Joint analysis of bipartite networks collection

JdS 2025

Louis Lacoste, Pierre Barbillon and Sophie Donnet
Laboratoire MIA Paris-Saclay

May 16, 2025

Why a network?

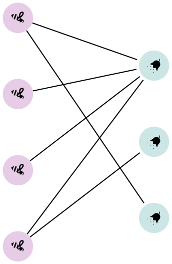


Figure 1: Example of a network

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

Associated adjacency matrix

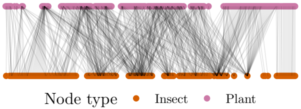


Figure 2: Plant-pollinator network from Bristol Baldock et al., 2019

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

Analysis methods for a network

TODO (Supprimable) Several methods :

- Metrics : degree, centrality, nesting . . .
- Network embedding with GNN
- *Clustering* of nodes with latent variable models

Analysis methods for a network

TODO (Supprimable) Several methods :

- Metrics : degree, centrality, nesting . . .
- Network embedding with GNN
- **Clustering of nodes with latent variable models**

Latent Block Model (LBM[1])

Govaert and Nadif, 2005.

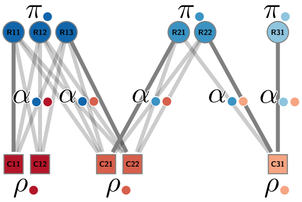


Figure 3: Example of LBM[1]

Hierarchical model

$$\forall q \in [\![1, Q_1]\!],\ \mathbb{P}(Z_i = q) = \pi_q$$
$$\forall r \in [\![1, Q_2]\!],\ \mathbb{P}(W_j = r) = \rho_r$$
$$Y_{ij} | Z_i, W_j \sim \mathcal{F}(\alpha_{Z_i, W_j})$$

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

Concise LBM formula

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

[1]Which I will henceforth call BiSBM

Latent Block Model (LBM[1])

Govaert and Nadif, 2005.

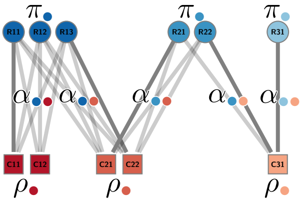


Figure 3: Example of LBM[1]

With

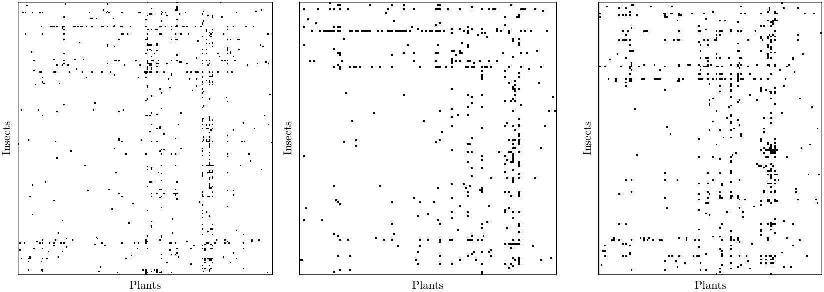
- $Q_1 = |\{\bullet, \bullet, \bullet\}|$ fixed row blocks
- $Q_2 = |\{\bullet, \bullet, \bullet\}|$ fixed column blocks

Parameters

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

[1]Which I will henceforth call BiSBM

Multiple networks



(a) Bristol (b) Edinburgh (c) Leeds

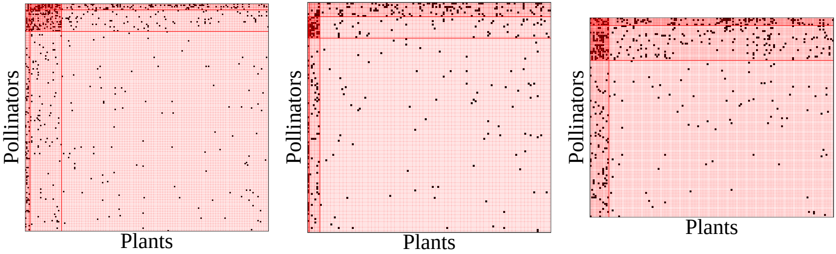
Figure 4: Adjacency matrices, Baldock et al., 2019

Bipartite collections different BiSBM

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

Bipartite collections different BiSBM

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



(a) Bristol (b) Edinburgh (c) Leeds

Figure 5: Reordered adjacency matrices, thanks to LBM

Several joint models

iid-colBiSBM

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

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

Several joint models

iid-colBiSBM

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

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

$\pi\rho$-colBiSBM

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

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

And intermediate models freeing π or ρ.

Parameter estimation

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

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

$\mathcal{R}_{Y^m,\tau}(\mathbf{Z}^m, \mathbf{W}^m) = \mathcal{R}^1_{Y^m,\tau}(\mathbf{Z}^m) \times \mathcal{R}^2_{Y^m,\tau}(\mathbf{W}^m) \Rightarrow$ 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 $\mathcal{Q}^m(\theta \mid \theta^{(t)}) = \mathbb{E}_{\mathbf{Z}^m, \mathbf{W}^m \sim \mathcal{R}_{Y^m,\tau}(.)} \left[\ell_c(Y^m, \mathbf{Z}^m, \mathbf{W}^m | \theta) \right]$

Problem of choosing (Q_1, Q_2)

Need to select Q_1 and Q_2. BIC-Like criterion[2]

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

Exploration problems

- Exploration of $\mathbb{N}^2$ costly.
- Sensitivity to initializations.

[2] ICL + Entropy + penalty

Problem of choosing (Q_1, Q_2)

Need to select Q_1 and Q_2. BIC-Like criterion[2]

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

Exploration problems

- Exploration of $\mathbb{N}^2$ costly. $\rightarrow$ **Greedy approach** and **sliding window**
- Sensitivity to initializations.

[2] ICL + Entropy + penalty

Problem of choosing (Q_1, Q_2)

Need to select Q_1 and Q_2. BIC-Like criterion[2]

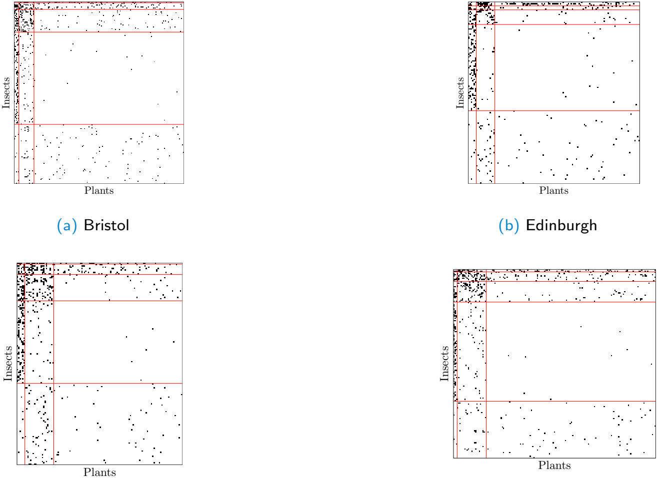
$$\begin{aligned}\text{BIC-L}(\mathbf{Y}, Q_1, Q_2) &= \max_{\theta} \mathbb{E}_{\mathcal{R}_{\mathbf{Y},\hat{\tau}}}[\ell_c(\mathbf{Y}, \mathbf{Z}, \mathbf{W}; \theta)] + \mathcal{H}(\mathcal{R}_{\mathbf{Y},\hat{\tau}}) - \frac{1}{2}\text{pen}(\theta, Q_1, Q_2) \\ &= \max_{\theta} \mathcal{J}(\mathcal{R}_{\mathbf{Y},\hat{\tau}},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

Exploration problems

- Exploration of $\mathbb{N}^2$ costly. $\rightarrow$ **Greedy approach** and **sliding window**
- Sensitivity to initializations. $\rightarrow$ **Spectral clustering** and **reuse of previous inits**

[2] ICL + Entropy + penalty

Results Baldock et al., 2019



(a) Bristol (b) Edinburgh (c) Leeds (d) Reading

Figure 6: Reordered adjacency matrices by *iid*-colBiSBM, Baldock et al., 2019

TODO Interesting structures detected, functional roles in the british networks.

Network clustering

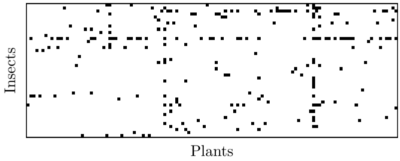
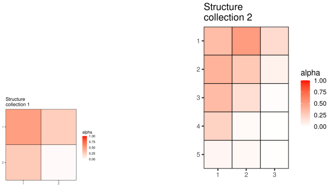
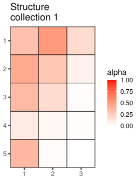


Figure 7: Adjacency matrix, Baldock et al., 2011

Application to Baldock et al., 2019, 2011 I



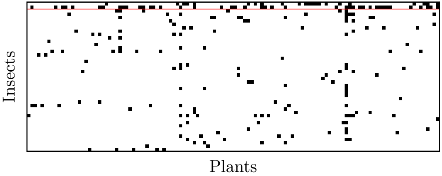
(a) Model *iid*,
separate African network and English networks



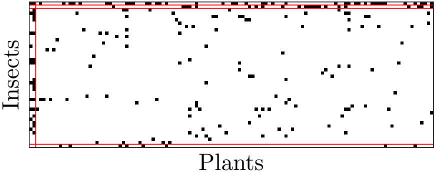
(b) Model $\pi\rho$,
merge African and English networks

Figure 8: Structures detected for networks of Baldock et al., 2019, 2011

Results



(a) Reordered by LBM



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

Figure 9: 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

- Investigate stability against randomness and local *optima*.

Package and applications

- Integration into the `colSBM` package, improvement of user interface and addition of ecologists' feedback
- 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

Thank you for your attention !

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

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

References II

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

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.

References III

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

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

$$\ell(\mathbf{Y};\theta) \geq \sum_{m=1}^{M} \Bigg(\sum_{i=1}^{n_1^m} \sum_{j=1}^{n_2^m} \sum_{q\in\mathcal{Q}_{1,m}} \sum_{r\in\mathcal{Q}_{2,m}} \tau_{i,q}^{1,m}\tau_{j,r}^{2,m} \log f(Y_{ij}^m;\alpha_{qr})$$
$$+ \sum_{i=1}^{n_1^m} \sum_{q\in\mathcal{Q}_{1,m}} \tau_{i,q}^{1,m} \log \pi_q{}^m + \sum_{j=1}^{n_2^m} \sum_{r\in\mathcal{Q}_{2,m}} \tau_{j,r}^{2,m} \log \rho_r{}^m$$
$$- \sum_{i=1}^{n_1} \tau_{i,q}^{1,m} \log \tau_{i,q}^{1,m} - \sum_{j=1}^{n_2} \tau_{j,r}^{2,m} \log \tau_{j,r}^{2,m} \Bigg) =: \mathcal{J}(\tau;\theta),$$

Variational approximation

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

Variational Expectation Step

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

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

[2] Initialization of $\widehat{\tau}$ with a *spectral clustering* on the networks.

Maximization Step

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

Connectivity parameters

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

Proportions for *iid*

$$\widehat{\pi}_q = \frac{\sum_{m=1}^{M} \sum_{i=1}^{n_1^m} \tau_{iq}^{1,m}}{\sum_{m=1}^{M} n_1^m} \qquad \widehat{\rho}_r = \frac{\sum_{m=1}^{M} \sum_{j=1}^{n_2^m} \tau_{jr}^{2,m}}{\sum_{m=1}^{M} n_2^m}$$

Maximization Step

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

Connectivity parameters

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

Proportions for $\pi\rho$

$$\widehat{\pi}^{m}{}_{q} = \frac{\sum_{i=1}^{n_1^m} \tau_{iq}^{1,m}}{n_1^m} \qquad \widehat{\rho}^{m}{}_{r} = \frac{\sum_{j=1}^{n_2^m} \tau_{jr}^{2,m}}{n_2^m}$$

Why does VE minimizes KL ?

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

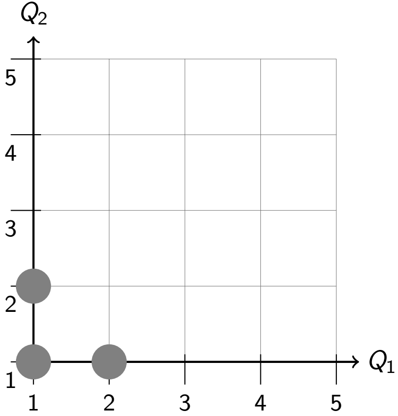
But $\mathbf{KL}[\mathcal{R}_{\mathbf{Y},\tau}, \log \mathbb{P}(\mathbf{Z},\mathbf{W}|\mathbf{Y};\theta)] = -\mathbb{E}_{\mathcal{R}_{\mathbf{Y},\tau}}[\log \frac{\mathbb{P}(\mathbf{Z},\mathbf{W}|\mathbf{Y};\theta)}{\mathcal{R}_{\mathbf{Y},\tau}}]$

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

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

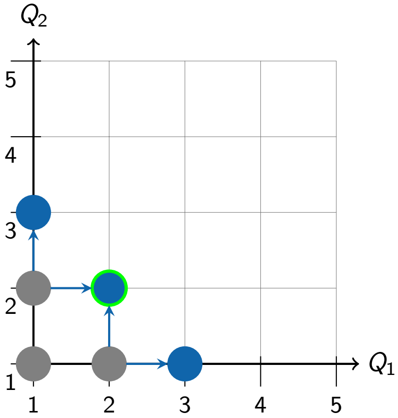
Thus $\ell(\mathbf{Y};\theta) - \mathbf{KL}[\mathcal{R}_{\mathbf{Y},\tau}, \log \mathbb{P}(\mathbf{Z},\mathbf{W}|\mathbf{Y};\theta)] = \mathcal{J}(\tau;\theta)$ □

Choice of (Q_1, Q_2) - Greedy approach



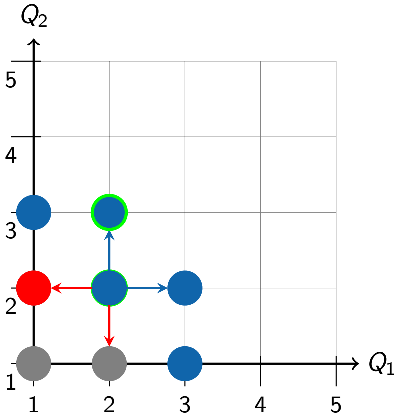
- Initial model :

Choice of (Q_1, Q_2) - Greedy approach



- Initial model : (grey disc)
- Model after *split* : (blue disc)
- Model maximizing the criterion : (green circle)

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

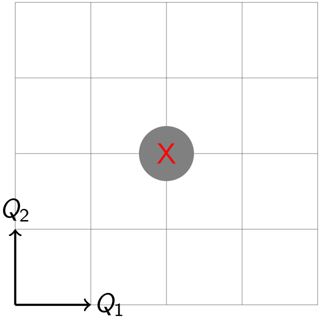


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

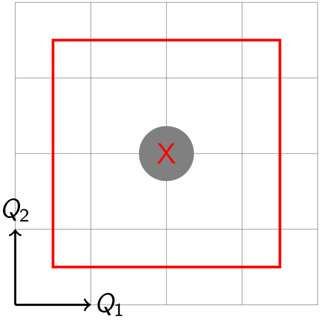


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

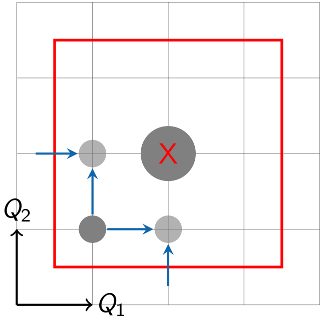


Figure 10: Sliding window

Initialization of the model if necessary

Choice of (Q_1, Q_2) - Sliding window

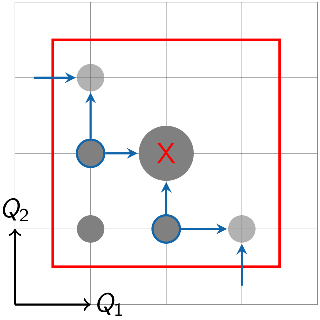


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

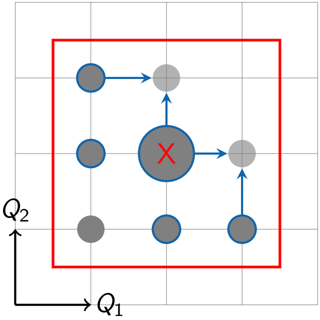


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

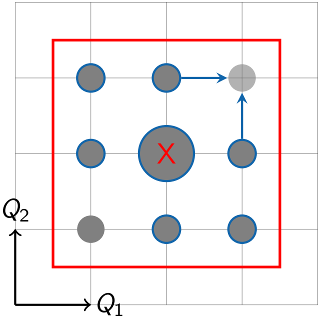


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

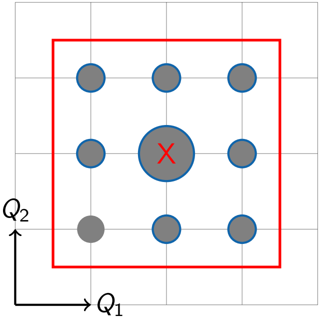


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

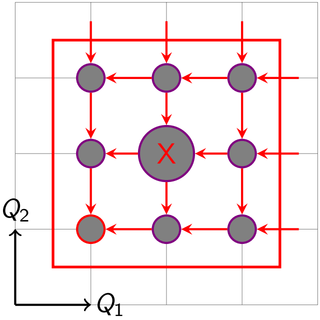


Figure 10: Sliding window

Choice of (Q_1, Q_2) - Sliding window

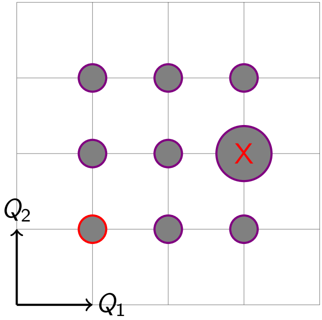


Figure 10: Sliding window

Localization of the new mode

Choice of (Q_1, Q_2) - Sliding window

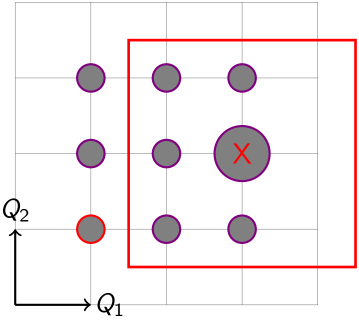
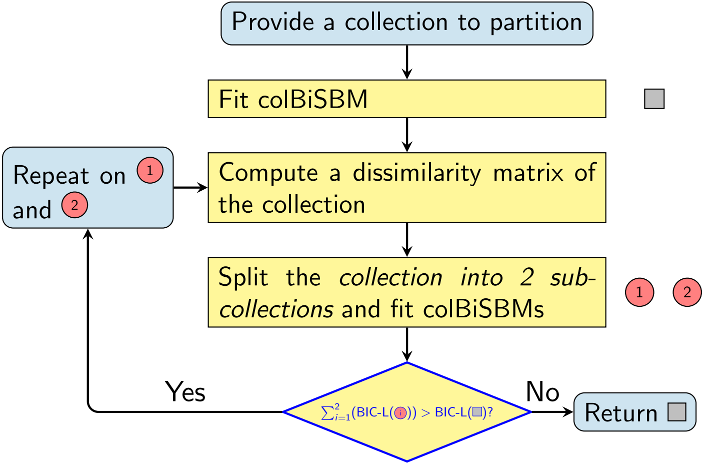


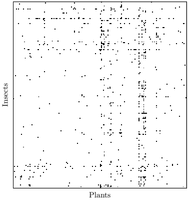
Figure 10: Sliding window

Move to the new mode then iterate

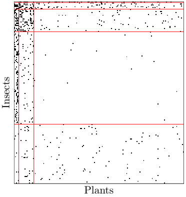
Clustering algorithm



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

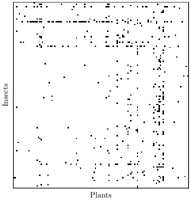


(a) Donnée

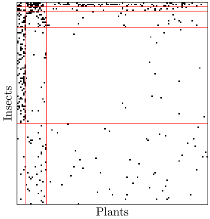


(b) Reordered

Figure 11: Bristol

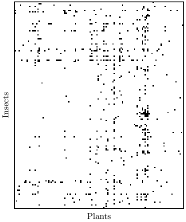


(a) Donnée

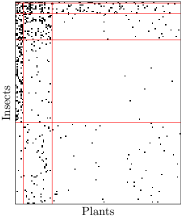


(b) Reordered

Figure 12: Edinburgh



(a) Donnée



(b) Réordonnée

Figure 13: Leeds

(a) Donnée (b) Réordonnée

Figure 14: Reading

Appendices references I