Graphical Gaussian Models with Symmetries and Regular Colourings

Helene Gehrmann
Department of Statistics, University of Oxford

WOGAS3, CRiSM
University of Warwick

Overview

1. Graphical Gaussian Models with Symmetries

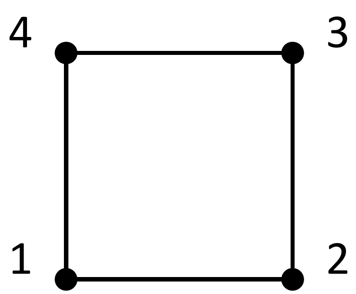
2. Need for Model Selection Methods Motivates Five Questions

3. Examples

Graphical Gaussian Models

- Concerned with the distribution of a multivariate Gaussian random vector
- Encode the independence structure in terms of edges in an undirected graph G=(V,E): vertices V = model variables, edges E defined by relation:

$$Y_\alpha \perp Y_\beta \mid Y_{V \setminus \{\alpha,\beta\}} \quad \Rightarrow \quad \alpha\beta \notin E$$

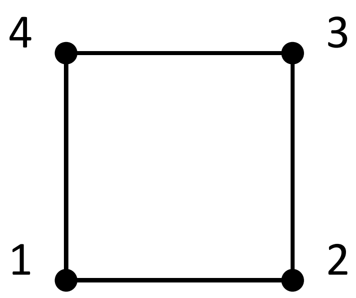


$$V = \{1,2,3,4\}$$
$$(Y_1, Y_2, Y_3, Y_4) \sim N_4(0, \Sigma)$$

Graphical Gaussian Models

- Concerned with the distribution of a multivariate Gaussian random vector
- Encode the independence structure in terms of edges in an undirected graph G=(V,E): vertices V = model variables, edges E defined by relation:

$$Y_\alpha \perp Y_\beta \mid Y_{V\setminus\{\alpha,\beta\}} \quad \Leftrightarrow \quad (\Sigma^{-1})_{\alpha\beta} = 0 \quad \Rightarrow \quad \alpha\beta \notin E$$

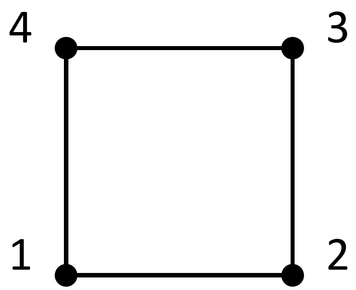


$$V = \{1,2,3,4\}$$
$$(Y_1, Y_2, Y_3, Y_4) \sim N_4(0, \Sigma)$$

Graphical Gaussian Models

- Concerned with the distribution of a multivariate Gaussian random vector
- Encode the independence structure in terms of edges in an undirected graph G=(V,E): vertices V = model variables, edges E defined by relation:

$$Y_\alpha \perp Y_\beta \mid Y_{V \setminus \{\alpha,\beta\}} \quad \Leftrightarrow \quad (\Sigma^{-1})_{\alpha\beta} = 0 \quad \Rightarrow \quad \alpha\beta \notin E$$



$$V = \{1,2,3,4\}$$

$$(Y_1, Y_2, Y_3, Y_4) \sim N_4(0, \Sigma)$$

$$\Sigma^{-1} = \begin{pmatrix} * & * & 0 & * \\ & * & * & 0 \\ & & * & * \\ & & & * \end{pmatrix}$$

positive definite

Graphical Gaussian Models

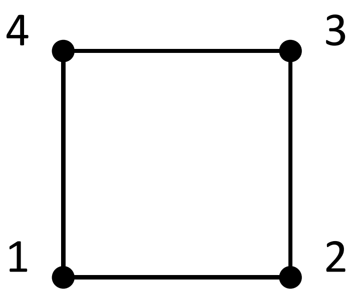
- Concerned with the distribution of a multivariate Gaussian random vector
- Encode the independence structure in terms of edges in an undirected graph G=(V,E): vertices V = model variables, edges E defined by relation:

$$Y_\alpha \perp Y_\beta \mid Y_{V\setminus\{\alpha,\beta\}} \quad \Leftrightarrow \quad (\Sigma^{-1})_{\alpha\beta} = 0 \quad \Rightarrow \quad \alpha\beta \notin E$$

- If |V|=p, then there are up to |V|+|E| = p(p+1)/2 parameters in the model (for mean = 0)
- Parsimony in number of parameters can be achieved through
  - ... sparsity in graph
  - ... symmetry constraints on model parameters

Graphical Gaussian Models with Symmetries

- Højsgaard and Lauritzen (2008):
  - RCON models: equality constraints on concentrations
  - RCOR models: equality constraints on partial correlations

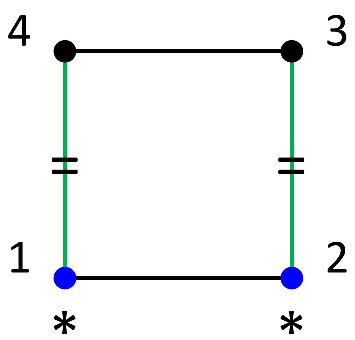


$$V = \{1,2,3,4\}$$
$$(Y_1, Y_2, Y_3, Y_4) \sim N_4(0, \Sigma)$$

$$K = \Sigma^{-1} = \begin{pmatrix} * & * & 0 & * \\ & * & * & 0 \\ & & * & * \\ & & & * \end{pmatrix} \quad \text{pd}$$

Graphical Gaussian Models with Symmetries

- Højsgaard and Lauritzen (2008):
  - RCON models: equality constraints on concentrations
  - RCOR models: equality constraints on partial correlations



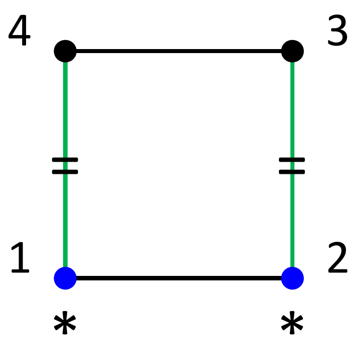
$V = \{1,2,3,4\}$

$(Y_1, Y_2, Y_3, Y_4) \sim N_4(0, \Sigma)$

$$K = \Sigma^{-1} = \begin{pmatrix} a & * & 0 & b \\ & a & b & 0 \\ & & * & * \\ & & & * \end{pmatrix} \quad \text{pd}$$

Graphical Gaussian Models with Symmetries

- Højsgaard and Lauritzen (2008):
  - RCON models: equality constraints on concentrations
  - RCOR models: equality constraints on partial correlations



$$V = \{1,2,3,4\}$$

$$(Y_1, Y_2, Y_3, Y_4) \sim N_4(0, \Sigma)$$

$$K = \Sigma^{-1} = \begin{pmatrix} a & * & 0 & * \\ & a & * & 0 \\ & & * & * \\ & & & * \end{pmatrix} \quad \text{pd}$$

$$c_{\alpha\beta} = -\rho_{\alpha\beta|V\setminus\{\alpha,\beta\}} = \frac{k_{\alpha\beta}}{\sqrt{k_{\alpha\alpha}k_{\beta\beta}}}$$

$$c_{14} = c_{23}$$

partial correlations

Graphical Gaussian Models with Symmetries

- Højsgaard and Lauritzen (2008):
  - RCON models: equality constraints on concentrations
    - Place linear constraints on natural parameter K of exponential family
    - Linear exponential families $\Longrightarrow$ MLE unique whenever it exists
  - RCOR models: equality constraints on partial correlations
    - Constrains not necessarily linear in natural parameter K
    - Curved exponential families $\Longrightarrow$ MLE not necessarily unique
    - Scale-invariance within vertex colour classes
  - MLE computation algorithms described in Højsgaard and Lauritzen (2008).

Graphical Gaussian Models with Symmetries

In order to make RCON and RCOR models widely applicable, model selection methods are required:

Q1: What is the structure of the set of RCON and RCOR models for a given V?

Q2: Can we design efficient model selection algorithms for RCON/RCOR models?

Q3: Are there statistically interesting model sub-classes?

Q4: If so, what is their structure?

Q5: Are they (better) suited for model selection?

RCON: symmetries in concentrations RCOR: symmetries in partial correlations

Q1: Structure of RCON and RCOR Models

Q1: What is the structure of the set of RCON and RCOR models for a given V?

- Graphical representation: coloured graph = uncoloured graph + colouring
  - Obtain larger model through ...
    - ... larger uncoloured graph (fewer 0's in K)
    - ... finer colouring (fewer symmetries)
  - Can go from any model to any model:
    - Moving up: add edge colour classes + split colour classes
    - Moving down: drop edge colour classes + merge colour classes

RCON: symmetries in concentrations

RCOR: symmetries in partial correlations

Q1: Structure of RCON and RCOR Models

> Q1: What is the structure of the set of RCON and RCOR models for a given V?

- Graphical representation: coloured graph = uncoloured graph + colouring
    - Obtain larger model through ...
        - ... larger uncoloured graph (fewer 0's in K)
        - ... finer colouring (fewer symmetries)
    - Can go from any model to any model:
        - Moving up: add edge colour classes + split colour classes
        - Moving down: drop edge colour classes + merge colour classes
    - Examining the graphical representation a bit closer: complete lattices!
        - Any two models have a unique supremum and infimum.
        - Smallest model = empty graph, all vertices same colour
        - Largest model = complete uncoloured graph

RCON: symmetries in concentrations

RCOR: symmetries in partial correlations

Q2: Model Selection in RCON and RCOR Models

Q2: Can we design efficient model selection algorithms for RCON/RCOR models?

- Stepwise search theoretically possible however very large search space: $2^{|V|}$ uncoloured models on V, colouring enlarges model space considerably!

- Edwards-Havránek model selection procedure for lattices:
  - Whenever a model is accepted, all supermodels accepted.
  - Whenever a model is rejected, all submodels rejected.
  - Returns minimally accepted models.

- Possible .... but:
  - Model space still large.
  - Models not always intuitively interpretable: properties of found model(s)?

RCON: symmetries in concentrations

RCOR: symmetries in partial correlations

Q3: Interesting Model Sub-Classes

Q3: Are there Statistically interesting model sub-classes?

RCON: symmetries in concentrations

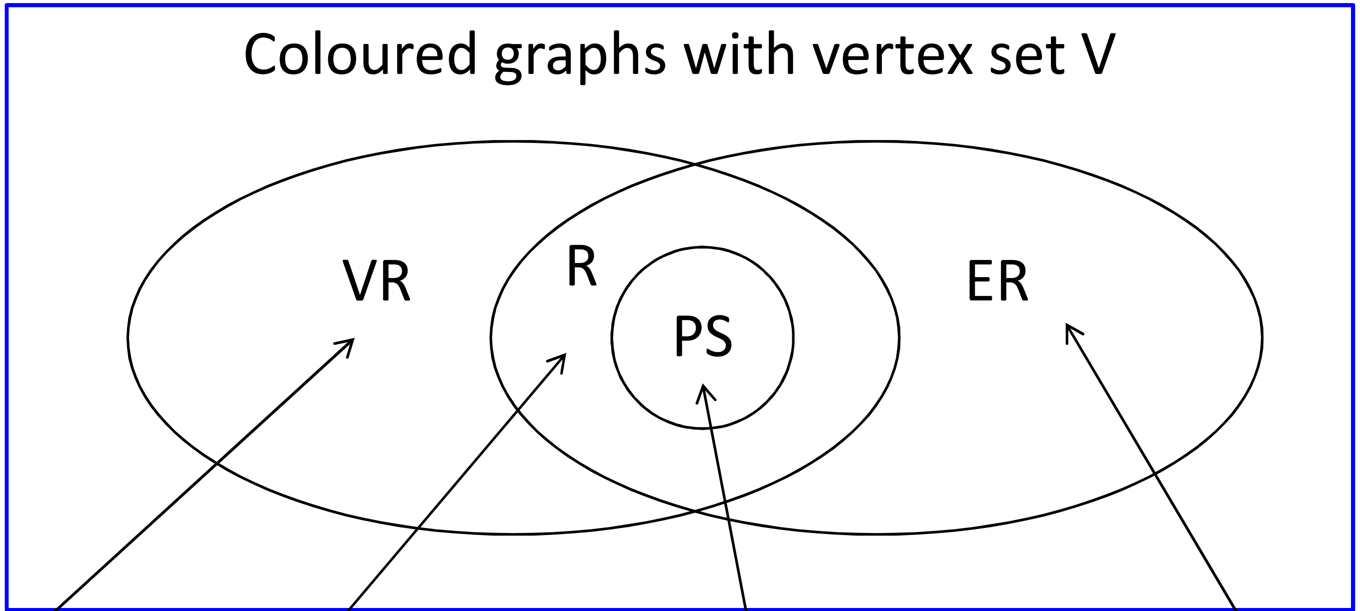
RCOR: symmetries in partial correlations

Q3: Interesting Model Sub-Classes

Q3: Are there Statistically interesting model sub-classes? YES!

They can be identified by their colouring:

(H&L, 2008; G&L, 2011)



Vertex regularity: estimability of mean

Regularity: vertex & edge regularity

Permutation symmetry: simplified MLE computation

Edge regularity: linear exponential families, scale-invariant (RCON = RCOR)

RCON: symmetries in concentrations

RCOR: symmetries in partial correlations

Q4 & Q5: Structure of Model Sub-Classes

Q4: If so, what is their structure?

RCON: symmetries in concentrations RCOR: symmetries in partial correlations

Q4 & Q5: Structure of Model Sub-Classes

Q4: If so, what is their structure?

- Each of ER, PG, VR and R is a complete lattice.
  - Any two models which lie in ER/PG/VR/R, they have unique infimum & supremum in the same model class. (non-trivial!)
  - Gehrmann (2011)

RCON: symmetries in concentrations

RCOR: symmetries in partial correlations

Q4 & Q5: Structure of Model Sub-Classes

Q4: If so, what is their structure?

- Each of ER, PG, VR and R is a complete lattice.
  - Any two models which lie in ER/PG/VR/R, they have unique infimum & supremum in the same model class. (non-trivial!)
  - Gehrmann (2011)

Q5: Are they (better) suited for model selection? YES!

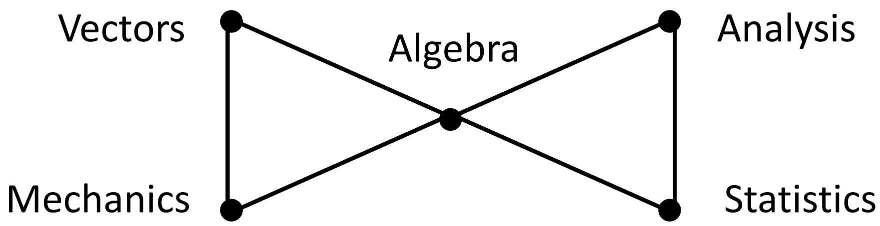
- Each model class qualifies for an Edwards-Havránek model search.
  - Faster than search in RCON/RCOR models.
  - Found model(s) guaranteed to have desirable properties.
  - Optionally: subsequent local search within RCON/RCOR models.

RCON: symmetries in concentrations RCOR: symmetries in partial correlations

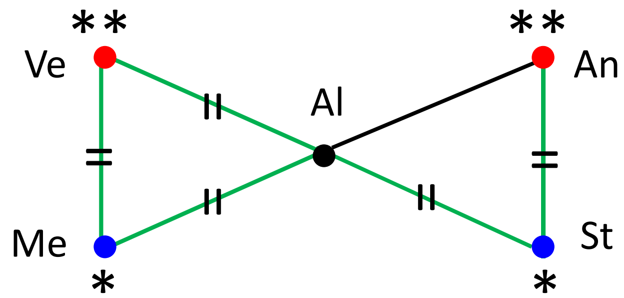
Example: Edge Regular Search

- Data: Examination marks of 88 students in 5 mathematical subjects (Mardia et al. 1979)

- Whittaker (1990) and Edwards (2000) :



- Højsgaard and Lauritzen (2008), RCON model:

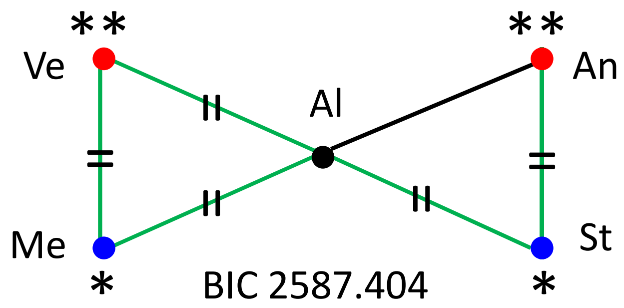


RCON: symmetries in concentrations

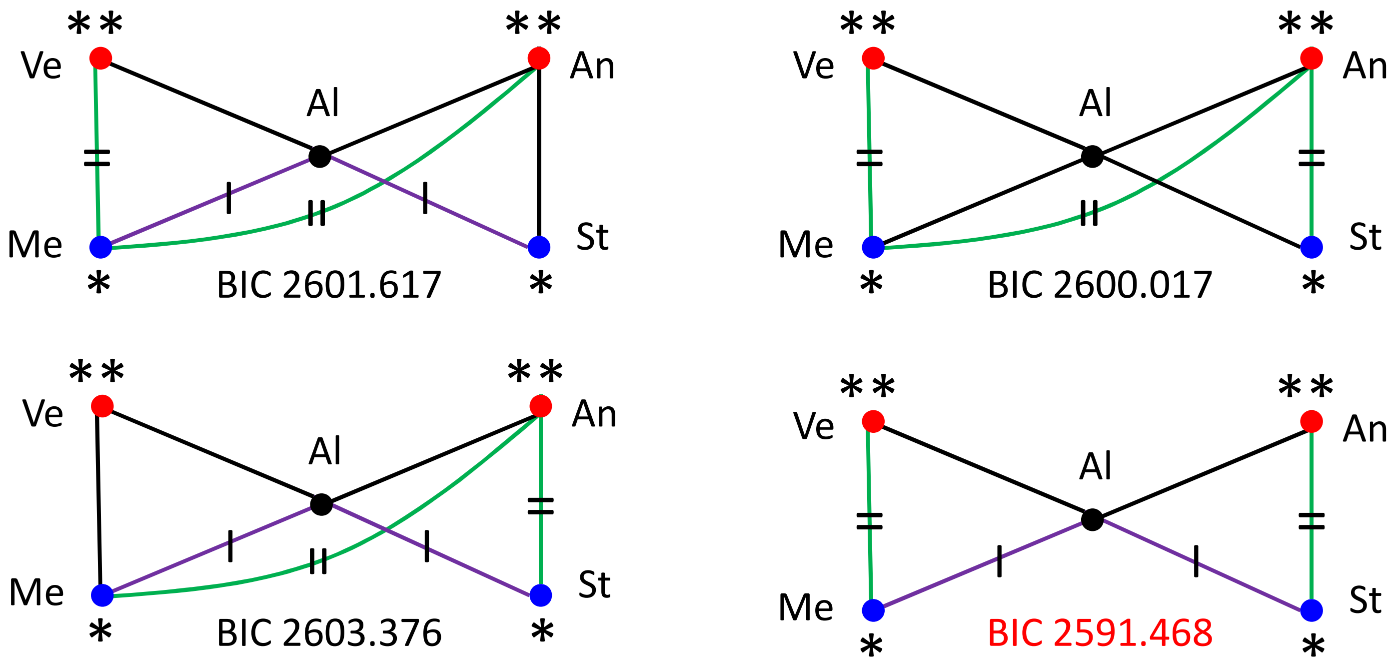
RCOR: symmetries in partial correlations

Example: Edge Regular Search

- Højsgaard and Lauritzen (2008) :



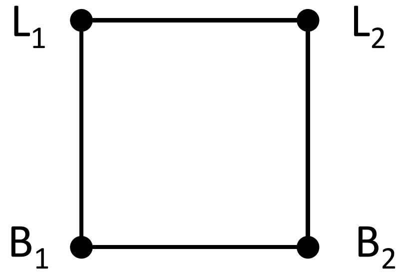
- Edwards-Havránek search, starting at saturated model, after 232 models:



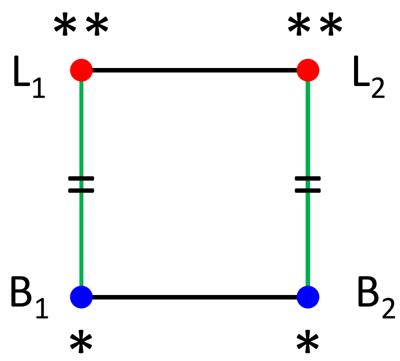
Example: Permutation Generated Search

- Data: Head dimensions of 25 pairs of 1st and 2nd sons, Frets (1921)

- Whittaker (1990):

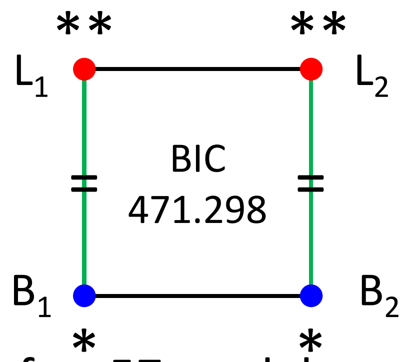


- Højsgaard and Lauritzen (2008),
  RCOP model (permutation of sons):

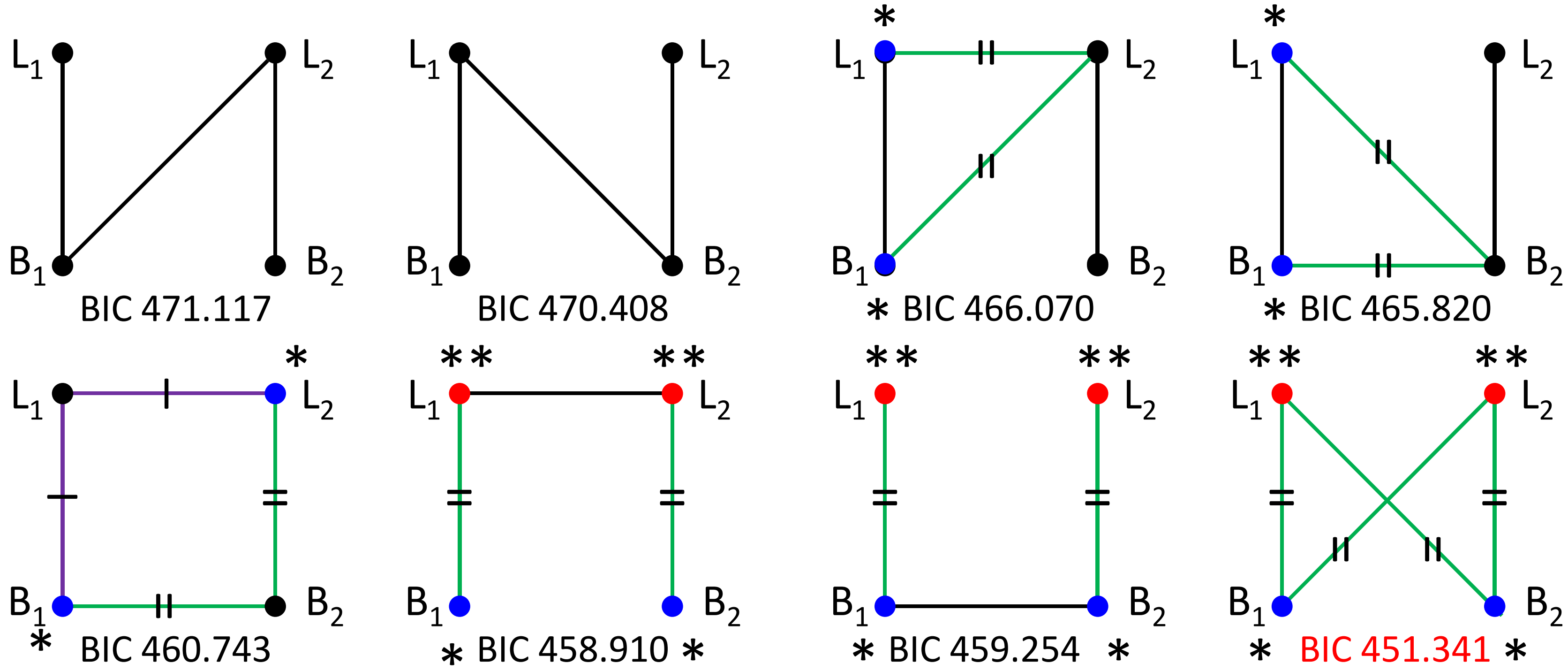


Example: Permutation Generated Search

- Højsgaard and Lauritzen (2008) :



- Edward-Havránek search, starting at saturated model, after 57 models:



Summary

- Model selection in RCON and RCOR models can be performed by reducing search to lattices of models represented by regular colourings.
- These are first model selection procedures for GGMs with symmetries.
- Examples suggest that search may be feasible in general.
- Next step: full implementation and performance analysis.

Thank you!

References

Edwards, D. *Introduction to Graphical Modelling.* Springer Verlag: New York, NY, USA, 2000.

Edwards, D. and Havránek, T. A fast model selection procedure for large families of models. J*ournal of the American Statistical Association.* 1987, 82: 205-213.

Frets, G.P. Heredity of head form in man. *Genetica*. 1921, 41: 193-400.

Gehrmann, H. and Lauritzen, S.L. (2011). Estimation of means in graphical Gaussian models with symmetries. Preprint available at http://arxiv.org/abs/1101.3709.

Gehrmann, H. (2011) Lattices of graphical Gaussian models with symmetries. Preprint available at http://arxiv.org/abs/1104.1608.

Højsgaard, S. and Lauritzen, S.L. (2008). Graphical Gaussian models with edge and vertex symmetries. *Journal of the Royal Statistical Society Series B. 70: 1*005-1027.

Mardia, K. V., Kent, J. T. and Bibby, J. M. *Multivariate Analysis.* Academic Press: New York, NY, USA, 1979.

Whittaker, J. *Graphical Models in Applied Multivariate Statistics*. Wiley: Chichester, UK, 1990.