

# Joint analysis of bipartite networks collection

JdS 2025

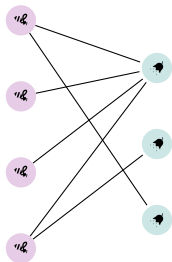
Louis Lacoste, Pierre Barbillon and Sophie Donnet

Laboratoire MIA Paris-Saclay



May 23, 2025

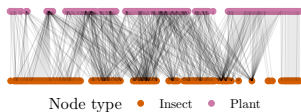
# Why a network?



$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

Associated  
bi-adjacency  
matrix

Figure 1: Example of a network



Node type ● Insect ● Plant

- Increasingly available
- Modeling of various interactions, here ecosystems
- Structure necessary for: biodiversity monitoring, robustness, risk of collapse

Figure 2: Plant-pollinator network from Bristol  
Baldock et al., 2019

# Analysis methods for a network

Several methods :

- Metrics at
  - ▶ node level: degree, centrality...
  - ▶ network level: density, nestedness...
- Node embedding and/or clustering with latent variable models  
Snijders and Nowicki, 1997; Hoff et al., 2002
- Node or network embedding with Graph Convolutional Networks  
Kipf and Welling, 2016

# Analysis methods for a network

Several methods :

- Metrics at
  - ▶ node level: degree, centrality. . .
  - ▶ network level: density, nestedness. . .
- **Node embedding and/or clustering with latent variable models**  
Snijders and Nowicki, 1997; Hoff et al., 2002
- Node or network embedding with Graph Convolutional Networks  
Kipf and Welling, 2016

# Latent Block Model (LBM<sup>1</sup>)

Govaert and Nadif, 2005.

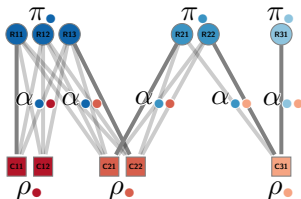


Figure 3: Example of LBM<sup>1</sup>

## Hierarchical model

$$\forall q \in \llbracket 1, Q_1 \rrbracket, \mathbb{P}(Z_i = q) = \pi_q$$

$$\forall r \in \llbracket 1, Q_2 \rrbracket, \mathbb{P}(W_j = r) = \rho_r$$

$$Y_{ij} | Z_i, W_j \sim \mathcal{F}(\alpha_{Z_i, W_j})$$

where

$$|\pi| = Q_1, |\rho| = Q_2, |\alpha| = Q_1 \times Q_2$$

## Concise LBM formula

$$Y \sim \mathcal{F}\text{-BiSBM}_{n_1, n_2}(Q_1, Q_2, \pi, \rho, \alpha)$$

<sup>1</sup>Which I will henceforth call BiSBM

# Latent Block Model (LBM<sup>1</sup>)

Govaert and Nadif, 2005.

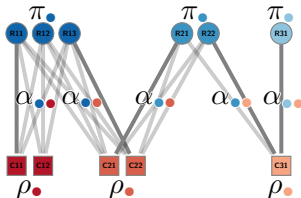


Figure 3: Example of LBM<sup>1</sup>

With

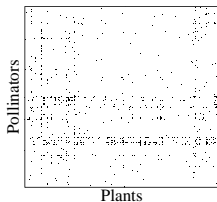
- $Q_1 = |\{\bullet, \bullet, \bullet\}|$  fixed row blocks
- $Q_2 = |\{\bullet, \bullet, \bullet\}|$  fixed column blocks

## Parameters

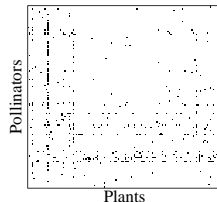
- $\pi_{\bullet} = \mathbb{P}(Z_i = \bullet)$
- $\rho_{\bullet} = \mathbb{P}(W_j = \bullet)$
- $\alpha_{\bullet\bullet} = \mathbb{P}(Y_{ij} = 1 | Z_i = \bullet, W_j = \bullet)$

<sup>1</sup>Which I will henceforth call BiSBM

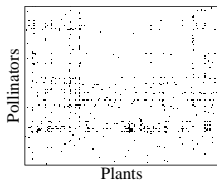
# Multiple networks



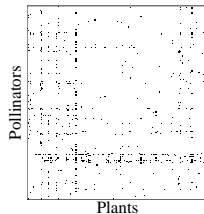
(a) Bristol



(b) Edinburgh



(c) Leeds



(d) Reading

Figure 4: Adjacency matrices, Baldock et al., 2019

# Multiple networks



Figure 5: Map of the four cities



## Model 0: sep-BiSBM

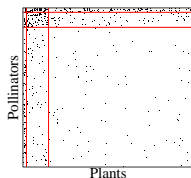
$$\forall m \in \{1 \dots M\}, Y^m \stackrel{ind}{\sim} \mathcal{F}\text{-BiSBM}_{n_1^m, n_2^m}(Q_1^m, Q_2^m, \pi^m, \rho^m, \alpha^m)$$

## Model 0: sep-BiSBM

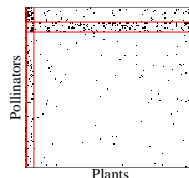
$$\forall m \in \{1 \dots M\}, Y^m \stackrel{ind}{\sim} \mathcal{F}\text{-BiSBM}_{n_1^m, n_2^m}(Q_1^{\textcolor{red}{m}}, Q_2^{\textcolor{red}{m}}, \pi^{\textcolor{red}{m}}, \rho^{\textcolor{red}{m}}, \alpha^{\textcolor{red}{m}})$$

## Model 0: sep-BiSBM

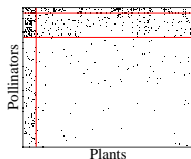
$$\forall m \in \{1 \dots M\}, Y^m \stackrel{\text{ind}}{\sim} \mathcal{F}\text{-BiSBM}_{n_1^m, n_2^m}(Q_1^m, Q_2^m, \pi^m, \rho^m, \alpha^m)$$



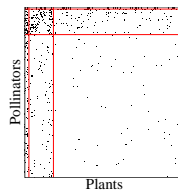
(a) Bristol



(b) Edinburgh



(c) Leeds



(d) Reading

Figure 6: Reordered adjacency matrices, using BiSBM for each network

# Several joint models

## *iid*-colBiSBM

$$\forall m \in \{1 \dots M\}, Y^m \stackrel{iid}{\sim} \mathcal{F}\text{-BiSBM}_{n_1^m, n_2^m}(Q_1, Q_2, \pi, \rho, \alpha)$$

with  $\theta = (\pi, \rho, \alpha)$ .

## Several joint models

### *iid*-colBiSBM

$$\forall m \in \{1 \dots M\}, Y^m \stackrel{iid}{\sim} \mathcal{F}\text{-BiSBM}_{n_1^m, n_2^m}(Q_1, Q_2, \pi, \rho, \alpha)$$

with  $\theta = (\pi, \rho, \alpha)$ .

### $\pi\rho$ -colBiSBM

$$\forall m \in \{1 \dots M\}, Y^m \stackrel{ind}{\sim} \mathcal{F}\text{-BiSBM}_{n_1^m, n_2^m}(Q_1, Q_2, \pi^m, \rho^m, \alpha)$$

with  $\theta = ((\pi^{\textcolor{red}{m}})_{m=1, \dots, M}, (\rho^{\textcolor{red}{m}})_{m=1, \dots, M}, \alpha)$ .

And intermediate models freeing  $\pi$  or  $\rho$ .

# Parameter estimation

By *Variational EM*, as proposed by Daudin et al., 2008; Chabert-Liddell et al., 2024.

Variational approximation of  $\mathbf{Z}, \mathbf{W} | \mathbf{Y}, \theta^{(t-1)}$

$\mathcal{R}_{Y^m, \tau}(\mathbf{Z}^m, \mathbf{W}^m) = \mathcal{R}_{Y^m, \tau}^1(\mathbf{Z}^m) \times \mathcal{R}_{Y^m, \tau}^2(\mathbf{W}^m) \Rightarrow$  independence rows, columns.

$$\ell(\mathbf{Y}; \theta) \geq \sum_{m=1}^M \left( \mathcal{Q}^m(\theta \mid \theta^{(t)}) + \mathcal{H}(\mathcal{R}_{Y^m, \theta^{(t)}}(\mathbf{Z}^m, \mathbf{W}^m)) \right) =: \mathcal{J}(\tau; \theta)$$

where  $\mathcal{Q}^m(\theta \mid \theta^{(t)}) = \mathbb{E}_{\mathbf{Z}^m, \mathbf{W}^m \sim \mathcal{R}_{Y^m, \tau}(\cdot)} [\ell_c(Y^m, \mathbf{Z}^m, \mathbf{W}^m | \theta)]$

## Problem of choosing $(Q_1, Q_2)$

Need to select  $Q_1$  and  $Q_2$ . BIC-Like criterion<sup>2</sup>

$$\begin{aligned}\text{BIC-L}(\mathbf{Y}, Q_1, Q_2) &= \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \hat{\tau}}}[\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \hat{\theta}^{\text{var}})] + \mathcal{H}(\mathcal{R}_{\mathbf{Y}, \hat{\tau}}) - \frac{1}{2}\text{pen}(Q_1, Q_2) \\ &= \mathcal{J}(\mathcal{R}_{\mathbf{Y}, \hat{\tau}}, \hat{\theta}^{\text{var}}) - \frac{1}{2}\text{pen}(Q_1, Q_2)\end{aligned}$$

### Exploration problems

- Exploration of a 2D grid is costly.
- Sensitivity to initializations.

---

<sup>2</sup>ICL + entropy - penalty

## Problem of choosing $(Q_1, Q_2)$

Need to select  $Q_1$  and  $Q_2$ . BIC-Like criterion<sup>2</sup>

$$\begin{aligned}\text{BIC-L}(\mathbf{Y}, Q_1, Q_2) &= \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \hat{\tau}}}[\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \hat{\theta}^{\text{var}})] + \mathcal{H}(\mathcal{R}_{\mathbf{Y}, \hat{\tau}}) - \frac{1}{2}\text{pen}(Q_1, Q_2) \\ &= \mathcal{J}(\mathcal{R}_{\mathbf{Y}, \hat{\tau}}, \hat{\theta}^{\text{var}}) - \frac{1}{2}\text{pen}(Q_1, Q_2)\end{aligned}$$

### Exploration problems

- Exploration of a 2D grid is costly. → **Greedy approach** and **sliding window**
- Sensitivity to initializations.

---

<sup>2</sup>ICL + entropy - penalty



## Problem of choosing $(Q_1, Q_2)$

Need to select  $Q_1$  and  $Q_2$ . BIC-Like criterion<sup>2</sup>

$$\begin{aligned}\text{BIC-L}(\mathbf{Y}, Q_1, Q_2) &= \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \hat{\tau}}}[\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \hat{\theta}^{\text{var}})] + \mathcal{H}(\mathcal{R}_{\mathbf{Y}, \hat{\tau}}) - \frac{1}{2}\text{pen}(Q_1, Q_2) \\ &= \mathcal{J}(\mathcal{R}_{\mathbf{Y}, \hat{\tau}}, \hat{\theta}^{\text{var}}) - \frac{1}{2}\text{pen}(Q_1, Q_2)\end{aligned}$$

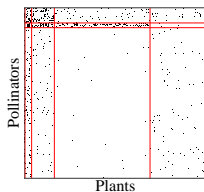
### Exploration problems

- Exploration of a 2D grid is costly. → **Greedy approach** and **sliding window**
- Sensitivity to initializations. → **Spectral clustering** and **reuse of previous inits**

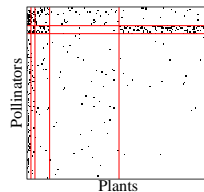
---

<sup>2</sup>ICL + entropy - penalty

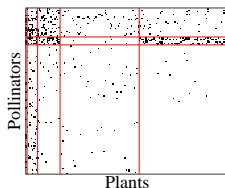
# Results Baldock et al., 2019



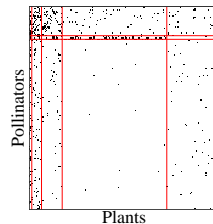
(a) Bristol



(b) Edinburgh



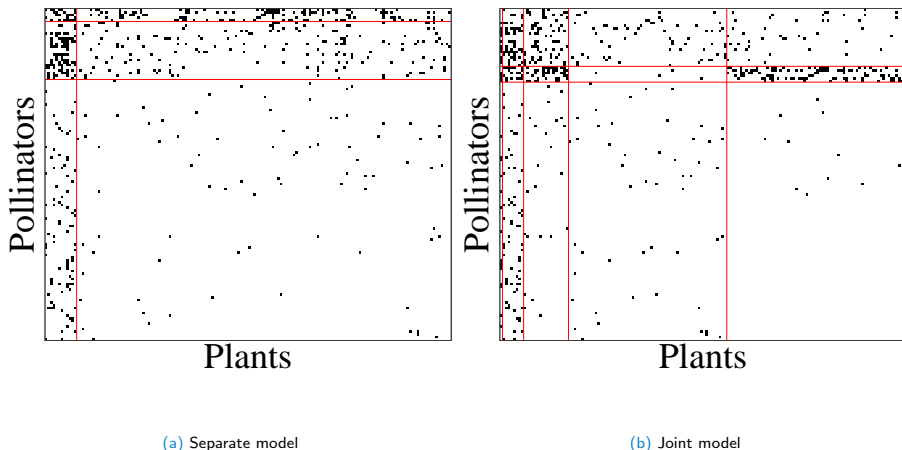
(c) Leeds



(d) Reading

Figure 7: Reordered adjacency matrices by *iid*-colBiSBM, Baldock et al., 2019

# Results Baldock et al., 2019 focus on Leeds



**Figure 8:** Reordered adjacency matrix by sep-BiSBM (left) and by *iid*-colBiSBM (right), Baldock et al., 2019

# Bombus



(a) *Bombus hortorum* or garden bumblebee



(b) *Bombus lapidarius* or red-tailed bumblebee

# Bombus

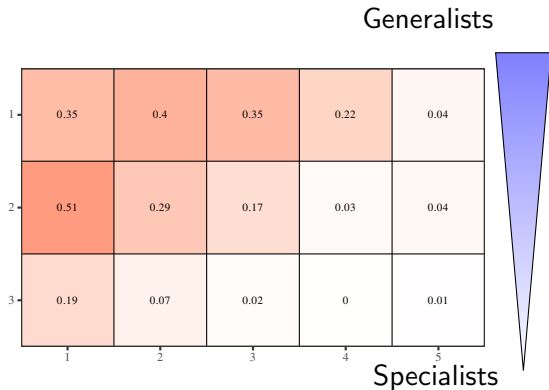


Figure 10: Shared structure of the 4 networks

# Bombus



(a) *Bombus hortorum* or garden bumblebee

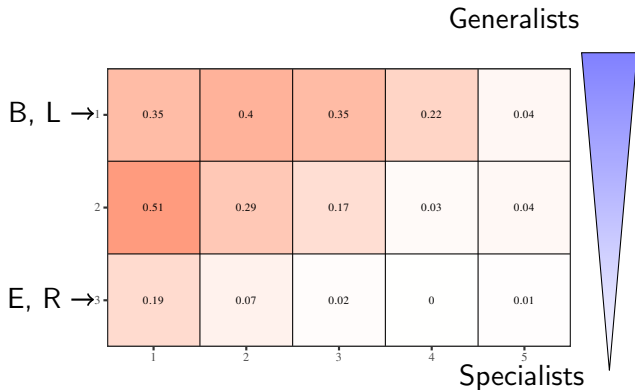


Figure 10: Shared structure of the 4 networks

# Bombus



(b) *Bombus Lapidarius* or red-tailed bumblebee

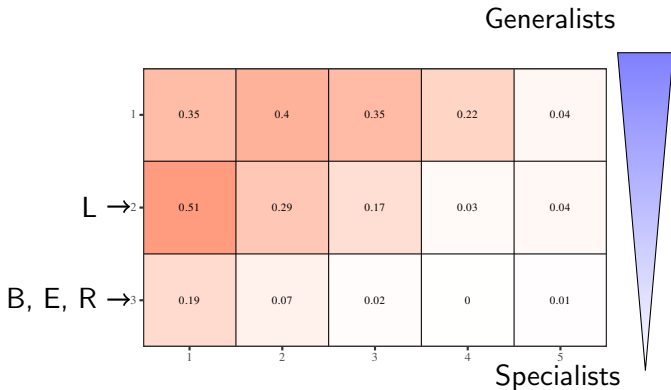


Figure 10: Shared structure of the 4 networks

# Network clustering

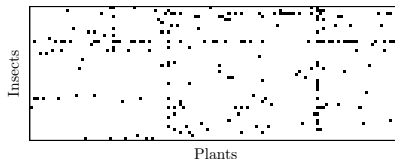


Figure 11: Adjacency matrix, Baldock et al., 2011



# Application to Baldock et al., 2019, 2011 I

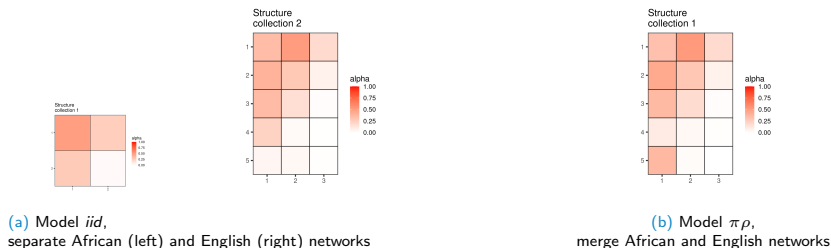
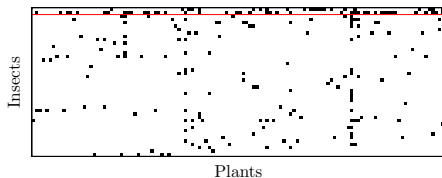
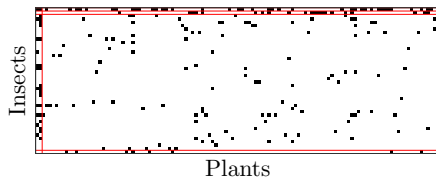


Figure 12: Structures detected for networks of Baldock et al., 2019, 2011

# Results



(a) Reordered by LBM



(b) Reordered by  $\pi\rho$ -colBiSBM

Figure 13: Reordered adjacency matrix by  $\pi\rho$ -colBiSBM, Baldock et al., 2011

# Conclusion and perspectives

## Capabilities

- 4 models including 3 with flexibility on at least one of the dimensions (adaptability to data).
- Detect classic and less classic structures in an agnostic way.
- Partition a set of networks according to their structures.

# Perspectives

## Future work

- Multi-layer networks (account for sampling bias, presence/absence)
- Graph Convolutional Network to allow for scalability

## Package and applications

- CRAN submission
- Integrate the possibility of an additional criterion for clustering (e.g. urbanization gradient [Fisogni et al., 2022](#))
- Apply clustering to data from [Pichon et al., 2024](#); [Doré et al., 2021](#)

# References I

- Baldock, K. C. R., Goddard, M. A., Hicks, D. M., Kunin, W. E., Mitschunas, N., Morse, H., Osgathorpe, L. M., Potts, S. G., Robertson, K. M., Scott, A. V., Staniczenko, P. P. A., Stone, G. N., Vaughan, I. P., & Memmott, J. (2019). A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nature Ecology & Evolution*, 3(3), 363–373.  
<https://doi.org/10.1038/s41559-018-0769-y>
- Snijders, T. A., & Nowicki, K. (1997). Estimation and Prediction for Stochastic Blockmodels for Graphs with Latent Block Structure. *Journal of Classification*, 14(1), 75–100.  
<https://doi.org/10.1007/s003579900004>
- Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent Space Approaches to Social Network Analysis. *Journal of the American Statistical Association*, 97(460), 1090–1098.  
<https://doi.org/10.1198/016214502388618906>

# References II

- Kipf, T. N., & Welling, M. (2016, November 21). *Variational Graph Auto-Encoders*. arXiv: 1611.07308 [stat].  
<https://doi.org/10.48550/arXiv.1611.07308>  
Read\_Status: New  
Read\_Status\_Date: 2025-05-09T11:54:37.094Z.
- Govaert, G., & Nadif, M. (2005). An EM algorithm for the block mixture model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4), 643–647.  
<https://doi.org/10.1109/TPAMI.2005.69>
- Daudin, J.-J., Picard, F., & Robin, S. (2008). A mixture model for random graphs. *Statistics and Computing*, 18(2), 173–183.  
<https://doi.org/10.1007/s11222-007-9046-7>
- Chabert-Liddell, S.-C., Barbillon, P., & Donnet, S. (2024). Learning common structures in a collection of networks. An application to food webs. *The Annals of Applied Statistics*, 18(2), 1213–1235.  
<https://doi.org/10.1214/23-AOAS1831>

## References III

- Baldock, K. C. R., Memmott, J., Ruiz-Guajardo, J. C., Roze, D., & Stone, G. N. (2011). Daily temporal structure in African savanna flower visitation networks and consequences for network sampling. *Ecology*, 92(3), 687–698. <https://doi.org/10.1890/10-1110.1>
- Fisogni, A., Hautekèete, N., Piquot, Y., Brun, M., Vanappelghem, C., Ohlmann, M., Franchomme, M., Hinnewinkel, C., & Massol, F. (2022). Seasonal trajectories of plant-pollinator interaction networks differ following phenological mismatches along an urbanization gradient. *Landscape and Urban Planning*, 226, 104512. <https://doi.org/10.1016/j.landurbplan.2022.104512>  
Read\_Status: New  
Read\_Status\_Date: 2025-05-14T20:18:00.025Z.
- Pichon, B., Le Goff, R., Morlon, H., & Perez-Lamarque, B. (2024). Telling mutualistic and antagonistic ecological networks apart by learning their multiscale structure. *Methods in Ecology and Evolution*, 15(6), 1113–1128. <https://doi.org/10.1111/2041-210X.14328>

## References IV

Doré, M., Fontaine, C., & Thébault, E. (2021). Relative effects of anthropogenic pressures, climate, and sampling design on the structure of pollination networks at the global scale. *Global Change Biology*, 27(6), 1266–1280. <https://doi.org/10.1111/gcb.15474>



## Developed formula of variational EM

$$\begin{aligned} \ell(\mathbf{Y}; \theta) \geq & \sum_{m=1}^M \left( \sum_{i=1}^{n_1^m} \sum_{j=1}^{n_2^m} \sum_{q \in Q_{1,m}} \sum_{r \in Q_{2,m}} \tau_{i,q}^{1,m} \tau_{j,r}^{2,m} \log f(Y_{ij}^m; \alpha_{qr}) \right. \\ & + \sum_{i=1}^{n_1^m} \sum_{q \in Q_{1,m}} \tau_{i,q}^{1,m} \log \pi_q^m + \sum_{j=1}^{n_2^m} \sum_{r \in Q_{2,m}} \tau_{j,r}^{2,m} \log \rho_r^m \\ & \left. - \sum_{i=1}^{n_1} \tau_{i,q}^{1,m} \log \tau_{i,q}^{1,m} - \sum_{j=1}^{n_2} \tau_{j,r}^{2,m} \log \tau_{j,r}^{2,m} \right) =: \mathcal{J}(\tau; \theta), \end{aligned}$$

### Variational approximation

$$\tau_{iq}^{1,m} = \mathcal{R}_{Y^m, \tau}^1(Z_{iq}^m = 1) \text{ and } \tau_{jr}^{2,m} = \mathcal{R}_{Y^m, \tau}^2(W_{jr}^m = 1)$$

# Variational Expectation Step

$$\hat{\tau}^{(t+1)} = \arg \max_{\tau} \mathcal{J}(\tau, \hat{\theta}^{(t)}) \Leftrightarrow \arg \min_{\tau \in \mathcal{T}} \mathbf{KL}[\mathcal{R}_{\mathbf{Y}, \tau}, \mathbb{P}(.|\mathbf{Y})]$$

$$\begin{cases} \hat{\tau}_{iq}^{1,m} \propto \hat{\pi}_q^{m(t)} \prod_{j=1}^{n_2^m} \prod_{r \in \mathcal{Q}_2^m} f(Y_{ij}^m; \hat{\alpha}_{qr}^{(t)})^{\hat{\tau}_{jr}^{2,m(t+1)}} & \forall i = 1, \dots, n_1^m, q \in \mathcal{Q}_1^m \\ \hat{\tau}_{jr}^{2,m} \propto \hat{\rho}_r^{m(t)} \prod_{i=1}^{n_1^m} \prod_{q \in \mathcal{Q}_1^m} f(Y_{ij}^m; \hat{\alpha}_{qr}^{(t)})^{\hat{\tau}_{iq}^{1,m(t+1)}} & \forall j = 1, \dots, n_2^m, r \in \mathcal{Q}_2^m \end{cases}$$

<sup>2</sup>Initialization of  $\hat{\tau}$  with a *spectral clustering* on the networks.

## Maximization Step

$$\hat{\theta}^{(t+1)} = \arg \max_{\theta} \mathcal{J}(\hat{\tau}^{(t+1)}, \theta)$$

### Connectivity parameters

$$\hat{\alpha}_{qr} = \frac{\sum_{m=1}^M \sum_{i=1}^{n_1^m} \sum_{j=1}^{n_2^m} \tau_{iq}^{1,m} \tau_{jr}^{2,m} Y_{ij}^m}{\sum_{m=1}^M \sum_{i=1}^{n_1^m} \sum_{j=1}^{n_2^m} \tau_{iq}^{1,m} \tau_{jr}^{2,m}}$$

### Proportions for *iid*

$$\hat{\pi}_q = \frac{\sum_{m=1}^M \sum_{i=1}^{n_1^m} \tau_{iq}^{1,m}}{\sum_{m=1}^M n_1^m}$$

$$\hat{\rho}_r = \frac{\sum_{m=1}^M \sum_{j=1}^{n_2^m} \tau_{jr}^{2,m}}{\sum_{m=1}^M n_2^m}$$

## Maximization Step

$$\hat{\theta}^{(t+1)} = \arg \max_{\theta} \mathcal{J}(\hat{\tau}^{(t+1)}, \theta)$$

### Connectivity parameters

$$\hat{\alpha}_{qr} = \frac{\sum_{m=1}^M \sum_{i=1}^{n_1^m} \sum_{j=1}^{n_2^m} \tau_{iq}^{1,m} \tau_{jr}^{2,m} \mathbf{Y}_{ij}^m}{\sum_{m=1}^M \sum_{i=1}^{n_1^m} \sum_{j=1}^{n_2^m} \tau_{iq}^{1,m} \tau_{jr}^{2,m}}$$

### Proportions for $\pi\rho$

$$\hat{\pi}^m_q = \frac{\sum_{i=1}^{n_1^m} \tau_{iq}^{1,m}}{n_1^m} \quad \hat{\rho}^m_r = \frac{\sum_{j=1}^{n_2^m} \tau_{jr}^{2,m}}{n_2^m}$$

## Why does VE minimizes KL ?

$$\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \theta) = \log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta) + \ell(\mathbf{Y}; \theta)$$

$$\Leftrightarrow \ell(\mathbf{Y}; \theta) = \ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \theta) - \log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)$$

$$\Leftrightarrow \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\ell(\mathbf{Y}; \theta)] = \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \theta)] - \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)]$$

$$\Leftrightarrow \ell(\mathbf{Y}; \theta) = \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \theta)] - \mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)]$$

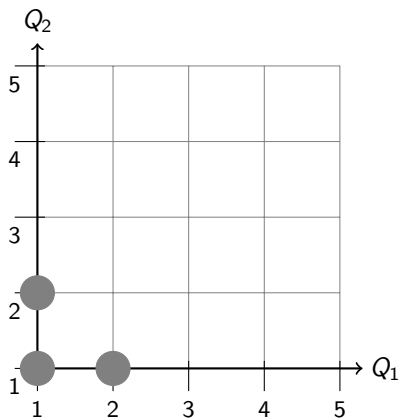
$$\text{But } \mathbf{KL}[\mathcal{R}_{\mathbf{Y}, \tau}, \log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)] = -\mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}\left[\log \frac{\mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)}{\mathcal{R}_{\mathbf{Y}, \tau}}\right]$$

$$= -\mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)] + \underbrace{\mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\log \mathcal{R}_{\mathbf{Y}, \tau}]}_{-\mathcal{H}(\mathcal{R}_{\mathbf{Y}, \tau})}$$

$$\Leftrightarrow \mathbf{KL}[\mathcal{R}_{\mathbf{Y}, \tau}, \log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)] + \mathcal{H}(\mathcal{R}_{\mathbf{Y}, \tau}) = -\mathbb{E}_{\mathcal{R}_{\mathbf{Y}, \tau}}[\log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)]$$

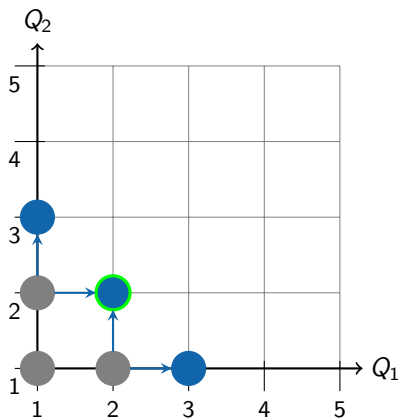
$$\text{Thus } \ell(\mathbf{Y}; \theta) - \mathbf{KL}[\mathcal{R}_{\mathbf{Y}, \tau}, \log \mathbb{P}(\mathbf{Z}, \mathbf{W} | \mathbf{Y}; \theta)] = \mathcal{J}(\tau; \theta) \quad \square$$




# Choice of $(Q_1, Q_2)$ - Greedy approach



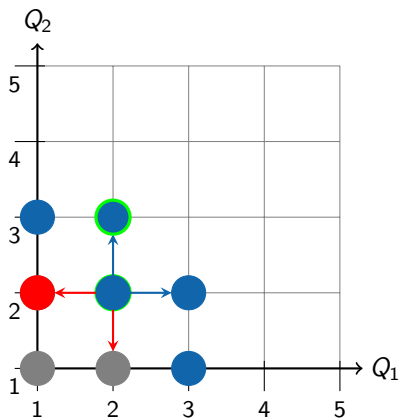
• Initial model :





# Choice of $(Q_1, Q_2)$ - Greedy approach



- Initial model : 
- Model after *split* : 
- Model maximizing the criterion : 

## Choice of $(Q_1, Q_2)$ - Greedy approach



- Initial model :  

- Model after *split* :  

- Model maximizing the criterion :  

- Model after *merge* :  




# Choice of $(Q_1, Q_2)$ - Sliding window

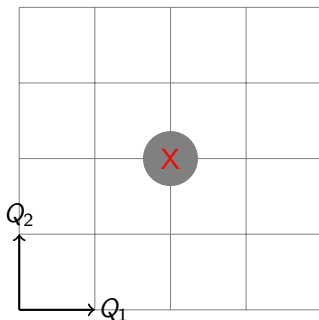


Figure 14: Sliding window

# Choice of $(Q_1, Q_2)$ - Sliding window

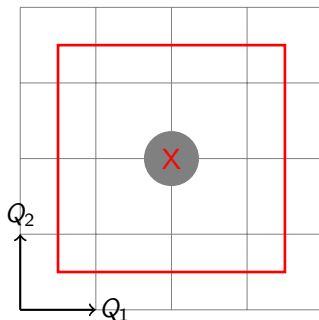


Figure 14: Sliding window

## Choice of $(Q_1, Q_2)$ - Sliding window

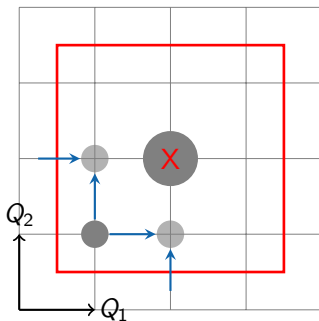


Figure 14: Sliding window

Initialization of the model if necessary

# Choice of $(Q_1, Q_2)$ - Sliding window

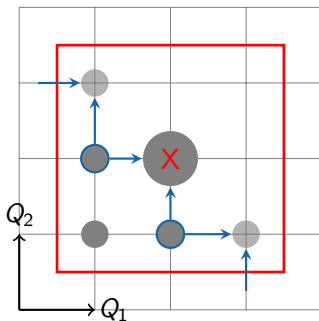


Figure 14: Sliding window

## Choice of $(Q_1, Q_2)$ - Sliding window

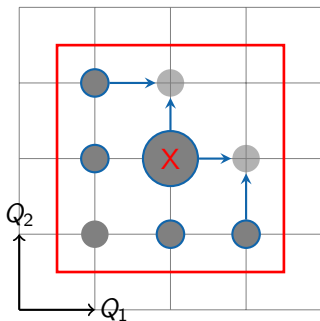


Figure 14: Sliding window

## Choice of $(Q_1, Q_2)$ - Sliding window

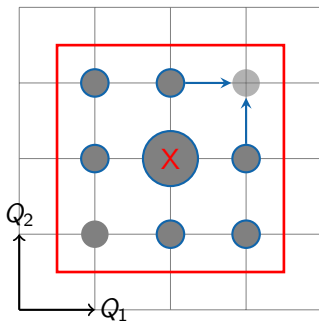


Figure 14: Sliding window

# Choice of $(Q_1, Q_2)$ - Sliding window

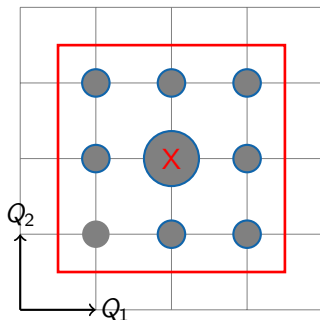


Figure 14: Sliding window

## Choice of $(Q_1, Q_2)$ - Sliding window

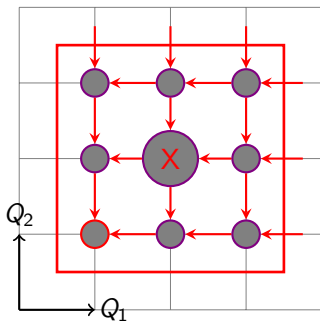


Figure 14: Sliding window



## Choice of $(Q_1, Q_2)$ - Sliding window

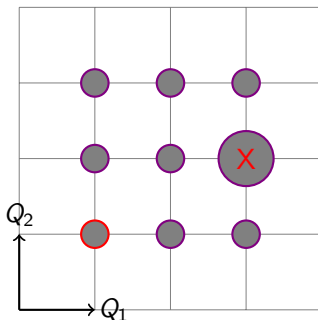


Figure 14: Sliding window

Localization of the new mode

## Choice of $(Q_1, Q_2)$ - Sliding window

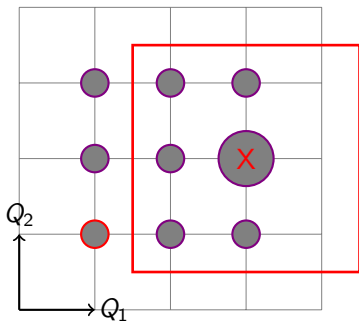
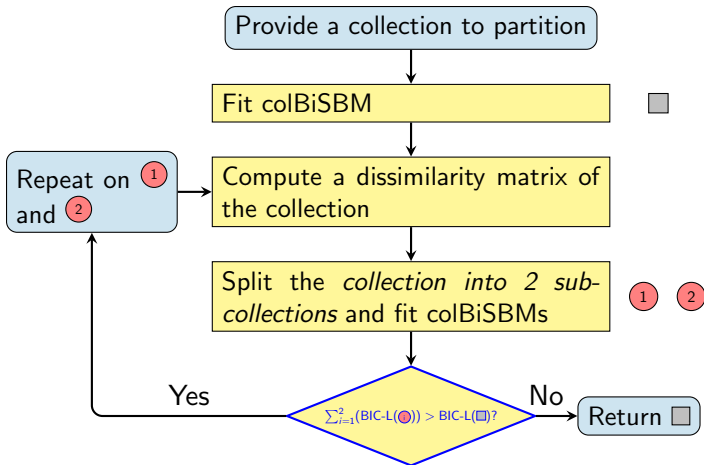


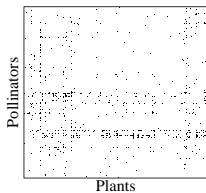
Figure 14: Sliding window

Move to the new mode then iterate

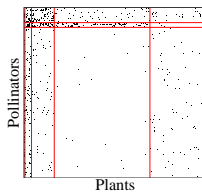
# Clustering algorithm



$$D_{\mathcal{M}}(m, m') = \sum_{q=1}^{Q_1} \sum_{r=1}^{Q_2} \max(\tilde{\pi}_q^m, \tilde{\pi}_q^{m'}) \left( \tilde{\alpha}_{qr}^m - \tilde{\alpha}_{qr}^{m'} \right)^2 \max(\tilde{\rho}_r^m, \tilde{\rho}_r^{m'})$$

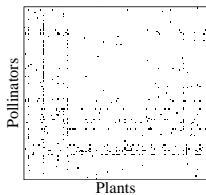


(a) Donnée

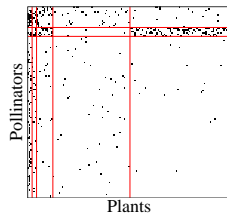


(b) Reordered

Figure 15: Bristol

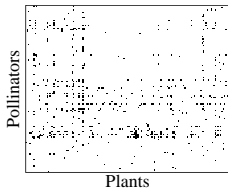


(a) Donnée

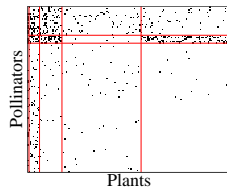


(b) Reordered

Figure 16: Edinburgh

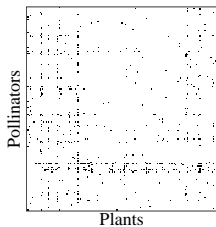


(a) Donnée

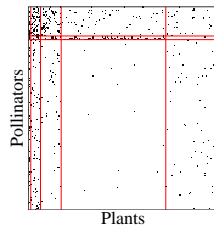


(b) Réordonnée

Figure 17: Leeds



(a) Donnée



(b) Réordonnée

Figure 18: Reading

## Appendices references I