread into Q clusters.
- Each node i have a latent representation $Z_i \in \mathbb{R}^{Q-1}$ such that:

$$Z_i \stackrel{\text{iid}}{\sim} \mathcal{N}_d(0, \mathbb{I}_{Q-1}).$$

- The probability for node i to be in cluster q is given by:

$$\eta_{iq} = \begin{cases} \frac{\exp(Z_{iq})}{1 + \sum_{r=1}^{Q-1} \exp(Z_{ir})} & \text{if } q \neq Q, \\ \frac{1}{1 + \sum_{r=1}^{Q-1} \exp(Z_{ir})} & \text{if } q = Q. \end{cases}$$

- The probability of nodes i and j to be linked given η_i and η_j is :

$$\mathbb{P}(A_{ij} = 1 | \eta_i, \eta_j) = \eta_i^T \Pi \eta_j,$$

where $\Pi^{(\ell)} = (\Pi_{qr}^{(\ell)})_{1 \leq q, r \leq Q}$ is a $Q \times Q$ symmetric matrix and refers to the connectivity matrix whose entry (q, r) is the probability that a node in block q is connected to a node in block r .

Therefore nodes are allowed to [partially belong](#) to many clusters providing them the possibility to have overlapping rules in many clusters.

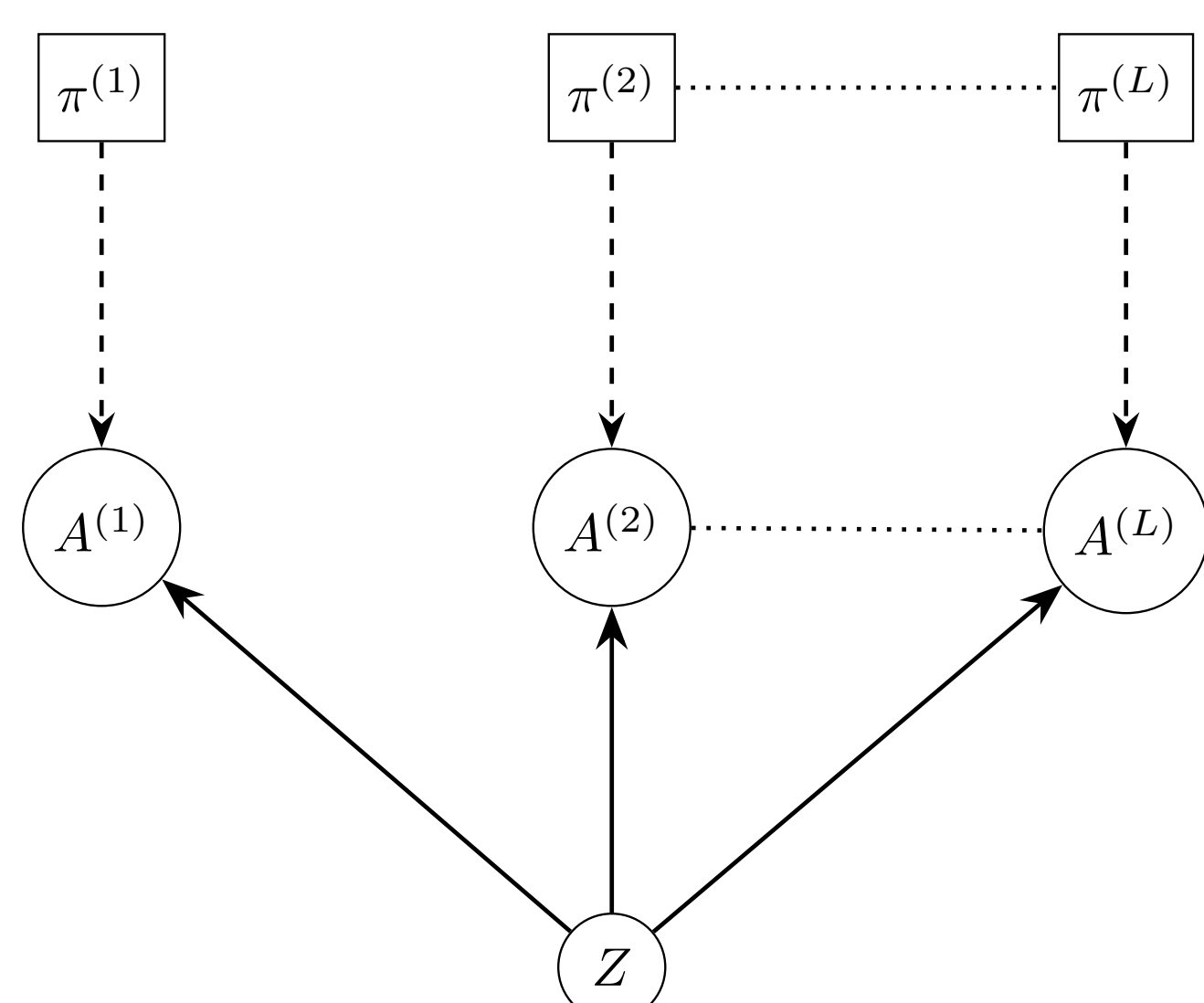
3. Our work: LPBMM

Data structure: Our observed [multidimensional](#) network is identified by a collection of L adjacency matrices $A = \{A^{(\ell)}\}^{\ell}$ each referring to a relationship between the same set of N individuals.

Our model assumes that:

- Nodes are spread into Q clusters.
- Each node i is supposed to have a single latent representation Z_i containing the clustering information of all the different networks.
- Each layer has its own deep-LPBM decoder with specific parameter $\Pi^{(\ell)}$ and the common latent variable Z_i , meaning that $\mathbb{P}(A_{ij}^{(\ell)} = 1 | \eta_i, \eta_j) = \eta_i^T \Pi^{(\ell)} \eta_j$
- The different networks are supposed to be independent given Z , resulting into the following joint distribution: $p(A, Z | \Pi) = \prod_{\ell=1}^L p(A^{(\ell)} | Z, \Pi^{(\ell)}) p(Z)$.

4. Graphical model



5. Inference

The observed data likelihood can be obtained by marginalizing over the latent variable Z

$$p(A) = \int_Z p(A, Z) dZ. \quad (1)$$

The direct evaluation of this integral is not [tractable](#) !

The impossibility to compute the joint posterior distribution of the latent variables given the data leads us to rely on the [variational inference](#).

To this end, we introduce a [variational posterior distribution](#) $q(\cdot)$ for which the following inequality holds:

$$\log p(A | \Pi) \geq \mathbb{E}_{q(Z|A)} \left[\log \frac{p(A, Z | \Pi)}{q(Z|A)} \right] := \mathcal{L}(\Pi, q(\cdot)),$$

and $\mathcal{L}(\Pi, q(\cdot))$ is known as the evidence lower born (ELBO) of the marginal log-likelihood. This [ELBO](#) is [maximized](#) when $q(\cdot)$ is equal to the [true posterior distribution](#) $p(Z|A, \Pi)$ which is itself non tractable.

Therefore, we suppose a mean-field factorization on $q(\cdot)$ in order to make the ELBO tractable:

$$q(Z|A) = \prod_{i=1}^N q_{\phi}(Z_i|A) = \prod_{i=1}^N \mathcal{N}(Z_i; \bar{\mu}_{\phi, \xi}(A)_i, \bar{\sigma}_{\phi, \omega}^2(A)_i \mathbb{I}_{Q-1}),$$

with

$\phi = (\phi^{(1)}, \dots, \phi^{(L)})$, $\bar{\mu}_{\phi, \xi}(A) = \text{MLP}_{\xi}(\mu_{\phi^{(1)}}^{(1)}, \dots, \mu_{\phi^{(L)}}^{(L)})$, $\bar{\sigma}_{\phi, \omega}(A) = \text{MLP}_{\omega}(\sigma_{\phi^{(1)}}^{(1)}, \dots, \sigma_{\phi^{(L)}}^{(L)})$, $[\mu_{\phi^{(\ell)}}^{(\ell)}, \sigma_{\phi^{(\ell)}}^{(\ell)}] = f_{\phi^{(\ell)}}^{(\ell)}(\bar{A}^{(\ell)})$. Here $f_{\phi^{(\ell)}}^{(\ell)}$ is a graph convolutional network and MLP refers to a multi-layer perceptron.

It has been established that the ELBO can be decomposed as follows:

$$\text{ELBO} = \underbrace{\mathbb{E}_{q(\cdot)} \left[\sum_{\ell} \log p(A^{(\ell)} | Z, \Pi) \right]}_{\text{Reconstruction term}} - \underbrace{\text{KL}(q(Z|A) || p(Z))}_{\text{Regularisation term}},$$

where the regularization term can be exactly computed and the reconstruction is estimated thanks to MCMC and stochastic gradient descent is done for all parameters thanks to the reparameterization trick.

Parameters initialisation: The parameters Π, ξ, ω, ϕ are initialized by fitting K-means on the k -th view $A^{(k)}$ to initialize cluster labels, which are then used to estimate initial latent positions through an inverse softmax transformation. It then optimizes an embedding by minimizing the reconstruction error with respect to the variational parameters and initializes the block connectivity matrix using edge counts between clusters in each network.

5. Experiment on Lazega dataset

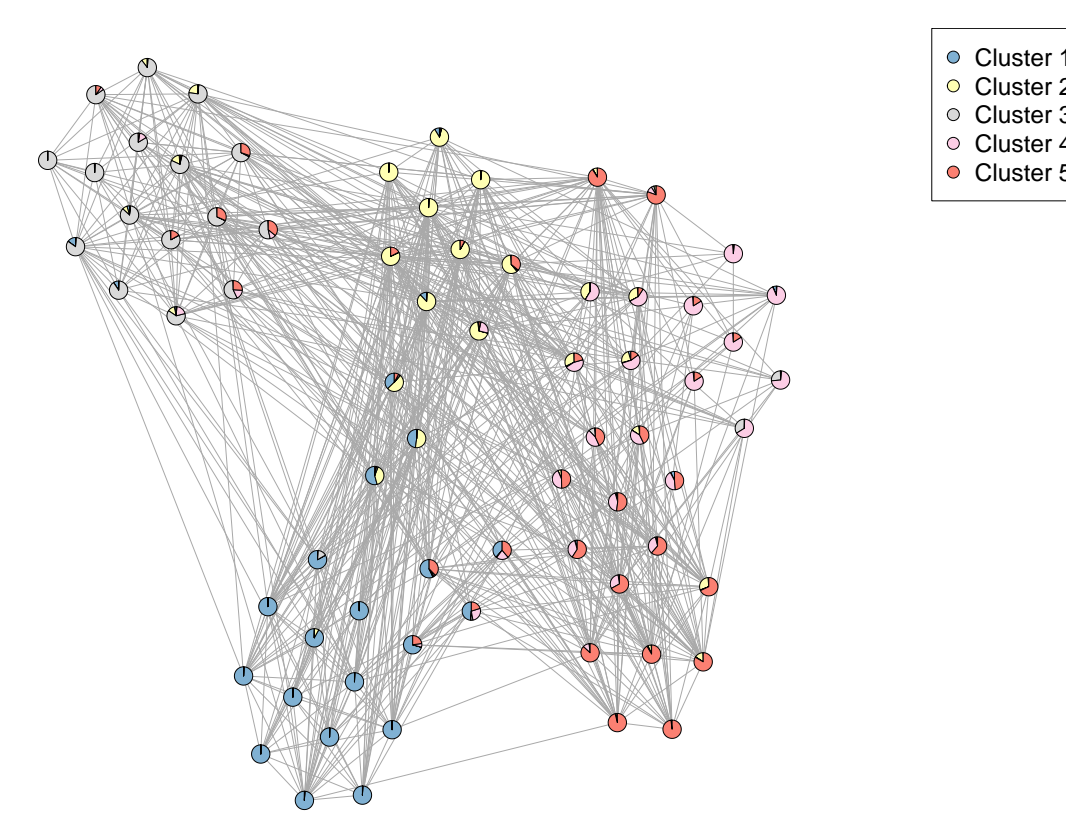


Figure 2: Networks of advice exchanges.

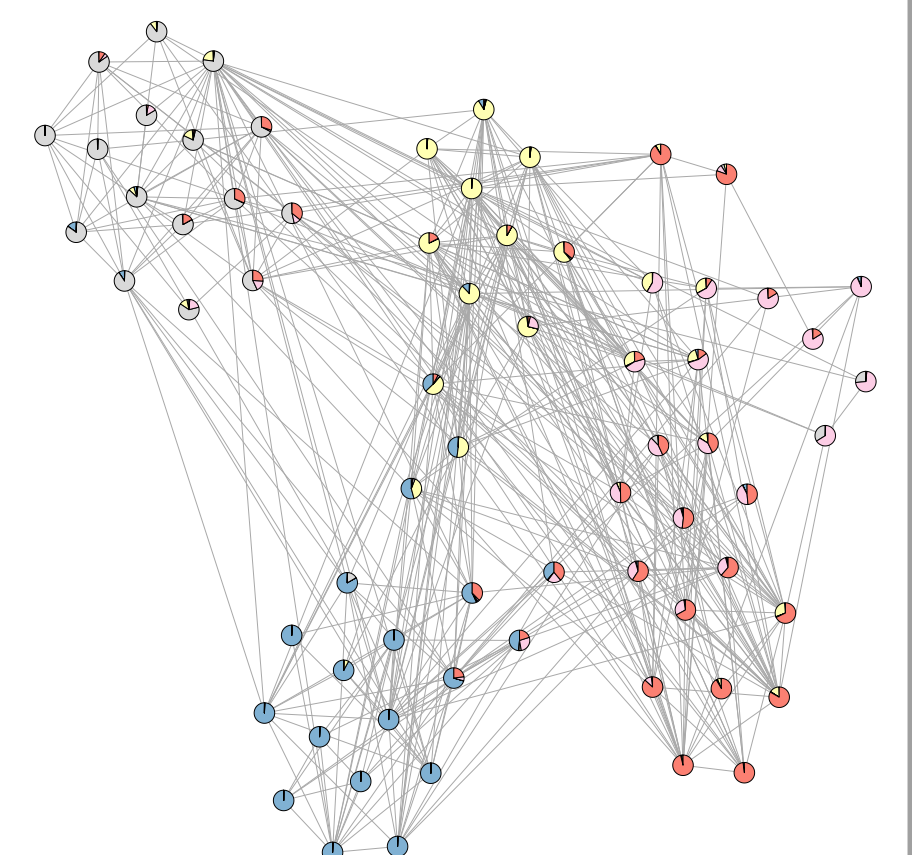


Figure 3: friendship Network.

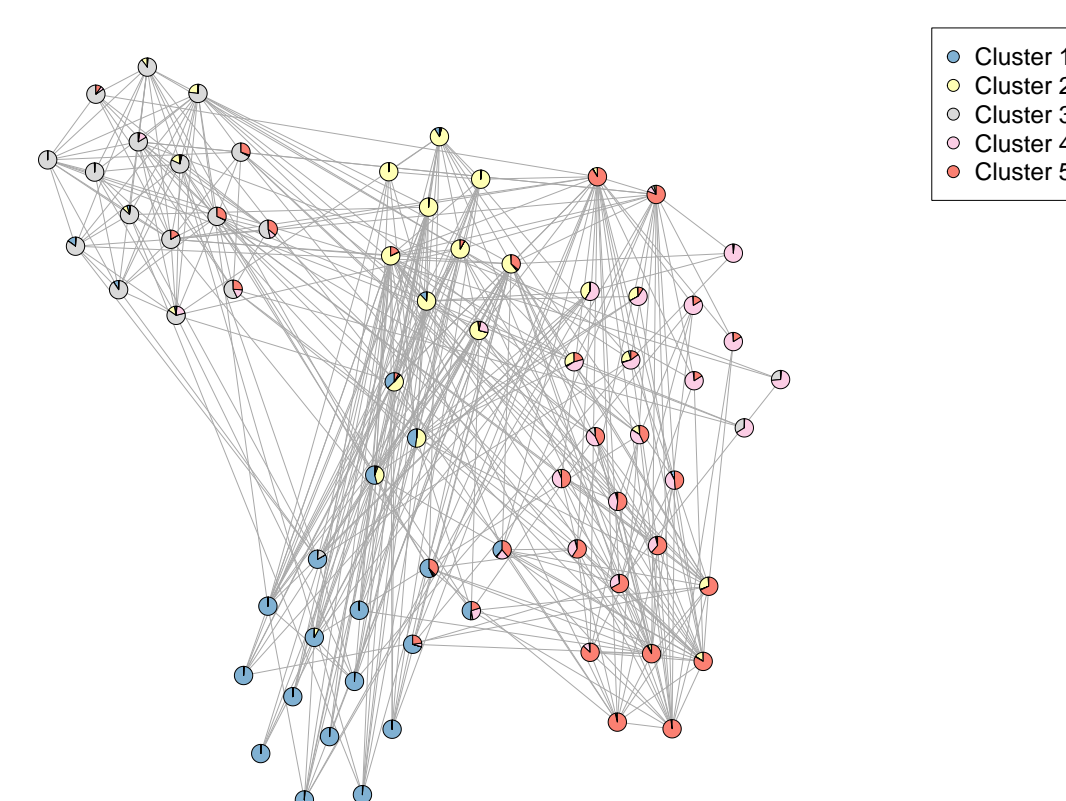


Figure 4: Co-work network.

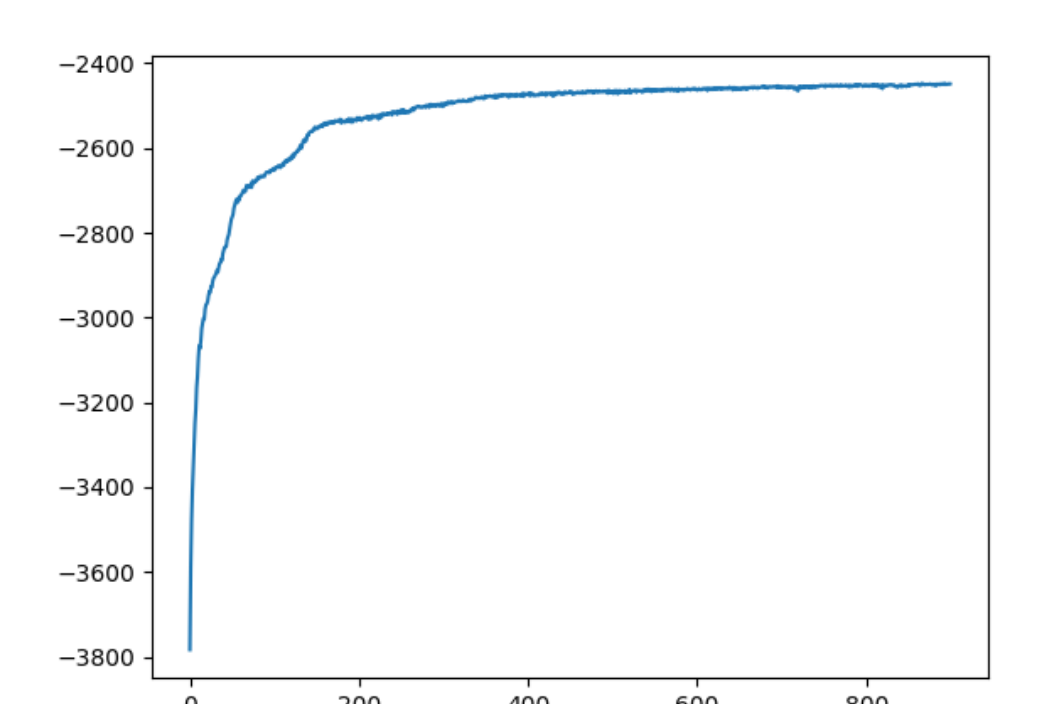


Figure 6: Evolution of the ELBO.

6. References

- [1] R. Boutin, P. LATOUCHE and C. BOUVEYRON (2024). Deep latent position block model for unsupervised node clustering.
- [2] T. Kipf and M. Welling (2017). Semi-supervised classification with graph convolutional networks