

# Stochastic block model for multilevel networks unravels structural interdependence between the social and economic networks in a TV program trade fair

Saint-Clair Chabert-Liddell<sup>†</sup>

Joint work with P. Barbillon<sup>†</sup>, S. Donnet<sup>†</sup> & E. Lazega<sup>\*</sup>

<sup>†</sup> UMR INRAe/AgroParisTech, MIA Paris  
Université Paris-Saclay

<sup>\*</sup> Institut d'Études Politiques de Paris

Sunbelt 2020

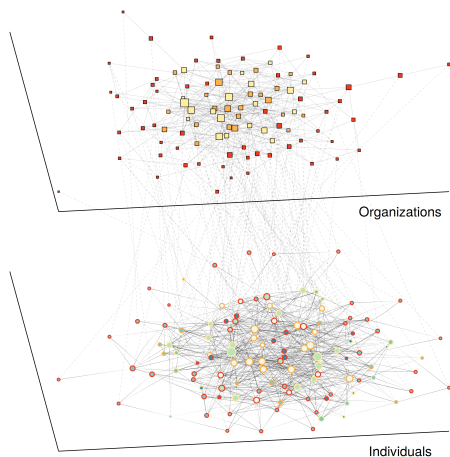
13 July 2020

# Outline

- 1 Modeling
- 2 Simulation Studies
- 3 Application to Television Program Trade Fair

# Motivation Dataset

Economic and social networks in a television trade fair <sup>1</sup>.



- Economic network: 109 organizations signing deals (undirected interactions)
- Represented on the trade fair by individuals
- Social network: 128 individuals sharing advice (directed interactions)

<sup>1</sup>Brailly, 2016

# Objective of this work

	$\overbrace{\hspace{1cm}}^{n_I}$		$\overbrace{\hspace{1cm}}^{n_O}$	
Individual 1	0	1	0	1
$\vdots$	$X_{ii'}^I$		$A_{ij}$	
Individual $n_I$	1	1	0	1
Organization 1			1	1
$\vdots$			$X_{jj'}^O$	
Organization $n_O$			0	1
	Individual 1	Individual $n_I$	Organization 1	Organization $n_O$

**Data :**

$X^I$  Interactions between individuals

$X^O$  Interactions between organizations

$A$  Affiliations of the individuals to the organizations

$A_{ij} = 1$  if  $i$  is affiliated to  $j$

Only one affiliation per individual

## Objectives

- Joint probabilistic model on  $\mathbf{X} = \{X^I, X^O\}$  given  $A$
- Evaluate the influence of the inter-organizational level on the inter-individual level

# Outline

- 1 Modeling
- 2 Simulation Studies
- 3 Application to Television Program Trade Fair

# Modeling of a Multilevel SBM



## Stochastic Block Model (SBM)<sup>a</sup>

<sup>a</sup>Snijders and Nowicki, 1997

- Mixture model for graphs
- Latent variables on nodes
- Model heterogeneity of connection

# Modeling of a Multilevel SBM

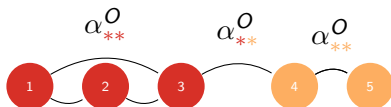


## Inter-organizational Level

- $n_O$  organizations into  $Q_O$  blocks
- Latent variables are independent
- $Z_j^O = l \Leftrightarrow j \in l, \quad l \in \{1, \dots, Q_O\}$

$$\mathbb{P}(Z_j^O = l) = \pi_l^O$$

# Modeling of a Multilevel SBM



## Inter-organizational Level

- $n_O$  organizations into  $Q_O$  blocks
- Latent variables are independent
- $Z_j^O = l \Leftrightarrow j \in l, \quad l \in \{1, \dots, Q_O\}$

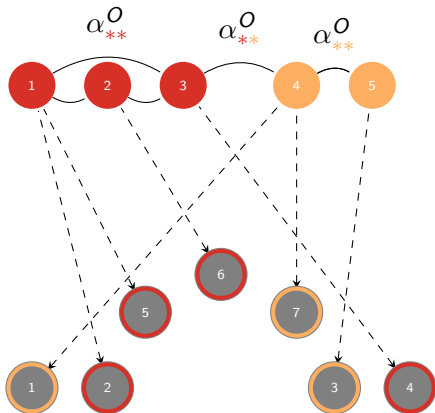
$$\mathbb{P}(Z_j^O = l) = \pi_l^O$$

- Connections are independent given the latent variables

$$\mathbb{P}(X_{jj'}^O = 1 | Z_j^O = l, Z_{j'}^O = l') = \alpha_{ll'}^O$$



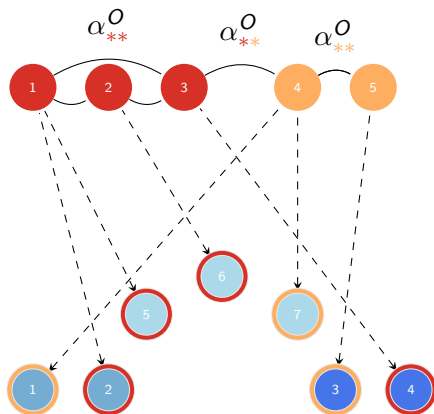
# Modeling of a Multilevel SBM



## Inter-individual Level

- $n_I$  individuals into  $Q_I$  blocks
- The block of an individual depends on the block of her/his organization

# Modeling of a Multilevel SBM

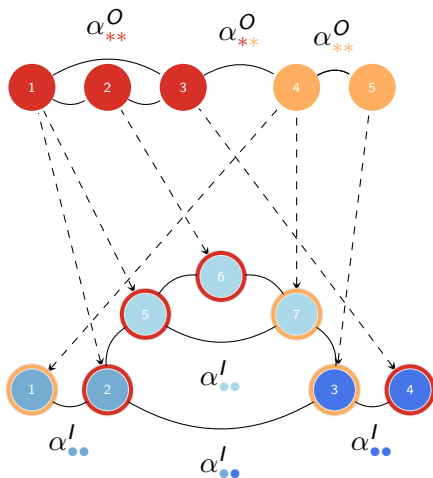


## Inter-individual Level

- $n_I$  individuals into  $Q_I$  blocks
- The block of an individual depends on the block of her/his organization
- $Z_i^I = k \Leftrightarrow i \in k, k \in \{1, \dots, Q_I\}$

$$\mathbb{P}(Z_i^I = k | A_i = j, Z_j^O = l) = \gamma_{kl}$$

# Modeling of a Multilevel SBM



## Inter-individual Level

- $n_I$  individuals into  $Q_I$  blocks
- The block of an individual depends on the block of her/his organization

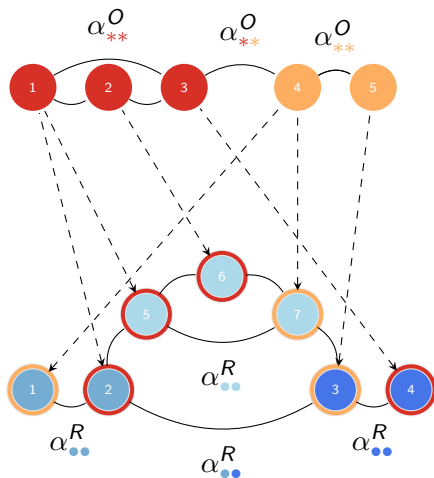
$$\bullet \quad Z_i^I = k \Leftrightarrow i \in k, k \in \{1, \dots, Q_I\}$$

$$\mathbb{P}(Z_i^I = k | A_i = j, Z_j^O = l) = \gamma_{kl}$$

- Connections are independent given the latent variables

$$\mathbb{P}(X_{ii'}^I = 1 | Z_i^I = k, Z_{i'}^O = k) = \alpha_{kk'}^I$$

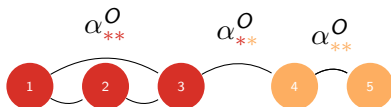
# Independence Between Levels



- $\pi^O$  is a probability vector
- Each column of  $\gamma$  as well
- If  $\gamma_{kl} = \gamma_{kl'} \quad \forall l, l'$

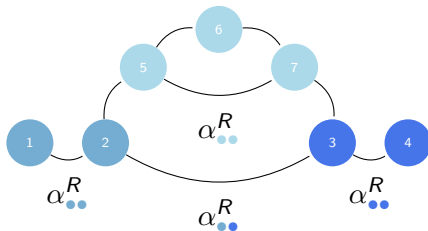
$$\mathcal{L}(X^I, X^O|A) = \mathcal{L}(X^I)\mathcal{L}(X^O)$$

# Independence Between Levels



- $\pi^O$  is a probability vector
- Each column of  $\gamma$  as well
- If  $\gamma_{kl} = \gamma_{kl'} \quad \forall l, l'$

$$\mathcal{L}(X^I, X^O|A) = \mathcal{L}(X^I)\mathcal{L}(X^O)$$



- Each level of the multilevel network is a SBM with  $\pi^I = \gamma_{\cdot 1}$
- Organizational structure has no influence on the connections of individuals

# Inference of the Multilevel SBM

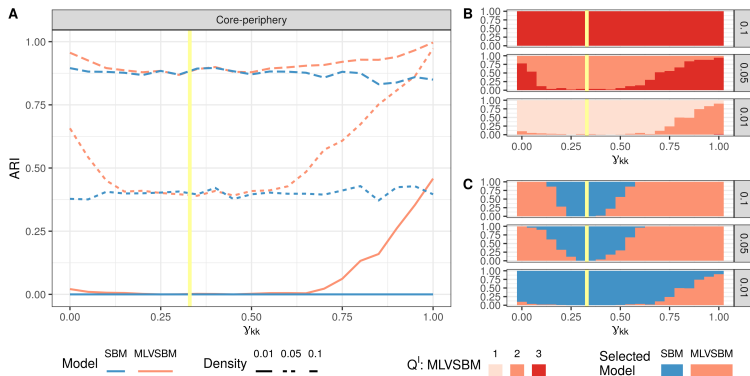
- Inference by maximum likelihood estimation
- Variational EM algorithm for a given number of blocks
- Step-wise procedure to navigate between models
- Model selection by a model based penalized criterion (ICL)
- ICL also used to state on the independence between the inter-individual and the inter-organizational levels

# Outline

- 1 Modeling
- 2 Simulation Studies**
- 3 Application to Television Program Trade Fair

# Simulation Studies

- Strong dependence between levels ( $\gamma_{kk}$  far from 1/3) helps recover the structure of the inter-individual level with the information of the inter-organizational level.
- ICL tends to select model of small size  $\Rightarrow$  Good for testing the interdependence.



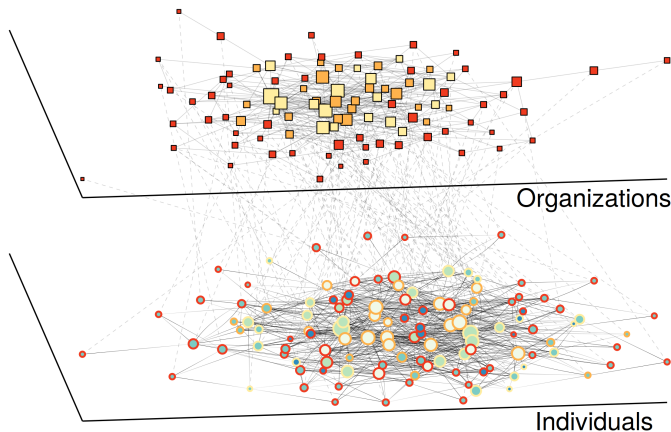


# Outline

- 1 Modeling
- 2 Simulation Studies
- 3 Application to Television Program Trade Fair**

# Application to a Television Program Trade Fair Dataset<sup>2</sup>

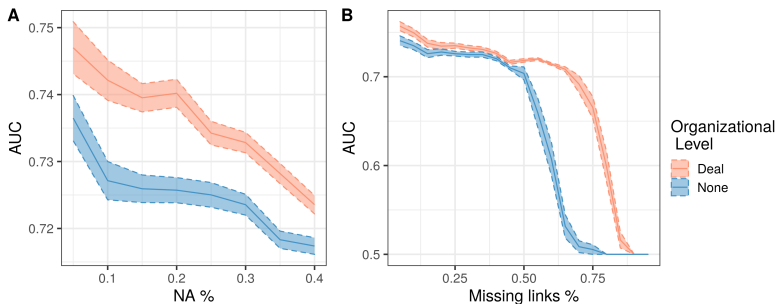
128 individuals (buyers and sellers) with directed interactions (advice) and 109 organizations with undirected interactions (deal).



<sup>2</sup>Brailly, 2016

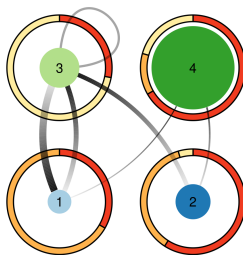
# Link Prediction

- The social network and the economic network are interdependent.
- Inter-organizational level helps predicting links on the inter-individual level.

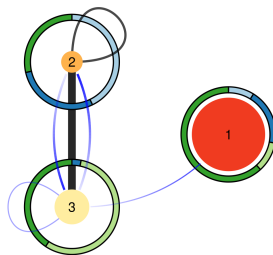


# Dataset analysis

**Individuals**



**Organizations**



- 4 blocks of individuals and 3 blocks of organizations
- Core-periphery structure for the inter-organizational level
- Mainly inter-block connections for individuals (except block 3, sub-group of sellers)
- Intra-block connection between individuals do not replicate the intra-block connections of their organizations (block 2 and 3)

- Preprint available on arXiv: <https://arxiv.org/abs/1910.10512>
- R package available at <https://chabert-liddell.github.io/MLVSBM/>
  - Simulation and inference of multilevel networks
  - Handling of missing data on  $X^I$  and  $X^O$
  - Prediction on missing dyads, missing links and spurious links
  - Extend to multi-affiliation datasets

Any question? [saint-clair.chabert-liddell@agroparistech.fr](mailto:saint-clair.chabert-liddell@agroparistech.fr)

**Thank you for your attention!**

# References

- Brailly, Julien (2016). “Dynamics of networks in trade fairs—A multilevel relational approach to the cooperation among competitors”. In: *Journal of Economic Geography* 16.6, pp. 1279–1301.
- Snijders, Tom A.B. and Krzysztof Nowicki (Jan. 1997). “Estimation and Prediction for Stochastic Blockmodels for Graphs with Latent Block Structure”. In: *Journal of Classification* 14.1, pp. 75–100. ISSN: 0176-4268. DOI: 10.1007/s003579900004. URL: <http://link.springer.com/10.1007/s003579900004>.