Model inclusion lattice of coloured Gaussian graphical models for paired data

Alberto Roverato alberto.roverato@unipd.it

Dung Ngoc Nguyen ngocdung.nguyen@unipd.it

Department of Statistical Sciences
University of Padova (Italy)

Abstract

We consider the problem of learning a graphical model when the observations come from two groups sharing the same variables but, unlike the usual approach to the joint learning of graphical models, the two groups do not correspond to different populations and therefore produce dependent samples. A Gaussian graphical model for paired data may be implemented by applying the methodology developed for the family of graphical models with edge and vertex symmetries, also known as coloured graphical models. We identify a family of coloured graphical models suited for the paired data problem and investigate the structure of the corresponding model space. More specifically, we provide a comprehensive description of the lattice structure formed by this family of models under the model inclusion order. Furthermore, we give rules for the computation of the join and meet operations between models, which are useful in the exploration of the model space. These are then applied to implement a stepwise model search procedure and an application to the identification of a brain network from fMRI data is given.

Keywords: Brain network, lattice, model search, poset, RCON model, search space.

1. Introduction

Statistical models associated with graphs, called graphical models, have become a popular tool for representing network structures in many modern applications; see Maathuis et al. (2019) for a recent collection of reviews. Let Y_V be a multivariate Gaussian random vector whose entries are indexed by a finite set $V = \{1, \dots, p\}$ and let $G = (V, E)$ be an undirected network. Every vertex of the graph G is associated with a variable in Y_V, and the distribution of Y_V is said to be Markov with respect to G if every missing edge between two vertices implies that the corresponding variables are conditionally independent given the remaining variables (Lauritzen, 1996). The Gaussian graphical model (GGM) for Y_V, represented by a graph $G = (V, E)$, is the family of Gaussian distributions which are Markov relative to G.

The seminal paper by Højsgaard and Lauritzen (2008) introduced GGMs with additional symmetry restrictions in the form of equality constraints on the parameters. These are called coloured GGMs because symmetries are usually depicted on the dependence graph of the model by colouring of edges and vertices. The application of coloured GGMs was motivated by the need of reducing the number of parameters in high dimensional settings. On the other hand, there exist applied contexts where symmetry restrictions naturally follow from substantive research hypotheses of interest. This is the case of paired data problems where variables can be naturally arranged into two blocks such that for every variable in the first block there is an homologous variable in the second block.

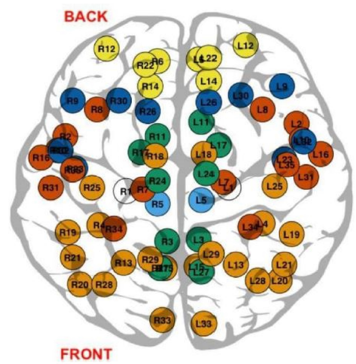


Figure 1: Example of ROI locations on the brain, every ROI on the left hemisphere is associated with a ROI on the right hemisphere, that in this case gives the pairs $(L_i; R_i)$ for $i = 1, \ldots, 35$. Different colours correspond to distinct brain regions.

Example 1 (Brain network) *Ranciati et al. (2021) considered the problem of identifying symmetries in brain networks from fMRI data. In this type of application, every variable is measured on a spatial regions of interest (ROI) and for every ROI on the left hemisphere of the brain there is a corresponding, or homologous, ROI on the right hemisphere; see Figure 1. Interest is for symmetries concerning both the structure of the brain network and the values of parameters associated with vertices and edges. This type of analysis involves the comparison of the two subnetworks relative to the left and right hemispheres, but also the comparison of pairs of edges across the two hemispheres in the case where the endpoins of the first edge are homologous to the endpoints of the second. See Sections 3 and 5 for details.*

In the following we provide a smaller instance of paired data problem that will be used as a running example throughout this paper.

Example 2 (Frets' Heads) *Whittaker (1990, Section 8.4) fitted a GGM to the Frets' Heads data which consist of measurements of the* head length*, or* height*, and* head breadth *of the first and second adult sons in a sample of 25 families. Variables are therefore naturally split into two blocks associated with the first son, denoted by H and B, and second son, denoted by H' and B'. Symmetries of interest concern, for instance, the between-block comparison of (H, B) with (H', B') as well as the across-block comparison (H, B') with (B, H'); see also Figure 2.*

The problem of learning a coloured GGM from data has been approached from different viewpoints. Li et al. (2021) and Ranciati et al. (2021) developed penalized likelihood methods whereas Li et al. (2018), Massam et al. (2018) and Li et al. (2020) implemented Bayesian methods. Finally, Gehrmann (2011) considered classical stepwise coherent procedures. However, Gehrmann (2011) mostly focused on an analysis of the structure of search spaces, that is of the spaces of candidate models. Indeed, many model search algorithms, both in the Bayesian and in the frequentist approaches to inference, require the exploration of the search space that is typically carried out by means of local moves between neighbouring models. It is therefore crucial to be able to rely on procedures that allow us to explore efficiently the space of models. However, the exploration of the space of coloured GGMs is much more challenging than for classical GGMs. One first difficulty concerns the dimensionality of the search space, which highly increases (Gehrmann, 2011). As an example, consider the complete graph on p vertices: there is only one GGM represented by this graph but the number of coloured GGMs for paired data, as formally defined in Section 3 below, is equal to $2^{(p/2)^2}$. Furthermore, it is well-established that the space of classical GGMs forms a complete distributive lattice with respect to model inclusion, where neighbouring models can be efficiently obtained by adding or removing single edges from the current graph. On the other hand, Gehrmann (2011) showed that, although also the family of coloured GGMs forms a complete lattice, in this case the identification of neighbouring structures is considerably more complex, and thus more computationally expensive, than in GGMs.

In this paper we consider the family of coloured GGMs called RCON models (Højsgaard and Lauritzen, 2008), where equality restrictions are imposed on the entries of the inverse covariance matrix of Y_V. Within this framework, we focus on RCON models for paired data (PD-RCON) and analyse the structure of the associated search space. More specifically, we show that, unlike other relevant subfamilies of RCON models, PD-RCON models form a sublattice of the lattice of RCON models. Furthermore, we deal with the implementation of the meet and join operations over the sublattice, which are useful for the computation of neighbouring structures, and provide explicit rules for the computation of the meet and the join of two PD-RCON models. These are considerably more efficient than the corresponding operations for arbitrary RCON models because they are based on simple set union and set intersection operations. We then apply these rules to implement a stepwise model search procedure and an application to the identification of a brain network from fMRI data is provided.

2. Background

2.1 Coloured graphical models

Let $Y \equiv Y_V$ be a continuous random vector indexed by a finite set $V = \{1, \ldots, p\}$. We denote by $\Sigma = (\sigma_{ij})_{i,j \in V}$ and $\Sigma^{-1} = \Theta = (\theta_{ij})_{i,j \in V}$ the covariance and the concentration matrix of Y, respectively. An *undirected graph* with vertex set V is a pair $G = (V, E)$ where E is an edge set that is a set of unordered pairs of distinct vertices. Notice that, when not clear from the context which graph is under consideration, we write V_G to denote the vertex set of G and, similarly, E_G for edges.

We say that the concentration matrix Θ is adapted to a graph $G = (V, E)$ if every missing edge of G corresponds to a zero entry in Θ; formally, $(i, j) \notin E$ implies that $\theta_{ij} = 0$ for any $i, j \in V$ with $i \neq j$. A *Gaussian graphical model* (GGM) with graph G is the family of multivariate normal distributions whose concentration matrix is adapted to G. These models are also known with the name of covariance selection models or concentration graph models; see (Lauritzen, 1996) and references therein. We denote by $M \equiv M(V)$ the family of GGMs for Y_V and by $M(G) \in M$ the GGM represented by the graph $G = (V, E)$.

A *colouring* of $G = (V, E)$ is a pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{V_1, \ldots, V_v\}$ is a partition of V into vertex colour classes and, similarly, $\mathcal{E} = \{E_1, \ldots, E_e\}$ is a partition of E into edge colour classes. Accordingly, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a *coloured graph*. Similarly to the notation used for uncoloured graphs, we may write $\mathcal{V}_\mathcal{G}$ and $\mathcal{E}_\mathcal{G}$ to denote the vertex and edge colour classes of $\mathcal{G}$, respectively. Furthermore, we set $V_\mathcal{G} = \cup_{j=1}^{v} V_j$ and $E_\mathcal{G} = \cup_{j=1}^{e} E_j$ so that $(V_\mathcal{G}, E_\mathcal{G})$ is the *uncoloured version* of $\mathcal{G}$. In the graphical representation, all the vertices belonging to a same colour class are depicted of the same colour, and similarly for edges. Furthermore, in order to make coloured graphs readable also in black and white printing, we put a common symbol next to every vertex or edge of the same colour. The only exception to this rule is for vertices and edges belonging to colour classes with a single element, called *atomic*, which are all depicted in black with no symbol next to them.

Højsgaard and Lauritzen (2008) introduced *coloured graphical models* which are GGMs with additional restrictions on the parameter space. The model are represented by coloured graphs, where parameters that are associated with edges or vertices of the same colour are restricted to being identical. In this paper we focus on the family of coloured graphical Gaussian models known as RCON models (Højsgaard and Lauritzen, 2008) which place equality restrictions on the entries of the concentration matrix. More specifically, in the RCON model with coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ every vertex colour class V_i, $i = 1, \ldots, v$, identifies a set of diagonal concentrations whose value is constrained to be equal, and similarly for edge colour classes which identify subsets of off-diagonal concentrations. We denote by $\mathcal{M} \equiv \mathcal{M}(V)$ the family of RCON models for Y_V and by $\mathcal{M}(\mathcal{G}) \in \mathcal{M}$ the RCON model represented by $\mathcal{G}$.

2.2 Search spaces and lattices

Typically, a model search procedure requires the exploration of the search space, that is of the space of candidate models. A statistical model is a family of probability distributions, and if a model is contained in another model then the former is called a submodel of the latter. We can also say that a model is "larger" than any of its submodels so that *model inclusion* can be used to embed a model class with a partial order, and we write $\langle M, \preceq \rangle$ to denote the partially ordered set, or *poset*, of GGMs. For the family of GGMs, model inclusion coincides with the subset relationship between edge sets, so that for every pair of graphs, G and H, with vertex set V, it holds that $M(G) \preceq M(H)$ if and only if $E_G \subseteq E_H$. It is well-known that the poset $\langle M, \preceq \rangle$ forms a complete *lattice*. We refer to Davey and Priestley (2002) for the basic elements of lattice theory required for this paper, such as *upper* and *lower bounds*, *completeness*, the *unit* and the *zero* elements, the *Hasse diagram*. Here we recall that every pair or elements of a lattice, $M(G)$ and $M(H)$ say, has a least upper bound, called a *join*, denoted by $M(G) \vee M(H)$ and, dually, a greatest lower bound, called a *meet*, denoted by $M(G) \wedge M(H)$. Accordingly, $\vee$ is called the *join operation* whereas $\wedge$ is the *meet operation*. The meet and the joint operations are used in structure learning of graphical models for the identification of neighbouring models, and it is important that they can be efficiently computed. This is the case, for instance, of GGMs where the meet and the join take an especially simple form because $M(G) \vee M(H)$ and $M(G) \wedge M(H)$ are the models represented by the graphs with edge sets $E_G \cup E_H$ and $E_G \cap E_H$, respectively. Furthermore, the lattice $\langle M, \preceq \rangle$ satisfies the distributivity property (Davey and Priestley, 2002, Chapter 4), that is the operations of meet and join distribute over each other. Distributivity is a useful property that facilitates the implementation of efficient procedures, and for its application in model selection we refer to Edwards and Havránek (1987), Antoch and Hanousek (2000), Gehrmann (2011) and references therein.

Consider now the family $\mathcal{M}$ of RCON models for Y_V and let $\mathcal{G}$ and $\mathcal{H}$ be two coloured graphs with vertex set V. It was shown by Gehrmann (2011) that $\mathcal{M}(\mathcal{G})$ is a submodel of $\mathcal{M}(\mathcal{H})$, i.e. $\mathcal{M}(\mathcal{G}) \preceq \mathcal{M}(\mathcal{H})$, if and only if all of the three following conditions hold true,

(S1) $E_{\mathcal{G}} \subseteq E_{\mathcal{H}}$

(S2) every colour class in $\mathcal{V}_{\mathcal{G}}$ is a union of colour classes in $\mathcal{V}_{\mathcal{H}}$;

(S3) every colour class in $\mathcal{E}_{\mathcal{G}}$ is a union of colour classes in $\mathcal{E}_{\mathcal{H}}$.

Furthermore, Gehrmann (2011) showed that $\langle \mathcal{M}, \preceq \rangle$ is a complete lattice, although non-distributive, and provided the following rules for the computation of the meet and the join in $\langle \mathcal{M}, \preceq \rangle$ of two graphs $\mathcal{G}$ and $\mathcal{H}$,

$$\mathcal{M}(\mathcal{G}) \wedge \mathcal{M}(\mathcal{H}) = (\mathcal{V}_{\mathcal{G}} \vee_p \mathcal{V}_{\mathcal{H}}, \mathcal{E}^*_{\mathcal{G}} \vee_p \mathcal{E}^*_{\mathcal{H}}) \quad \text{and} \quad \mathcal{M}(\mathcal{G}) \vee \mathcal{M}(\mathcal{H}) = (\mathcal{V}_{\mathcal{G}} \wedge_p \mathcal{V}_{\mathcal{H}}, \mathcal{E}^{**}_{\mathcal{G}} \wedge_p \mathcal{E}^{**}_{\mathcal{H}}) \quad (1)$$

where $\mathcal{E}^*_{\mathcal{G}} \subseteq \mathcal{E}_{\mathcal{G}}$ and $\mathcal{E}^*_{\mathcal{H}} \subseteq \mathcal{E}_{\mathcal{H}}$ are maximal with the property that they are partitions of the same set of edges inside $E_{\mathcal{G}} \cap E_{\mathcal{H}}$, $\mathcal{E}^{**}_{\mathcal{G}} = \mathcal{E}_{\mathcal{G}} \cup \{\{E_{\mathcal{H}} \setminus E_{\mathcal{G}}\}\}$ and $\mathcal{E}^{**}_{\mathcal{H}} = \mathcal{E}_{\mathcal{H}} \cup \{\{E_{\mathcal{G}} \setminus E_{\mathcal{H}}\}\}$. We remark that the operations $\wedge_p$ and $\vee_p$ in (1) are the meet and the join, respectively, of the so-called lattice of partitions, as described in Canfield (2001), and that the implementation of the meet and the join in $\langle \mathcal{M}, \preceq \rangle$ is more involved, and thus computationally expensive, than in $\langle M, \preceq \rangle$. Gehrmann (2011) also considered some relevant subclasses of $\langle \mathcal{M}, \preceq \rangle$ and described the properties of the corresponding lattices. Here we recall that all the subclasses considered by Gehrmann (2011) form complete non-distributive lattices which are *stable* under the meet operation in $\langle \mathcal{M}, \preceq \rangle$, in the sense that their meet operation is induced by that in $\langle \mathcal{M}, \preceq \rangle$, but they are not stable under the join operation. Hence, they are not sublattices of $\langle \mathcal{M}, \preceq \rangle$.

We close this section by noticing that every RCON model $\mathcal{M}(\mathcal{G}) \in \mathcal{M}$ is uniquely represented by a coloured graph $\mathcal{G}$, and in the rest of this paper, with a slight abuse of notation, we will not make an explicit distinction between sets of models and sets of graphs, thereby equivalently writing, for example, $\mathcal{M}(\mathcal{G}) \in \mathcal{M}$ and $\mathcal{G} \in \mathcal{M}$.

3. RCON models for paired data

In this section, we formally introduce coloured graphs for paired data, which are then used to represent the subfamily of RCON models for paired data.

We consider the case where there exists a natural matching between pairs of variables so that the vector Y_V can be naturally split into two subvectors $Y_V = (Y_L, Y_R)^\top$ with $|L| = |R| = p/2 = q$ and, without loss of generality, we assume $L = \{1, \ldots, q\}$ and $R = \{q+1, \ldots, p\}$. Furthermore we set $i' = i + q$, for $i \in L$, so that $R = \{1', \ldots, q'\}$. Every variable Y_i with $i \in L$ is associated with an homologous, or *twin*, variable $Y_{i'}$, and we say that a vertex colour class is *twin-pairing* if it is formed by one pair of twin vertices, so that it can be formally written as $\{i, i'\}$ with $i \in L$. Similarly, a twin-pairing edge colour class is formed by one pair of twin edges which can either link vertices within the "L"eft and "R"ight part of the network, such as $\{(i, j), (i', j')\}$, or link vertices across L and R, such as $\{(i, j'), (j, i')\}$, with $i, j \in L$ and, consequently, $i', j' \in R$.

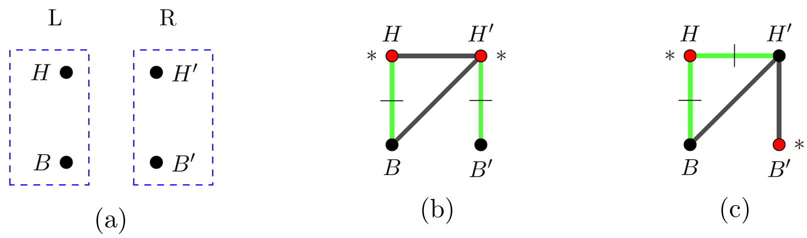


Figure 2: Frets' Heads example: (a) Partition of the vertex set into two blocks, (b) a PD-CG and (c) a coloured graph containing two non-twin-pairing colour classes.

Example 3 (Frets' Heads continued) *In the Frets' Head example the vertex set can be split into the variables associated with the first son,* Y_H *and* Y_B*, and the second son,* $Y_{H'}$ *and* $Y_{B'}$*, so that* $L = \{H, B\}$ *and* $R = \{H', B'\}$ *as depicted in Figure 2a. Furthermore, Figure 2b shows a graph with two twin-pairing colour classes, specifically* $\{H, H'\}$ *and* $\{(H, B), (H', B')\}$ *whereas all the remaining colour classes are atomic. On the other hand, the graph in Figure 2c has two colour classes which are neither atomic nor twin-pairing, specifically* $\{H, B'\}$ *and* $\{(H, H'), (H, B)\}$*.*

Definition 1 *Assume that* $Y_V = (Y_L, Y_R)^\top$ *so that every variable in* Y_L *is uniquely paired with a variable in* Y_R*, and let* $\mathcal{G}$ *be a coloured graph with vertex set* V*. We say that* $\mathcal{G}$ *is a* coloured graph for paired data (PD-CG) *if all its colour classes are either atomic or twin-pairing. Accordingly, we say that* $\mathcal{M}(\mathcal{G})$ *is a* RCON model for paired data (PD-RCON)*, and we denote by* $\mathcal{P} \equiv \mathcal{P}(L, R)$ *the subfamily of PD-RCON models for* (Y_L, Y_R)*.*

If $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a PD-CG then we can write $\mathcal{V} = \mathcal{V}^{(a)} \cup \mathcal{V}^{(t)}$, with $\mathcal{V}^{(a)} \cap \mathcal{V}^{(t)} = \emptyset$, where $\mathcal{V}^{(a)}$ is the set of atomic vertex colour classes and $\mathcal{V}^{(t)}$ is the set of twin-pairing vertex colour classes and, similarly, $\mathcal{E} = \mathcal{E}^{(a)} \cup \mathcal{E}^{(t)}$, with $\mathcal{E}^{(a)} \cap \mathcal{E}^{(t)} = \emptyset$. It follows that, for a given partition $V = L \cup R$ of the vertex set, every PD-CG $\mathcal{G}$ is uniquely identified by the triplet $(E_{\mathcal{G}}, \mathcal{V}^{(a)}_{\mathcal{G}}, \mathcal{E}^{(a)}_{\mathcal{G}})$.

Example 4 (Frets' Heads continued) *The graph in Figure 2b is a PD-CG with* $\mathcal{V}^{(a)} = \{\{B\}, \{B'\}\}$ *so that* $\mathcal{V}^{(t)} = \{\{H, H'\}\}$ *whereas* $\mathcal{E}^{(a)} = \{\{(H, H')\}, \{(B, H')\}\}$*, and thus,* $\mathcal{E}^{(t)} = \{\{(H, B), (H', B')\}\}$*. On the other hand, the graph in Figure 2c is not a PD-CG.*

We finally introduce two specific colour classes which are required for the derivation of the results in the next section: if $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a PD-CG and $E' \subseteq E_{\mathcal{G}}$ is a set of edges, then we set $\mathcal{E}_{\mathcal{G}}(E') = \{C \in \mathcal{E}_{\mathcal{G}} \mid C \subseteq E'\}$ and also write $\mathcal{E}_{\mathcal{G}}(E') = \mathcal{E}^{(a)}_{\mathcal{G}}(E') \cup \mathcal{E}^{(t)}_{\mathcal{G}}(E')$. Furthermore, $\mathcal{E}(E')$ is defined as the collection of atomic and twin-pairing colour classes obtained from the edges in E' in such a way that all possible twin-pairing classes are formed, that is if the edge atomic colour class $\{e\}$ is such that $\{e\} \in \mathcal{E}(E')$ then the twin e' of e is such that $e' \notin E'$, and also in this case may use the partition of $\mathcal{E}(E')$ into $\mathcal{E}^{(a)}(E')$ and $\mathcal{E}^{(t)}(E')$.

4. Lattice of RCON models for paired data

In this section we show that the subset $\mathcal{P} \subseteq \mathcal{M}$ forms a sublattice of $\langle \mathcal{M}, \preceq \rangle$ and provide simple rules to deal efficiently with this model space. Such rules rely on the representation

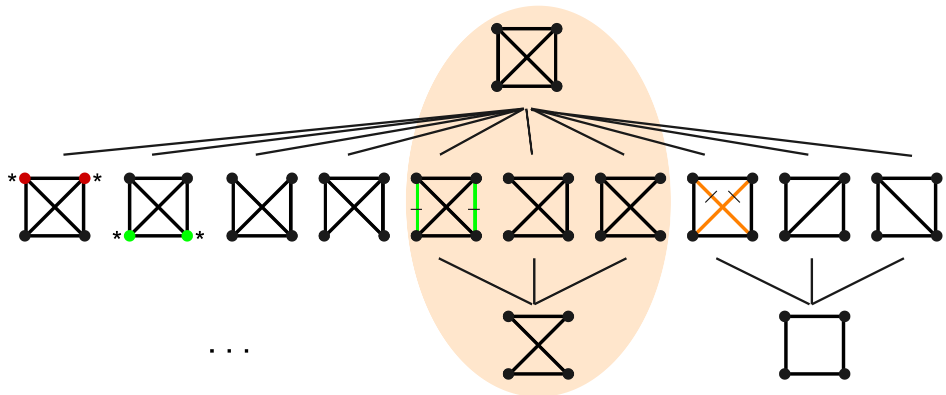


Figure 3: A part of Hasse diagram of the lattice $\langle \mathcal{P}, \preceq \rangle$ for the Frets' Heads example. The highlighted graphs form the so-called diamond structure.

of a graph $\mathcal{G} \in \mathcal{P}$ through the triplet $(E_{\mathcal{G}}, \mathcal{V}_{\mathcal{G}}^{(a)}, \mathcal{E}_{\mathcal{G}}^{(a)})$ because all the relevant operations can be carried out by considering an element of this triplet at the time.

Example 5 (Frets' Heads continued) *Figure 3 gives a part of the Hasse diagram for the Frets' Heads lattice $\langle \mathcal{P}, \preceq \rangle$. On the top there is the largest graph, also called the unit, forming the first layer. The second layer gives the ten neighbours of the unit. Only two graphs of the third layer are depicted, and the four highlighted graphs form the structure known as diamond, which shows that this lattice is non-distributive, see Theorem 5 below.*

The following proposition shows that the model inclusion relationship between models belonging to $\mathcal{P} \subseteq \mathcal{M}$ can be straightforwardly checked by comparing the respective elements of the triplets identifying the two models.

Proposition 2 *If $\mathcal{G}, \mathcal{H} \in \mathcal{P}$ then it holds that $\mathcal{G} \preceq \mathcal{H}$, as specified by conditions (S1) to (S3), if and only if*

$$\text{(P1)}\ \ E_{\mathcal{G}} \subseteq E_{\mathcal{H}}, \qquad \text{(P2)}\ \ \mathcal{V}_{\mathcal{G}}^{(a)} \subseteq \mathcal{V}_{\mathcal{H}}^{(a)}, \qquad \text{(P3)}\ \ \mathcal{E}_{\mathcal{G}}^{(a)} \subseteq \mathcal{E}_{\mathcal{H}}^{(a)},$$

and, furthermore, (P3) is equivalent to (P3′) $\tilde{\mathcal{E}}_{\mathcal{G}}^{(a)} \subseteq \mathcal{E}_{\mathcal{H}}^{(a)}$, where $\tilde{\mathcal{E}}_{\mathcal{G}}^{(a)} = \mathcal{E}_{\mathcal{G}}^{(a)} \cup \mathcal{E}^{(a)}(E_{\mathcal{H}} \setminus E_{\mathcal{G}})$.

Proof We first show that if conditions (S1), (S2) and (S3) are satisfied and $\mathcal{G}, \mathcal{H} \in \mathcal{P}$ then also conditions (P1), (P2) and (P3) hold true. Clearly, (P1) is trivially implied by (S1) because they coincide. Condition (P2) follows immediately from (S2) because the latter implies that every atomic vertex colour class in $\mathcal{V}_{\mathcal{G}}$ is an atomic vertex colour class in $\mathcal{V}_{\mathcal{H}}$. Likewise, condition (P3) follows from the fact that, by (S1) and (S3), every atomic edge class in $\mathcal{E}_{\mathcal{G}}^{(a)}$ is an atomic edge class also in $\mathcal{H}$. We now turn to the reverse implication, i.e. that if conditions (P1), (P2) and (P3) are satisfied and $\mathcal{G}, \mathcal{H} \in \mathcal{P}$ then also conditions (S1), (S2) and (S3) hold true. As remarked above, (S1) is trivially implied by (P1). Condition (S2) is satisfied because (P2) implies that every atomic vertex class in $\mathcal{G}$ also belongs to $\mathcal{H}$. Furthermore, the fact that $\mathcal{G} \in \mathcal{P}$ implies that the remaining vertex colour classes of $\mathcal{G}$ are all twin-pairing and, because also $\mathcal{H} \in \mathcal{P}$, it follows that every vertex twin-pairing class in $\mathcal{G}$ either belongs to $\mathcal{H}$ or is the union of a pair of atomic classes in $\mathcal{H}$. In a similar way, (S3) is satisfied because (P3) implies that every edge atomic colour class in $\mathcal{G}$ is an atomic colour class in $\mathcal{H}$. The fact that $\mathcal{G} \in \mathcal{P}$ implies that the remaining edge colour classes of $\mathcal{G}$ are all twin-pairing and, because also $\mathcal{H} \in \mathcal{P}$ and by (P1) every edge in $\mathcal{G}$ is also present in $\mathcal{H}$, it follows that every edge twin-pairing class in $\mathcal{G}$ either belongs to $\mathcal{H}$ or is the union of a pair of atomic classes in $\mathcal{H}$. We now show the equivalence of (P3) and (P3$'$). Clearly (P3$'$) implies (P3) because, by construction, $\mathcal{E}_{\mathcal{G}}^{(a)} \subseteq \tilde{\mathcal{E}}_{\mathcal{G}}^{(a)}$. We now show that (P3) implies that $\mathcal{E}^{(a)}(E_{\mathcal{H}} \setminus E_{\mathcal{G}}) \subseteq \mathcal{E}_{\mathcal{H}}^{(a)}$ and therefore (P3$'$). Assume that $\mathcal{E}^{(a)}(E_{\mathcal{H}} \setminus E_{\mathcal{G}}) \not\subseteq \mathcal{E}_{\mathcal{H}}^{(a)}$, this is possible only if there exists an edge $e \in E_{\mathcal{H}} \setminus E_{\mathcal{G}}$ which forms a twin-pairing class in $\mathcal{H}$, say $\{e, e'\} \in \mathcal{E}_{\mathcal{H}}^{(t)}$, and such that $e' \notin E_{\mathcal{H}} \setminus E_{\mathcal{G}}$. In turn, this implies that $e' \in E_{\mathcal{H}}$ with $\{e'\} \notin \mathcal{E}_{\mathcal{H}}^{(a)}$ and that both $e \notin E_{\mathcal{G}}$ and $e' \in E_{\mathcal{G}}$, but in this case $\{e'\} \in \mathcal{E}_{\mathcal{G}}^{(a)}$ thereby contradicting the assumption that (P3) is satisfied. ■

We now need to state the following preliminary result.

Lemma 3 *Let $\mathcal{G}_1, \ldots, \mathcal{G}_k$, with $k \geq 1$, and $\mathcal{F}$ be coloured graphs belonging to $\mathcal{M}$. If, for every $i = 1, \ldots, k$, the graphs $\mathcal{G}_i$ are such that both $\mathcal{G}_i \in \mathcal{P}$ and $\mathcal{G}_i \preceq \mathcal{F}$ and, furthermore, $E_{\mathcal{F}} \subseteq \cup_{i=1}^{k} E_{\mathcal{G}_i}$ then it holds that $\mathcal{F} \in \mathcal{P}$.*

Proof We show that assuming $\mathcal{F} \notin \mathcal{P}$ contradicts the assumption that $\mathcal{G}_i \in \mathcal{P}$ for every $i = 1, \ldots, k$. If $\mathcal{F} \notin \mathcal{P}$ then $\mathcal{F}$ has a colour class that is neither atomic nor twin-pairing. Here, we deal with the case where the latter is an edge colour class, but we remark that the same argument applies also if it is a vertex colour class. Hence, we denote by $E_{\mathcal{F}}^{*}$ a non-atomic non-twin-pairing edge colour class of $\mathcal{F}$ and let $e \in E_{\mathcal{F}}^{*}$. Because $E_{\mathcal{F}} \subseteq \cup_{i=1}^{k} E_{\mathcal{G}_i}$ then there exists at least a graph $\mathcal{G}_i$ such that $e \in E_{\mathcal{G}_i}$ and we assume, without loss of generality, that $e \in E_{\mathcal{G}_1}$. Hence, there exists a colour class $E_{\mathcal{G}_1}^{*}$ of $\mathcal{G}_1$ such that $e \in E_{\mathcal{G}_1}^{*}$ and furthermore, by (S3) and the fact that $\mathcal{G}_1 \preceq \mathcal{F}$, it holds that $E_{\mathcal{F}}^{*} \subseteq E_{\mathcal{G}_1}^{*}$. However, the latter inclusion implies that $E_{\mathcal{G}_1}^{*}$ is neither atomic nor twin-pairing thereby contradicting the assumption that $\mathcal{G}_1 \in \mathcal{P}$. ■

The following theorem provides the rules for the implementation of the operations of meet and the join between models belonging to $\mathcal{P}$. It should be noticed that these involve simple set union and intersection operations and are therefore considerably easier and more efficient to implement than the rules given in (1) for arbitrary models in $\mathcal{M}$.

Theorem 4 *If $\mathcal{G}, \mathcal{H} \in \mathcal{P}$ then the join of $\mathcal{G}$ and $\mathcal{H}$ from (1) belongs to $\mathcal{P}$, that is $\mathcal{G} \vee \mathcal{H} \in \mathcal{P}$, and it is identified by*

$$\text{(J1)}\;\; E_{\mathcal{G}\vee\mathcal{H}} = E_{\mathcal{G}} \cup E_{\mathcal{H}}, \qquad \text{(J2)}\;\; \mathcal{V}_{\mathcal{G}\vee\mathcal{H}}^{(a)} = \mathcal{V}_{\mathcal{G}}^{(a)} \cup \mathcal{V}_{\mathcal{H}}^{(a)}, \qquad \text{(J3)}\;\; \mathcal{E}_{\mathcal{G}\vee\mathcal{H}}^{(a)} = \tilde{\mathcal{E}}_{\mathcal{G}}^{(a)} \cup \tilde{\mathcal{E}}_{\mathcal{H}}^{(a)}.$$

Furthermore, the twin-pairing colour classes of $\mathcal{G} \vee \mathcal{H}$ can be computed as $\mathcal{V}_{\mathcal{G}\vee\mathcal{H}}^{(t)} = \mathcal{V}_{\mathcal{G}}^{(t)} \cap \mathcal{V}_{\mathcal{H}}^{(t)}$ and $\mathcal{E}_{\mathcal{G}\vee\mathcal{H}}^{(t)} = \tilde{\mathcal{E}}_{\mathcal{G}}^{(t)} \cap \tilde{\mathcal{E}}_{\mathcal{H}}^{(t)}$, where, as above, $\tilde{\mathcal{E}}_{\mathcal{G}}^{(\cdot)} = \mathcal{E}_{\mathcal{G}}^{(\cdot)} \cup \mathcal{E}^{(\cdot)}(E_{\mathcal{H}} \setminus E_{\mathcal{G}})$ and, accordingly, $\tilde{\mathcal{E}}_{\mathcal{H}}^{(\cdot)} = \mathcal{E}_{\mathcal{H}}^{(\cdot)} \cup \mathcal{E}^{(\cdot)}(E_{\mathcal{G}} \setminus E_{\mathcal{H}})$.

Likewise, the meet of $\mathcal{G}$ and $\mathcal{H}$ from (1) belongs to $\mathcal{P}$, that is $\mathcal{G} \wedge \mathcal{H} \in \mathcal{P}$, and it is identified by

$$\text{(M1)}\;\; E_{\mathcal{G}\wedge\mathcal{H}} = (E_{\mathcal{G}} \cap E_{\mathcal{H}}) \setminus E^{*}, \qquad \text{(M2)}\;\; \mathcal{V}_{\mathcal{G}\wedge\mathcal{H}}^{(a)} = \mathcal{V}_{\mathcal{G}}^{(a)} \cap \mathcal{V}_{\mathcal{H}}^{(a)}, \qquad \text{(M3)}\;\; \mathcal{E}_{\mathcal{G}\wedge\mathcal{H}}^{(a)} = \mathcal{E}_{\mathcal{G}}^{(a)} \cap \mathcal{E}_{\mathcal{H}}^{(a)},$$

where E^* is the set of the edges belonging to the colour classes of $\mathcal{E}^{(a)}(E_\mathcal{G} \cap E_\mathcal{H}) \setminus (\mathcal{E}^{(a)}_\mathcal{G} \cap \mathcal{E}^{(a)}_\mathcal{H})$. Furthermore, the twin-pairing colour classes of $\mathcal{G} \wedge \mathcal{H}$ can be computed as $\mathcal{V}^{(t)}_{\mathcal{G}\wedge\mathcal{H}} = \mathcal{V}^{(t)}_\mathcal{G} \cup \mathcal{V}^{(t)}_\mathcal{H}$ and $\mathcal{E}^{(t)}_{\mathcal{G}\wedge\mathcal{H}} = \mathcal{E}^{(t)}_\mathcal{G}(E_\mathcal{H} \cap E_\mathcal{G}) \cup \mathcal{E}^{(t)}_\mathcal{H}(E_\mathcal{H} \cap E_\mathcal{G})$.

Proof We first consider the join operation and set $\mathcal{U}_\vee = \mathcal{G} \vee \mathcal{H}$ from (1). Furthermore, we let $\mathcal{U} = (\mathcal{V}_\mathcal{U}, \mathcal{E}_\mathcal{U})$ be the graph obtained from (J1), (J2) and (J3), and we show that $\mathcal{U}_\vee \in \mathcal{P}$ and that $\mathcal{U}_\vee = \mathcal{U}$. Firstly, we prove that $\mathcal{G} \preceq \mathcal{U}$ by showing that the conditions of Proposition 2 are satisfied. Indeed, in this case (P1), (P2) and (P3′) are true by construction because (J1) immediately implies that $E_\mathcal{G} \subseteq E_\mathcal{U}$, (J2) that $\mathcal{V}^{(a)}_\mathcal{G} \subseteq \mathcal{V}^{(a)}_\mathcal{U}$ and (J3) that $\tilde{\mathcal{E}}^{(a)}_\mathcal{G} \subseteq \mathcal{E}^{(a)}_\mathcal{U}$. In the same way, one can show that $\mathcal{H} \preceq \mathcal{U}$ and from the definition of the join operation and the fact the both $\mathcal{G}, \mathcal{H} \preceq \mathcal{U}$ it follows that $\mathcal{U}_\vee \preceq \mathcal{U}$. Hence, the required result that $\mathcal{U}_\vee \in \mathcal{P}$ follows from Lemma 3 because $\mathcal{G}, \mathcal{H} \preceq \mathcal{U}_\vee \preceq \mathcal{U}$ so that by (S1) one has $E_{\mathcal{U}_\vee} \subseteq E_\mathcal{U} = E_\mathcal{G} \cup E_\mathcal{H}$. We can now apply Proposition 2 to show that $\mathcal{U} \preceq \mathcal{U}_\vee$. By definition, the join $\mathcal{U}_\vee$ of $\mathcal{G}$ and $\mathcal{H}$ is such that both $\mathcal{G} \preceq \mathcal{U}_\vee$ and $\mathcal{H} \preceq \mathcal{U}_\vee$ and therefore from (P1), (P2) and (P3′) it holds that (i) $E_{\mathcal{U}_\vee} \supseteq E_\mathcal{G} \cup E_\mathcal{H}$, (ii) $\mathcal{V}^{(a)}_{\mathcal{U}_\vee} \supseteq \mathcal{V}^{(a)}_\mathcal{G} \cup \mathcal{V}^{(a)}_\mathcal{H}$, (iii) $\mathcal{E}^{(a)}_{\mathcal{U}_\vee} \supseteq \tilde{\mathcal{E}}^{(a)}_\mathcal{G} \cup \tilde{\mathcal{E}}^{(a)}_\mathcal{H}$, and therefore that $\mathcal{U} \preceq \mathcal{U}_\vee$. We have thus shown that both $\mathcal{U}_\vee \preceq \mathcal{U}$ and $\mathcal{U} \preceq \mathcal{U}_\vee$ and therefore that $\mathcal{U}_\vee = \mathcal{U}$.

We consider now the meet operation and set $\mathcal{L}_\wedge = \mathcal{G} \wedge \mathcal{H}$ from (1). Furthermore, we let $\mathcal{L} = (\mathcal{V}_\mathcal{L}, \mathcal{E}_\mathcal{L})$ be the graph obtained from (M1), (M2) and (M3), and we show that $\mathcal{L}_\wedge \in \mathcal{P}$ and that $\mathcal{L}_\wedge = \mathcal{L}$. Firstly, we prove that $\mathcal{L} \preceq \mathcal{G}$ by showing that the conditions of Proposition 2 are satisfied. Indeed, in this case (P1), (P2) and (P3) are true by construction because (M1) implies that $E_\mathcal{L} \subseteq E_\mathcal{G}$, (M2) that $\mathcal{V}^{(a)}_\mathcal{L} \subseteq \mathcal{V}^{(a)}_\mathcal{G}$ and (M3) that $\mathcal{E}^{(a)}_\mathcal{L} \subseteq \mathcal{E}^{(a)}_\mathcal{G}$. In the same way, one can show that $\mathcal{L} \preceq \mathcal{H}$ and from the definition of the meet operation and the fact the both $\mathcal{L} \preceq \mathcal{G}, \mathcal{H}$ it follows that $\mathcal{L} \preceq \mathcal{L}_\wedge$. We now show that $E_\mathcal{L} = E_{\mathcal{L}_\wedge}$. Because $\mathcal{L} \preceq \mathcal{L}_\wedge \preceq \mathcal{G}, \mathcal{H}$ then (S1) implies that $E_\mathcal{L} \subseteq E_{\mathcal{L}_\wedge} \subseteq E_\mathcal{G} \cap E_\mathcal{H}$ and we can use (M1) to write $(E_\mathcal{G} \cap E_\mathcal{H}) \setminus E^* \subseteq E_{\mathcal{L}_\wedge} \subseteq E_\mathcal{G} \cap E_\mathcal{H}$. We note, however, that there cannot exist any edge $e \in E_{\mathcal{L}_\wedge}$ such that $e \in E^*$. Indeed, by construction of E^*, in that case one would have that $\{e\} \in \mathcal{E}^{(a)}_{\mathcal{L}_\wedge}$ and e would belong to a twin-pairing colour class in either $\mathcal{G}$ or in $\mathcal{H}$, thereby contradicting the assumption that $\mathcal{L}_\wedge \preceq \mathcal{G}, \mathcal{H}$. Hence we can write $(E_\mathcal{G} \cap E_\mathcal{H}) \setminus E^* \subseteq E_{\mathcal{L}_\wedge} \subseteq (E_\mathcal{G} \cap E_\mathcal{H}) \setminus E^*$ so that $E_{\mathcal{L}_\wedge} = E_\mathcal{L}$. This allows us to prove the required result that $\mathcal{L}_\wedge \in \mathcal{P}$, because it holds that $\mathcal{L} \preceq \mathcal{L}_\wedge$ with $E_{\mathcal{L}_\wedge} = E_\mathcal{L}$ an thus Lemma 3 applies. We can now apply Proposition 2 to show that $\mathcal{L}_\wedge \preceq \mathcal{L}$. We have shown above that (i) $E_{\mathcal{L}_\wedge} = (E_\mathcal{G} \cap E_\mathcal{H}) \setminus E^*$. Furthermore, by definition, the meet $\mathcal{L}_\wedge$ of $\mathcal{G}$ and $\mathcal{H}$ is such that both $\mathcal{L}_\wedge \preceq \mathcal{G}$ and $\mathcal{L}_\wedge \preceq \mathcal{H}$ and therefore from (P2) and (P3) it holds that (ii) $\mathcal{V}^{(a)}_{\mathcal{L}_\wedge} \subseteq \mathcal{V}^{(a)}_\mathcal{G} \cap \mathcal{V}^{(a)}_\mathcal{H}$, (iii) $\mathcal{E}^{(a)}_{\mathcal{L}_\wedge} \subseteq \mathcal{E}^{(a)}_\mathcal{G} \cap \mathcal{E}^{(a)}_\mathcal{H}$, and therefore that $\mathcal{L}_\wedge \preceq \mathcal{L}$. We have thus shown that both $\mathcal{L} \preceq \mathcal{L}_\wedge$ and $\mathcal{L}_\wedge \preceq \mathcal{L}$ and therefore that $\mathcal{L}_\wedge = \mathcal{L}$.

We omit for brevity the derivation of the twin-pairing colour classes. ■

The statements of Theorem 4 can be readily applied to show the following.

Theorem 5 *The subset $\mathcal{P} \subseteq \mathcal{M}$ is stable under the meet and the join operations in $\langle \mathcal{M}, \preceq \rangle$ and therefore $\langle \mathcal{P}, \preceq \rangle$ forms a complete sublattice of $\langle \mathcal{M}, \preceq \rangle$. The unit is the uncoloured complete graph and the zero is the empty graph with all twin vertices being symmetric. Furthermore, $\langle \mathcal{P}, \preceq \rangle$ is non-distributive.*

Proof The stability of the meet and the join operation follows from Theorem 4 where it is shown that both the meet and the join of $\mathcal{G}, \mathcal{H} \in \mathcal{P}$ belong to $\mathcal{P}$. In turn, this implies that $\langle \mathcal{P}, \preceq \rangle$ forms a sublattice of $\langle \mathcal{M}, \preceq \rangle$. The unit $\mathcal{U}$ is the PD-CG that is larger than any element in $\mathcal{P}$, hence, the edge set $E_{\mathcal{U}}$ is the edge set of the complete graph and $\mathcal{V}_{\mathcal{U}}^{(t)} = \mathcal{E}_{\mathcal{U}}^{(t)} = \emptyset$. Therefore, $\mathcal{U}$ is the uncoloured complete graph. Moreover, the zero $\mathcal{L}$ is the PD-CG that is smaller than any element in $\mathcal{P}$, so that $E_{\mathcal{L}} = \mathcal{V}_{\mathcal{L}}^{(a)} = \emptyset$ and therefore also $\mathcal{E}_{\mathcal{L}}^{(a)} = \emptyset$. Finally, as shown for instance in Theorem 4.10 of Davey and Priestley (2002), in order to show that $\langle \mathcal{P}, \preceq \rangle$ is non-distributive it is sufficient to show that its Hasse diagram may contain a so-called diamond structure, as highlighted in Figure 3. ■

5. Application to the selection of brain networks from fMRI data

In this section, we apply a stepwise backward elimination procedure to the selection of a PD-RCON model from functional MRI data, in order to investigate the dynamic activity of brain regions between hemispheres.

The stepwise procedure we use is implemented as follows: at every step one model is selected as starting point, which we denote by $\mathcal{M}(\mathcal{G}^{(i)})$ for step i. Then, a set of candidate models is constructed by considering all the neighbouring submodels of $\mathcal{M}(\mathcal{G}^{(i)})$ and then removing from this set the weakly-rejected models, that is, the models which are submodels of previously rejected models. Each of the candidate models is then classified as either "accepted" or "rejected" on the basis of the likelihood ratio test at the 0.05 level, and the accepted model with the highest p-value is selected as starting point, $\mathcal{M}(\mathcal{G}^{(i+1)})$, for the next step. The starting point $\mathcal{M}(\mathcal{G}^{(1)})$ of the first step is the largest model in $\langle \mathcal{P}, \preceq \rangle$, i.e. the unit, as given in Theorem 5, and the procedure terminates when all the candidate models are rejected. Although for space reason we omit the technical details, it is worth remarking that the implementation of this procedure relies on the results of Proposition 2 and Theorem 4. More specifically, a reduction in the computational cost is obtained by computing the candidate models as the meet in Proposition 2 between the selected model in the current step and the accepted models in the previous step.

The data come from a pilot study of the Enhanced Nathan Kline Institute-Rockland Sample project. A detailed description of the project, scopes, and technical aspects can be found at `http://fcon_1000.projects.nitrc.org/indi/enhanced/`. The fMRI time series are recorded on 70 spatial Region of Interests (ROIs) based on the Desikan atlas at 404 equally spaced time, and are available at `http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html`. The data we used are residuals estimated from the vector autogression models, carried out to remove the temporal dependence (see Ranciati et al., 2021). According to the available information on subjects, we focus on two participants, indexed as subjects 14 and 15, who have the same psychological traits with no neuropsychiatric diseases and right-handedness. However, subject 14 is 19 years old whereas subject 15 is 57. For each subject, we studied 22 cortical regions in the frontal lobe, and 14 brain regions in the anterior temporal lobe. The graphs of the models resulting from the model selection procedure are given in Table 1, whereas some summary statistics concerning the identified symmetries are given in Table 2. In particular, we see that subject 15 seems to present

Table 1: PD-CGs resulting form the application of the search procedure.

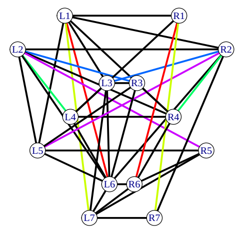
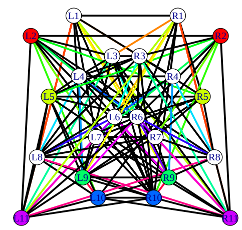
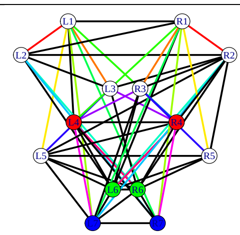
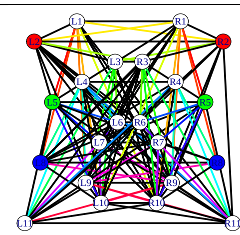
Subject	**Anterior temporal lobe**	**Frontal lobe**
14		
15		

Table 2: Summary statistics for the graphs of Table 1. From the left: number of edges, graph density, number of twin-pairing vertex colour classes, number of pairs of symmetric edges, number of twin-pairing edge colour classes.

Subject	Lobe	# edges	den.(%)	# pairs of sym. vertices	# pairs of edges	# pairs of sym. edges
14	Anterior temp.	40	43.96	0	9	5
	Frontal	127	54.98	5	36	23
15	Anterior temp.	58	63.74	3	20	13
	Frontal	136	58.87	3	44	26

denser graphs and, furthermore, for both subjects the graphs relative to the temporal lobe seem to present a higher level of symmetry.

6. Conclusions and future research directions

We have defined the family of RCON models for paired data as a subfamily of the RCON models from Højsgaard and Lauritzen (2008), and the main results of this paper concern the lattice structure of this class of models. We have shown that, unlike other relevant subclasses of coloured graphical models, PD-RCON models form a sublattice of RCON models and we have provided explicit rules for dealing with this lattice. These have been applied to implement a stepwise model search procedure. An important challenge in this framework concerns the dimension of the model space that is very large even when the number of variables is relatively small. Indeed, the space of PD-RCON models is much larger than the space of classical GGMs and, consequently, model search in the high-dimensional setting much more challenging. Future research directions involve the identification of strategies to obtain a more efficient exploration of the models space. A related problem concerns the non-distributivity of the lattice of PD-RCON models and the possible identification of distributive sublattices which may facilitate the computation of neighbouring models and therefore improve efficiency.

References

J. Antoch and J. Hanousek. Model selection and simplification using lattices. *CERGE-EI Working Paper*, (164), 2000.

E. R. Canfield. Meet and join within the lattice of set partitions. *The Electronic Journal of Combinatorics*, 8(1):R5, 2001.

B. A. Davey and H. A. Priestley. *Introduction to lattices and order*. Cambridge university press, second edition, 2002.

D. Edwards and T. Havránek. A fast model selection procedure for large families of models. *Journal of the American Statistical Association*, 82(397):205–213, 1987.

H. Gehrmann. Lattices of graphical Gaussian models with symmetries. *Symmetry*, 3(3): 653–679, 2011.

S. Højsgaard and S. L. Lauritzen. Graphical Gaussian models with edge and vertex symmetries. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70 (5):1005–1027, 2008.

S. L. Lauritzen. *Graphical models*, volume 17. Clarendon Press, 1996.

Q. Li, X. Gao, and H. Massam. Approximate bayesian estimation in large coloured graphical Gaussian models. *Canadian Journal of Statistics*, 46(1):176–203, 2018.

Q. Li, X. Gao, and H. Massam. Bayesian model selection approach for coloured graphical Gaussian models. *Journal of Statistical Computation and Simulation*, 90(14):2631–2654, 2020.

Q. Li, X. Sun, N. Wang, and X. Gao. Penalized composite likelihood for colored graphical Gaussian models. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 14(4):366–378, 2021.

M. Maathuis, M. Drton, S. Lauritzen, and M. Wainwright. *Handbook of Graphical Models*. CRC Press, 2019.

H. Massam, Q. Li, and X. Gao. Bayesian precision and covariance matrix estimation for graphical Gaussian models with edge and vertex symmetries. *Biometrika*, 105(2):371–388, 2018.

S. Ranciati, A. Roverato, and A. Luati. Fused graphical lasso for brain networks with symmetries. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 70(5): 1299–1322, 2021.

J. Whittaker. *Graphical Models in Applied Multivariate Analysis*. John Wiley & Sons, Chichester, 1990.