# Node Clustering in Edge-Exchangeable Networks based on Exchangeable Partition Processes

**Tâm Le Minh**[1], **Louise Alamichel**[2]

[1] *Univ. Grenoble Alpes, Inria, France* [2] *Bocconi University, Italy*

## I. Introduction

**Bipartite networks**

- Inbound (row) and outbound (column) nodes represent **different types of nodes**, from sets $\mathcal{U}$ and $\mathcal{V}$.
- **Examples:** ecological networks, recommendation systems, affiliation networks, semantic graphs.
- **Results can be generalized** to other network types.

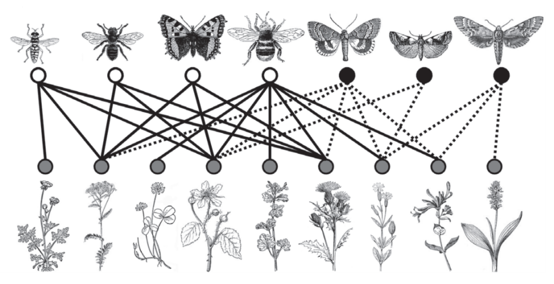
**Interactions**

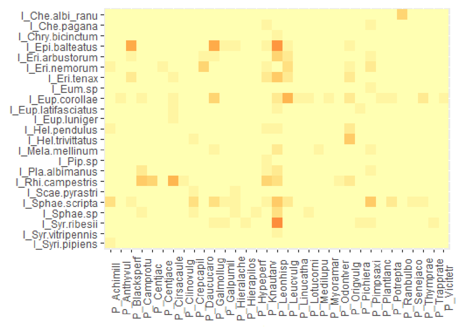
- Interaction $X_i = (X_i^{(1)}, X_i^{(2)})$ = network edge, given by a pair of node labels.
- Adjacency matrix defined by $Y_{uv} = \sum_{i=1}^{n} \mathbb{1}\{X_i = (u, v)\}$.

**Community detection in networks**

- **Task:** Group the row and column nodes in blocks $\{\mathcal{U}_1, \mathcal{U}_2, \dots\}$ and $\{\mathcal{V}_1, \mathcal{V}_2, \dots\}$.
- **Problem:** Traditional approaches offer **limited interpretability** (spectral clustering) or **unsuited** to interaction sampling processes (stochastic block model).

| COMPLEX | YEAR | DATE | PLANT_SP | INSECT_SP |
|---|---|---|---|---|
| Bunchgrass | 2011 | 24/07/2011 | ACHIMILL | EPICPUNC |
| Bunchgrass | 2011 | 24/07/2011 | ACHIMILL | EVYLS5M |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | MUSCGEN3 |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | MUSCGEN3 |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | TRICORNA |
| Bunchgrass | 2011 | 24/07/2011 | ACHIMILL | ANTHASP1 |
| Bunchgrass | 2011 | 24/07/2011 | ACHIMILL | ERISHIRT |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | SYRPRIBE |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | BOMBMIXT |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | DIALSPX9 |
| Bunchgrass | 2011 | 24/07/2011 | ERIGFOLI | MUSCGEN1 |





**Objective:** Exploit **generative network models** that

- reproduce **realistic network characteristics**,
- support **inference of latent block structures** underlying interactions.

Our approach exploits the relationship between the representation of *block edge-exchangeable networks* and *exchangeable random partitions*.

## II. Block edge-exchangeable networks

Given a partition of the node labels, the network is **block edge-exchangeable** [2] if

$$(\boldsymbol{X}^\sigma, \boldsymbol{Z}^\sigma) \overset{\mathcal{D}}{=} (\boldsymbol{X}, \boldsymbol{Z}), \text{ for all permutation } \sigma,$$

where $Z_i = (Z_i^{(1)}, Z_i^{(2)})$ are the block assignments of $X_i = (X_i^{(1)}, X_i^{(2)})$.

**Representation theorem**

There exist distributions $H$ on $\mathbb{N}^2$, and $(P_{k\ell})_{k,\ell}$ on $\mathcal{U}_k \times \mathcal{V}_\ell$, such that

$$Z_i \mid H \sim H,$$
$$X_i \mid Z_i = (k, \ell), P_{k\ell} \sim P_{k\ell}, \text{ for all } (k, \ell).$$

## III. Exchangeable random partitions

The discrete measures $H$ and $P_{k\ell}$ induce (bivariate) **marginally exchangeable random partition processes**. Marginal exchangeability ensures that row and column clusters can be interpreted separately while retaining joint dependencies.

**Example:** Clustering with Independence Centring (CLIC) process [1]

$$F \mid \pi, G^{(1)}, G^{(2)} \sim \text{DP}(\pi, G^{(1)} \times G^{(2)}),$$
$$G^{(j)} \mid \alpha^{(j)}, H^{(j)} \sim \text{DP}(\alpha^{(j)}, H^{(j)}) \text{ for } j \in \{1, 2\},$$

yielding a bivariate partition $(\boldsymbol{C}^{(1)}, \boldsymbol{C}^{(2)}) \sim \text{CLIC}(\pi, \alpha^{(1)}, \alpha^{(2)})$ that is marginally exchangeable.

## References

[1] A. Dombowsky and D. B. Dunson. Product centred Dirichlet processes for Bayesian multiview clustering. *JRSSB*, 2025.

[2] Y. Zhang and W. Dempsey. Node-level community detection within edge exchangeable models for interaction processes. *JASA*, 2024.

## IV. Nested CLIC model

The joint partition $(\boldsymbol{Z}, \boldsymbol{X})$ is described by

- $\boldsymbol{Z} \sim \text{CLIC}(\pi, \alpha^{(1)}, \alpha^{(2)})$,
- $\boldsymbol{X}$ is the partition induced by $\tilde{\boldsymbol{X}}$, with

$$\tilde{X}_i \mid Z_i = (k, \ell), P_{k\ell} \sim P_{k\ell}, \qquad P_{k\ell} \mid Q_k^{(1)}, Q_\ell^{(2)} \sim \text{DP}(\rho_{k\ell}, Q_k^{(1)} \times Q_\ell^{(2)}),$$
$$Q_k^{(j)} \mid Q_0^{(j)} \sim \text{DP}(\beta_k^{(j)}, Q_0^{(j)}), \qquad \text{for } k \in \mathbb{N}, j \in \{1, 2\}.$$

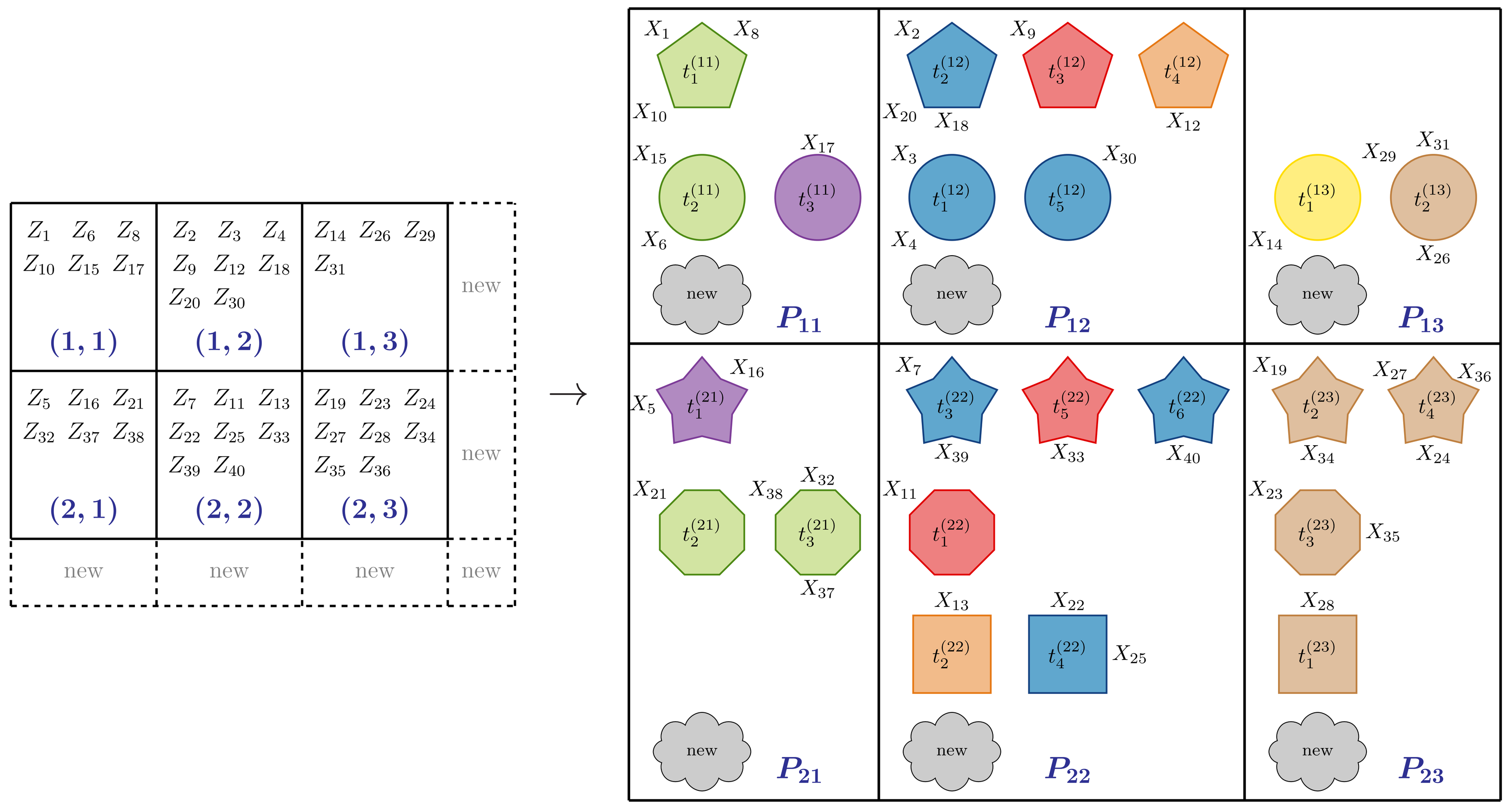
**Properties**

- **Nested structure:** There is a partition process for the blocks, and within each block, a partition process for the species.
- **Atom sharing between blocks:** Blocks in the same row share the same row species; blocks in the same column share the same column species.
- **Multiview dependency:** For both blocks and species, row and column selections are dependent.

## V. Restaurant metaphor

To build intuition, we illustrate the generative model through a metaphor involving restaurants and tables.

- Consider restaurants located on a 2D grid. These restaurants have tables, characterized by their shape and color, where clients can sit.
- The **latent partition** $\boldsymbol{Z}$ of the individuals is induced by the **restaurant coordinates** in which they sit.
- The **observed partition** $\boldsymbol{X}$ of the individuals is induced by the **shapes and colors** of the tables at which they sit (regardless of the restaurant).




**Process**

- **Clients** are assigned a **restaurant** according to **CLIC**$(\pi, \alpha^{(1)}, \alpha^{(2)})$,
- In each restaurant $(k, \ell)$, **clients** are assigned a **table**, according to **CRP**$(\rho_{k\ell})$,
- Tables are assigned a shape and a color: restaurants on the same row share table shapes, restaurants on the same column share table colors.
  1. **Shape:** tables in restaurants on row $k$ assigned according to **CRP**$(\beta_k^{(1)})$.
  2. **Color:** tables in restaurants on column $\ell$ assigned according to **CRP**$(\beta_\ell^{(2)})$.

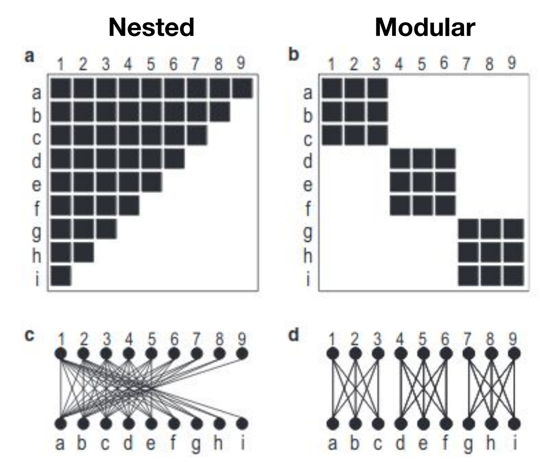
## VIII. Future work

- Investigate **consistency** of the inferred block structure.
- Clarify the **identifiability** and **interpretation** of the inferred parameters.
- Extend the framework to **hypergraphs** via exchangeable **feature allocation** processes (e.g., IBP-based priors).

## VI. Network properties

Real-life networks often exhibit

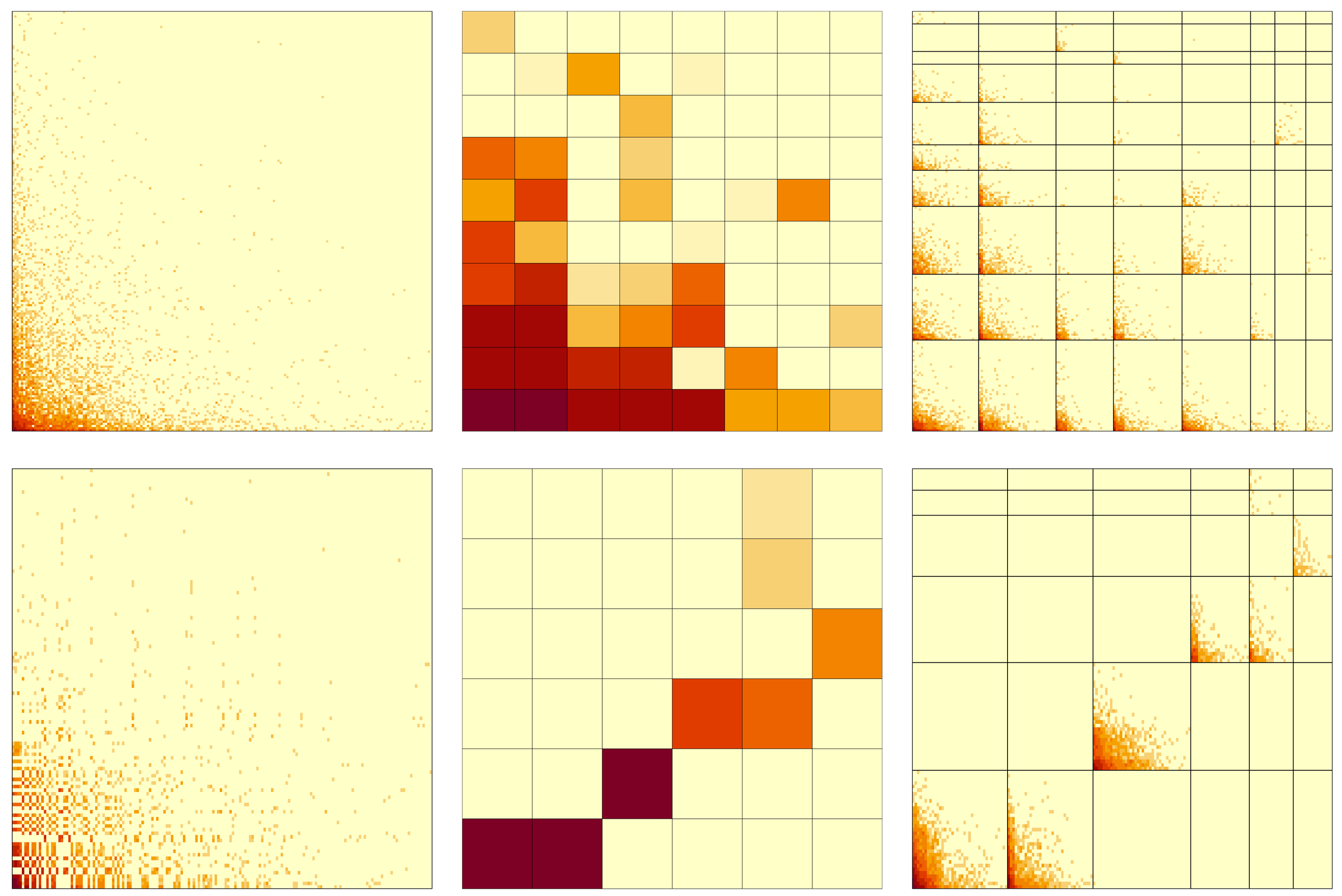
- sparsity,
- power-law degree distributions,
- nestedness,
- modularity.




**Simulations**

Two networks with $\rho_{k\ell} = 100, \beta_k^{(1)} = \beta_\ell^{(2)} = 5$, and:

1. $\pi = 100, \alpha^{(1)} = \alpha^{(2)} = 0.8$,
2. $\pi = 1, \alpha^{(1)} = \alpha^{(2)} = 8$.




▶ Our model is flexible enough to depict **classic network structures**.

## VII. Inference

We aim to compute the posterior distribution for the block assignments $Z$ and the parameters $\pi$, $\boldsymbol{\rho}$, $\boldsymbol{\beta}^{(1)}$ and $\boldsymbol{\beta}^{(2)}$.

**Partition probability**

For a more general model, replacing the distributions DP$(\gamma)$ by (proper) species sampling models SSM$(\gamma)$ with EPPF $\phi_k^n(n_1, \dots, n_k; \gamma)$.

$$P(\boldsymbol{X} \mid \boldsymbol{Z}) = \sum_{\boldsymbol{r}} \prod_{k=1}^{K} \phi_{|\mathcal{U}_k|}^{|\boldsymbol{r}_1^{(k)}|}\left(\boldsymbol{r}_1^{(k)}; \beta_k^{(1)}\right) \prod_{\ell=1}^{L} \phi_{|\mathcal{V}_\ell|}^{|\boldsymbol{r}_2^{(\ell)}|}\left(\boldsymbol{r}_2^{(\ell)}; \beta_\ell^{(2)}\right)$$
$$\times \prod_{(k,\ell)} \left\{ \sum_{\boldsymbol{w}^{(k\ell)}} \prod_{(u,v) \in \mathcal{U}_k \times \mathcal{V}_\ell} \frac{1}{r_{uv}!} \binom{n_{uv}}{w_{uv,1}, \dots, w_{uv,r_{uv}}} \times \phi_{|\boldsymbol{r}^{(k\ell)}|}^{n^{(k\ell)}}\left(\boldsymbol{w}^{(k\ell)}; \rho_{k\ell}\right) \right\},$$

where $\boldsymbol{r}_1^{(k)} := (r_{u\cdot} : u \in \mathcal{U}_k)$, $\boldsymbol{r}_2^{(\ell)} := (r_{\cdot v} : v \in \mathcal{V}_\ell)$, $\boldsymbol{w}^{(k\ell)} = (w_{uv,t} : t \in [r_{uv}], (u, v) \in \mathcal{U}_k \times \mathcal{V}_\ell)$.

**Metaphor interpretation**

- First term handles shape counts, second handles color counts, third handles seating arrangements within restaurants,
- $r_{uv}$ is the number of tables with shape $u$ and color $v$,
- $w_{uv,t}$ is the number of clients sitting at the $t$-th table with shape $u$ and color $v$.

▶ We perform inference using a **Gibbs sampler** with finite approximations and HDP data augmentation techniques.