

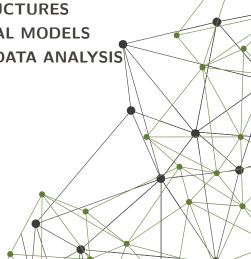
LEARNING BLOCK STRUCTURES
IN GAUSSIAN GRAPHICAL MODELS
FOR SPECTROMETRIC DATA ANALYSIS

SISBAYES 2025 - Padova

Alessandro Colombi[†]

 † University of Milano-Bicocca

05/09/2025



COLLABORATORS



(a) Raffaele Argiento



(d) Laura Codazzi



(b) Alessia Pini



(e) Matteo Gianella

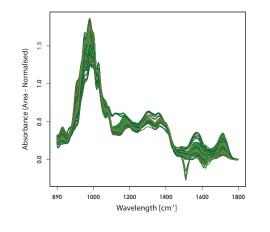


(c) Lucia Paci



(f) Alessandro Colombi

FRUIT PUREES DATASET



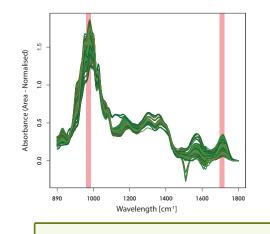
Data: Spectrum of absorbance of strawberry purees

Curves: 351

Wavelengths: 235 measurements

Goal: Investigating relationships among the substances within the compound

FRUIT PUREES DATASET



Data: Spectrum of absorbance of strawberry purees

Curves: 351

Wavelengths: 235 measurements

Goal: Investigating relationships among the substances within the compound

B-SPLINE NONPARAMETRIC REGRESSION

Non Parametric Regression

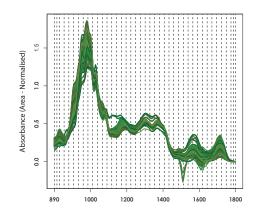
$$Y_{i}(s) = \sum_{j=1}^{p} \beta_{ij} \varphi_{j}(s) + \varepsilon_{i}(s)$$

- $Y_i(s) \in L^2(\mathbb{R})$, for $i = 1, \dots, n$
- $\varphi_{j}\left(s\right)$, B-spline basis of order 3, for $j=1,\ldots,p$
- $\varepsilon_i \stackrel{\text{iid}}{\sim} N\left(0, \tau_{\varepsilon}^{-2}\right)$

Among the priors for β s:

- → First and second order random walks (Lang and Brezger 2004)
- → Gaussian-Wishart process having a Matérn covariance kernel (Yang et al. 2017)

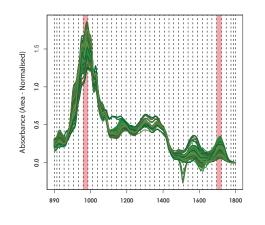
INTERPRETATION OF eta



- B-spline basis have compact support
- Each spline represents a portion of the domain

Our task is equivalent to learn the dependence structure among the basis expansion coefficients (β s)

INTERPRETATION OF eta



- B-spline basis have compact support
- Each spline represents a portion of the domain

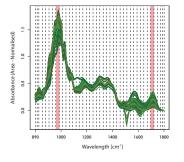
Our task is equivalent to learn the dependence structure among the basis expansion coefficients (β s)

GAUSSIAN GRAPHICAL PRIOR FOR eta

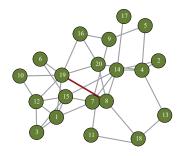
- graph G = (V, E), $V = \{1, \dots, p\}$ and $E \subseteq V \times V$
- $\bullet \quad \beta_i = (\beta_{i1}, \dots, \beta_{ip})^\top, \quad i = 1, \dots, n$

$$\beta_1, \dots, \beta_n \mid \mu, K \stackrel{\text{iid}}{\sim} N_p \left(\mu, K^{-1}\right)$$

 $\beta_i \perp \!\!\!\perp \beta_j \mid \beta_{V \setminus \{i,j\}} \iff K_{ij} = 0 \iff (i,j) \notin E$







$$K \mid G \sim \text{G-Wishart}(b,D)$$

$$P(K \mid G) = I_G(b, D)^{-1} |K|^{\frac{b-2}{2}} \exp\left\{-\frac{1}{2}tr(KD)\right\}, K \in \mathbb{P}_G$$

- → Same kernel of Wishart distribution
- $\longrightarrow \mathbb{P}_G = \{ K \text{ precision matrix } : K_{lm} = 0, \forall (i,j) \notin E \}$
- \longrightarrow $I_G(b,D)$ is, in general, not known in closed and tractable analytic form. More in Atay-Kayis and Massam (2005), Uhler et al. (2018) and Mohammadi et al. (2021)

$$\boldsymbol{K} \mid G \sim \text{G-Wishart}(b,D)$$

$$P(K \mid G) = I_G(b, D)^{-1} |K|^{\frac{b-2}{2}} \exp\left\{-\frac{1}{2}tr(KD)\right\}, K \in \mathbb{P}_G$$

- → Same kernel of Wishart distribution
- $\longrightarrow \mathbb{P}_G = \{ K \text{ precision matrix } : K_{lm} = 0, \forall (i,j) \notin E \}$
- \longrightarrow $I_G(b,D)$ is, in general, not known in closed and tractable analytic form. More in Atay-Kayis and Massam (2005), Uhler et al. (2018) and Mohammadi et al. (2021)

$$\boldsymbol{K} \mid G \sim \text{G-Wishart}(b,D)$$

$$P(K \mid G) = I_G(b, D)^{-1} |K|^{\frac{b-2}{2}} \exp\left\{-\frac{1}{2}tr(KD)\right\}, K \in \mathbb{P}_G$$

- → Same kernel of Wishart distribution
- $\longrightarrow \mathbb{P}_G = \{ K \text{ precision matrix } : K_{lm} = 0, \forall (i,j) \notin E \}$
- \longrightarrow $I_G(b,D)$ is, in general, not known in closed and tractable analytic form. More in Atay-Kayis and Massam (2005), Uhler et al. (2018) and Mohammadi et al. (2021)

$$\boldsymbol{K} \mid G \sim \text{G-Wishart}(b,D)$$

$$P(K \mid G) = I_G(b, D)^{-1} |K|^{\frac{b-2}{2}} \exp\left\{-\frac{1}{2}tr(KD)\right\}, K \in \mathbb{P}_G$$

- → Same kernel of Wishart distribution.
- $\longrightarrow \mathbb{P}_G = \{ K \text{ precision matrix } : K_{lm} = 0, \forall (i,j) \notin E \}.$
- \longrightarrow $I_G(b,D)$ is, in general, not known in closed and tractable analytic form. More in Atay-Kayis and Massam (2005), Uhler et al. (2018) and Mohammadi et al. (2021).

GRAPH PRIOR

Let $\mathcal G$ be the space of all possible undirected graphs having p nodes, hence $|\mathcal G|=2^{\binom{p}{2}}$

Uniform prior:

$$\pi_U(G) = \frac{1}{|\mathcal{G}|}$$

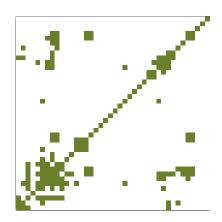
Bernoulli prior:

$$P((i,j) \in E) = \theta, 1 \le i < j \le p$$

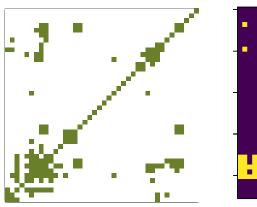
 $\pi_B(G) = \theta^{|E|} (1 - \theta)^{\binom{p}{2} - |E|}$

- → Assume a priori independence between the edges
- \longrightarrow Jones et al. (2005) suggest $\theta = 2/(p-1)$
- \longrightarrow Multiplicity correction prior: place a Beta hyperprior on θ (Scott and Carvalho 2008)

ESTIMATED QUANTITIES



ESTIMATED QUANTITIES



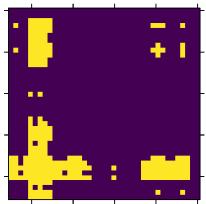
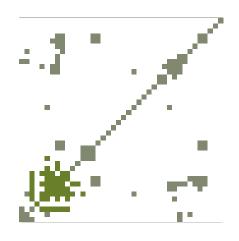
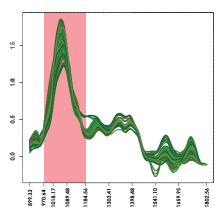


Figure 5 in Waghmare and Panaretos (2024)

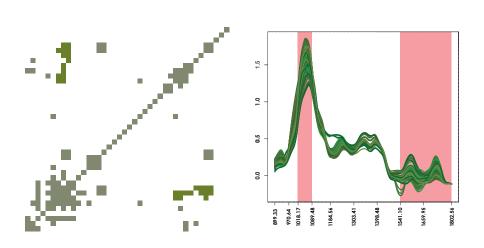
Kartik G Waghmare, Victor M Panaretos, Continuously indexed graphical models, Journal of the Royal Statistical Society Series B: Statistical Methodology, Volume 87, Issue 1, February 2025, Pages 211–231,

IDENTIFIED CONNECTIONS

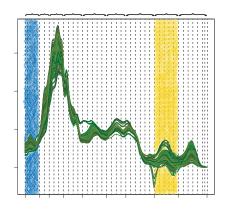




IDENTIFIED CONNECTIONS

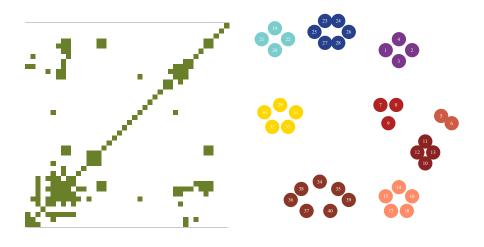


NODE PARTITION

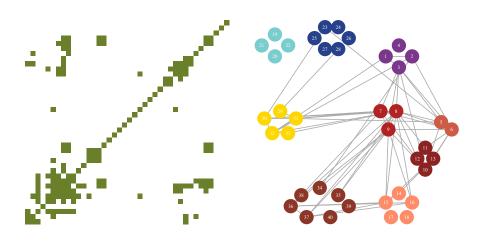


- Experts identified 9 intervals of scientific interest
- Nodes are partitioned in 9 contiguous sub-groups
- In Colombi et al. (2024), we:
 - → include such prior knowledge in the graphical model
 - → look for connections among the groups

NETWORK REPRESENTATION



NETWORK REPRESENTATION



LARGE-SCALE STRUCTURE

Following van den Boom et al. (2023)

- Currently: focus on single-edge estimation
- Next: set up a graphical model that also incorporates the identification of structures on a larger scale

We want to identify groups of nodes which form block structures in the graph

LARGE-SCALE STRUCTURE

Following van den Boom et al. (2023)

- Currently: focus on single-edge estimation
- Next: set up a graphical model that also incorporates the identification of structures on a larger scale

We want to identify groups of nodes which form block structures in the graph

- Let M be the (unknown) number of groups
- $z = (z_1, \ldots, z_p) \in \{1, \ldots, M\}^p : i \sim j \iff z_i = z_j$

$$P(G, K, z) \propto P(K \mid G)P(G \mid z)P(z)$$

STOCHASTIC BLOCK MODELS PRIOR

- Borrow ideas from network analysis
- SBM (Holland et al. 1983) are the most widely used models to discover latent structures

STOCHASTIC BLOCK MODELS PRIOR

- Borrow ideas from network analysis
- SBM (Holland et al. 1983) are the most widely used models to discover latent structures

Stochastic Equivalence

The probability of an edge only depends on the group membership of the nodes.

$$i \sim j \iff P((i,s) \in E) = P((j,s) \in E)$$

STOCHASTIC BLOCK MODELS PRIOR

- Borrow ideas from network analysis
- SBM (Holland et al. 1983) are the most widely used models to discover latent structures

Stochastic Equivalence

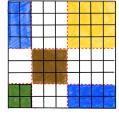
The probability of an edge only depends on the group membership of the nodes.

$$i \sim j \iff P((i,s) \in E) = P((j,s) \in E)$$

• Let Q be a $M \times M$ symmetric matrix, $Q_{rs} \in (0,1)$

$$P((i,j) \in E \mid \mathbf{z}, Q) \stackrel{\text{ind}}{=} Q_{z_i z_j}$$

$$Q_{rs} \stackrel{\text{ind}}{\sim} \text{Beta}(a,b)$$





PARTITION PRIOR

- We want to learn z and M from the data
- van den Boom et al. (2023) builds upon the degree-corrected SBM proposed in Tan and Iorio (2019)
- The Extended SBM framework (Legramanti et al. 2022) encompasses any Gibbs-type prior ranging from the *Infinite* relational model (Kemp et al. 2006) to the MFM-SBM (Geng et al. 2019)

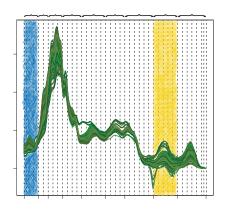
PARTITION PRIOR

- We want to learn z and M from the data
- van den Boom et al. (2023) builds upon the degree-corrected SBM proposed in Tan and Iorio (2019)
- The Extended SBM framework (Legramanti et al. 2022) encompasses any Gibbs-type prior ranging from the *Infinite* relational model (Kemp et al. 2006) to the MFM-SBM (Geng et al. 2019)

Node Exchangeability

Gibbs-type priors assume that the nodes can be arbitrarily relabeled

NODE PARTITION



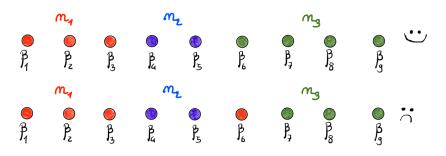
- Each node represents a band of the spectrum
- The bands have a precise meaning and a natural ordering
- We work under the constraint that the blocks must be contiguous

ADMISSIBLE PARTITIONS PRIOR

- Martínez and Mena (2014) introduced the notion of admissible partitions
- Let $\rho_p = \{C_1, \dots, C_M\}$ be a partition of p nodes in M groups \rightarrow where $C_m = \{j: z_j = m\}$
- ρ_p is admissible $\iff \forall x \in C_i, y \in C_j, i < j \text{ implies } x < y$

ADMISSIBLE PARTITIONS PRIOR

- Martínez and Mena (2014) introduced the notion of admissible partitions
- Let $\rho_p = \{C_1, \dots, C_M\}$ be a partition of p nodes in M groups \rightarrow where $C_m = \{j : z_j = m\}$
- ρ_p is admissible $\iff \forall x \in C_i, y \in C_j, i < j \text{ implies } x < y$



ADMISSIBLE PARTITIONS PRIOR

$$P\left(\rho_{p}\right) = p'(n_{1},\ldots,n_{M}) = \begin{cases} \binom{n}{n_{1},\ldots,n_{M}} \frac{1}{M!}p(n_{1},\ldots,n_{M}), & \rho_{p} \text{ admissible} \\ 0, & \rho_{p} \text{ not admissible} \end{cases}$$

- $p(n_1, \ldots, n_M)$ can be any eppf
- Employing a Pitman–Yor process (Perman et al. 1992)

$$\begin{split} p'(n_1,\ldots,n_M) \\ &= \frac{p!}{M!} \frac{\prod_{i=1}^{M-1} (\theta+i\sigma)}{(\theta+1)_{(p-1)}} \prod_{i=1}^{M} \frac{(1-\sigma)_{(n_j-1)}}{n_j} \mathbf{1}_{(\rho_p \text{ is admissible})} \end{split}$$

GIBBS SAMPLING

 \longrightarrow Exploit conjugacy to get rid of Q

$$P(\boldsymbol{G}\mid\boldsymbol{z}) = \prod_{u=1}^{M} \prod_{v=u}^{M} \frac{B(\alpha + S_{uv}, \beta + S_{uv}^{\star})}{B(\alpha,\beta)}$$

GIBBS SAMPLING

 \longrightarrow Exploit conjugacy to get rid of Q

$$P(\mathbf{G} \mid \mathbf{z}) = \prod_{u=1}^{M} \prod_{v=u}^{M} \frac{B(\alpha + S_{uv}, \beta + S_{uv}^{\star})}{B(\alpha, \beta)}$$

→ Sampling Graph and Precision Matrix

G and K, given z, are sampled using a modified version of a Birth-and-Death chain (Mohammadi and Wit (2015))

Birth rate
$$\propto \frac{P(\mathbf{G}^{+e} \mid \mathbf{z})}{P(\mathbf{G} \mid \mathbf{z})} = \frac{S_{uv} + \alpha}{S_{uv}^{\star} + \beta}$$

Death rate $\propto \frac{P(\mathbf{G}^{-e} \mid \mathbf{z})}{P(\mathbf{G} \mid \mathbf{z})} = \frac{S_{uv}^{\star} + \beta}{S_{uv} + \alpha}$

GIBBS SAMPLING

 \longrightarrow Exploit conjugacy to get rid of Q

$$P(\boldsymbol{G}\mid\boldsymbol{z}) = \prod_{u=1}^{M} \prod_{v=u}^{M} \frac{B(\alpha + S_{uv}, \beta + S_{uv}^{\star})}{B(\alpha,\beta)}$$

→ Sampling Graph and Precision Matrix

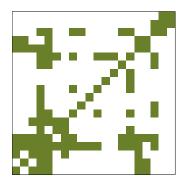
G and K, given z, are sampled using a modified version of a Birth-and-Death chain (Mohammadi and Wit (2015))

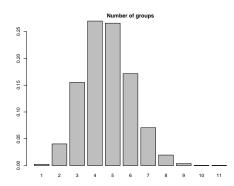
$$\begin{aligned} & \text{Birth rate} \propto \frac{P(\,\boldsymbol{G}^{+e} \mid \boldsymbol{z})}{P(\,\boldsymbol{G} \mid \boldsymbol{z})} = \frac{S_{uv} + \alpha}{S_{uv}^{\star} + \beta} \\ & \text{Death rate} \propto \frac{P(\,\boldsymbol{G}^{-e} \mid \boldsymbol{z})}{P(\,\boldsymbol{G} \mid \boldsymbol{z})} = \frac{S_{uv}^{\star} + \beta}{S_{uv} + \alpha} \end{aligned}$$

→ Sampling the Random Partition

Conditionally on G, we can sample z through an adaptive *split and merge* sampler (Benson et al. (2018))

RESULTS





- 4 groups detected
- 5 groups is also supported by the data
- The case of one single community has almost zero posterior probability mass
 - → We are improving the previous analysis of Codazzi et al. (2022) and Colombi et al. (2024)

WRAPPING UP

- Graphical representation of the conditional independence structure of the basis expansion coefficients
- → Novel framework for SBM prior in Gaussian graphical model with ordered nodes
- Improved interpretability of short- and long-term interactions among portions of the spectrum

What's next

 \longrightarrow Achieving better scaling as the number of nodes increases



Codazzi et al. (2022)



Colombi et al. (2024)



Working paper

REFERENCES

- Amongero M, De Blasi P (2024) A gibbs sampler for community detection in assortative stochastic block model. In: Scientific Meeting of the Italian Statistical Society, Springer, pp 50–55
- Atay-Kayis A, Massam H (2005) A Monte Carlo method for computing the marginal likelihood in nondecomposable Gaussian graphical models. Biometrika 92:317–335
- Benson A, Friel N, et al. (2018) Adaptive mcmc for multiple changepoint analysis with applications to large datasets. Electronic Journal of Statistics 12(2):3365–3396
- van den Boom W, De Iorio M, Beskos A (2023) Bayesian learning of graph substructures. Bayesian Analysis 18(4):1311-1339
- Codazzi L, Colombi A, Gianella M, Argiento R, Paci L, Pini A (2022) Gaussian graphical modeling for spectrometric data analysis. Comput Statist Data Anal
- Colombi A, Argiento R, Paci L, Pini A (2024) Learning block structured graphs in gaussian graphical models. Journal of Computational and Graphical Statistics 33(1):152–165
- Geng J, Bhattacharya A, Pati D (2019) Probabilistic community detection with unknown number of communities. Journal of the American Statistical Association 114(526):893–905
- Holland PW, Laskey KB, Leinhardt S (1983) Stochastic blockmodels: First steps. Social Networks 5(2):109-137
- Jones B, Carvalho C, Dobra A, Hans C, Carter C, West M (2005) Experiments in stochastic computation for high-dimensional graphical models. Statistical Science 20:388–400
- Kemp C, Tenenbaum JB, Griffiths TL, Yamada T, Ueda N (2006) Learning systems of concepts with an infinite relational model. In: AAAI, vol 3, p 5
- Lang S, Brezger A (2004) Bayesian P-splines. Journal of Computational and Graphical Statistics 13(1):183–212
- Larremore D, Clauset A, Buckee C (2013) A network approach to analyzing highly recombinant malaria parasite genes. PLoS computational biology 9
- Lee C, Wilkinson D (2019) A review of stochastic block models and extensions for graph clustering. Applied Network Science 4

REFERENCES

- Legramanti S, Rigon T, Durante D, Dunson DB (2022) Extended stochastic block models with application to criminal networks. The annals of applied statistics 16(4)
- Martínez AF, Mena RH (2014) On a Nonparametric Change Point Detection Model in Markovian Regimes. Bayesian Analysis 9(4):823-858, DOI 10.1214/14-BA878
- Mohammadi A, Wit EC (2015) Bayesian structure learning in sparse Gaussian graphical models. Bayesian Analysis 10(1):109–138
- Mohammadi R, Massam H, Letac G (2021) Accelerating bayesian structure learning in sparse gaussian graphical models. J Amer Statist Assoc 0(0):1–14
- Perman M, Pitman J, Yor M (1992) Size-biased sampling of poisson point processes and excursions. Probability Theory and Related Fields 92(1):21–39
- Roverato A (2002) Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. Scandinavian Journal of Statistics 29(3):391-411
- Scott J, Carvalho C (2008) Feature-inclusion stochastic search for Gaussian graphical models. Journal of Computational and Graphical Statistics 17(4):790–808
- Tan LS, Iorio MD (2019) Dynamic degree-corrected blockmodels for social networks: A nonparametric approach. Statistical Modelling 19(4):386–411
- Uhler C, Lenkoski A, Richards D (2018) Exact formulas for the normalizing constants of Wishart distributions for graphical models. The Annals of Statistics 46(1):90-118
- Waghmare KG, Panaretos VM (2024) Continuously indexed graphical models. Journal of the Royal Statistical Society Series B: Statistical Methodology 87(1):211–231
- Yang J, Cox DD, Lee JS, Ren P, Choi T (2017) Efficient bayesian hierarchical functional data analysis with basis function approximations using Gaussian–Wishart processes. Biometrics 73(4):1082–1091

"SBM are employed to cluster the nodes into groups with sparser inter-group connections compared to more dense within-group connectivities"

"SBM are employed to cluster the nodes into groups with sparser inter-group connections compared to more dense within-group connectivities"

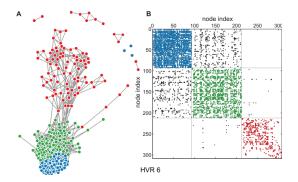


Figure 5 in Larremore et al. (2013)

 "The concept of stochastic equivalence in itself does not require that the nodes in the same group are more connected with themselves, than with nodes in other groups" (Lee and Wilkinson 2019)

- "The concept of stochastic equivalence in itself does not require that the nodes in the same group are more connected with themselves, than with nodes in other groups" (Lee and Wilkinson 2019)
- \bullet The diagonal elements of ${\it Q}$ are not necessarily higher than the off-diagonal ones

- "The concept of stochastic equivalence in itself does not require that the nodes in the same group are more connected with themselves, than with nodes in other groups" (Lee and Wilkinson 2019)
- \bullet The diagonal elements of ${\it Q}$ are not necessarily higher than the off-diagonal ones
- SBM are a general framework that accommodates for many latent structure such as dis-assortative, core-periphery, weak-community (Legramanti et al. 2022)
- See Amongero and De Blasi (2024) for a Bayesian approach for Assortative SMB

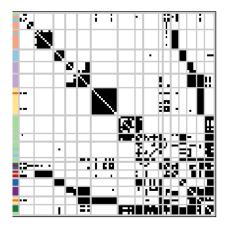


Figure 2c in Legramanti et al. (2022)

Sirio Legramanti, Tommaso Rigon, Daniele Durante, David B. Dunson Extended stochastic block models with application to criminal networks, *The Annals of Applied Statistics*