Geometric clustering of a random graph

Antoine Channarond, Jean-Jacques Daudin, Stéphane Robin

4^{th} Stochastic Geometry Days - August 27^{th} 2015

Interaction networks

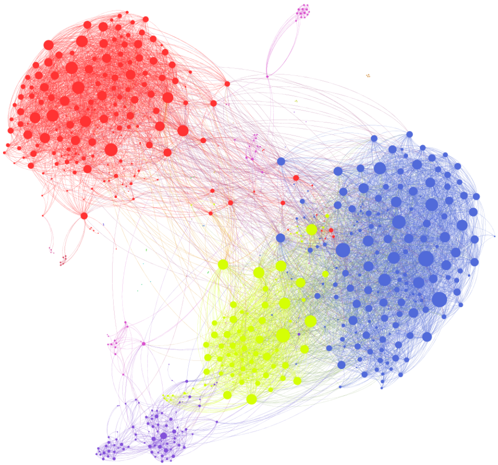
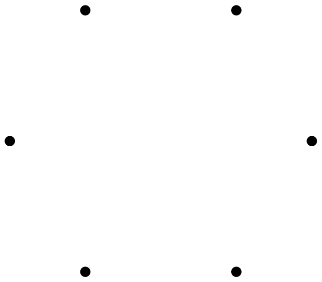


Figure: Friendship network (Source: griffsgraphs.com)

Random graphs

Graph:

- n nodes (individuals)



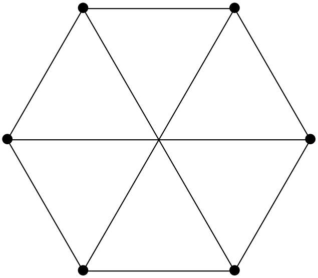
$n = 6$

Random graphs

Graph:

- n nodes (individuals)
- edges (interactions)
  - $X_{ij} = 1$ if nodes i and j interact.
  - $X_{ij} = 0$ if not.
- undirected edges: $X_{ij} = X_{ji}$

$X = (X_{ij})$ symmetric adjacency matrix.



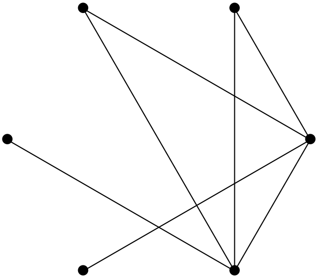
$n = 6$

Random graphs

Random Graph:

- n nodes (individuals)
- random edges (interactions)
  - $X_{ij} = 1$ if nodes i and j interact.
  - $X_{ij} = 0$ if not.
- undirected edges: $X_{ij} = X_{ji}$

$X = (X_{ij})$ random symmetric adjacency matrix.



$n = 6$

Random graphs

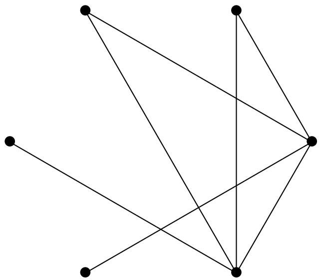
Random Graph:

- n nodes (individuals)
- random edges (interactions)
  - $X_{ij} = 1$ if nodes i and j interact.
  - $X_{ij} = 0$ if not.
- undirected edges: $X_{ij} = X_{ji}$

$X = (X_{ij})$ random symmetric adjacency matrix.

Erdős-Rényi model:

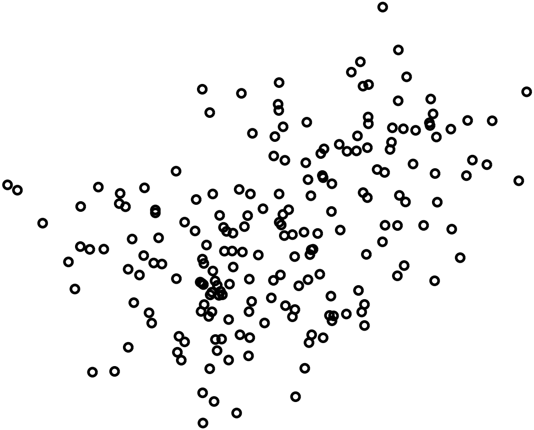
$$X_{ij} \text{ i.i.d. } \sim \mathcal{B}(p)$$



$n = 6$; $p = 0.5$

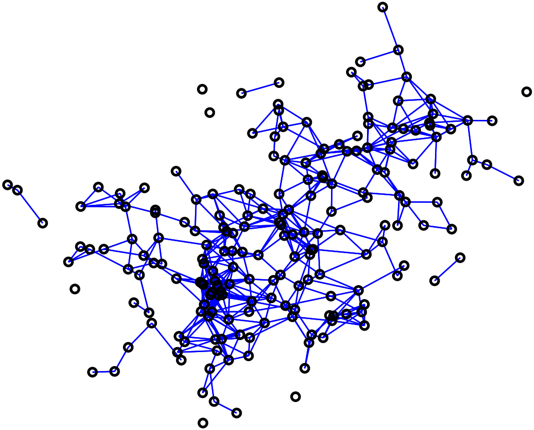
Paradigm: interactions induced by (unobserved) closeness

- Embedding of the individuals in a metric space; here the euclidean space $\mathbb{R}^2$



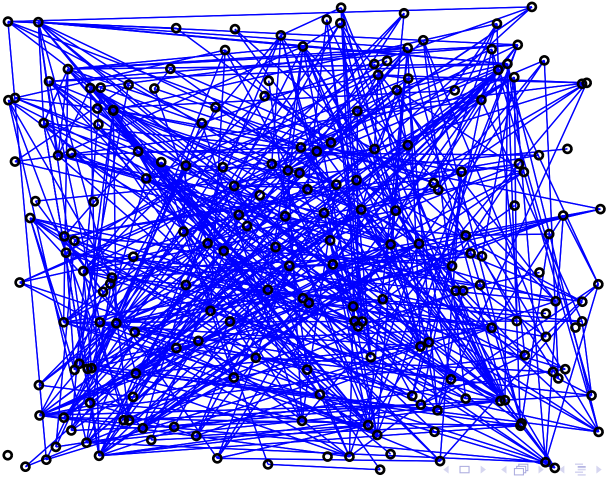
Paradigm: interactions induced by (unobserved) closeness

- Embedding of the individuals in a metric space; here the euclidean space $\mathbb{R}^2$
- Probability of connection decreasing with respect to the euclidean distance



Paradigm: interactions induced by (unobserved) closeness

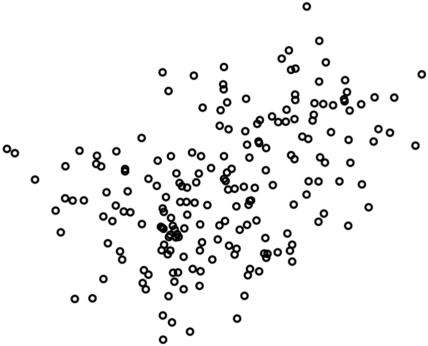
- Embedding of the individuals in a metric space; here the euclidean space $\mathbb{R}^2$
- Probability of connection decreasing with respect to the euclidean distance
- Positions/distances are not observed.



Latent Position Models

Latent Positions Latent space $\mathcal{S} = \mathbb{R}^d$

- Z_i: random position of node i
- (Z_i) i.i.d. $\sim f$ (unknown density)



Latent Position Models

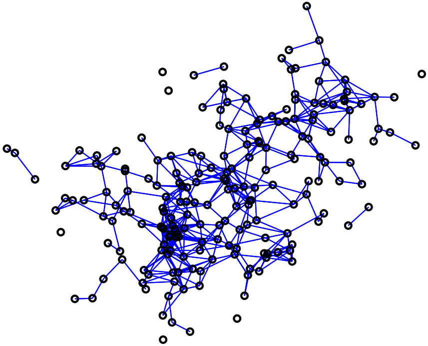
Latent Positions Latent space $\mathcal{S} = \mathbb{R}^d$

- Z_i: random position of node i
- (Z_i) i.i.d. $\sim f$ (unknown density)

Observed Graph (X_{ij}) independent conditionally on Z and:

$$X_{ij} \,|Z_i, Z_j \quad \sim \; \mathcal{B}\left(k_n\left(\|Z_i - Z_j\|\right)\right)$$

with $k_n : \mathbb{R}^+ \to [0,1]$ decreasing w.r.t. norm $\|\cdot\|$.



Properties:

- Homophily
- Transitivity

Latent Position Models

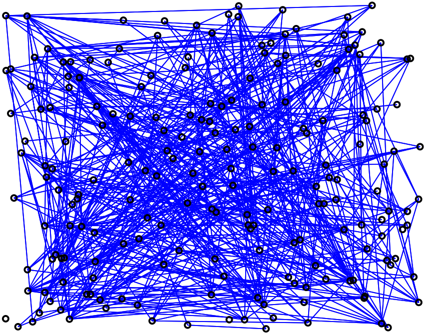
Latent Positions Latent space $\mathcal{S} = \mathbb{R}^d$

- Z_i: random position of node i
- (Z_i) i.i.d. $\sim f$ (unknown density)

Observed Graph (X_{ij}) independent conditionally on Z and:

$$X_{ij} \,|Z_i, Z_j \quad \sim \; \mathcal{B}\left(k_n\left(\|Z_i - Z_j\|\right)\right)$$

with $k_n : \mathbb{R}^+ \to [0,1]$ decreasing w.r.t. norm $\|\cdot\|$.



Properties:

- Homophily
- Transitivity

Outline

Two ideas of clustering

Latent positions $(Z_i)_{i\in[n]}$ i.i.d. $\sim f$

Parametric clustering

- Gaussian Mixture with Q components: $f = \sum_{q=1}^{Q} \pi_q \mathcal{N}_d(\mu_q, \sigma_q^2 Id)$.
- Clusters defined by components of the mixture.

Two ideas of clustering

Latent positions $(Z_i)_{i\in[n]}$ i.i.d. $\sim f$

Parametric clustering

- Gaussian Mixture with Q components: $f = \sum_{q=1}^{Q} \pi_q \mathcal{N}_d(\mu_q, \sigma_q^2 Id)$.
- Clusters defined by components of the mixture.

Non-parametric clustering

- f just assumed to be nice (regular).
- Clusters defined using level sets of f:

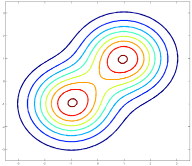


Figure: Level lines of a gaussian mixture

Statistical questions

How to get clustering information on the latent variables Z and their distribution, from the observed graph X only ?

Test	Is the distribution of Z clustered ?
Model choice	How many clusters are there ?
Classification	Which nodes are in which cluster ?
Estimation	What are the characteristics of each cluster ?

Latent Position Cluster Model (Handcock et al, 2007)

Positions Latent space $\mathbb{R}^d$. Parametric Gaussian mixture:

$$(Z_i)_{i\in[n]} \text{ i.i.d. } \sim f = \sum_{q=1}^{Q} \pi_q \mathcal{N}_d(\mu_q, \sigma_q^2 Id)$$

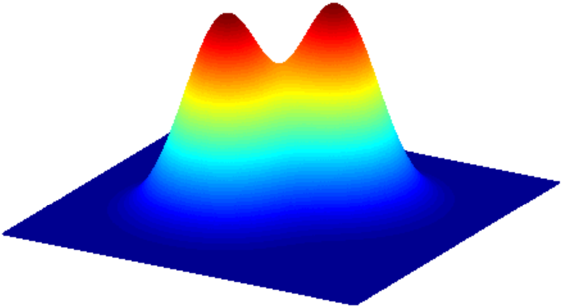
Graph Logistic regression for the connection probability:

$$\text{log-odds}(X_{ij} = 1 \mid Z_i, Z_j) = -\beta \|Z_i - Z_j\| \quad (\beta > 0)$$

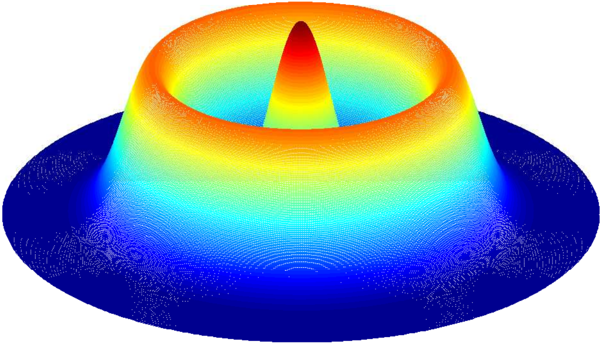
Statistics

- Clustering structure defined by mixture components
- Unsupervised classification of the nodes:
  1. ML-Estimation of the distances $(\|Z_i - Z_j\|)_{i,j\in[n]}$,
  2. Multidimensional scaling: estimating $(Z_i)_{i\in[n]}$ up to isometries,
  3. EM-algorithm: estimating mixture parameters $\alpha_q, \mu_q, \sigma_q$ and constructing a classification rule.
- (Estimation of the mixture parameters)
- Model selection for Q.

From parametric to non-parametric setting



From parametric to non-parametric setting



Definition of clustering in a non-parametric setting

Definition

Let $t > 0$ and $\mathcal{L}(t) = \{f \geq t\}$ be the t-level set of a function f. A connected component of $\mathcal{L}(t)$ is called t-cluster (Hartigan, 1975). $Q(t)$ denotes the number of t-clusters.

- Clusters: connected regions of "high" density, i.e. higher than some level t
- Estimation of $Q(t)$, number of such regions

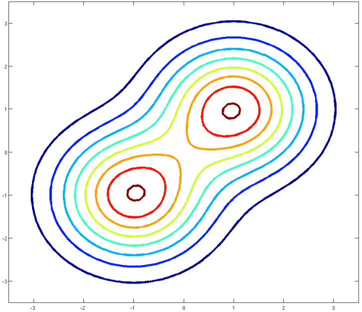
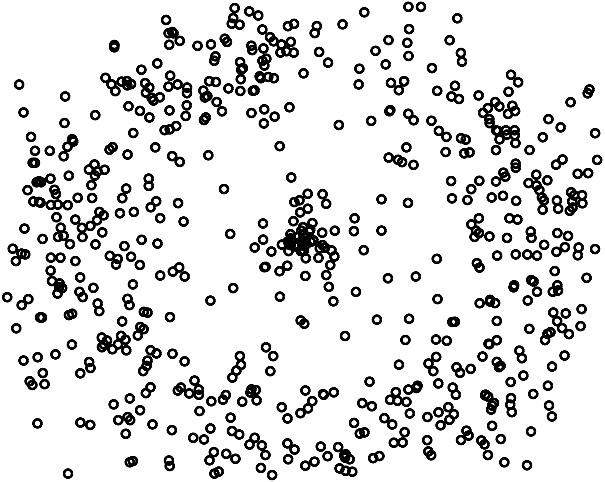
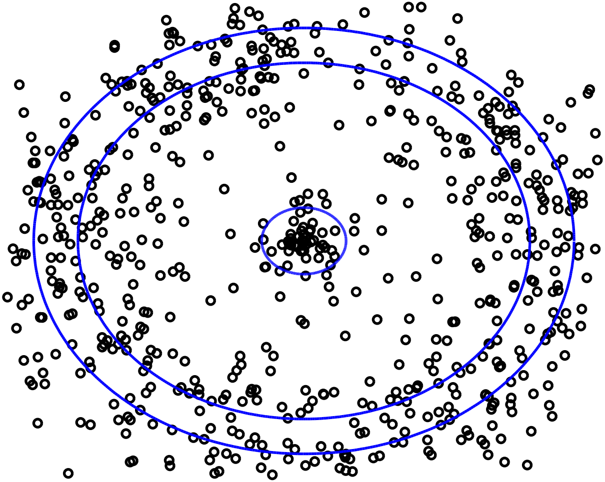
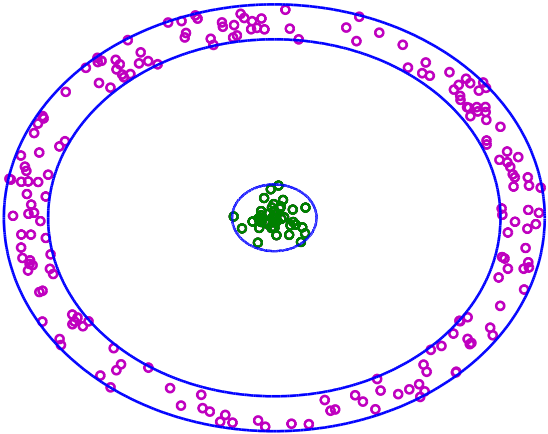


Figure: Level lines of a gaussian mixture





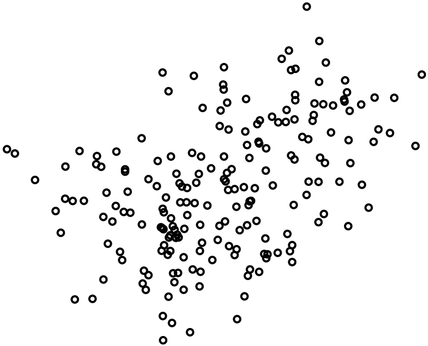


Outline

1. Introduction: from interaction networks to random graphs
2. Discussion about clustering
3. Fast count of the number of clusters in a non-parametric clustering setting

Non-parametric latent position model

Latent Positions $Z = (Z_i)$ i.i.d. $\sim f$ non-parametric density



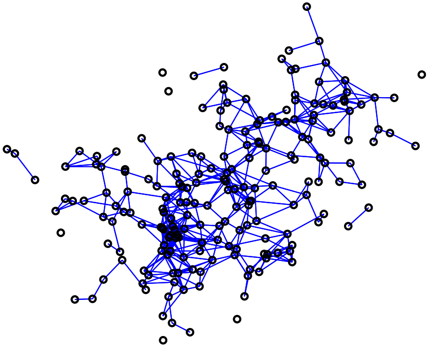
Non-parametric latent position model

Latent Positions $Z = (Z_i)$ i.i.d. $\sim f$ non-parametric density

Observed Graph (X_{ij}) independent conditionally on Z and:

$$X_{ij} \,|Z_i, Z_j \quad \sim \mathcal{B}\left(k\left(\frac{Z_i - Z_j}{h_n}\right)\right)$$

with $k : \mathbb{R}^d \to [0,1]$ isotropic and decreasing w.r.t. $\|\cdot\|$, with support $B(0,1)$, and $h_n > 0$ (connection radius).

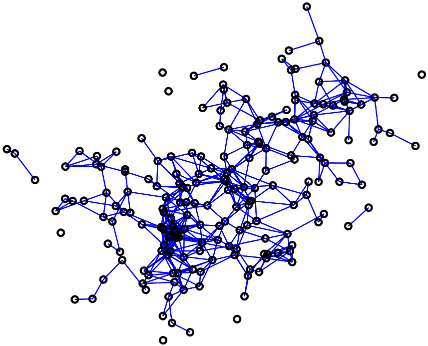


Non-parametric latent position model

Latent Positions $Z = (Z_i)$ i.i.d. $\sim f$ non-parametric density
Observed Graph (X_{ij}) independent conditionally on Z and:

$$X_{ij} \,|Z_i, Z_j \quad \sim \mathcal{B}\left(k\left(\frac{Z_i - Z_j}{h_n}\right)\right)$$

with $k : \mathbb{R}^d \to [0,1]$ isotropic and decreasing w.r.t. $\|\cdot\|$, with support $B(0,1)$, and $h_n > 0$ (connection radius).



Graph properties:

- Homophily
- Transitivity
- Sparsity: $\zeta_n = \dfrac{2}{n(n-1)} \displaystyle\sum_{1 \le i,j \le n} X_{ij}$

$$\mathbb{E}(\zeta_n) \sim_{h_n \to 0} h_n^d \int_{\mathbb{R}^d} f^2(z)dz$$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

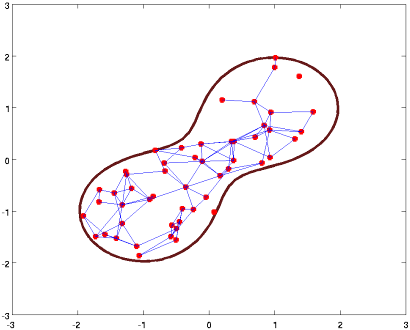


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

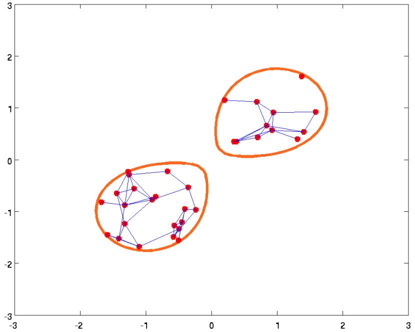


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*

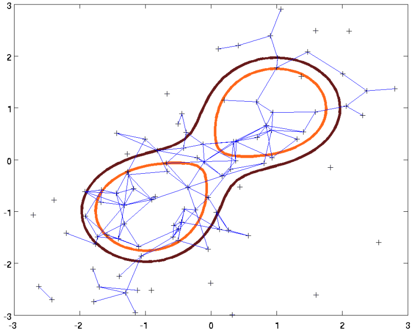


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*
2. *Find the set $\widehat{J_n}(t)$ of nodes i such that $T_i^{h_n} \geq t$*

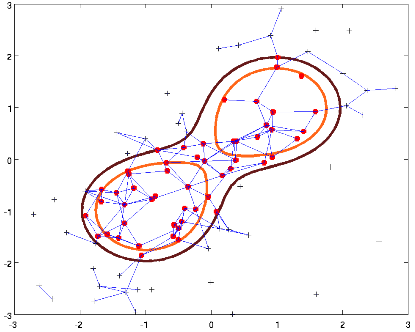


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*
2. *Find the set $\widehat{J}_n(t)$ of nodes i such that $T_i^{h_n} \geq t$*
3. *Remove other nodes from X*

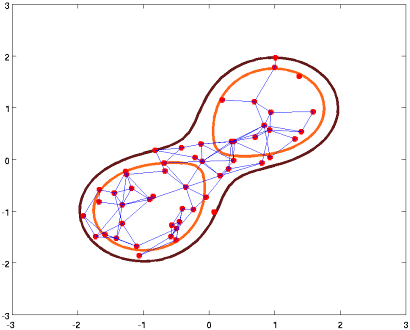


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*
2. *Find the set $\widehat{J_n}(t)$ of nodes i such that $T_i^{h_n} \geq t$*
3. *Remove other nodes from X*
4. *Count the number $\widehat{Q_n}(t)$ of connected components of the graph $\widehat{X} = X_{\widehat{J_n}(t)}$ with DFS (linear)*

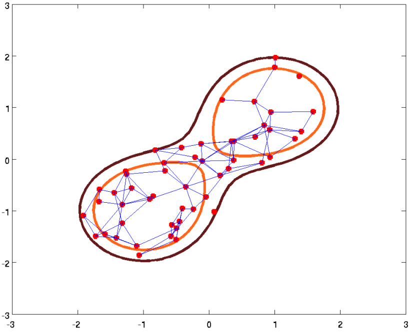


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*
2. *Find the set $\widehat{J}_n(t)$ of nodes i such that $T_i^{h_n} \geq t$*

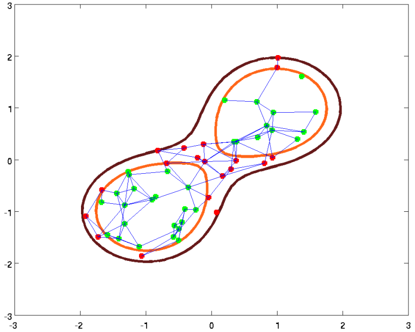


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*
2. *Find the set $\widehat{J_n}(t)$ of nodes i such that $T_i^{h_n} \geq t$*
3. *Remove other nodes from X*

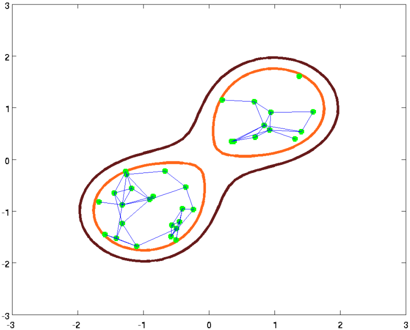


Figure: —: $t = 0.06$ —: $t = 0.05$

Counting the t-clusters with a covering graph

- Find a graph $\widehat{X}$ covering $\mathcal{L}(t)$ to make a link between clusters of $\mathcal{L}(t)$ and connected components of the graph $\widehat{X}$: Biau, Cadre, Pelletier (2007).
- Extract $\widehat{X}$ from X by removing low degree nodes.
- nh_n^d-normalized degrees : $T_i^{h_n} = \dfrac{D_i}{nh_n^d}$

Algorithm

1. *Compute nh_n^d-normalized degrees*
2. *Find the set $\widehat{J}_n(t)$ of nodes i such that $T_i^{h_n} \geq t$*
3. *Remove other nodes from X*
4. *Count the number $\widehat{Q_n}(t)$ of connected components of the graph $\widehat{X} = X_{\widehat{J}_n(t)}$ with DFS (linear)*

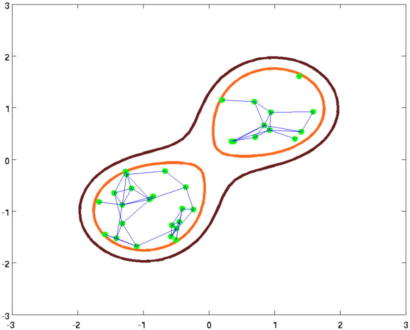


Figure: —: $t = 0.06$ —: $t = 0.05$

Biau, Cadre, Pelletier (2007)

Model X is deterministic w.r.t. Z: $k(x) = \mathbb{1}_{\|x\| \leq 1}$

Generalization

Model k is a kernel function with support in the unit ball

Biau, Cadre, Pelletier (2007)

Model X is deterministic w.r.t. Z: $k(x) = \mathbb{1}_{\|x\| \leq 1}$

Context not latent, positions Z are observed

Generalization

Model k is a kernel function with support in the unit ball

Context latent, positions Z are not observed

Biau, Cadre, Pelletier (2007)

Model X is deterministic w.r.t. Z: $k(x) = \mathbb{1}_{\|x\| \leq 1}$

Context not latent, positions Z are observed

Estimator of the density at the node positions $\widehat{f_n}(Z_i)$, with a kernel estimator:

$$\widehat{f_n}(z) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K\left(\frac{z - Z_i}{h_n}\right) \text{ where } K \text{ is a kernel function}$$

Generalization

Model k is a kernel function with support in the unit ball

Context latent, positions Z are not observed

Estimator of the density at the node positions, nh_n^d-normalized degrees of X:

$$T_i^{h_n} = \frac{1}{nh_n^d} \sum_{i=1}^{n} X_{ij}$$

Biau, Cadre, Pelletier (2007)

Model X is deterministic w.r.t. Z: $k(x) = \mathbb{1}_{\|x\| \leq 1}$

Context not latent, positions Z are observed

Estimator of the density at the node positions $\widehat{f}_n(Z_i)$, with a kernel estimator:

$$\widehat{f}_n(z) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K\left(\frac{z - Z_i}{h_n}\right) \text{ where } K \text{ is a kernel function}$$

Generalization

Model k is a kernel function with support in the unit ball

Context latent, positions Z are not observed

Estimator of the density at the node positions, nh_n^d-normalized degrees of X:

$$T_i^{h_n} = \frac{1}{nh_n^d} \sum_{i=1}^{n} X_{ij}$$

Proposition: $\mathbb{E}\left(T_i^{h_n} | Z\right) = \widehat{f}_n(Z_i)$ where $\widehat{f}_n(z) = \frac{1}{nh_n^d} \sum_{j=1}^{n} k\left(\frac{z - Z_j}{h_n}\right)$

Hypotheses and result (article to be submitted)

- f uniformly continuous and of class $\mathcal{C}^1$ on the neighborhood of $\{f = t\}$.
- $df \neq 0$ on $\{f = t\}$.
- h_n small enough and $\dfrac{nh_n^d}{\ln n} \xrightarrow[n\to\infty]{} +\infty$

Theorem

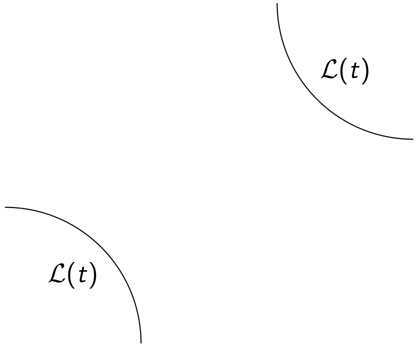
Non-underestimation *$\widehat{Q_n}(t)$=number of connected components of $X_{\widehat{J_n}(t)}$. For some ε_n, h_n small enough:*

$$P\left(\widehat{Q_n}(t) < Q(t)\right) \leq 3n \exp\left(-K_0 \varepsilon_n^2 n h_n^d\right)$$

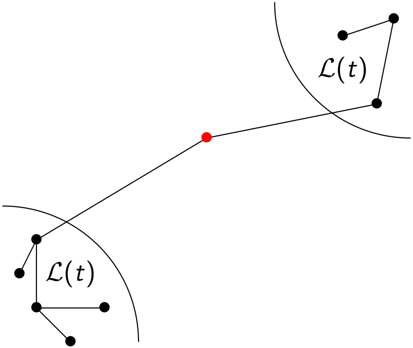
Non-overestimation *$J_n(t)$ set of the nodes i such that $f(Z_i) \geq t$. $\widetilde{Q}_n(t)$ is the number of connected components of $X_{J_n(t)}$. For h_n small enough:*

$$P\left(\widetilde{Q}_n(t) > Q(t)\right) \leq K_1 n \exp(-K_2 n h_n^d)$$

Non-underestimation of $Q(t)$



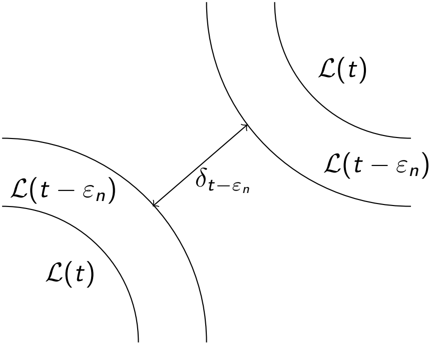
Non-underestimation of $Q(t)$



Non-underestimation of $Q(t)$

- On the event $\left\{ \sup_{i\in[n]} |T_i^{h_n} - f(Z_i)| \le \varepsilon_n \right\}$ with $\varepsilon_n > 0$

$$T_i^{h_n} \ge t \Rightarrow Z_i \in \mathcal{L}(t - \varepsilon_n)$$



Non-underestimation of $Q(t)$

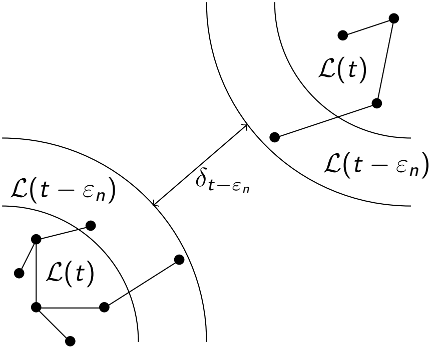
- On the event $\left\{ \sup_{i \in [n]} |T_i^{h_n} - f(Z_i)| \leq \varepsilon_n \right\}$ with $\varepsilon_n > 0$

$$T_i^{h_n} \geq t \Rightarrow Z_i \in \mathcal{L}(t - \varepsilon_n)$$

- δ_t = distance between two t-clusters.

$$h_n < \delta_{t-\varepsilon_n}$$

No connection between distinct clusters.



Non-underestimation of $Q(t)$

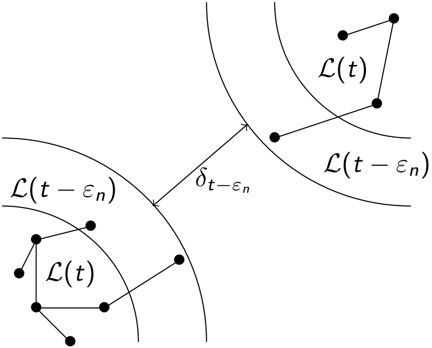
- On the event $\left\{ \sup_{i \in [n]} |T_i^{h_n} - f(Z_i)| \le \varepsilon_n \right\}$ with $\varepsilon_n > 0$

$$T_i^{h_n} \ge t \Rightarrow Z_i \in \mathcal{L}(t - \varepsilon_n)$$

- δ_t = distance between two t-clusters.

$$h_n < \delta_{t-\varepsilon_n}$$

No connection between distinct clusters.

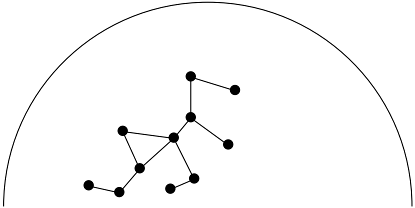


Concentration inequality (Bernstein):

$$P\left(\sup_{i \in [n]} |T_i^{h_n} - f(Z_i)| > \varepsilon_n \mid Z_i \right) \le 2n \exp(-K_3 \varepsilon_n^2 n h_n^d)$$

Non-overestimation of $Q(t)$

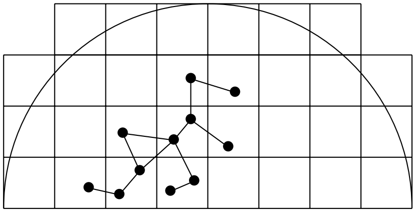
Connectivity of the subgraph induced by each cluster:



Non-overestimation of $Q(t)$

Connectivity of the subgraph induced by each cluster:

- Cover of each cluster with hypercubes of side $h_n/2$

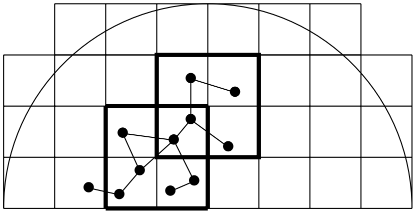


Non-overestimation of $Q(t)$

Connectivity of the subgraph induced by each cluster:

- Cover of each cluster with hypercubes of side $h_n/2$
- Local connectivity:
  - In hypercubes of side h_n, any two nodes can be connected
  - Comparison to Erdős-Rényi. Let C be one hypercube of the cover:

$$P(X_C \text{ is connected}) \geq P(\mathcal{ER} \text{ is connected})$$



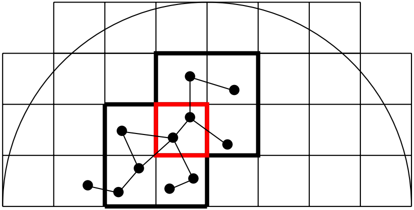
Non-overestimation of $Q(t)$

Connectivity of the subgraph induced by each cluster:

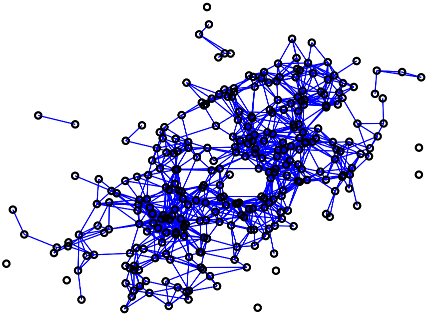
- Cover of each cluster with hypercubes of side $h_n/2$
- Local connectivity:
  - In hypercubes of side h_n, any two nodes can be connected
  - Comparison to Erdős-Rényi. Let C be one hypercube of the cover:

$$P(X_C \text{ is connected}) \geq P(\mathcal{ER} \text{ is connected})$$

- From local to global connectivity:
  - Filling of the hypercubes



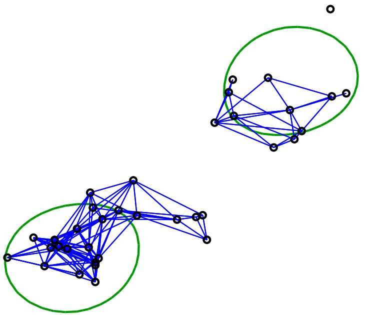
Classification



Two kinds of error controlled in the theorems :

- support error (thresholding error)
- classification error

Classification



Two kinds of error controlled in the theorems :

- support error (thresholding error)
- classification error

Simulation design

- $R_i \sim 0.15\mathcal{N}(0,1) + 0.85\mathcal{N}(5,1)$
- $\theta_i \sim \mathcal{U}([0, 2\pi])$

- k Epanechnikov kernel
- $h_n = h = 1$
- $t = 0.005$
- 300 graphs drawn

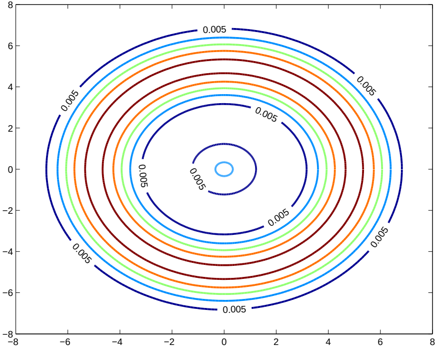


Figure: Isolines of the density

Illustration of the consistency of $\widehat{Q_n}(t)$

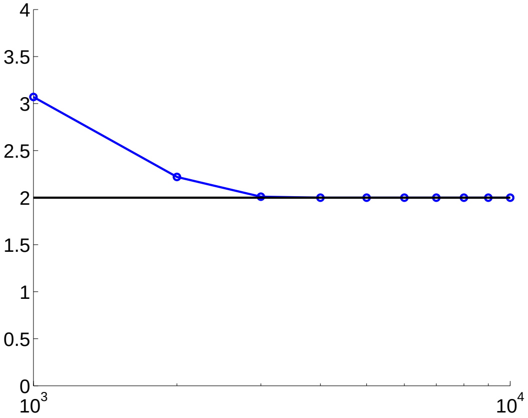
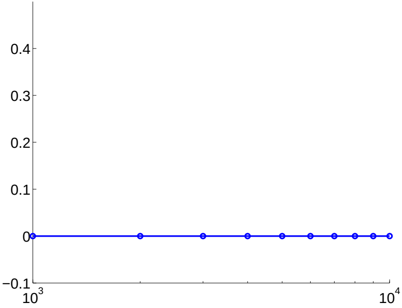
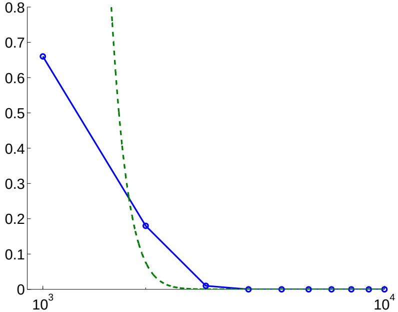


Figure: —: Estimator $\widehat{Q_n}(t)$ as a function of n (averaged over 300 graphs); —: Objective

Under- and overestimation frequency



—: Underestimation frequency of $\widehat{Q_n}(t)$ as a function of n



—: Overestimation frequency of $\widehat{Q_n}(t)$ as a function of n.

—: Overestimation frequency of $\widetilde{Q_n}(t)$.

- - -: Theoretical bound.

Clustering profile: a practical implementation

Invariance under similarity transformations of the latent space

If R is a similarity transformation of $\mathbb{R}^d$ with scale factor λ:

- $Z_i \longrightarrow Z'_i = R(Z_i)$
- $h_n \longrightarrow \lambda h_n$

$$P(X_{ij} = 1 \mid Z'_i, Z'_j) = k\left(\frac{\|R(Z_i) - R(Z_j)\|}{\lambda h_n}\right) = k\left(\frac{\|Z_i - Z_j\|}{h_n}\right)$$

Practical algorithm

- Compute n-normalized degrees: $T_i = \frac{D_i}{n}$. (threshold: $u = th_n^d$)
- Sort (T_i) without ex-aequo values $\longrightarrow (T_{(k)})_{k\in[m]}$
- Run the algorithm: remove nodes i such that $T_i \geq u = T_{(k)}$ for each k
- Plot $\widehat{Q}_n(u)$ as a function of u.

Clustering profile: simulation

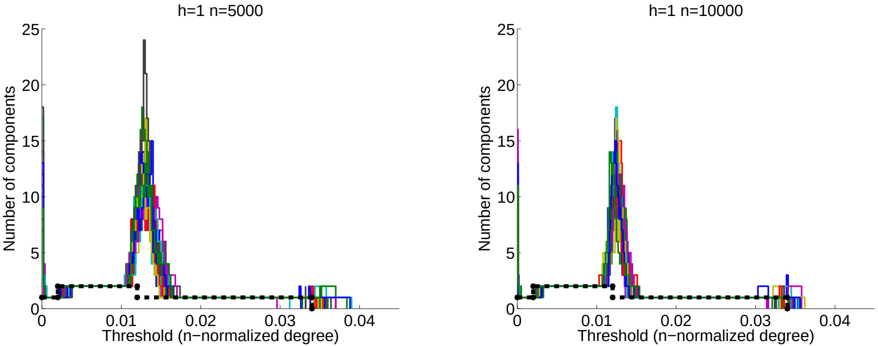


Figure: Clustering profiles ($\widehat{Q}_n(u)$ as a function of u) of 30 graphs with $h = 1$, $n = 5000$ and $n = 10000$

Clustering profile: simulation

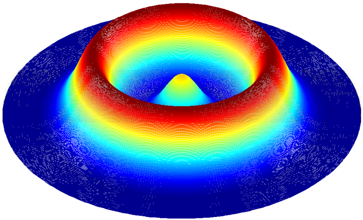
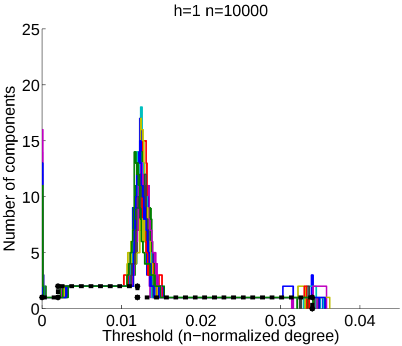



Figure: Clustering profiles ($\widehat{Q}_n(u)$ as a function of u) of 30 graphs with $h = 1$, $n = 5000$ and $n = 10000$

Conclusions and perspectives

Conclusions

- No likelihood-based strategy
- Behaviour of the degree distribution of the graph model
- Fast algorithm, able to process large graphs
- Theoretical guarantees: consistency proof

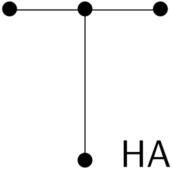
Conclusions and perspectives

Conclusions

- No likelihood-based strategy
- Behaviour of the degree distribution of the graph model
- Fast algorithm, able to process large graphs
- Theoretical guarantees: consistency proof

Perspectives

- Application to real-world networks
- Complete consistency of $\widehat{Q}_n(t)$
- Statistical properties of clustering profiles
- Robustness: small components filtering and theoretical arguments
- First step to test " latent distribution clustered" vs. "not clustered"



HANKS !

References

G. Biau, B. Cadre, and B. Pelletier.
A graph-based estimator of the number of clusters.
ESAIM: P&S, 11:272–280, 2007.

B. Bollobás, S. Janson, and O. Riordan.
The phase transition in inhomogeneous random graphs.
Random Structures & Algorithms, 31(1):3–122, 2007.

M.S. Handcock, A.E. Raftery, and J.M. Tantrum.
Model-based clustering for social networks.
Journal of the Royal Statistical Society: Series A (Statistics in Society), 170(2):301–354, 2007.

P.D. Hoff, A.E. Raftery, and M.S. Handcock.
Latent space approaches to social network analysis.
Journal of the American Statistical Association, 97(460):1090–1098, 2002.

K. Nowicki and T.A.B. Snijders.
Estimation and prediction for stochastic blockstructures.
Journal of the American Statistical Association, 96(455):1077–1087, 2001.