Graphical Gaussian Models with Symmetries

Helene Gehrmann
Jesus College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity Term 2011

Acknowledgements

I am greatly indebted to my supervisor Professor Steffen L. Lauritzen for his support during the work on this thesis. His expert advice and mentorship have been invaluable.

I am very grateful to the former and current members of the Graphical Models group at the Department of Statistics, University of Oxford. Particular mention deserve Sofia Massa and Therese Graversen, to whom I am thankful for many fruitful discussions. This thesis has further benefited from comments by Professor Michael Perlman and three anonymous referees on paper drafts containing part of the material.

My thanks go to the Engineering and Physical Research Council (EPSRC) and the Department of Statistics, University of Oxford for financially supporting me with a Doctoral Training Award and a Research Travel Allowance, as well as to Steffen Lauritzen for funding me during the last six months of my DPhil work.

I am also grateful to Jesus College, University of Oxford for the award of a Jesus College Graduate Scholarship in the academic year 2009-2010, and for their generous Graduate Research Allowance which allowed me to participate in a number of conferences and workshops.

Finally I would like to thank the Department of Statistics, St Hilda's College, University of Oxford and Steffen Lauritzen for giving me the opportunity to be engaged in undergraduate and graduate teaching in Probability and Statistics.

Abstract

This thesis is concerned with graphical Gaussian models with equality constraints on the concentration or partial correlation matrix introduced by Højsgaard and Lauritzen (2008) as RCON and RCOR models. The models can be represented by vertex and edge coloured graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where parameters associated with equally coloured vertices or edges are restricted to being identical.

In the first part of this thesis we study the problem of estimability of a non-zero model mean μ if the covariance structure Σ is restricted to satisfy the constraints of an RCON or RCOR model but is otherwise unknown. Exploiting results in Kruskal (1968), we obtain a characterisation of suitable linear spaces Ω such that if Σ is restricted as above, the maximum likelihood estimator $\hat{\mu}$ and the least squares estimator μ^* of μ coincide for $\mu \in \Omega$, thus allowing μ and Σ to be estimated independently. For the special case of Ω being specified by equality relations among the entries of μ according to a partition $\mathcal{M}$ of the model variables V, our characterisation translates into a necessary and sufficient regularity condition on $\mathcal{M}$ and $(\mathcal{V}, \mathcal{E})$.

In the second part we address model selection for RCON and RCOR models. Due to the large number of models, we study the structure of four model classes lying strictly within the sets of RCON and RCOR models, each of which is defined by desirable statistical properties corresponding to colouring regularity conditions. Two of these appear in Højsgaard and Lauritzen (2008), while the other two arise from the regularity condition ensuring equality of estimators $\hat{\mu} = \mu^*$ we find in the first part. We show each of the colouring classes to form complete lattices, which qualifies the corresponding model spaces for an Edwards-Havránek model selection procedure (Edwards and Havránek, 1987). We develop a corresponding algorithm for one of the model classes and give an algorithm for a systematic search in accordance with the Edwards-Havránek principles for a second class. Both are applied to data sets previously analysed in the literature, with very encouraging performances.

Contents

1. Introduction

Graphical models use graphs to represent dependencies between stochastic variables, which Markov properties interpret in terms of conditional independence relations between the variables. Together with distributional assumptions the represented dependency relations define a statistical model. Monographs which give an overview of the field include Whittaker (1990), Lauritzen (1996) and Edwards (2000).

We consider graphical models for random vectors $Y = (Y_\alpha)_{\alpha \in V}$ which are assumed to follow a $\mathcal{N}_V(\mu, \Sigma)$ multivariate Gaussian distribution. The models, also known as covariance selection models (Dempster, 1972), can be represented by undirected graphs $G = (V, E)$. A vertex $\alpha \in V$ represents variable Y_α and a pair of vertices $\alpha, \beta \in V$ are joined by an edge in E unless Y_α and Y_β are conditionally independent given the remaining variables. As, for the Gaussian distribution, Y_α and Y_β are conditionally independent given the rest if and only if $(\Sigma^{-1})_{\alpha\beta} = 0$, missing edges in E correspond to zero entries in the inverse covariance matrix Σ^{-1}.

In addition to providing a concise form of visualisation of the conditional independence structure, the graphical representation can be exploited to make statistical inference computations more efficient, which has made graphical Gaussian models a popular modelling tool. In modern day applications of the models, such as mathematical genetics (Wu et al., 2003; Schäfer and Strimmer, 2005), Statisticians are often faced with a very large number of variables p while having relatively few observations n (e.g. Castelo and Roverato, 2006), which makes Statistical modelling challenging. The difficulty arises as the maximum likelihood estimate $\hat{\Sigma}$ of Σ exists with probability one if and only if the rank of the sample covariance matrix is sufficiently large, with the bound depending on the ratio of p and n. If this is too small, estimation of Σ may not be feasible.

The models studied in this thesis address this difficulty by placing additional equality constraints on the model parameters, thus reducing the number of parameters to be estimated (Højsgaard and Lauritzen, 2008). More precisely, RCON and RCOR models combine equality constraints on the diagonal of Σ^{-1} with constraints on a further set of parameters. For RCON models, this is given by the non-zero off-diagonal entries of Σ^{-1}, while RCOR models place equality constraints on the non-zero partial correlations $\rho_{\alpha\beta|V\setminus\{\alpha\beta\}}$. As $(\Sigma^{-1})_{\alpha\beta} = 0$ if and only if $\rho_{\alpha\beta|V\setminus\{\alpha\beta\}} = 0$, both constraint types can be represented by a vertex and edge coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V}$ depicting the constraints on the diagonal entries of Σ^{-1} and $\mathcal{E}$ corresponding to the constraints on off-diagonal concentrations or partial correlations respectively.

The exact relationship between equality constraints and $n_{\min}$, the minimal number of required observations for maximum likelihood estimates to exist with probability one, is not known in general, however first studies (Højsgaard and Lauritzen, 2008; Uhler, 2010) show that equality restrictions generally reduce $n_{\min}$. This makes RCON and RCOR models a promising model class for modern day applications.

A topic not addressed in the literature is estimation of a restricted model mean μ if Σ is to satisfy the constraints of an RCON or RCOR model but is otherwise unknown. Further, suitable model selection methods are required for the models to become widely applicable in practice. These are the main topics of this thesis.

1.1. Thesis Overview

We begin by introducing graphical models and reviewing past treatments of symmetry constraints in graphical Gaussian models in Chapter 2, and devote Chapter 3 to the definition and first properties of RCON and RCOR models. Chapters 4-6 present our own results.

Chapter 4

When Højsgaard and Lauritzen (2008) introduced RCON and RCOR models, their primary interest lay in maximum likelihood estimation of the covariance matrix Σ and they therefore restricted the model mean to be zero throughout. We lift this constraint and consider maximum likelihood estimation of a non-zero mean μ if Σ is subject to the constraints of an RCON or RCOR model but otherwise unknown.

For the Gaussian distribution, in the case of an unknown covariance matrix Σ, maximum likelihood estimation for μ is non-trivial as for Σ known the maximum likelihood estimator $\hat{\mu}$ of μ may depend on Σ. The least-squares estimator μ^* however is defined by minimizing the sum of squares and has no relation to Σ, so that in case of equality of estimators $\hat{\mu} = \mu^*$, μ and Σ may be estimated independently.

Exploiting Kruskal (1968), in Section 4.3 we obtain a characterisation of all linear spaces Ω such that if Σ is restricted to satisfy the constraints of a given RCON or RCOR model, $\hat{\mu}$ and μ^* coincide if μ is restricted to lie in Ω (Theorem 4.2).

Section 4.4 treats the special case that Ω is specified by equality relations among the entries of μ. We let $\mathcal{M}$ denote the partition of V such that whenever two vertices lie in the same set in $\mathcal{M}$ the corresponding means are restricted to being equal, and show that the previously obtained chracterisation translates into a two-fold necessary and sufficient regularity condition on $\mathcal{M}$ and $(\mathcal{V}, \mathcal{E})$ (Theorem 4.3).

In Section 4.5 we illustrate the last mentioned result by applying it to comparative experiments according to Latin square designs (Bailey, 2008) if Σ^{-1} is restricted to lie in a homogeneous nearest-neighbour lattice scheme (Besag, 1974) or a homogeneous nearest-neighbour torus scheme, which we define following Besag and Moran (1975).

Chapter 5

Whereas every undirected graph $G = (V, E)$ gives rise to one unrestricted graphical Gaussian model, there are as many RCON and RCOR models with conditional independence structure represented by G as there are combinations of a vertex colouring $\mathcal{V}$ with an edge colouring $\mathcal{E}$. This large number of models makes model selection challenging and motivates the study of the structure of model spaces lying strictly within the sets of RCON and RCOR models. Particularly favourable are classes of models which place equality constraints in a pattern which makes them more readily interpretable. Chapter 5 is concerned with the study of the structure of four such model classes, each corresponding to a regularity condition on $(\mathcal{V}, \mathcal{E})$.

Two model classes appear in Højsgaard and Lauritzen (2008) and we term the corresponding graph colouring classes to be *edge regular* and *permutation generated* respectively. Edge regular colourings represent constraints which yield the same model specification whether they represent an RCON or an RCOR model, and permutation generated colourings, special cases of the former, represent constraints induced by permutation symmetry in the variable labels. Further, we term colourings which for $\mathcal{M} = \mathcal{V}$ satisfy the regularity condition ensuring $\hat{\mu} = \mu^*$ found in Chapter 4 *vertex regular*, and, following Siemons (1983), term colourings which are both vertex and edge regular to be *regular*.

Our main results are that each of the four corresponding model classes forms a complete non-distributive lattice with respect to model inclusion with meet and join operations we identify (Theorems 5.2, 5.3, 5.4, 5.5).

Chapter 6

Given that each of the four model classes studied in Chapter 5 forms a complete lattice, their model spaces qualify for an Edwards-Havránek model selection procedure (Edwards and Havránek, 1987). Due to greatest tractability of model structure and most intuitive interpretability of constraints, we focus on edge regular and permutation generated colourings.

In Section 6.2 we develop an Edwards-Havránek model selection algorithm for models represented by edge regular colourings and illustrate the procedure by applying it to a data set with five variables. After fitting 232 models out of $1.3 \cdot 10^6$ models in the search space, our algorithm arrived at 4 minimally accepted models, each in very good agreement with previous analyses (Edwards, 2000).

The intricate relationship between permutation groups acting on the vertex set of an undirected graph and their vertex and edge orbits makes it non-trivial to develop an explicit Edwards-Havránek model selection procedure for graphical Gaussian models represented by permutation generated colourings. Nevertheless, we provide a description of the structure of the model class in Section 6.3 which makes it possible to conduct a model search obeying the Edwards-Havránek principles.

Applied to a data set on four variables, commonly known as *Frets' heads* (Frets, 1921; Mardia et al., 1979), our search procedure arrived at 9 minimally accepted models after fitting 48 models, out of a total of 251, collectively in strong agreement with previous analyses (Whittaker, 1990).

1.2. Contributions

In summary, the contributions in this thesis include:

- Characterisation of all linear spaces Ω such that if Σ is restricted to satisfy the constraints of a given RCON or RCOR model but is otherwise unknown, the maximum likelihood estimator $\hat{\mu}$ and least squares estimator μ^* of μ coincide if μ is restricted to lie in Ω, thus allowing μ and Σ to be estimated independently.
- Identification of a colouring criterion giving a necessary and sufficient condition for equality of estimators $\hat{\mu} = \mu^*$ if Σ is restricted to satisfy the constraints of a given RCON or RCOR model and μ is subject to equality constraints.
- Description of the lattice structure of four model classes strictly within the sets of RCON and RCOR models each defined by statistically desirable properties corresponding to colouring regularities in $(\mathcal{V}, \mathcal{E})$, two of which are first identified in this thesis.
- Construction of an Edwards-Havránek model selection procedure for one of the model classes above.
- Construction of a systematic search procedure in accordance with the Edwards-Havránek principles for a second class.

2. Background

2.1. Notation and Terminology

2.1.1. Undirected Graphs

An *undirected graph* is a pair $G = (V, E)$, where V is a finite set of *vertices* and the set of *edges* $E \subseteq \{\{\alpha, \beta\} \mid \alpha, \beta \in V\}$ is a set of unordered pairs of vertices. We assume all of the considered graphs to be *simple*, meaning that no vertex is connected to itself through an edge and that there are no multiple edges. For ease of exposition we write $\alpha\beta$ for the edge $\{\alpha, \beta\}$ connecting vertices α and β. $\alpha\beta$ and $\beta\alpha$ thus refer to the same edge. A graph is *complete* if all possible edges are present and *empty* if it has no edges. $V = \{1, 2, \ldots, d\}$ is abbreviated as $V = [d]$.

A *path* of length n from α to β is a sequence $\alpha = \alpha_0, \ldots, \alpha_n = \beta$ of distinct vertices such that $\alpha_{i-1}\alpha_i \in E$ for all $i = 1, \ldots, n$. For $A, B, C \subseteq V$, C is said to *separate* A *from* B if for all pairs $\alpha \in A$ and $\beta \in B$, all paths from α to β contain a vertex in C. $G' = (V, E')$ is a subgraph of $G = (V, E)$ if $E' \subseteq E$. For A as above, we say that $G_A = (A, E_A)$ is a subgraph *induced by* A if E_A contains all edges in E which connect vertices in A.

If there is an edge $\alpha\beta$ between α and β, then $\alpha\beta$ is said to be *incident* with α and β and α and β are *adjacent* or *neighbours*. The set of neighbours of a vertex α is denoted $\mathrm{ne}(\alpha)$, with $\{\alpha\} \cup \mathrm{ne}(\alpha) = \mathrm{cl}(\alpha)$ denoting the *closure* of α. $\mathrm{ne}(A)$ denotes the set $\bigcup_{\alpha\in A} \mathrm{ne}(\alpha) \setminus A$, and similarly $\mathrm{cl}(A)$=$\bigcup_{\alpha\in A} \mathrm{cl}(\alpha)$. For $G = (V, E)$ and $\alpha \in V$ we call the size of $\mathrm{ne}(\alpha)$ the *degree* of α in G. If all $\alpha \in V$ have the same degree in G, we say that G is *regular*.

G is *bipartite with vertex components* V_1 and V_2 if V_1 and V_2 partition V and G_{V_1} and G_{V_2} are both empty. If G is bipartite with vertex components V_1 and V_2 such that all $\alpha \in V_1$ have the same degree in G and all $\beta \in V_2$ have the same degree in G, we call G *bi-regular*.

2.1.2. Vectors and Matrices

We generally index vectors and matrices by discrete finite sets. For a set V, a vector in $\mathbb{R}^V$ is denoted by $y = (y_\alpha)_{\alpha\in V}$ and for $a \subseteq V$, $y_a = (y_\alpha)_{\alpha\in a}$. A matrix A in $\mathbb{R}^{V\times V}$ is written as $A = (a_{\alpha\beta})_{\alpha,\beta\in V}$.

For $A = (a_{\alpha\beta})_{\alpha,\beta\in V}$ and an undirected graph $G = (V, E)$ we let $A(G)$ denote the $\mathbb{R}^{V\times V}$ matrix given by

$$A(G)_{\alpha\beta} = \begin{cases} a_{\alpha\beta} & \text{if } \alpha\beta \in E \\ 0 & \text{otherwise} \end{cases}.$$

For a set of matrices M, $M(G) = \{A(G) \mid A \in M\}$, and M^+ denotes the set of positive-definite matrices inside M. The set of symmetric matrices is denoted by $\mathcal{S}$. If not stated explicitly, the indexing set of a matrix is determined by the context.

We indicate that a matrix is symmetric by only writing its elements on the diagonal and above. Asterisks as matrix entries indicate that the corresponding entry is unconstrained apart from any restrictions stated explicitly.

2.1.3. Graph Colouring

For an undirected graph G with vertex set V, a *vertex colouring* of G is a partition $\mathcal{V} = \{V_1, \ldots, V_k\}$ of V, where we refer to $V_1, \ldots, V_k$ as the *vertex colour classes*. Similarly, an *edge colouring* of G is a partition $\mathcal{E} = \{E_1, \ldots, E_l\}$ of the edge set into l *edge colour classes* $E_1, \ldots, E_l$. A colour class consisting of just one element is called a *singleton*, and a colour class with more than one element *composite*. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called a *coloured graph* and $(\mathcal{V}, \mathcal{E})$ a *graph colouring*.

For $\mathcal{V}$ and $\mathcal{E}$ as above, and $u \in \mathcal{V}$, we let T^u denote the $V \times V$ diagonal matrix with $T^u_{\alpha\alpha} = 1$ if and only if $\alpha \in u$ and zero otherwise. Similarly for $u \in \mathcal{E}$, we let T^u be the symmetric $V \times V$ matrix with $T^u_{\alpha\beta} = 1$ if and only if $\alpha\beta \in u$ and zero otherwise. We call $\{T^u\}_{u \in \mathcal{V} \cup \mathcal{E}}$ the *generating matrices* of $\mathcal{V}$ and $\mathcal{E}$.

For a partition $\mathcal{P}$ of a set S and $a, b \in S$ we write $a \equiv b \ (\mathcal{P})$ to denote that a and b are elements of the same set in $\mathcal{P}$. For two partitions $\mathcal{P}_1, \mathcal{P}_2$ of the same set we write $\mathcal{P}_1 \leq \mathcal{P}_2$ if each set in $\mathcal{P}_2$ is a union of sets in $\mathcal{P}_1$ and say that $\mathcal{P}_1$ is *finer* than $\mathcal{P}_2$, or equivalently that $\mathcal{P}_2$ is *coarser* than $\mathcal{P}_1$.

In our display of coloured graphs, we let black vertices and edges represent singleton colour classes. Whenever a vertex receives a colour, for clarity of exposition, we indicate its colour class by the number of asterisks we place next to it. Similarly, we indicate the colour class of an edge by dashes.

2.1.4. Permutation Groups

For a discrete set A, we let $S(A)$ denote the symmetric group of A, which contains all permutations σ acting on A. We denote permutations in cycle notation, so that $\sigma = (i_1 \ i_2 \ldots i_r)$ maps i_k to i_{k+1} for $1 \leq k < r$ and i_r to i_1. It is a group theoretic fact that every permutation $\sigma \in S(A)$ can be uniquely expressed as a product of disjoint cycles (up to their order). For $\sigma_1, \sigma_2 \in S(A)$, we write $\sigma_1\sigma_2$ to mean $\sigma_1 \circ \sigma_2$.

For a group $\Gamma \subseteq S(A)$ and $a \in A$, we call the set $\{b \mid \exists\, \sigma \in \Gamma \text{ such that } \sigma(a) = b\}$ the Γ-*orbit* of a in A, which we sometimes simply abbreviate by *orbit* if it is clear which group the orbit is generated by. The relation of being in the same Γ-orbit is an equivalence relation on A. Thus Γ partitions A into disjoint orbits and we call this partition Γ-*orbit partition* or simply *orbit partition* of A.

$\Gamma' \subseteq S(A)$ is a *subgroup* of $\Gamma \subseteq S(A)$, denoted $\Gamma' \leq \Gamma$, if Γ' is a subset of Γ and Γ' is itself a group with the same binary operation as Γ. Γ' is a *proper* subgroup of Γ if Γ' is a subgroup of Γ and $\Gamma' \neq \Gamma$. A *generating set* $\{g_1, \ldots, g_k\}$, with *generators* $g_1, \ldots, g_k$, of Γ is a subset of Γ which is not contained in any proper subgroup of Γ. Equivalently, $\{g_1, \ldots, g_k\}$ is a generating set of Γ if it is a subset of Γ and every element of Γ can be expressed as a combination of the g_i and their inverses. We then write $\Gamma = \langle g_1, \ldots, g_k \rangle$ and say that Γ is *generated* by $g_1, \ldots, g_k$. A generating set does not have to be unique.

Each permutation $\sigma \in S(A)$ can be represented by an $A \times A$ *permutation matrix* $G(\sigma)$ with $G(\sigma)_{\alpha\beta} = 1$ if and only if σ maps β to α and zero otherwise for $\alpha, \beta \in A$. For all $\sigma \in S(A)$,

$$G(\sigma)^T = G(\sigma^{-1}) = G(\sigma)^{-1}$$

so that for $Y \in \mathbb{R}^A$, $\Sigma \in \mathbb{R}^{A \times A}$, $\sigma \in S(A)$ and $\alpha, \beta \in A$,

$$\big(G(\sigma)Y\big)_\alpha = Y_{\sigma(\alpha)} \qquad \text{and} \qquad \big(G(\sigma)\Sigma G(\sigma)^{-1}\big)_{\alpha\beta} = \Sigma_{\sigma(\alpha)\sigma(\beta)}. \tag{2.1}$$

The *direct product* $\Gamma = \Gamma_1 \times \Gamma_2$ of two groups Γ_1 and Γ_2 is given by $\Gamma = \{(\sigma_1, \sigma_2) \mid \sigma_1 \in \Gamma_1,\ \sigma_2 \in \Gamma_2\}$ and binary operation $(\sigma_1, \sigma_2)(\sigma_1', \sigma_2') = (\sigma_1\sigma_1', \sigma_2\sigma_2')$. For a graph $G = (V, E)$, we let $\mathrm{Aut}(G)$ denote the automorphism group of G, containing all permutations $\sigma \in S(V)$ which leave E invariant. For a comprehensive introduction to group theory see Rotmann (1995).

2.2. Graphical Gaussian Models

We closely follow parts of Chapters 3 and 5 in Lauritzen (1996) in our exposition.

2.2.1. The Multivariate Gaussian Distribution

Let $Y = (Y_\alpha)_{\alpha \in V}$ be a random vector in $\mathbb{R}^V$, $\mu \in \mathbb{R}^V$ and $\Sigma \in \mathbb{R}^{V \times V} \cap \mathcal{S}^+$. Y follows a *multivariate normal* or *Gaussian distribution* with *mean* μ and *covariance matrix* Σ, denoted by $Y \sim \mathcal{N}_V(\mu, \Sigma)$ if the density of Y is given by

$$\begin{aligned} f(y \mid \mu, \Sigma) &= \frac{1}{(2\pi)^{|V|/2}(\det \Sigma)^{1/2}} \exp\{-(y-\mu)^T \Sigma^{-1}(y-\mu)/2\} \\ &\propto \left(\det \Sigma^{-1}\right)^{1/2} \exp\{-y^T\Sigma^{-1}y/2 + \mu^T\Sigma^{-1}y - \mu^T\Sigma^{-1}\mu/2\} \\ &= \left(\det \Sigma^{-1}\right)^{1/2} \exp\{-\mathrm{tr}(\Sigma^{-1}yy^T/2) + \mu^T\Sigma^{-1}y - \mu^T\Sigma^{-1}\mu/2\} \quad (2.2) \end{aligned}$$

where we refer to Σ^{-1} as the *concentration matrix*. If we let

$$\theta = (\Sigma^{-1}, \Sigma^{-1}\mu) = (\Sigma^{-1}, \xi) \in \Theta = (\mathcal{S}^+, \mathbb{R}^V) \tag{2.3}$$

and

$$t(y) = (-yy^T/2, y) \tag{2.4}$$

then for the inner product on Θ

$$\langle (S_1, v_1), (S_2, v_2) \rangle = \mathrm{tr}(S_1 S_2) + v_1^T v_2,$$

the expression in the exponential in (2.2) can be written as

$$\langle \theta, t(y) \rangle - \mu^T \Sigma^{-1} \mu / 2.$$

Since the integral

$$\int_{\mathbb{R}^V} e^{-(y-\mu)^T \Sigma^{-1} (y-\mu)/2} dy$$

is finite if and only if Σ^{-1} is positive definite, it follows that the multivariate Gaussian distribution $\mathcal{N}_V(\mu, \Sigma)$ is a regular exponential family (Brown, 1986) with canonical parameter θ as in (2.3) and canonical statistic $t(y)$ as in (2.4).

The entries in Σ^{-1} have a simple interpretation. The diagonal elements $(\Sigma^{-1})_{\alpha\alpha}$ are reciprocals of the conditional variances given the remaining variables,

$$(\Sigma^{-1})_{\alpha\alpha} = \mathrm{Var}(Y_\alpha | Y_{V \setminus \{\alpha\}})^{-1} \tag{2.5}$$

for all $\alpha \in V$, and letting

$$C_{\alpha\beta} = \frac{(\Sigma^{-1})_{\alpha\beta}}{\sqrt{(\Sigma^{-1})_{\alpha\alpha} (\Sigma^{-1})_{\beta\beta}}} \tag{2.6}$$

for $\alpha, \beta \in V$, $\alpha \neq \beta$, gives the negative partial correlation coefficients

$$\rho_{\alpha\beta | V \setminus \{\alpha,\beta\}} = \frac{\mathrm{Cov}(Y_\alpha, Y_\beta | Y_{V \setminus \{\alpha,\beta\}})}{\mathrm{Var}(Y_\alpha | Y_{V \setminus \{\alpha,\beta\}})^{1/2} \mathrm{Var}(Y_\beta | Y_{V \setminus \{\alpha,\beta\}})^{1/2}} = -C_{\alpha\beta} \tag{2.7}$$

for $\alpha, \beta \in V$, $\alpha \neq \beta$. It follows that $(\Sigma^{-1})_{\alpha\beta} = 0$ if and only if Y_α and Y_β are conditionally independent given the remaining variables. Adopting the notation introduced by Dawid (1979), we write

$$(\Sigma^{-1})_{\alpha\beta} = 0 \quad \Longleftrightarrow \quad Y_\alpha \perp\!\!\!\perp Y_\beta \mid Y_{V \setminus \{\alpha,\beta\}}. \tag{2.8}$$

2.2.2. Markov Properties for Undirected Graphs

Definition 2.1 Let $Y = (Y_\alpha)_{\alpha \in V}$ and let $G = (V, E)$ be an undirected graph. Then the distribution of Y is said to obey

(P) the *pairwise Markov property* (relative to G) if for all $\alpha, \beta \in V$,

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

(L) the *local Markov property* (relative to G) if for all $\alpha \in V$,

$$Y_\alpha \perp\!\!\!\perp Y_{V \setminus \mathrm{cl}(\alpha)} \mid \; Y_{\mathrm{bd}(\alpha)},$$

(G) the *global Markov property* (relative to G) if for all $A, B, S \subseteq V$,

$$A \perp_G B \mid S \quad \Rightarrow \quad Y_A \perp\!\!\!\perp Y_B \mid Y_S$$

where $\perp_G$ denotes graph separation. A model is said to obey the above Markov properties if all its distributions do.

□

A further property associated with G and $(Y_\alpha)_{\alpha \in V}$ is the factorisation property.

Definition 2.2 For $Y = (Y_\alpha)_{\alpha \in V}$ and $G = (V, E)$, a density or distribution satisfies the *factorisation property* (F) (relative to G) if it may be written as the product $\Pi_{a \in \mathcal{A}} \psi_a(y)$, where $\mathcal{A}$ is the set of subsets of V which induce complete subgraphs of G, and $\psi_a(y)$ depends on y_a only. A model satisfies the factorisation property if all its distributions do.

□

The factorisation property can be visualised by a *factor graph* (Frey et al., 1998).

Definition 2.3 (Frey) Let $y = (y_\alpha)_{\alpha \in V}$ and let $g(y)$ be a function in y which factorises as $g(y) = \Pi_{i \in I} g_{A_i}(y)$ where $A_i \subseteq V$ and g_{A_i} cannot be factorised further for $i \in I$. The *factor graph* of g is the graph $G_F = (V \cup F, E_F)$ with F being the set of factor nodes $F = \{g_{A_i}\}_{i \in I}$ and $E_F = \{\alpha g_{A_i} \mid \alpha \in V, g_{A_i} \in F \text{ with } \alpha \in A_i\}$.

□

The properties satisfy the following structural relations. The first part of the result is due to Pearl and Paz (1987), while the second part is usually attributed to Hammersley and Clifford (1971).

Proposition 2.1 *If a probability distribution on $Y = (Y_\alpha)_{\alpha \in V}$ satisfies*

$$Y_A \perp\!\!\!\perp Y_B \mid Y_{C \cup D} \quad \text{and} \quad Y_A \perp\!\!\!\perp Y_C \mid Y_{B \cup D} \quad \Longrightarrow \quad Y_A \perp\!\!\!\perp Y_{(B \cup C)} \mid Y_D$$

for all disjoint subsets A, B, C and D of V, then

$$(G) \Longleftrightarrow (L) \Longleftrightarrow (P).$$

Further, for positive probability distributions with continuous positive density,

$$(F) \Longleftrightarrow (G) \Longleftrightarrow (L) \Longleftrightarrow (P).$$

□

2.2.3. Graphical Gaussian Models

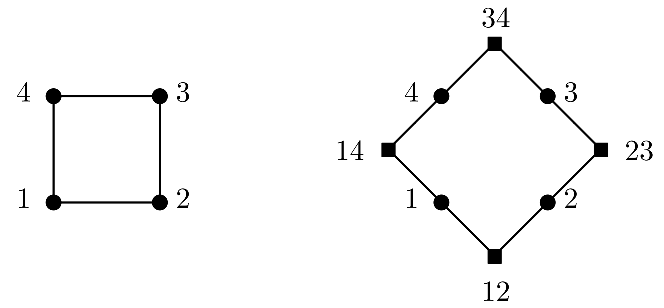
Definition 2.4 If $G = (V, E)$ is an undirected graph and $Y = (Y_\alpha)_{\alpha \in V}$ is a random variable taking values in $\mathbb{R}^V$, the *graphical Gaussian model* for Y with graph G is given by assuming that Y follows a $\mathcal{N}_V(\mu, \Sigma)$ distribution with $\mu \in \mathbb{R}^V$ and $\Sigma \in \mathcal{S}^+$ which obeys the Markov properties with respect to G.

□

As the density of a multivariate normal distribution is positive and continuous, by Proposition 2.1, graphical Gaussian models satisfy the factorisation property. It follows that relation (2.8) is equivalent to assuming the concentrations $(\Sigma^{-1})_{\alpha\beta}$ to be equal to zero for all pairs $\alpha, \beta \in V$ which are not adjacent in G. The graphical Gaussian model for Y can thus be compactly described as

$$Y \sim \mathcal{N}_V(\mu, \Sigma), \quad \mu \in \mathbb{R}^V, \quad \Sigma^{-1} \in \mathcal{S}^+(G).$$

Further, in factor graphs of graphical Gaussian models, in the notation of Definition 2.3, either $A_i = \{\alpha\}$ or $A_i = \{\alpha, \beta\}$ for $\alpha, \beta \in V$, with a factor being present if and only if the corresponding entry in Σ^{-1} is not restricted to zero. Thus the factor graph of a graphical Gaussian model with graph $G = (V, E)$ equals $G_F = (V \cup F, E_F)$ with $F = \{e \mid e \in E\}$ and $E_F = \{\alpha e \mid \alpha \in V,\ e \in E \text{ is incident with } \alpha \text{ in } G\}$. An example is given in Figure 2.1 for $V = [4]$, displayed as circles, and $F = \{12, 14, 23, 34\}$, displayed as squares.



(a) $G = (V, E)$ (b) $G_F = (V \cup F, E_F)$

Figure 2.1.: Graph $G = (V, E)$ and factor graph $G_F = (V \cup F, E_F)$.

As the constraints of graphical Gaussian models on the multivariate distribution are linear in Σ^{-1}, they form regular exponential families, with

$$\begin{aligned}\theta_0 &= (\Sigma^{-1}(G), \xi) \in \Theta_0 = \Theta \cap (\mathcal{S}(G), \mathbb{R}^V) = (\mathcal{S}^+(G), \mathbb{R}^V) \\ t_0(y) &= (-yy^T(G)/2, y)\end{aligned}$$

where Θ is as in (2.3).

2.3. Symmetry in the Gaussian Distribution

As the elegant principle of symmetry plays a significant role in many disciplines describing natural phenomena, symmetry studies and methodology incorporating inherent symmetry appear throughout the Statistics literature. Classical examples are designs of comparative experiments (Bailey, 2008) and analysis of structured data (Dawid, 1988; Viana, 2008). Further, invariance arguments appear in estimation and decision problems (Eaton, 1983, 1989).

The Gaussian distribution has received attention in many applications with intrinsic symmetries. Examples are interchangeable tests, referred to as parallel forms (Wilks, 1946; Votaw, 1948), signal transmission with receivers arranged on the vertices of a regular polygon (Olkin and Press, 1969; Olkin, 1972), responses at election polls (Arnold, 1973) and observations between related measurements in biology and the medical sciences (Votaw et al., 1950; Arnold, 1973; Olkin, 2000).

Below, we distinguish between permutation and general group induced symmetry. In this thesis we only treat permutation symmetry, however include general group symmetry here for completeness.

2.3.1. Permutation Symmetry

A covariance structure $\Sigma \in \mathcal{S}^+$ is said to be subject to permutation induced symmetry if there exists a group $\Gamma \leq S(V)$ such that

$$G(\sigma)\Sigma G(\sigma)^{-1} = \Sigma \qquad \text{for all } \sigma \in \Gamma \tag{2.9}$$

and $\Sigma_{\alpha\beta} = \Sigma\gamma\delta$ only if there exists $\sigma \in \Gamma$ such that σ maps $\{\alpha, \beta\}$ to $\{\gamma, \delta\}$. Γ is then said to *generate* the covariance structure.

Permutation group induced symmetry structures in the covariance which have received particular attention are the *spherical* covariance structure,

$$\Sigma = \sigma^2 I, \qquad \sigma^2 > 0, \tag{2.10}$$

generated by $\Gamma = S(V)$, the *intraclass* covariance structure, also referred to as *completely symmetric* or *uniform*,

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} \in \mathcal{S}^+, \tag{2.11}$$

also generated by $\Gamma = S(V)$, the *circular* covariance structure,

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{12} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1k} & \rho_{2k} & \cdots & 1 \end{pmatrix} \in \mathcal{S}^+, \qquad \rho_{ij} = \rho_{|i-j|}, \tag{2.12}$$

with $\Gamma = \langle(1\ 2\ 3\ \ldots\ |V|)\rangle$, and the *compound symmetry* covariance structure,

$$\Sigma = \begin{pmatrix} \Sigma_1 & \sigma_{12}J_{12} & \cdots & \sigma_{1k}J_{1k} \\ \sigma_{12}J_{12} & \Sigma_2 & \cdots & \sigma_{2k}J_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1k}J_{1k} & \sigma_{2k}J_{2k} & \cdots & \Sigma_k \end{pmatrix} \in \mathcal{S}^+, \tag{2.13}$$

with Σ_i having the intraclass covariance structure and J_{ij} being a matrix of suitable dimensions whose entries are all one for $1 \leq i, j \leq k$. The symmetry is generated by $\Gamma = S(V_1) \times S(V_2) \times \ldots \times S(V_k)$, V_i being the index set of Σ_i. Each of the above covariance structures contains the previously stated ones as special cases.

Past treatments are primarily concerned with estimation and hypothesis testing. Wilks (1946) considered testing the null hypotheses

$$\begin{aligned} H_{mvc}: &\quad \text{intraclass structure in } \Sigma,\ \mu_\alpha = \mu \text{ for all } \alpha \in V \\ H_{vc}: &\quad \text{intraclass structure in } \Sigma \\ H_m: &\quad \mu_\alpha = \mu \text{ for all } \alpha \in V,\ \text{given that } \Sigma \text{ has an intraclass structure} \end{aligned}$$

based on one sample and developed the likelihood ratio test statistics to test the above hypotheses against the general alternative. Exact moments and null distributions of the statistics were found to be given by products of independent beta distributions and expressions involving their logarithms were shown to have χ^2 asymptotic distributions with appropriate numbers of degrees of freedom.

Votaw (1948) generalised the results in Wilks (1946) for the compound symmetry structure for more than one sample by repeating the analysis for 12 null hypotheses, which contained H_{mvc}, H_{vc} and H_m as special cases.

Olkin and Press (1969) studied the spherical, intraclass and circular covariance structures and considered the following statistical tests:

H_1 : sphericity versus circularity (with unconstrained mean vector)

H_2 : intraclass versus circularity (with unconstrained mean vector)

H_3 : circularity versus general alternative (with unconstrained mean vector)

H_4 : $\mu = 0$, given that Σ has circular structure

H_5 : $\mu_\alpha = \mu$ for all $\alpha \in V$, given that Σ has circular structure

The corresponding likelihood ratio tests were shown to be distributed as products of suitable independent beta distributions and different approximations of null and non-null distributions of the test statistics were obtained.

A further generalisation of previous results was obtained by Olkin (1972) who considered block symmetric models, in which Σ is partitioned into blocks that are arranged according to a spherical, intraclass or circular structure. The analysis in Olkin and Press (1969) was repeated for $H'_1, \ldots, H'_5$, with H'_i denoting the block variant of H_i, and analogous results were obtained. Further developments can be found in Olkin and Press (1969); Olkin (1972); Young (1975); Olkin (2000), and the references therein.

2.3.2. General Group Induced Symmetry

Dropping the assumption of a fixed basis for the observations, Andersson (1975) unified all previously obtained results on Gaussian models invariant under a permutation group by considering them as special cases of *invariant normal models*, which he defined to be Gaussian models invariant under a general group Γ acting on the observation space. He showed all invariant normal models to be given by products of models with real, complex or quaternion covariance structures, and derived the maximum likelihood estimator of μ and Σ and its distribution in full generality.

Andersson et al. (1983) studied ten hypothesis testing problems involving the above covariance structures and obtained their likelihood ratio test statistics. It was shown that all testing problems in which both the null and alternative hypothesis are an invariant normal model can be decomposed into simpler problems, each of which has the form of one of the ten problems referred to above. A further covariance structure, parametrised by a special Clifford Algebra, was found and analysed in the same fashion by Jensen (1988). Estimation and hypothesis testing for the complex covariance structure were further considered in Andersen et al. (1995).

2.3.3. Graphical Symmetry Models

A first treatment of a class of graphical Gaussian models with symmetry constraints generated by permutation is given in Hylleberg et al. (1993), who studied models which they termed *graphical symmetry models*. These combine permutation symmetry induced restrictions on Σ^{-1} with conditional independence constraints, and thus, in particular, contain the covariance structures (2.9)-(2.13) as special cases.

Put formally, if for a permutation group $\Gamma \leq S(V)$

$$\mathcal{S}(\Gamma) = \{S \in \mathbb{R}^{V \times V} \mid S^T = S,\ G(\sigma)SG(\sigma)^{-1} = S,\ \forall\, \sigma \in \Gamma\} \tag{2.14}$$

then the graphical symmetry model with dependence structure represented by $G = (V, E)$ and symmetry constraints on Σ^{-1} induced by Γ is given by assuming

$$Y \sim \mathcal{N}_V(0, \Sigma), \quad \Sigma^{-1} \in \mathcal{S}(\Gamma) \cap \mathcal{S}^+(G). \tag{2.15}$$

As for all $\sigma \in S(V)$ and all $S \in \mathbb{R}^{V \times V}$ such that S is non-singular,

$$\begin{aligned} G(\sigma) S G(\sigma)^{-1} = S \quad &\iff \quad G(\sigma) S^{-1} G(\sigma)^{-1} = S^{-1} \\ S^T = S \quad &\iff \quad \left(S^{-1}\right)^T = S^{-1} \end{aligned}$$

the condition on Σ^{-1} in (2.15) may be equivalently expressed as

$$\Sigma^{-1} \in \{\Sigma^{-1} \mid \Sigma^T = \Sigma,\ G(\sigma)\Sigma G(\sigma)^{-1} = \Sigma,\ \forall\ \sigma \in \Gamma\} \cap \mathcal{S}^+(G)$$

which exhibits the constraints of graphical symmetry models to be combinations of the previously studied permutation symmetry induced constraints on Σ given in (2.9) with conditional independence constraints.

In direct analogy to previous work on the Gaussian distribution, Hylleberg et al. (1993) studied maximum likelihood estimation of Σ, paying particular attention to the special case of compound covariance structure, given in (2.13). Explicit likelihood ratio tests for additional conditional independence constraints were developed for decomposable graphs, i.e. for graphs which do not contain cycles of length greater than 3 as induced subgraphs, and possible tests for additional symmetry constraints were described and characterised.

It will be shown in Section 3.4.1 that graphical symmetry models are a special case of the models studied in this thesis.

For completeness we mention that symmetry studies for other Gaussian conditional independence models, not considered in this thesis, include Andersson and Madsen (1998) and Madsen (2000).

3. Model Definitions and First Properties

This chapter is devoted to the formal introduction of the models this thesis is concerned with, as they were defined in Højsgaard and Lauritzen (2008).

3.1. RCON Models: Equality Restrictions on Concentrations

RCON models place equality restrictions on the concentration matrix Σ^{-1}. They restrict off-diagonal elements separately from those lying on the diagonal, so that the models can be represented by coloured graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The corresponding set of positive definite matrices is denoted by $\mathcal{S}^+(\mathcal{V}, \mathcal{E})$. Put formally:

Definition 3.1 Let $G = (V, E)$ be an undirected graph, let $(\mathcal{V}, \mathcal{E})$ be a colouring of G and let Y be a random vector in $\mathbb{R}^V$. Then the *RCON model* for Y *represented by* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is given by assuming

$$Y \sim \mathcal{N}_V(0, \Sigma), \quad \Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$$

with

$$\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \left\{ \sum_{u \in \mathcal{V} \cup \mathcal{E}} \lambda_u T^u, \ \lambda \in \mathbb{R}^{\mathcal{V} \cup \mathcal{E}} \right\}^+. \tag{3.1}$$

□

Thus in brief, RCON models restrict the entries of Σ^{-1} according to

$$\begin{aligned}
\alpha \equiv \beta\ (\mathcal{V}) \quad &\Longrightarrow \quad (\Sigma^{-1})_{\alpha\alpha} = (\Sigma^{-1})_{\beta\beta} \quad \text{for } \alpha, \beta \in V \\
\alpha\beta \equiv \gamma\delta\ (\mathcal{E}) \quad &\Longrightarrow \quad (\Sigma^{-1})_{\alpha\beta} = (\Sigma^{-1})_{\gamma\delta} \quad \text{for } \alpha\beta, \gamma\delta \in E \\
\alpha\beta \notin E \quad &\Longrightarrow \quad (\Sigma^{-1})_{\alpha\beta} = 0 \quad \text{for } \alpha, \beta \in V,\ \alpha \neq \beta
\end{aligned}$$

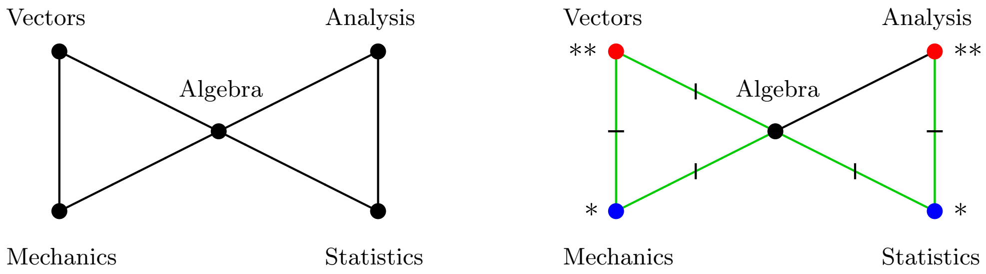
which exhibits RCON models as regular exponential families. For all $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $\mathcal{S}^{+}(\mathcal{V}, \mathcal{E})$ is an open convex cone. The following RCON model will be considered further in Section 6.2.3.

Example 3.1 The data consist of the examination marks of 88 students in five mathematical subjects (Mardia et al., 1979). Whittaker (1990) and Edwards (2000) previously demonstrated an excellent fit to the unconstrained model represented by the graph shown in Figure 3.1(a). Højsgaard and Lauritzen (2008) show the data to also support the RCON model represented by the graph in Figure 3.1(b). The model specifications are

$$Y \sim \mathcal{N}_V(0, \Sigma), \qquad \Sigma^{-1} \in M,$$

with M given below the corresponding graphs. If the subjects are indexed in alphabetical order, the graph colouring representing the constraints of the RCON model is given by $(\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{\{1\}, \{2, 5\}, \{3, 4\}\}$ and $\mathcal{E} = \{\{12\}, \{13, 14, 15, 24, 35\}\}$. Note that the number of model parameters has been reduced from 11 to 5.

□



$$M = \left\{ \begin{pmatrix} * & * & * & * & * \\ & * & 0 & * & 0 \\ & & * & 0 & * \\ & & & * & 0 \\ & & & & * \end{pmatrix} \right\}^{+}$$

(a) Conditional independence structure.

$$M = \left\{ \begin{pmatrix} * & * & \lambda_{13} & \lambda_{13} & \lambda_{13} \\ & \lambda_{22} & 0 & \lambda_{13} & 0 \\ & & \lambda_{33} & 0 & \lambda_{13} \\ & & & \lambda_{33} & 0 \\ & & & & \lambda_{22} \end{pmatrix} \right\}^{+}$$

(b) RCON model.

Figure 3.1.: Mathematics marks example.

3.2. RCOR Models: Equality Restrictions on Partial Correlations

RCOR models place equality restrictions on the diagonal elements of the concentration matrix Σ^{-1} and on the non-zero partial correlations $\rho_{\alpha\beta|V\setminus\{\alpha,\beta\}}$, given in (2.7). If for a given Σ^{-1}, we let $A, C \in \mathbb{R}^{V\times V}$ be given by

$$A_{\alpha\beta} = \begin{cases} \sqrt{(\Sigma^{-1})_{\alpha\alpha}} & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases} \qquad C_{\alpha\beta} = \begin{cases} 1 & \alpha = \beta \\ -\rho_{\alpha\beta|V\setminus\{\alpha,\beta\}} & \alpha \neq \beta \end{cases}, \tag{3.2}$$

then, by (2.6), $\Sigma^{-1} = ACA$, and

$$\begin{aligned} (\Sigma^{-1})_{\alpha\alpha} = (\Sigma^{-1})_{\beta\beta} \quad &\Longleftrightarrow \quad A_{\alpha\alpha} = A_{\beta\beta} \quad \text{for } \alpha, \beta \in V \\ \rho_{\alpha\beta|V\setminus\{\alpha,\beta\}} = \rho_{\gamma\delta|V\setminus\{\gamma,\delta\}} \quad &\Longleftrightarrow \quad C_{\alpha\beta} = C_{\gamma\delta} \quad \text{for } \alpha\beta, \gamma\delta \in E \end{aligned}$$

leading to the following definition for RCOR models.

Definition 3.2 Let $G = (V, E)$ be an undirected graph, let $(\mathcal{V}, \mathcal{E})$ be a colouring of G and let Y be a random vector in $\mathbb{R}^V$. Then the *RCOR model* for Y *represented by* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is given by assuming

$$Y \sim \mathcal{N}_V(0, \Sigma), \quad \Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$$

where

$$\mathcal{R}^+(\mathcal{V}, \mathcal{E}) = \left\{ ACA \mid A = \sum_{u \in \mathcal{V}} \eta_u T^u, \ \eta \in \mathbb{R}_+^{\mathcal{V}}, \ C = I + \sum_{u \in \mathcal{E}} \tau_u T^u, \ \tau \in (-1, 1)^{\mathcal{E}} \right\}^+ . \tag{3.3}$$

□

Thus in brief, RCOR models restrict the entries of Σ^{-1} according to

$$\begin{aligned}
\alpha \equiv \beta \ (\mathcal{V}) \quad &\Longrightarrow \quad (\Sigma^{-1})_{\alpha\alpha} = (\Sigma^{-1})_{\beta\beta} \quad \text{for } \alpha, \beta \in V \\
\alpha\beta \equiv \gamma\delta \ (\mathcal{E}) \quad &\Longrightarrow \quad C_{\alpha\beta} = C_{\gamma\delta} \quad \text{for } \alpha\beta, \gamma\delta \in E \\
\alpha\beta \notin E \quad &\Longrightarrow \quad (\Sigma^{-1})_{\alpha\beta} = 0 \quad \text{for } \alpha, \beta \in V, \ \alpha \neq \beta.
\end{aligned}$$

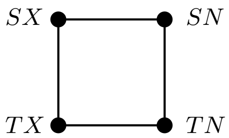
Note that as, by (2.6) and (2.7), for $\alpha \neq \beta$, $\Sigma^{-1}_{\alpha\beta} = 0$ if and only if $\rho_{\alpha\beta|V\setminus\{\alpha,\beta\}} = 0$, the above constraints are indeed consistent.

$\mathcal{R}^+(\mathcal{V}, \mathcal{E})$ forms a differentiable manifold in $\mathcal{S}^+$ which is not generally a cone, so that RCOR models are only curved exponential families (Brown, 1986). RCOR models are scale invariant if variables inside the same vertex colour class are rescaled in the same way (Højsgaard and Lauritzen, 2008). Thus they are particularly suitable for variables measured on different scales.

Example 3.2 The data is concerned with anxiety and anger in a trait and state version of 684 students (Cox and Wermuth, 1993) and strongly support the conditional independence model displayed in Figure 3.2(a). As shown in Højsgaard and Lauritzen (2008), they also support the RCOR model represented by the coloured graph in Figure 3.2(b), parametrised by 6 parameters rather than 8. The variable names are combinations between T or S, for 'trait' and 'state', and X or N, standing for 'anxiety' and 'anger'. The model specifications are

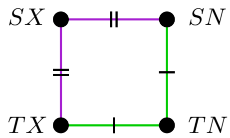
$$Y \sim \mathcal{N}_V(0, \Sigma), \qquad \Sigma^{-1} \in M$$

with M given below the graphs. The variables are indexed anti-clockwise starting from TX.



$$M = \left\{ \begin{pmatrix} * & * & 0 & * \\ & * & * & 0 \\ & & * & * \\ & & & * \end{pmatrix} \right\}^+$$

(a) Conditional independence structure.



$$M = \left\{ \begin{pmatrix} \eta_1^2 & \eta_1\tau_1\eta_2 & 0 & \eta_1\tau_2\eta_4 \\ & \eta_2^2 & \eta_2\tau_1\eta_3 & 0 \\ & & \eta_3^2 & \eta_3\tau_2\eta_4 \\ & & & \eta_4^2 \end{pmatrix} \right\}^+$$

(b) RCOR model.

Figure 3.2.: Personality characteristics example.

$\square$

3.3. Coloured Factor Graphs

Just as unconstrained graphical Gaussian models may be represented by factor graphs, RCON and RCOR models can be represented by coloured factor graphs.

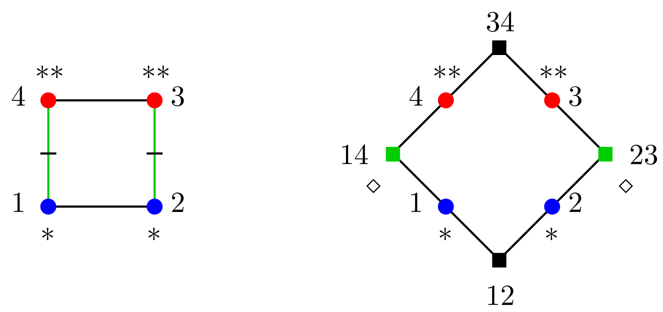
Definition 3.3 For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ representing an RCON or RCOR model, let N be a set of nodes with each node representing an edge in E and let $\mathcal{N}$ be the colouring of N in which nodes receive the same colour if and only if the corresponding edges are equally coloured in $\mathcal{E}$. The *coloured factor graph* of the model is the vertex and node coloured graph

$$\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F)$$

where $E_F = \{\alpha n \mid \alpha \in V, n \in N$ and n represents an edge incident with $\alpha\}$.

□

We provide an example for illustration in Figure 3.3.



(a) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (b) $\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F)$

Figure 3.3.: Coloured graph $\mathcal{G}$ and coloured factor graph $\mathcal{G}_F$.

3.4. Structural Relations Between Model Types

While RCON and RCOR models do not place equality constraints on the same parameter sets, an RCON and RCOR model represented by the same coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ may still yield the same model constraints. The corresponding necessary and sufficient condition on $(\mathcal{V}, \mathcal{E})$ is given in Højsgaard and Lauritzen (2008). First we define *edge regularity* of a graph colouring.

Definition 3.4 Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a coloured graph. We say that $(\mathcal{V}, \mathcal{E})$ is *edge regular* if any unordered pair of edges in the same colour class in $\mathcal{E}$ connects the same vertex colour classes in $\mathcal{V}$.

□

Proposition 3.1 (Højsgaard and Lauritzen) *The RCON and RCOR models that are determined by the coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ yield identical restrictions*

$$\mathcal{S}^{+}(\mathcal{V}, \mathcal{E}) = \mathcal{R}^{+}(\mathcal{V}, \mathcal{E})$$

if and only if $(\mathcal{V}, \mathcal{E})$ is edge regular.

□

3.4.1. Permutation Symmetry

A special case of the models characterised in Proposition 3.1 are models with equality constraints which are induced by permutation symmetry, i.e. models which were termed graphical symmetry models in Hylleberg et al. (1993). This will be shown below in Proposition 3.2. First we introduce some notation.

Recall that for a permutation group $\Gamma \leq S(V)$ and an undirected graph $G = (V, E)$, the graphical symmetry model with dependence structure represented by G and symmetry constraints induced by Γ is given by assuming $\Sigma^{-1} \in \mathcal{S}(\Gamma) \cap \mathcal{S}^{+}(G)$, with $S(\Gamma)$ as in (2.14). Let

$$V^{(2)} = \{\alpha\beta \mid \alpha, \beta \in V\} = \{\alpha\beta \mid \alpha, \beta \in V, \alpha \neq \beta\} \cup \{\alpha\alpha \mid \alpha \in V\} = V^{(2)}_{\neq} \cup V^{(2)}_{=}$$

and for $\sigma \in \Gamma$, let $s_\sigma : V^{(2)} \to V^{(2)}$ be defined by $s_\sigma : \alpha\beta \mapsto \sigma(\alpha)\sigma(\beta)$. Then $\tilde{\Gamma} = \{s_\sigma \mid \sigma \in \Gamma\}$ is a group with binary operation given by composition of functions.

For $\sigma \in \Gamma \leq S(V)$ let $\sim_\sigma$ be the binary relation on $V^{(2)}$ given by

$$\alpha\beta \sim_\sigma \gamma\delta \quad \text{if} \quad s_\sigma(\alpha\beta) = \gamma\delta \tag{3.4}$$

and define the binary relation $\sim_\Gamma$ by

$$\alpha\beta \sim_\Gamma \gamma\delta \quad \text{if} \quad \exists\, \sigma \in \Gamma \text{ such that } \alpha\beta \sim_\sigma \gamma\delta. \tag{3.5}$$

As Γ is a group, $\sim_\Gamma$ is an equivalence relation on $V^{(2)}$ and the orbits of $\tilde{\Gamma}$ in $V^{(2)}$ are the equivalence classes of $\sim_\Gamma$. Further, since all $\sigma \in \Gamma$ are bijections on V, $\alpha\beta \not\sim_\Gamma \alpha\alpha$ for all $\alpha, \beta \in V$ with $\alpha \neq \beta$, so that the orbits of $\tilde{\Gamma}$ in $V^{(2)}$ either contain only elements from $V^{(2)}_{\neq}$ or only elements from $V^{(2)}_{=}$. We obtain that the $\tilde{\Gamma}$-orbit partition of $V^{(2)}$ is given by the union of a partition of $V^{(2)}_{\neq}$ and a partition of $V^{(2)}_{=}$.

Clearly the two sets $V^{(2)}_{=}$ and V are isomorphic and may therefore be identified, so that the $\tilde{\Gamma}$-orbit partition of $V^{(2)}$ induces a partition of $V^{(2)}_{\neq}$ and a partition of V. We denote these by $\mathcal{E}_\Gamma$ and $\mathcal{V}_\Gamma$ respectively, which, by construction, are defined by

$$\alpha \equiv \beta\ (\mathcal{V}_\Gamma) \qquad \Longleftrightarrow \qquad \alpha\alpha \sim_\Gamma \beta\beta, \quad \alpha, \beta \in V \tag{3.6}$$

$$\alpha\beta \equiv \gamma\delta\ (\mathcal{E}_\Gamma) \qquad \Longleftrightarrow \qquad \alpha\beta \sim_\Gamma \gamma\delta, \quad \alpha\beta,\ \gamma\delta \in V^{(2)}_{\neq}. \tag{3.7}$$

It is natural to associate the vertex and edge sets of the complete graph $K_V = (V, E_V)$ with $V^{(2)}$. $\alpha\beta \in V^{(2)}_{\neq}$ then corresponds to an edge in E_V and $\alpha\alpha \in V^{(2)}_{=}$ corresponds to a vertex in V. The partitions $\mathcal{V}_\Gamma$ and $\mathcal{E}_\Gamma$ may then be visualised by the coloured graph $\mathcal{K}_\Gamma = (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$.

Example 3.3 Let $V = [5]$ and

$$\Gamma = \langle (12), (345) \rangle = \{Id, (12), (345), (543), (12)(345), (12)(543)\}.$$

Then $\mathcal{V}_\Gamma = \{\{1,2\},\{3,4,5\}\}$ and $\mathcal{E}_\Gamma = \{\{12\},\{13,14,15,23,24,25\},\{34,35,45\}\}$. The corresponding coloured complete graph $\mathcal{K}_\Gamma = (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$ is given in Figure 3.4.

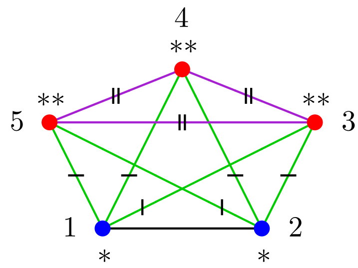


Figure 3.4.: $\mathcal{K}_\Gamma = (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$

□

Finally, for two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$, let

$$\mathcal{P}_1 \Delta \mathcal{P}_2 = \{C \mid C \in \mathcal{P}_1,\ C \subseteq D \text{ for some } D \in \mathcal{P}_2\}.$$

Lemma 3.1 *Let $G = (V, E)$ be an undirected graph and let $\Gamma \leq S(V)$ with $\mathcal{V}_\Gamma$ and $\mathcal{E}_\Gamma$ being the orbit partitions of V and $V_{\neq}^{(2)}$ induced by $\tilde{\Gamma}$. Then*

$$\mathcal{S}(\Gamma) \cap \mathcal{S}^+(G) = \mathcal{S}^+(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma \Delta \{E\}).$$

□

Proof: Let $\Sigma^{-1} \in \mathcal{S}(\Gamma) \cap \mathcal{S}^+(G)$. By definitions of $\mathcal{S}(\Gamma)$ in (2.14) and of $\sim_\Gamma$ in (3.5), we can deduce that for all $\alpha\beta, \gamma\delta \in V^{(2)}$

$$\alpha\beta \sim_\Gamma \gamma\delta \qquad \Longrightarrow \qquad \left(\Sigma^{-1}\right)_{\alpha\beta} = \left(\Sigma^{-1}\right)_{\gamma\delta}.$$

By (3.6) and (3.7), this is equivalent to

$$\alpha \equiv \beta \ (\mathcal{V}_\Gamma) \quad \Longrightarrow \quad \left(\Sigma^{-1}\right)_{\alpha\alpha} = \left(\Sigma^{-1}\right)_{\beta\beta}, \quad \alpha, \beta \in V \tag{3.8}$$

$$\alpha\beta \equiv \gamma\delta \ (\mathcal{E}_\Gamma) \quad \Longrightarrow \quad \left(\Sigma^{-1}\right)_{\alpha\beta} = \left(\Sigma^{-1}\right)_{\gamma\delta}, \quad \alpha\beta, \ \gamma\delta \in V_{\neq}^{(2)}. \tag{3.9}$$

Further, $\Sigma^{-1} \in \mathcal{S}^+(G)$ if and only if

$$\alpha\beta \notin E \quad \Longrightarrow \quad \left(\Sigma^{-1}\right)_{\alpha\beta} = 0 \tag{3.10}$$

and Σ^{-1} is positive definite, so that, if we let $E_\Gamma \Delta E$ denote the set of edges which lie in some set of $\mathcal{E}_\Gamma \Delta \{E\}$, then by (3.8), (3.9) and (3.10), $\Sigma^{-1} \in \mathcal{S}(\Gamma) \cap \mathcal{S}^+(G)$ if and only if

$$\begin{aligned}
\alpha \equiv \beta \ (\mathcal{V}_\Gamma) \quad &\Longrightarrow \quad \left(\Sigma^{-1}\right)_{\alpha\alpha} = \left(\Sigma^{-1}\right)_{\beta\beta}, \quad \alpha, \beta \in V \\
\alpha\beta \equiv \gamma\delta \ (\mathcal{E}_\Gamma \Delta \{E\}) \quad &\Longrightarrow \quad \left(\Sigma^{-1}\right)_{\alpha\beta} = \left(\Sigma^{-1}\right)_{\gamma\delta}, \quad \alpha\beta, \ \gamma\delta \in V_{\neq}^{(2)} \\
\alpha\beta \notin E_\Gamma \Delta E \quad &\Longrightarrow \quad \left(\Sigma^{-1}\right)_{\alpha\beta} = 0, \quad \alpha\beta \in V_{\neq}^{(2)},
\end{aligned}$$

and Σ^{-1} is positive definite, or equivalently

$$\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma \Delta \{E\}).$$

■

Example 3.4 Let V and Γ be as in Example 3.3 and let $G = (V, \{12, 15, 23, 34, 35, 45\})$ as displayed in Figure 3.5.

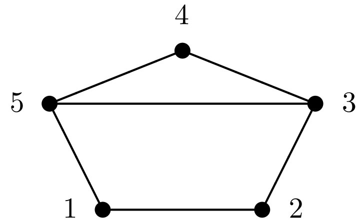


Figure 3.5.: $G = (V, \{12, 15, 23, 34, 35, 45\})$

Then the covariance structure of the graphical symmetry model with dependence structure represented by G and symmetry induced by Γ is given by assuming

$$\Sigma^{-1} \in \left\{ \begin{pmatrix} \lambda_{\mathbf{11}} & * & \lambda_{\mathbf{23}} & \lambda_{\mathbf{23}} & \lambda_{\mathbf{23}} \\ & \lambda_{\mathbf{11}} & \lambda_{\mathbf{23}} & \lambda_{\mathbf{23}} & \lambda_{\mathbf{23}} \\ & & \lambda_{\mathbf{33}} & \lambda_{\mathbf{34}} & \lambda_{\mathbf{34}} \\ & & & \lambda_{\mathbf{33}} & \lambda_{\mathbf{34}} \\ & & & & \lambda_{\mathbf{33}} \end{pmatrix} \right\} \cap \left\{ \begin{pmatrix} * & * & 0 & 0 & * \\ & * & * & 0 & 0 \\ & & * & * & * \\ & & & * & * \\ & & & & * \end{pmatrix} \right\}^{+}$$

$$= \left\{ \begin{pmatrix} \lambda_{\mathbf{11}} & * & 0 & 0 & 0 \\ & \lambda_{\mathbf{11}} & 0 & 0 & 0 \\ & & \lambda_{\mathbf{33}} & \lambda_{\mathbf{34}} & \lambda_{\mathbf{34}} \\ & & & \lambda_{\mathbf{33}} & \lambda_{\mathbf{34}} \\ & & & & \lambda_{\mathbf{33}} \end{pmatrix} \right\}^{+}$$

which is precisely $\mathcal{S}^{+}(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma \Delta \{E\})$ with $\mathcal{E}_\Gamma \Delta \{E\} = \{\{12\}, \{34, 35, 45\}\}$. The corresponding graph is displayed in Figure 3.6.

$\square$

We are now ready to prove the claim stipulated at the beginning of this section.

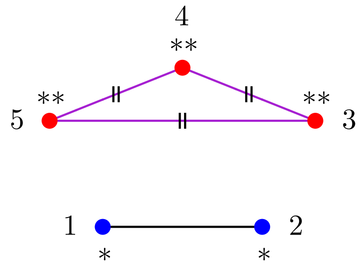


Figure 3.6.: $\mathcal{G} = (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma \Delta \{E\})$

Proposition 3.2 *Let $G = (V, E)$ and let $\Gamma \leq S(V)$ with $\mathcal{V}_\Gamma$ and $\mathcal{E}_\Gamma$ being the orbit partitions of V and $V_{\neq}^{(2)}$ induced by $\tilde{\Gamma}$. Then*

$$\mathcal{S}(\Gamma) \cap \mathcal{S}^+(G) = \mathcal{S}^+(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma \Delta \{E\}) = \mathcal{R}^+(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma \Delta \{E\}).$$

□

Proof: The first equality has been shown in Lemma 3.1, so in order to prove the claim, by Proposition 3.1, it suffices to show that whenever two edges in $\mathcal{E}_\Gamma \Delta \{E\}$ are of equal colour they connect the same vertex colour classes in $\mathcal{V}_\Gamma$.

So let $\alpha\beta, \gamma\delta \in V_{\neq}^{(2)}$ be such that $\alpha\beta \equiv \gamma\delta$ $(\mathcal{E}_\Gamma \Delta \{E\})$. Then, by definition of $\mathcal{E}_\Gamma \Delta \{E\}$, it must also hold that $\alpha\beta \equiv \gamma\delta$ $(\mathcal{E}_\Gamma)$. As, by (3.4) and (3.5),

$$\alpha\beta \equiv \gamma\delta \; (\mathcal{E}_\Gamma) \qquad \Longleftrightarrow \qquad \gamma\delta = \sigma(\alpha)\sigma(\beta)$$

for some $\sigma \in \Gamma$, we must have either

$$\alpha \equiv \gamma \; (\mathcal{V}_\Gamma) \text{ and } \beta \equiv \delta \; (\mathcal{V}_\Gamma) \qquad \text{or} \qquad \alpha \equiv \delta \; (\mathcal{V}_\Gamma) \text{ and } \beta \equiv \gamma \; (\mathcal{V}_\Gamma).$$

■

In Example 3.4, $E_\Gamma \Delta E \neq E$ so that some additional conditional independence constraints have been introduced by the symmetry. This happens if and only if there exists an edge colour class in $\mathcal{E}_\Gamma$ which partially lies in E and partially doesn't, or equivalently if and only if $\Gamma \not\leq \mathrm{Aut}(G)$. It is therefore sensible to only consider graphical symmetry models specified by constraints of the form

$$\Sigma^{-1} \in \mathcal{S}(\Gamma) \cap \mathcal{S}^+(G), \quad \Gamma \leq \mathrm{Aut}(G)$$

which is precisely the definition of covariance structures of *RCOP* models given in Højsgaard and Lauritzen (2008).

Definition 3.5 Let $G = (V, E)$ be an undirected graph, let $\Gamma \leq \mathrm{Aut}(G)$ be a permutation group and let Y be a random vector in $\mathbb{R}^V$. Then the *RCOP model for Y generated by* Γ with graph G is given by assuming

$$Y \sim \mathcal{N}_V(0, \Sigma), \quad \Sigma^{-1} \in \mathcal{S}(\Gamma) \cap \mathcal{S}^+(G)$$

with $\mathcal{S}(\Gamma)$ as in (2.14).

□

If for $G = (V, E)$ and Γ as in Definition 3.5 we let $\mathcal{V}$ denote the orbits of Γ in V and let $\mathcal{E}$ denote the partition of E consisting of $\tilde{\Gamma}$-orbits in $V_{\neq}^{(2)}$, then the constraints of the RCOP model generated by Γ with graph G can be represented by the coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We define graph colourings of this type representing RCOP models to be *permutation generated* below.

Definition 3.6 Let $\Gamma \leq S(V)$ and let $\mathcal{V}_\Gamma$ and $\mathcal{E}_\Gamma$ denote the vertex orbit partition of V and edge orbit partition of $V_{\neq}^{(2)}$ generated by $\tilde{\Gamma}$ respectively. A graph colouring $(\mathcal{V}, \mathcal{E})$ is *permutation generated with generating group* Γ if

$$\mathcal{V} = \mathcal{V}_\Gamma, \qquad \mathcal{E} \subseteq \mathcal{E}_\Gamma.$$

□

For RCOP models, Proposition 3.2 becomes the following, also to be found in Højsgaard and Lauritzen (2008).

Corollary 3.1 (Højsgaard and Lauritzen) *Let* $G = (V, E)$ *and let* $\Gamma \leq \mathrm{Aut}(G)$ *with orbits* $\mathcal{V}$ *and* $\mathcal{E}$ *in* V *and* E *respectively. Then*

$$\mathcal{S}(\Gamma) \cap \mathcal{S}^+(G) = \mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E}).$$

□

From now on, we will work with RCOP rather than graphical symmetry models. In particular, we will consider the following RCOP model further in Section 6.3.2.

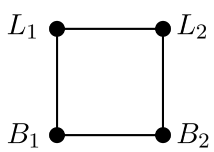
Example 3.5 The data, commonly referred to as *Fret's heads*, is concerned with the head dimensions of 25 pairs of first and second sons (Frets, 1921; Mardia et al., 1979). Previous analyses (Whittaker, 1990) support a model represented by the graph in Figure 3.7(a), where L_i and B_i denote the head length and head breadth of son i for $i = 1, 2$. Højsgaard and Lauritzen (2008) showed the model generated by $\Gamma = \langle (B_1 B_2)(L_1 L_2) \rangle$, corresponding to permuting the two sons, represented by the graph in Figure 3.7(b) to be an excellent fit. The model specifications are

$$Y \sim \mathcal{N}_V(0, \Sigma), \quad \Sigma^{-1} \in M$$

with M given in Figure 3.7. For $i = 1, 2$, L_i stands for the head length of son i and B_i stands for the head breadth of son i.

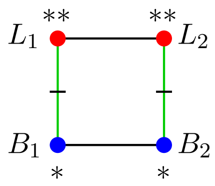
□

As RCOP models place the same equality constrains on concentrations as they do on partial correlations, they clearly combine the favourable properties of RCON and



$$M = \left\{ \begin{pmatrix} * & * & 0 & * \\ & * & * & 0 \\ & & * & * \\ & & & * \end{pmatrix} \right\}^{+}$$

(a) Conditional independence structure.



$$M = \left\{ \begin{pmatrix} \lambda_{11} & * & 0 & \lambda_{14} \\ & \lambda_{11} & \lambda_{14} & 0 \\ & & \lambda_{33} & * \\ & & & \lambda_{33} \end{pmatrix} \right\}^{+}$$

(b) RCOP model.

Figure 3.7.: Frets' heads example.

RCOR models. Thus RCOP models are RCON models which are scale-invariant if variables in the same vertex colour class are manipulated in the same way.

A particularly attractive feature of RCOP models is that the maximum likelihood estimate $\hat{\Sigma}$ of Σ, if it exists, may be obtained as for an unconstrained graphical Gaussian model after replacing yy^T by

$$\overline{yy^T} = \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} G(\sigma) yy^T G(\sigma)^{-1}, \tag{3.11}$$

i.e. by replacing the entries in yy^T by the average of all entries in yy^T which correspond to the same colour class in $\mathcal{V} \cup \mathcal{E}$ (Lauritzen, 1989).

4. Estimation of the Mean under Linear Constraints

When Højsgaard and Lauritzen (2008) introduced RCON and RCOR models, they assumed the model mean μ to be zero throughout. We extend their study by considering maximum likelihood estimation of a non-zero mean μ subject to linear constraints based on a (column vector) sample $Y^1 = y^1, \ldots, Y^n = y^n$ assumed to be drawn from a $\mathcal{N}_V(\mu, \Sigma)$ distribution, with Σ being restricted to satisfy the constraints of an RCON or RCOR model. More precisely, we let

$$\Omega = \mathrm{Im}(X) = \{\mu \mid \mu = X\beta, \beta \in \mathbb{R}^k\} \subseteq \mathbb{R}^V \tag{4.1}$$

for some fixed *design matrix* $X \in \mathbb{R}^{V \times k}$, using the obvious notation, of rank k and consider inference about μ based on assuming

$$Y^i = \mu + \epsilon_i, \qquad \mu \in \Omega, \qquad \epsilon_i \sim_{iid} \mathcal{N}_V(0, \Sigma) \tag{4.2}$$

with Σ being subject to $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ or $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ for some given graph colouring $(\mathcal{V}, \mathcal{E})$, but is otherwise unknown.

For the Gaussian distribution, maximum likelihood estimation of μ under an unknown covariance structure is generally non-trivial, as the maximiser of the likelihood function in μ for fixed Σ may depend on Σ. The least squares estimator μ^* is defined by minimizing the sum of squares so that in case of equality of estimators $\hat{\mu}$ and μ^*, the former is independent of Σ. Using Kruskal (1968), we characterise all vector spaces Ω which ensure equality of estimators $\hat{\mu} = \mu^*$ in (4.2) for both constraint types $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ or $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$.

Conditions for equality of estimators have been considered in the literature before. Eaton (1970) showed that for

$$Y^i = \mu_i + \epsilon_i, \quad \epsilon_i \sim_{iid} \mathcal{N}_V(0, \Sigma), \quad \Sigma^{-1} \in \mathcal{S}^+ \tag{4.3}$$

and $\mu = (\mu_1, \ldots, \mu_n)^T$ the only vector space W ensuring that $\hat{\mu} = \mu^*$ for $\mu \in W$ is of the form

$$W = \{\mu \mid \mu = ZB, Z \in \mathbb{R}^{n \times k} \text{ fixed}, B \in \mathbb{R}^{k \times V}\}, \tag{4.4}$$

with (4.3) and (4.4) jointly defining a multivariate analysis of variance (MANOVA) model. Using results in Kruskal (1968), Eaton (1983) further characterised all such spaces W ensuring equality of estimators if Σ is assumed to have a diagonal block structure, and if Σ is assumed to have an intraclass covariance structure, see (2.11).

Our work is similar in spirit, however while the results in Eaton (1970) and Eaton (1983) is coordinate free, we assume the observations to be given in a fixed basis. This makes the constraints $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ or $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$, and $\mu^i = \mu$ for $1 \leq i \leq n$ meaningful.

We pay special attention to the case when Ω in (4.1) is specified by equality constraints on the entries of μ. Such hypotheses in the mean have been considered for the Gaussian distribution with symmetries and without conditional independence relations, see Section 2.3.1, however not for models combining the two.

We let $\mathcal{M}$ denote the partition of V placing two vertices in the same set whenever the corresponding means are restricted to being identical and write $\Omega = \Omega(\mathcal{M})$. For $(\mathcal{V}, \mathcal{E})$ as above, the derived characterisation of all viable Ω translates into a twofold necessary and sufficient condition on $\mathcal{M}$ and $(\mathcal{V}, \mathcal{E})$: (i) $\mathcal{M}$ must be finer than $\mathcal{V}$; and (ii) $\mathcal{M}$ must be *equitable* with respect to every edge colour class in $\mathcal{E}$ as defined by Sachs (1966). An example where condition (i) is violated is the *Behrens-Fisher problem*, concerned with testing the equality of two means under the assumption of unequal variances (Behrens, 1929; Fisher, 1939). Most of the presented results concerning equality constraints on the mean can also be found in Gehrmann and Lauritzen (2011).

We conclude the chapter with examples from comparative experiments according to Latin square designs (Bailey, 2008). The two models we consider are given by a *homogeneous nearest-neighbour lattice scheme* (Besag, 1974), which restricts Σ^{-1} to lie inside $\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ with $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ being a coloured lattice, and a *homogeneous nearest-neighbour torus scheme*, represented by a coloured torus, which we define following Besag and Moran (1975).

4.1. Estimation of the Mean

Estimators for μ other than the maximum likelihood estimator $\hat{\mu}$ are the *least squares estimator* μ^* and the *Gauss-Markov estimator* $\tilde{\mu}$. Below we formally define the estimators and state relations between them, first obtained by Kruskal (1968), which can be exploited to ensure uniqueness of $\hat{\mu}$.

4.1.1. The Least Squares Estimator

If $Y \in \mathbb{R}^V$ is a random vector satisfying $\mathbb{E}(Y) = \mu \in \Omega$ with Ω as in (4.1), then the least squares estimator μ^* of $\mu \in \Omega$ takes the value of $\mu \in \Omega$ which minimises

$$g(\mu) = (Y - \mu)^T (Y - \mu)$$

over Ω. It directly follows that μ^* is equal to the orthogonal projection of Y on Ω with respect to the standard inner product in $\mathbb{R}^V$, $(x, y) = x^T y$, and is given by

$$\mu^* = X(X^T X)^{-1} X^T Y. \tag{4.5}$$

Thus μ^* exists for all Y and all Ω, and in particular is always uniquely determined by the two. Further, μ^* is independent of any distributional assumptions about Y, as well as the covariance structure Σ.

4.1.2. The Gauss-Markov Estimator

If $Y \in \mathbb{R}^V$ is as above and $\mathrm{Cov}(Y) = \Sigma \in \mathcal{S}^+$ is known up to a multiplicative constant, i.e. $\Sigma = \sigma^2 \Sigma_0$ with $\Sigma_0 \in \mathcal{S}^+$ known and $\sigma^2 > 0$ unknown, then the Gauss-Markov estimator $\tilde{\mu}$ of μ is the linear unbiased estimator of $\mu \in \Omega$ which has minimal variance. It is given by the orthogonal projection of Y on Ω with respect to the inner product

$$(x, y)_{\Sigma_0} = x^T \Sigma_0^{-1} y \tag{4.6}$$

(Kruskal, 1968) and equals

$$\tilde{\mu} = \tilde{\mu}_{\Sigma_0} = X(X^T \Sigma_0^{-1} X)^{-1} X^T \Sigma_0^{-1} Y. \tag{4.7}$$

The fact that $\tilde{\mu}$ is unique is usually referred to as the Gauss-Markov Theorem (e.g. Theorem 4.1 in Eaton, 1983).

While the Gauss-Markov estimator, just as the least squares estimator, does not depend on any distributional assumptions, it directly depends on the covariance structure of Y, which, in general, is required to be known up to a multiplicative constant for the estimate to be computed. The following necessary and sufficient condition on Ω and Σ_0 for the two estimators μ^* and $\tilde{\mu}$ to agree has first been shown by Kruskal (1968). It can also be found in Haberman (1975) and Eaton (1983).

Proposition 4.1 (Kruskal) *Let $Y \in \mathbb{R}^V$ be a random vector, $\Omega \leq \mathbb{R}^V$ and $\Sigma_0 \in \mathcal{S}^+$. Then if $\mathbb{E}(Y) = \mu$ is restricted to lie in Ω and $\mathrm{Cov}(Y) = \Sigma = \sigma^2\Sigma_0$ with $\sigma^2 > 0$ unknown, the two estimators μ^* and $\tilde{\mu}$ agree if and only if Ω is invariant under Σ_0^{-1}, i.e. if and only if $\Sigma_0^{-1}\Omega \subseteq \Omega$.*

□

Lemma 4.1 *For $\Sigma \in \mathcal{S}^+$ and $\Omega \leq \mathbb{R}^V$,*

$$\Sigma\Omega \subseteq \Omega \quad \Longleftrightarrow \quad \Sigma\Omega = \Omega \quad \Longleftrightarrow \quad \Sigma^{-1}\Omega = \Omega \quad \Longleftrightarrow \quad \Sigma^{-1}\Omega \subseteq \Omega.$$

□

Proof of Proposition 4.1: The two defining conditions of the estimators μ^* and $\tilde{\mu}$ are

$$(Y - \mu^*, x) = (Y - \tilde{\mu}, \Sigma_0^{-1}x) = 0 \quad \text{for all } x \in \Omega.$$

By Lemma 4.1, if $\Sigma_0^{-1}\Omega \subseteq \Omega$ then $\Sigma_0^{-1}\Omega = \Omega$ so that the two defining relations are the same, and therefore, by uniqueness, $\mu^* = \tilde{\mu}$. For the converse, let $y \in \Omega^\perp$. Then $\mu^* = 0$, and so $\tilde{\mu} = 0$, giving that $y = y - \tilde{\mu} \in \left(\Sigma_0^{-1}\Omega\right)^\perp$. We obtain that

$\Omega^\perp \subseteq \left(\Sigma_0^{-1}\Omega\right)^\perp$, which implies $\Sigma_0^{-1}\Omega \subseteq \Omega$, which by Lemma 4.1 is equivalent to $\Sigma_0^{-1}\Omega = \Omega$.

■

We conclude:

Corollary 4.1 *Let $Y^1 = y^1, \ldots, Y^n = y^n \in \mathbb{R}^V$ be a random sample drawn from a distribution whose mean value μ is restricted to lie in a vector space $\Omega \leq \mathbb{R}^V$ and whose concentration matrix Σ^{-1} is restricted to lie in $\mathcal{R} \subseteq \mathcal{S}^+$. Then for all $\Sigma^{-1} \in \mathcal{R}$, the Gauss-Markov estimator $\tilde{\mu}$ of μ and the least squares estimator μ^* of μ based on $Y^1, \ldots, Y^n$ are identical if and only if $\Sigma^{-1}\Omega \subseteq \Omega$.*

□

4.1.3. The Maximum Likelihood Estimator

For a random vector $Y \in \mathbb{R}^V$ assumed to follow a $\mathcal{N}_V(\mu, \Sigma)$ distribution with $\mu \in \Omega \leq \mathbb{R}^V$ and $\Sigma^{-1} \in \mathcal{R} \subseteq \mathcal{S}^+$, the maximum likelihood estimator $\hat{\mu}$ is defined by

$$\begin{aligned}
(\hat{\mu}, \hat{\Sigma}) &= \arg \max_{\substack{\mu\in\Omega,\\ \Sigma^{-1}\in\mathcal{R}}} L(\mu, \Sigma \mid y) \\
&= \arg \max_{\substack{\mu\in\Omega,\\ \Sigma^{-1}\in\mathcal{R}}} \frac{1}{(2\pi)^{|V|/2}(\det\Sigma)^{1/2}} \exp\{-(y-\mu)^T\Sigma^{-1}(y-\mu)/2\}.
\end{aligned}$$

Thus in order to obtain $\hat{\mu}$, the likelihood function $L(\mu, \Sigma \mid y)$ needs to be jointly maximised in μ and Σ. If $\Sigma^{-1} \in \mathcal{R}$ is fixed, maximising the likelihood in μ subject to $\mu \in \Omega$ is equivalent to minimising

$$h(\mu) = (y-\mu)^T\Sigma^{-1}(y-\mu)$$

over Ω. The value taken by $\hat{\mu}$ which minimises $h(\mu)$ over Ω is precisely the orthogonal projection of Y on Ω with respect to the inner product in (4.6) with $\Sigma_0 = \Sigma$, so that for Σ fixed we have that the maximum likelihood estimator $\hat{\mu}$ of μ is equal to the Gauss-Markov estimator $\tilde{\mu} = \tilde{\mu}_\Sigma$ of μ. Thus if $\Sigma^{-1} \in \mathcal{R}$ is unknown, $\hat{\mu}$, if it exists, can be found by maximising $L(\tilde{\mu}_\Sigma, \Sigma \mid y)$ in Σ over $\mathcal{R}$:

$$\hat{\mu} = \tilde{\mu}_{\hat{\Sigma}} \quad \text{with} \quad \hat{\Sigma} = \arg \max_{\Sigma^{-1} \in \mathcal{R}} L(\tilde{\mu}_\Sigma, \Sigma \mid y)$$

with $L(\tilde{\mu}_{\hat{\Sigma}}, \Sigma \mid y)$ usually referred to as the profile likelihood of Σ. Combining this with Corollary 4.1 we obtain:

Proposition 4.2 *Let $Y^1 = y^1, \ldots, Y^n = y^n$ be a random sample drawn from a $\mathcal{N}_V(\mu, \Sigma)$ distribution with $\mu \in \Omega$ for some $\Omega \leq \mathbb{R}^V$ and $\Sigma^{-1} \in \mathcal{R} \subseteq \mathcal{S}^+$. Then for all $\Sigma^{-1} \in \mathcal{R}$, the least squares estimator μ^* and the maximum likelihood estimator $\hat{\mu}$ of μ based on $Y^1, \ldots, Y^n$ are identical if and only if $\Sigma^{-1}\Omega \subseteq \Omega$.*

□

4.2. Invariance under Generator Matrices

Below we show that for RCON and RCOR models, invariance of Ω under Σ^{-1} is equivalent to invariance under the generator matrices $\{T^u\}_{u \in \mathcal{V} \cup \mathcal{E}}$.

Proposition 4.3 *For a random sample drawn from a $\mathcal{N}_V(\mu, \Sigma)$ distribution with $\mu \in \Omega \leq \mathbb{R}^V$ and $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ for some coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$,*

$$\hat{\mu} = \mu^* \quad \iff \quad T^u\Omega \subseteq \Omega \quad \forall u \in \mathcal{V} \cup \mathcal{E}.$$

□

Proof: By Proposition 4.2, $\hat{\mu} = \mu^*$ if and only if $\Sigma^{-1}\Omega \subseteq \Omega$ for all $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$, so that it suffices to show that $\Sigma^{-1}\Omega \subseteq \Omega$ for all $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ if and only if

$T^u\Omega \subseteq \Omega$ for all $u \in \mathcal{V} \cup \mathcal{E}$.

By Definition 3.1, all $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ can be written in the form $\Sigma^{-1} = \sum_{u \in \mathcal{V}\cup\mathcal{E}} \lambda_u T^u$ with $\lambda \in \mathbb{R}^{\mathcal{V}\cup\mathcal{E}}$ such that Σ^{-1} is positive definite.

Suppose first that

$$T^u\Omega \subseteq \Omega \quad \forall u \in \mathcal{V} \cup \mathcal{E}.$$

Since Ω is an affine space, we directly obtain that $\Sigma^{-1}\Omega \subseteq \Omega$ for all $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$.

Conversely, suppose $\Sigma^{-1}\Omega \subseteq \Omega$ for all $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$. As $\mathcal{S}^+(\mathcal{V}, \mathcal{E})$ is a relatively open cone, for all $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ and $u \in \mathcal{V} \cup \mathcal{E}$ there exists $\nu_u \in \mathbb{R} \setminus \{0\}$ such that

$$S_u = \Sigma^{-1} + \nu_u T^u \in \mathcal{S}^+(\mathcal{V}, \mathcal{E}).$$

By assumption, $\Sigma^{-1}\Omega \subseteq \Omega$ and $S_u\Omega \subseteq \Omega$, so that

$$(S_u - \Sigma^{-1})\Omega = \nu_u T^u \Omega \subseteq \Omega.$$

■

The same result holds for $\mathcal{R}^+(\mathcal{V}, \mathcal{E})$:

Proposition 4.4 *For a random sample drawn from a $\mathcal{N}_V(\mu, \Sigma)$ distribution with $\mu \in \Omega \leq \mathbb{R}^V$ and $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ for some coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$,*

$$\hat{\mu} = \mu^* \quad \iff \quad T^u\Omega \subseteq \Omega \quad \forall u \in \mathcal{V} \cup \mathcal{E}.$$

□

Proof: We again exploit Proposition 4.2, and prove the claim by showing that $\Sigma^{-1}\Omega \subseteq \Omega$ for all $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ if and only if $T^u\Omega \subseteq \Omega$ for all $u \in \mathcal{V} \cup \mathcal{E}$.

By Definition 3.2, all $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ can be written as

$$\Sigma^{-1} = ACA = \sum_{v,w \in \mathcal{V}} a_v a_w T^v T^w \; + \sum_{\substack{v,w \in \mathcal{V}, \\ u \in \mathcal{E}}} c_u a_v a_w T^v T^u T^w. \tag{4.8}$$

where A and C are as in (3.2) such that Σ^{-1} is positive definite, and for $v \in \mathcal{V}$ and $u \in \mathcal{E}$ we write a_v for $a_{\alpha\alpha}$ if $\alpha \in v$ and c_u for $c_{\alpha\beta}$ if $\alpha\beta \in u$. As for $v, w \in \mathcal{V}$, $T^v T^w = T^v$ if and only if $v = w$ and $T^v T^w = 0$ otherwise, (4.8) simplifies to

$$\Sigma^{-1} = \sum_{v \in \mathcal{V}} a_v^2 T^v \; + \sum_{\substack{v,w \in \mathcal{V}, \\ u \in \mathcal{E}}} c_u a_v a_w T^v T^u T^w. \tag{4.9}$$

As before, invariance of Ω under T^u for all $u \in \mathcal{V} \cup \mathcal{E}$ directly implies invariance of Ω under all $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$. So suppose $\Sigma^{-1}\Omega \subseteq \Omega$ for all $\Sigma^{-1} \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$. Then, in particular, Ω is invariant under all Σ^{-1} inside the model which are of the form

$$\Sigma^{-1} = ACA \text{ with } A = \sigma^2 I, \; \sigma^2 > 0,$$

represented by the graph colouring $(\{V\}, \mathcal{E})$, which is clearly edge regular as all edges connect the only vertex colour class back to itself. By Proposition 3.1 this implies $\mathcal{S}^+(\{V\}, \mathcal{E}) = \mathcal{R}^+(\{V\}, \mathcal{E})$ so that by Proposition 4.3,

$$T^u \Omega \subseteq \Omega \quad \forall u \in \mathcal{E}.$$

For the vertex colouring, consider the submodel

$$\Sigma^{-1} = ACA \text{ with } C = I,$$

represented by the empty graph with edge regular colouring $(\mathcal{V}, \emptyset)$. By Proposition 4.3,

$$T^v\Omega \subseteq \Omega \quad \forall v \in \mathcal{V}.$$

$\blacksquare$

Definition 4.1 For a coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, let $\mathcal{T}(\mathcal{G}) = \{T^u \mid u \in \mathcal{V} \cup \mathcal{E}\}$. We say that $\Omega \leq \mathbb{R}^V$ is Ω is *$\mathcal{G}$-stable* if $T\Omega \subseteq \Omega$ for all $T \in \mathcal{T}(\mathcal{G})$ and let the set of $\mathcal{G}$-stable Ω be denoted by $\Phi(\mathcal{G})$.

$\square$

Then Propositions 4.3 and 4.4 may be equivalently phrased as:

Proposition 4.5 *For a random sample drawn from a $\mathcal{N}_V(\mu, \Sigma)$ distribution with $\mu \in \Omega \leq \mathbb{R}^V$ and $\Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$ for some coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $\hat{\mu} = \mu^*$ if and only if $\Omega \in \Phi(\mathcal{G})$. The same holds if $\mathcal{S}^+(\mathcal{V}, \mathcal{E})$ is replaced by $\mathcal{R}^+(\mathcal{V}, \mathcal{E})$.*

$\square$

4.3. Characterisation of $\Phi(\mathcal{G})$

The goal of this section is a full characterisation of $\Phi(\mathcal{G})$, given in Theorem 4.2 below.

4.3.1. Preliminaries

Notation

For a finite dimensional vector space W and $W_1, \ldots, W_k \leq W$ we let

$$+_{1 \leq i \leq k} W_i = W_1 + \ldots + W_k = \{w \mid \exists\, w_i \in W_i \text{ for } 1 \leq i \leq k \text{ such that } w = \sum_{i=1}^{k} w_i\}.$$

If $W = +_{1\le i\le k} W_i$ and $W_j \cap (+_{1\le i\le j-1} W_i) = \{0\}$ for all $2 \le j \le k$, W is a direct sum of $W_1, W_2, \ldots, W_k$ and we write

$$W = W_1 \oplus \ldots \oplus W_k = \bigoplus_{1\le i\le k} W_i. \tag{4.10}$$

For $R \le W$, $R^\perp$ denotes the orthogonal complement of R in W with respect to the standard inner product

$$R^\perp = \{\omega \in W \mid (x, \omega) = 0 \;\forall\; x \in R\}.$$

We say that $W_1, \ldots, W_k$ provide an *orthogonal decomposition* of W in (4.10) if $\left(\bigoplus_{i\ne j} W_i\right)^\perp = W_j$ for $1 \le j \le k$.

For $b \subseteq A \subseteq V$, we let $\mathcal{I}_b^A : \mathbb{R}^A \hookrightarrow \mathbb{R}^b$ be the embedding

$$\mathcal{I}_b^A : \omega \mapsto x \quad \text{with } x_\alpha = \omega_\alpha \text{ for } \alpha \in b.$$

$\Pi_b^A : \mathbb{R}^A \to \mathbb{R}^A$ denotes the orthogonal projection

$$\Pi_b^A : \omega \mapsto x \quad \text{with } x_\alpha = \begin{cases} \omega_\alpha & \text{for } \alpha \in b \\ 0 & \text{otherwise} \end{cases}$$

so that $\left(\Pi_b^A\right)^2 = \Pi_b^A$ for all $b \subseteq A$ and

$$a \cap b = \emptyset \implies \Pi_a^A \Pi_b^A = 0, \quad \text{and} \qquad a \cap b = \emptyset,\; A = a \cup b \implies \Pi_a^A + \Pi_b^A = I_A \tag{4.11}$$

where I_A denotes the identity matrix in $\mathbb{R}^A$, which we sometimes also write as $I_{|A|}$.

For $\Omega \leq \mathbb{R}^V$, we let

$$\Omega_b = \{x \in \mathbb{R}^b \mid \exists\, \omega \in \Omega \text{ such that } x_\alpha = \omega_\alpha \text{ for } \alpha \in b\} \leq \mathbb{R}^b$$

denote the image of Ω under $\mathcal{I}_b^V$, also abbreviated by $\mathcal{I}_b$, and let $\tilde{\Omega}_b \leq \mathbb{R}^V$ the image of Ω under Π_b^V, also abbreviated by Π_b.

$$\tilde{\Omega}_b^A = \{x \in \Omega_A \mid \exists\, \omega \in \Omega_A \text{ such that } x_\alpha = \omega_\alpha \text{ for } \alpha \in b, x_\alpha = 0 \text{ for } \alpha \in A \setminus b\} \leq \mathbb{R}^A$$

denotes the image of Ω_A under Π_b^A.

Linear Algebra

We make repeated use of the Spectral Theorem (e.g. Theorem 7.27 in Friedberg et al. (1979) or Theorem 1 in §79 in Halmos (1958)):

Theorem 4.1 (Spectral Theorem) *Let T be a symmetric real matrix in $\mathbb{R}^{W \times W}$ and let Λ denote the spectrum of T, containing all eigenvalues of T. Then $\Lambda \subset \mathbb{R}$, and, if for $\lambda \in \Lambda$, E_λ denotes the eigenspace corresponding to λ and Π_λ the orthogonal projection on E_λ,*

(i) $\mathbb{R}^W = \bigoplus_{\lambda \in \Lambda} E_\lambda$

(ii) $\left(\bigoplus_{\lambda \in \Lambda \setminus \{\lambda'\}} E_\lambda\right)^\perp = E_{\lambda'}$, $\lambda' \in \Lambda$

(iii) $\Pi_\lambda \Pi_{\lambda'} = 0$ *if* $\lambda \neq \lambda'$, $\lambda, \lambda' \in \Lambda$

(iv) $I = \sum_{\lambda \in \Lambda} \Pi_\lambda$

(v) $T = \sum_{\lambda \in \Lambda} \lambda \Pi_\lambda$.

$\square$

The following is a direct consequence of the Spectral Theorem:

Lemma 4.2 *Let T be a symmetric real matrix in $\mathbb{R}^{W\times W}$ with spectrum Λ and for $\lambda \in \Lambda$ let E_λ denote the corresponding eigenspace. Then for $A \leq \mathbb{R}^W$, $TA \subseteq A$ if and only if there exist $A_\lambda \leq E_\lambda$ for $\lambda \in \Lambda$ such that*

$$A = \bigoplus_{\lambda\in\Lambda} A_\lambda. \tag{4.12}$$

□

Proof: We observe first that $TA \subseteq A$ if and only if whenever $\{\omega_1, \dots, \omega_k\}$ is a basis for A, then $\operatorname{span}(\{T\omega_1, \dots, T\omega_k\}) \subseteq \operatorname{span}(\{\omega_1, \dots, \omega_k\}) = A$, and that A satisfies (4.12) if and only if A has a basis of the form

$$\omega_i \in E_{\lambda_i} \qquad \lambda_i \in \Lambda \qquad \forall\ 1 \leq i \leq k. \tag{4.13}$$

This directly implies that if A is given by (4.12), then $TA \subseteq A$ since the basis in (4.13) satisfies $\operatorname{span}(\{T\omega_1, \dots, T\omega_k\}) = \operatorname{span}(\{\lambda_1\omega_1, \dots, \lambda_k\omega_k\}) \subseteq \operatorname{span}(\{\omega_1, \dots, \omega_k\})$.

Conversely, suppose $TA \subseteq A$. Let S denote the restriction of T operating on A. Then as T is a self-adjoint linear transformation, so is S. Applying The Spectral Theorem to S then directly gives (4.12).

■

Lemma 4.3 *Let $G = (V, E)$ be a bipartite graph with vertex components a and b and let T denote the adjacency matrix of G. Let Λ denote the spectrum of T and for $\lambda \in \Lambda$ let E_λ be the corresponding eigenspace. Then*

$$(\Pi_a - \Pi_b)E_\lambda = (\Pi_b - \Pi_a)E_\lambda = E_{-\lambda}.$$

□

Proof: As E_λ is a vector space, clearly

$$(\Pi_a - \Pi_b)E_\lambda = (\Pi_b - \Pi_a)E_\lambda.$$

For the second equality, let $g \in E_\lambda$. Then, as T is the adjacency matrix of a bipartite graph connecting a to b, it has the form

$$T = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \quad \text{with} \quad B \in \mathbb{R}^{a\times b}, \quad B_{\alpha\beta} = \begin{cases} 1 & \text{if } \alpha\beta \in E, \alpha \in a, \beta \in b \\ 0 & \text{otherwise} \end{cases} \tag{4.14}$$

so that

$$T\left(\Pi_a g\right) = \lambda \Pi_b g \qquad \text{and} \qquad T\left(\Pi_b g\right) = \lambda \Pi_a g.$$

Hence

$$\begin{aligned} T\left[\left(\Pi_a - \Pi_b\right) g\right] &= T\left(\Pi_a g\right) - T\left(\Pi_b g\right) \\ &= \lambda \Pi_b g - \lambda \Pi_a g \\ &= -\lambda \left(\Pi_a - \Pi_b\right) g, \end{aligned}$$

giving

$$(\Pi_a - \Pi_b)E_\lambda \subseteq E_{-\lambda}.$$

By symmetry we also have

$$(\Pi_a - \Pi_b)E_{-\lambda} \subseteq E_\lambda$$

so that

$$(\Pi_a - \Pi_b)(\Pi_a - \Pi_b)E_{-\lambda} = E_{-\lambda} \subseteq (\Pi_a - \Pi_b)E_\lambda.$$

■

4.3.2. First Properties of $\Phi(\mathcal{G})$

Proposition 4.6 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a coloured graph. Then*

(i)

$$\Omega \in \Phi(\mathcal{G}) \quad \iff \quad \Omega^{\perp} \in \Phi(\mathcal{G})$$

(ii)

$$\Omega_1, \Omega_2 \in \Phi(\mathcal{G}) \quad \Longrightarrow \quad \Omega_1 \cap \Omega_2,\ \Omega_1 + \Omega_2 \in \Phi(\mathcal{G}).$$

□

Proof: For (i), suppose $\Omega \in \Phi(\mathcal{G})$ and let $x \in \Omega^{\perp}$. Then $(x, \omega) = 0$ for all $\omega \in \Omega$. So let $\omega \in \Omega$ and consider (Tx, ω) for $T \in \mathcal{T}(\mathcal{G})$:

$$(Tx, \omega) = (x, T^T\omega) = (x, T\omega) = (x, \omega') = 0$$

where the last equality follows from the fact that $T\omega = \omega' \in \Omega$ by stability of Ω under T. This gives that $Tx \in \Omega^{\perp}$ for all $T \in \mathcal{T}(\mathcal{G})$ and all $x \in \Omega^{\perp}$, or equivalently that $\Omega^{\perp} \in \Phi(\mathcal{G})$. The other direction follows directly from the fact that $\left(\Omega^{\perp}\right)^{\perp} = \Omega$.

For (ii), let $\Omega_1, \Omega_2 \in \Phi(\mathcal{G})$ and suppose $x \in \Omega_1 \cap \Omega_2$ and $y \in \Omega_1 + \Omega_2$. Then $Tx \in \Omega_1$ and $Tx \in \Omega_2$, so that $Tx \in \Omega_1 \cap \Omega_2$, giving $\Omega_1 \cap \Omega_2 \in \Phi(\mathcal{G})$. Further, $Ty = T(\omega_1 + \omega_2) = \omega_1' + \omega_2'$ for $\omega_1, \omega_1' \in \Omega_1$ and $\omega_2, \omega_2' \in \Omega_2$, giving that $Ty \in \Omega_1 + \Omega_2$, so that $\Omega_1 + \Omega_2 \in \Phi(\mathcal{G})$. ■

Proposition 4.6 only exploits that all $T \in \mathcal{T}(\mathcal{G})$ are symmetric. Below we describe the implications of stability under the generating matrices T^u for $u \in \mathcal{V}$ and $u \in \mathcal{E}$.

Vertex Colour Classes

Lemma 4.4 *For a coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\Omega \leq \mathbb{R}^V$,*

$$T^v\Omega \subseteq \Omega \quad \forall v \in \mathcal{V} \qquad \Longleftrightarrow \qquad \Omega = \bigoplus_{v\in\mathcal{V}} \tilde{\Omega}_v.$$

□

Proof: Suppose that $T^v\Omega \subseteq \Omega$ for all $v \in \mathcal{V}$. Then $\tilde{\Omega}_v \subseteq \Omega$ for all $v \in \mathcal{V}$ so that certainly $+_{v\in\mathcal{V}} \tilde{\Omega}_v \subseteq \Omega$. Since clearly $\Omega \subseteq +_{v\in\mathcal{V}} \tilde{\Omega}_v$, we have $\Omega = +_{v\in\mathcal{V}} \tilde{\Omega}_v$. As further $\tilde{\Omega}_v$ is constantly zero on $\mathbb{R}^{V\setminus v}$ for all $v \in \mathcal{V}$, we obtain

$$\Omega = \bigoplus_{v\in\mathcal{V}} \tilde{\Omega}_v. \tag{4.15}$$

Conversely, suppose (4.15) holds. It follows directly that

$$T^v\Omega = \tilde{\Omega}_v \subseteq \Omega \quad \forall v \in \mathcal{V}.$$

■

Edge Colour Classes

For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $u \in \mathcal{E}$ and $v, w \in \mathcal{V}$, let $E^{u\{vw\}} \subseteq u$ denote the set of edges inside u which connect v to w and call the coloured subgraph

$$\mathcal{G}^{u\{vw\}} = (\mathcal{V}, \{E^{u\{vw\}}\})$$

which only contains edges in $E^{u\{vw\}}$ the *$u\{vw\}$-subgraph* of $\mathcal{G}$. Distinguishing between the two cases $v \neq w$ and $v = w$, the adjacency matrix $T^{u\{vw\}}$ of $\mathcal{G}^{u\{vw\}}$ can be directly expressed in terms of T^u, T^v and T^w,

$$T^{u\{vw\}} = \begin{cases} T^vT^uT^w + T^wT^uT^v & \text{for } v \neq w \\ T^vT^uT^v & \text{for } v = w \end{cases}.$$

If we order V such that the vertices in v appear in the ordering first, followed by the vertices in w, then $T^{u\{vw\}}$ and $T^{u\{vv\}}$ equal

$$T^{u\{vw\}} = \begin{pmatrix} T^{[u\{vw\}]} & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad T^{u\{vv\}} = \begin{pmatrix} T^{[u\{vv\}]} & 0 \\ 0 & 0 \end{pmatrix} \tag{4.16}$$

with $T^{[u\{vw\}]} \in \mathbb{R}^{(v \cup w) \times (v \cup w)}$ being the adjacency matrix of the $u\{vw\}$-*component*

$$\mathcal{G}^{[u\{vw\}]} = (\{v, w\}, \{E^{u\{vw\}}\})$$

of $\mathcal{G}$, which has the same edge set as $\mathcal{G}^{u\{vw\}}$ but only $v \cup w$ as its vertex set. Similarly, $T^{[u\{vv\}]} \in \mathbb{R}^{v \times v}$ is the adjacency matrix of the $u\{vv\}$-*component*

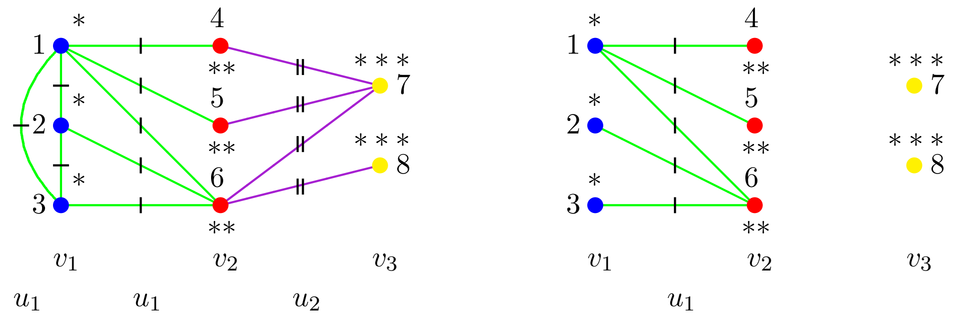
$$\mathcal{G}^{[u\{vv\}]} = (\{v\}, \{E^{u\{vv\}}\})$$

of $\mathcal{G}$, which has the same edge set as $\mathcal{G}^{u\{vv\}}$ but only v as its vertex set. The zeros in (4.16) represent matrices of suitable dimensions whose entries are all zero.

Example 4.1 Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the graph shown in Figure 4.1(a) with $V = [8]$, $\mathcal{V} = \{v_1, v_2, v_3\}$ and $\mathcal{E} = \{u_1, u_2\}$ with $v_1 = [3]$ $v_2 = \{4, 5, 6\}$, $v_3 = \{7, 8\}$, $u_1 = \{12, 13, 23, 14, 15, 16, 26, 36\}$ and $u_2 = \{47, 57, 67, 68\}$. The $u_1\{v_1 v_2\}$-subgraph of $\mathcal{G}$

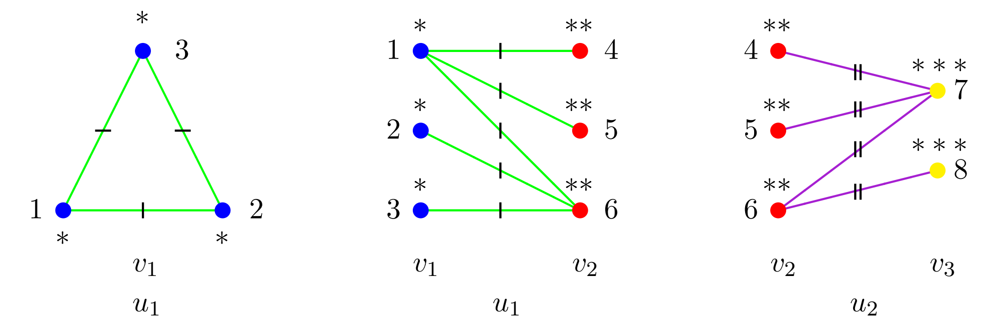
$$\mathcal{G}^{u_1\{v_1 v_2\}} = (\{\{1, 2, 3\}, \{4, 5, 6\}, \{7, 8\}\}, \{\{14, 15, 16, 26, 36\}\})$$

is displayed in Figure 4.1(b) and the three components of $\mathcal{G}$ are given in Figure 4.2. □



(a) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ (b) $\mathcal{G}^{u_1\{v_1v_2\}} = (\{v_1, v_2\}, \{E^{u_1\{v_1v_2\}}\})$

Figure 4.1.: Vertex and edge coloured graph $\mathcal{G}$ with one of its $u\{vw\}$-subgraphs.



(a) $u_1\{v_1v_1\}$-component. (b) $u_1\{v_1v_2\}$-component. (c) $u_2\{v_2v_3\}$-component.

Figure 4.2.: Components of $\mathcal{G}$.

Lemma 4.5 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a coloured graph and let $\Omega \leq \mathbb{R}^V$ be of the form $\Omega = \bigoplus_{v\in\mathcal{V}} \tilde{\Omega}_v$. Then for $u \in \mathcal{E}$,*

$$\begin{aligned} T^u\Omega \subseteq \Omega \quad &\iff \quad T^{u\{vw\}}\Omega \subseteq \Omega \quad \forall\, v, w \in \mathcal{V} \\ &\iff \quad T^{[u\{vw\}]}\Omega_{v\cup w} \subseteq \Omega_{v\cup w} \quad \forall\, v, w \in \mathcal{V}. \end{aligned}$$

□

Proof: We consider the first equivalence claim first. Since for $u \in \mathcal{E}$,

$$T^u = IT^uI \quad = \quad \left(\sum_{v\in\mathcal{V}} T^v\right) T^u \left(\sum_{w\in\mathcal{V}} T^w\right)$$

$$\begin{aligned} &= \sum_{\{v,w\}\subseteq V} (T^v T^u T^w + T^w T^u T^v) + \sum_{v\in\mathcal{V}} T^v T^u T^v \\ &= \sum_{\{v,w\}\subseteq V} T^{u\{vw\}} + \sum_{v\in\mathcal{V}} T^{u\{vv\}}, \end{aligned}$$

invariance of Ω under $T^{u\{vw\}}$ for $v, w \in \mathcal{V}$ clearly implies invariance of Ω under T^u.

So suppose that Ω is invariant under T^u. Since, by construction, Ω is invariant under T^v for all $v \in \mathcal{V}$, Ω is also invariant under $T^v T^u T^w$ and $T^w T^u T^v$ for all $v, w \in \mathcal{V}$. This directly gives that Ω is invariant under $(T^v T^u T^w + T^w T^u T^v)$ for all $v, w \in \mathcal{V}$, or equivalently under $T^{u\{vw\}}$ for all $v, w \in \mathcal{V}$.

The second equivalence claim follows from the relation between $T^{u\{vw\}}$ and $T^{[u\{vw\}]}$ in (4.16).

■

4.3.3. $\mathcal{G}$-Invariance

Combined, Lemma 4.4 and Lemma 4.5 imply:

Proposition 4.7 *For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\Omega \leq \mathbb{R}^V$, $\Omega \in \Phi(\mathcal{G})$ if and only if*

$$\Omega = \bigoplus_{v\in\mathcal{V}} \tilde{\Omega}_v \tag{4.17}$$

and for all $u \in \mathcal{E}$, $v, w \in \mathcal{V}$,

$$T^{[u\{vw\}]}\Omega_{v\cup w} \subseteq \Omega_{v\cup w}. \tag{4.18}$$

□

Proof: By definition, $\Omega \in \Phi(\mathcal{G})$ if and only if $T^u\Omega \subseteq \Omega$ for all $u \in \mathcal{V} \cup \mathcal{E}$. By Lemma 4.4, $T^u\Omega \subseteq \Omega$ for $u \in \mathcal{V}$ if and only if (4.17) holds. By Lemma 4.5, $T^u\Omega \subseteq \Omega$ for $u \in \mathcal{E}$ if and only if (4.18). ■

Thus $\Phi(\mathcal{G})$ is uniquely determined by the components of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. In particular, $\Phi(\mathcal{G})$ is invariant under graph modifications which leave the components of $\mathcal{G}$ invariant. These modifications consist precisely of merging and splitting of edge colour classes such that edges which lie in the same component and

(a) have the same colour in $\mathcal{E}$ never receive different colours

(b) have different colours in $\mathcal{E}$ never receive the same colour.

The graph representing the largest model one may obtain in this way never contains two equally coloured edges between different pairs of vertex colour classes and thus has an edge regular colouring. By Proposition 3.1, it represents a model placing the same constraints on the concentrations and partial correlations.

Corollary 4.2 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a coloured graph and let $\mathcal{E}'$ be an edge colouring obtained from $\mathcal{E}$ by splitting and merging of edge colour classes obeying the principles in (a) and (b) above. Then for $\mathcal{G}' = (\mathcal{V}, \mathcal{E}')$*

$$\Phi(\mathcal{G}) = \Phi(\mathcal{G}').$$

$\mathcal{G}'$ with the finest such colouring is given by $\mathcal{G}' = (\mathcal{V}, \mathcal{E}^*)$ where in $\mathcal{E}^*$ two edges have the same colour if and only if they have the same colour in $\mathcal{E}$ and connect the same vertex colour classes in $\mathcal{V}$. $(\mathcal{V}, \mathcal{E}^*)$ then satisfies

$$\mathcal{S}^+(\mathcal{V}, \mathcal{E}^*) = \mathcal{R}^+(\mathcal{V}, \mathcal{E}^*)$$

and represents the largest RCON model and the largest RCOR model with the same components as $\mathcal{G}$.

□

Distinguishing between the two cases $v = w$ and $v \neq w$, below we obtain a necessary and sufficient condition on Ω_v for $v \in \mathcal{V}$ for (4.18) to hold.

For $u \in \mathcal{E}$, we let $\Lambda^{[u\{vw\}]}$ denote the spectrum of $T^{[u\{vw\}]}$ and let $\Lambda_{\geq 0}^{[u\{vw\}]}$ contain all non-negative elements of $\Lambda^{[u\{vw\}]}$. For eigenvalue $\lambda \in \Lambda^{[u\{vw\}]}$, $E_\lambda^{[u\{vw\}]} \leq \mathbb{R}^{v\cup w}$ is to denote the corresponding eigenspace.

Proposition 4.8 *Let $u \in \mathcal{E}$ and $v, w \in \mathcal{V}$. Then*

$$T^{[u\{vw\}]}\Omega_{v\cup w} \subseteq \Omega_{v\cup w} \tag{4.19}$$

for $v = w$ if and only if there exist $A_\lambda^{[u\{vv\}]} \leq E_\lambda^{[u\{vv\}]}$ for $\lambda \in \Lambda^{[u\{vv\}]}$ such that

$$\Omega_v = \bigoplus_{\lambda\in\Lambda^{[u\{vv\}]}} A_\lambda^{[u\{vv\}]},$$

and for $v \neq w$ if and only if there exist $A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda_{\geq 0}^{[u\{vw\}]}$ such that

$$\Omega_v = \mathcal{I}_v^{v\cup w}\left(\bigoplus_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]}\right) \qquad \textit{and} \qquad \Omega_w = \mathcal{I}_w^{v\cup w}\left(\bigoplus_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]}\right). \tag{4.20}$$

$\square$

Proof: For $v = w$, the claim is a direct consequence of Lemma 4.2.

For $v \neq w$, suppose first that $\Omega_{v\cup w}$ is as in (4.20). It follows that for $g \in \Omega_{v\cup w}$,

$$g = \Pi_v^{v\cup w} g + \Pi_w^{v\cup w} g = \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} v_\lambda + \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} w_\lambda$$

for some $v_\lambda, w_\lambda \in A_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda_{\geq 0}^{[u\{vw\}]}$. Now

$$\begin{aligned}
T^{[u\{vw\}]}g &= T^{[u\{vw\}]}\left[\Pi_v^{v\cup w}\left(\sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} v_\lambda\right) + \Pi_w^{v\cup w}\left(\sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} w_\lambda\right)\right] \\
&= T^{[u\{vw\}]}\left(\sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \Pi_v^{v\cup w} v_\lambda + \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \Pi_w^{v\cup w} w_\lambda\right) \\
&= \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} T^{[u\{vw\}]}\left(\frac{I_{v\cup w} + (\Pi_v^{v\cup w} - \Pi_w^{v\cup w})}{2}\right) v_\lambda \\
&\qquad + \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} T^{[u\{vw\}]}\left(\frac{I_{v\cup w} - (\Pi_v^{v\cup w} - \Pi_w^{v\cup w})}{2}\right) w_\lambda
\end{aligned}$$

where the last equality follows from (4.11). As $v_\lambda, w_\lambda \in A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$, by Lemma 4.3, we have $(\Pi_v^{v\cup w} - \Pi_w^{v\cup w})v_\lambda, (\Pi_v^{v\cup w} - \Pi_w^{v\cup w})w_\lambda \in E_{-\lambda}^{[u\{vw\}]}$, giving

$$\begin{aligned}
T^{[u\{vw\}]}g &= \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \left(\frac{\lambda I_{v\cup w} - \lambda(\Pi_v^{v\cup w} - \Pi_w^{v\cup w})}{2}\right) v_\lambda \\
&\qquad + \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \left(\frac{\lambda I_{v\cup w} + \lambda(\Pi_v^{v\cup w} - \Pi_w^{v\cup w})}{2}\right) w_\lambda \\
&= \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \lambda\left(\frac{I_{v\cup w} - (\Pi_v^{v\cup w} - \Pi_w^{v\cup w})}{2}\right) v_\lambda \\
&\qquad + \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \lambda\left(\frac{I_{v\cup w} + (\Pi_v^{v\cup w} - \Pi_w^{v\cup w})}{2}\right) w_\lambda \\
&= \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \lambda \Pi_w^{v\cup w} v_\lambda + \sum_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \lambda \Pi_v^{v\cup w} w_\lambda \\
&\in \Omega_{v\cup w}
\end{aligned}$$

where we again used (4.11). As g and p are chosen arbitrarily, this establishes (4.19).

Conversely, suppose (4.19) holds. Then, as $\Omega_{v\cup w} = \tilde{\Omega}_v^{v\cup w} \oplus \tilde{\Omega}_w^{v\cup w}$, the Spectral Theorem applied to $T^{[u\{vw\}]}$ gives that for all $\omega \in \tilde{\Omega}_v^{v\cup w}$ there exist unique $\omega_\lambda \in E_\lambda^{[u\{vw\}]}$ such that

$$\omega = \sum_{\lambda\in\Lambda^{[u\{vw\}]}} \omega_\lambda,$$

and as, by definition of $\tilde{\Omega}_v^{v\cup w}$, for all $x \in \tilde{\Omega}_v^{v\cup w}$, $x_\alpha = 0$ for $\alpha \in w$, we must have for $\omega \in \tilde{\Omega}_v^{v\cup w}$,

$$\begin{aligned}
\omega &= \left(\Pi_v^{v\cup w} - \Pi_w^{v\cup w}\right)\omega \\
&= \left(\Pi_v^{v\cup w} - \Pi_w^{v\cup w}\right)\left(\sum_{\lambda\in\Lambda^{[u\{vw\}]}} \omega_\lambda\right) \\
&= \sum_{\lambda\in\Lambda^{[u\{vw\}]}} \left(\Pi_v^{v\cup w} - \Pi_w^{v\cup w}\right)\omega_\lambda.
\end{aligned}$$

By Lemma 4.3,

$$\left(\Pi_v^{v\cup w} - \Pi_w^{v\cup w}\right)\omega_\lambda \in E_{-\lambda}^{[u\{vw\}]},$$

so that

$$\omega_{-\lambda} = \left(\Pi_v^{v\cup w} - \Pi_w^{v\cup w}\right)\omega_\lambda$$

and therefore

$$\tilde{\Omega}_v^{v\cup w} = \bigoplus_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} \left(A_\lambda^{[u\{vw\}]} \oplus \left(\Pi_v^{v\cup w} - \Pi_w^{v\cup w}\right) A_\lambda^{[u\{vw\}]}\right),$$

for some $A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda_{\geq 0}^{[u\{vw\}]}$, which gives

$$\Omega_v = \mathcal{I}_v^{v\cup w}\left(\tilde{\Omega}_v^{v\cup w}\right) = \mathcal{I}_v^{v\cup w}\left(\bigoplus_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]}\right).$$

Similarly,

$$\Omega_w = \mathcal{I}_w^{v\cup w}\left(\bigoplus_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]}\right).$$

■

We are now ready to give our main result in this chapter, which follows from Proposition 4.7 and Proposition 4.8:

Theorem 4.2 *For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\Omega \leq \mathbb{R}^V$, $\Omega \in \Phi(\mathcal{G})$ if and only if*

$$\Omega = \bigoplus_{v\in\mathcal{V}} \tilde{\Omega}_v \tag{4.21}$$

and for all $u \in \mathcal{E}$ and all $v, w \in \mathcal{V}$ there exist $A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda^{[u\{vw\}]}$ such that simultaneously

$$\Omega_v = \mathcal{I}_v^{v\cup w}\left(\bigoplus_{\lambda\in\Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]}\right) = \bigoplus_{\lambda\in\Lambda^{[u\{vv\}]}} A_\lambda^{[u\{vv\}]}. \tag{4.22}$$

□

Proof: To prove the claim, by Proposition 4.7, it suffices to show that condition (4.22) is equivalent to condition (4.18). This follows from applying Proposition 4.8 to every combination of an edge colour class $u \in \mathcal{E}$ and a vertex colour class pair $v, w \in \mathcal{V}$.

■

Note the principal difference between Theorem 4.2 and Proposition 4.8. While Proposition 4.8 gives necessary and sufficient conditions for invariance for only one component, Theorem 4.2 does so for the entire graph, which is a union of components. It precisely states that the necessary and sufficient conditions derived in

Proposition 4.8 must hold for every component in the graph. Thus while Proposition 4.8 gives a condition for $\Omega \leq \mathbb{R}^V$ to be stable *within a single* component, Theorem 4.2 provides for stability *within all* components and *across all* components.

Finally we describe the construction of suitable design matrices X of spaces $\Omega \in \Phi(\mathcal{G})$. For a sequence of real-valued matrices $(M_j)_{j\in J}$, let $\operatorname{diag}((M_j)_{j\in J})$ denote the block-diagonal matrix whose j^{th} block is the j^{th} matrix in the sequence. Then:

Proposition 4.9 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a coloured graph and let $\Omega \in \Phi(\mathcal{G})$. Let $\mathcal{V}_+ \subseteq \mathcal{V}$ be the set of vertex colour classes v in $\mathcal{V}$ for which $\Omega_v \neq 0$ and let $V_+ \subseteq V$ be the set partitioned by $\mathcal{V}_+$. For $v \in \mathcal{V}_+$, let $d^v = \dim(\Omega_v) > 0$ and let $X^v \in \mathbb{R}^{v\times d^v}$ be a matrix whose columns are given by a basis for Ω_v. Then*

$$\Omega_{V_+} = \operatorname{Im}(X) \quad \text{where} \quad X = \operatorname{diag}((X^v)_{v\in\mathcal{V}_+}).$$

□

Proof: By Theorem 4.2, if $\Omega \in \Phi(\mathcal{G})$, then

$$\Omega = \bigoplus_{v\in\mathcal{V}} \tilde{\Omega}_v.$$

For $v \in \mathcal{V}_+$, by construction of X^v, $\Omega_v = \operatorname{Im}(X^v)$, so that, as

$$\Omega_{V_+} = \bigoplus_{v\in\mathcal{V}_+} \tilde{\Omega}_v,$$

$\Omega_{V_+} = \operatorname{Im}\left(\operatorname{diag}((X^v)_{v\in\mathcal{V}_+})\right)$, as claimed.

■

4.3.4. Example

We illustrate our results by applying them to the graph $\mathcal{G}$ displayed in Figure 4.1(a) in Example 4.1. Applying Proposition 4.8, we first give all stable spaces Ω for each component, displayed in Figure 4.2, individually and then, using Theorem 4.2, give all spaces in $\Phi(\mathcal{G})$. We abbreviate $\mathrm{span}(\{\omega_1,\dots,\omega_k\})$ by $\langle\omega_1,\dots,\omega_k\rangle$.

Example 4.2 The graph component $\mathcal{G}^{[u_1\{v_1v_1\}]}$ displayed in Figure 4.2(a) in Example 4.1 has adjacency matrix

$$T^{[u_1\{v_1v_1\}]} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

which has spectrum and eigenspaces

$$\Lambda^{[u_1\{v_1v_1\}]} = \{-1,2\}, \qquad E_2^{[u_1\{v_1v_1\}]} = \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle, \qquad E_{-1}^{[u_1\{v_1v_1\}]} = \left(\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \right)^{\perp}.$$

If $\{r,s\} = \left\{ \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}, \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \right\}$ is an orthonormal basis for $E_{-1}^{[u_1\{v_1v_1\}]}$, by Proposition 4.8, $\Omega_{v_1} \leq \mathbb{R}^{v_1}$ is stable under $T^{[u_1\{v_1v_1\}]}$ if and only if

$$\Omega_{v_1} = \alpha\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \oplus \beta\langle \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \rangle \oplus \gamma\langle \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \rangle$$

with $\alpha,\beta,\gamma \in \{0,1\}$. Thus Ω_{v_1} is $T^{[u_1\{v_1v_1\}]}$-invariant if and only if it contains the unit vector or is orthogonal to it. All 8 $T^{[u_1\{v_1v_1\}]}$-invariant spaces Ω_{v_1} for general r and s together with X^{v_1} are given in Table A.1 in Appendix A.

If $\mathcal{G}^{[u_1\{v_1v_1\}]}$ is regarded as a vertex and edge coloured graph with one component only, then we have found all spaces in $\Phi(\mathcal{G}^{[u_1\{v_1v_1\}]})$. In full agreement with Proposition 4.6, all of them can be grouped into pairs of orthogonal complements in $\mathbb{R}^{v_1}$. The diagram in Figure 4.3 summarises the structure of $\Phi(\mathcal{G}^{[u_1\{v_1v_1\}]})$ by representing each $\Omega \in \Phi(\mathcal{G}^{[u_1\{v_1v_1\}]})$ by their index in Table A.1 and connecting two indices whenever one space forms a maximal subset of the other. (In standard lattice theory terminology, introduced in Section 5.1.2, Figure 4.3 shows the Hasse diagram of the complete lattice $\Phi(\mathcal{G}^{[u_1\{v_1v_1\}]})$. We have not shown the last mentioned fact, which however is not difficult to do using Proposition 4.6.)

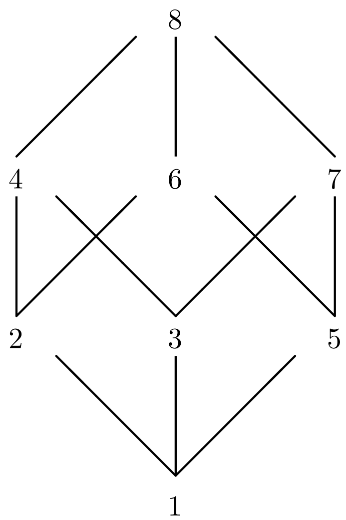


Figure 4.3.: Structure of $\Phi(\mathcal{G}^{[u_1\{v_1v_1\}]})$.

$\square$

Example 4.3 The graph component $\mathcal{G}^{[u_1\{v_1v_2\}]}$ displayed in Figure 4.2(b) has adjacency matrix

$$T^{[u_1\{v_1v_2\}]} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

which, omitting the superscript $[u_1\{v_1v_2\}]$ for the sake of brevity, has spectrum and eigenspaces

$$\Lambda = \{0, \pm 1, \pm 2\},$$

$$E_0 = \langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle,\ E_1 = \langle \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \rangle,\ E_{-1} = \langle \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \rangle,\ E_2 = \langle \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} \rangle,\ E_{-2} = \langle \begin{pmatrix} -2 \\ -1 \\ -1 \\ 1 \\ 1 \\ 2 \end{pmatrix} \rangle.$$

By Proposition 4.8, $\Omega_{v_1 \cup v_2}$ is invariant under $T^{[u_1\{v_1v_2\}]}$ if and only if

$$\begin{aligned} \Omega_{v_1} &= \alpha B_0 \oplus \gamma B_1 \oplus \delta B_2 \\ \Omega_{v_2} &= \beta C_0 \oplus \gamma C_1 \oplus \delta C_2 \end{aligned}$$

with $\alpha, \beta, \gamma, \delta \in \{0, 1\}$ and

$$B_0 = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle,\quad B_1 = \langle \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \rangle,\quad B_2 = \langle \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \rangle,\quad C_0 = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle,\quad C_1 = \langle \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \rangle,\quad C_2 = \langle \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \rangle.$$

All 16 possible pairs of Ω_{v_1} and Ω_{v_2} are given in Table A.2 in Appendix A. Table A.3 in Appendix A gives the relations between $\mu_1, \ldots, \mu_6$ corresponding to the spaces in Table A.2, together with design matrices X^{v_1} and X^{v_2}. If $\mathcal{G}^{[u_1\{v_1v_2\}]}$ is regarded as a vertex and edge coloured graph with one component only, then we have found all spaces in $\Phi(\mathcal{G}^{[u_1\{v_1v_2\}]})$, again in full agreement with Proposition 4.6. The structure of $\Phi(\mathcal{G}^{[u_1\{v_1v_2\}]})$ is summarised in Figure 4.4.

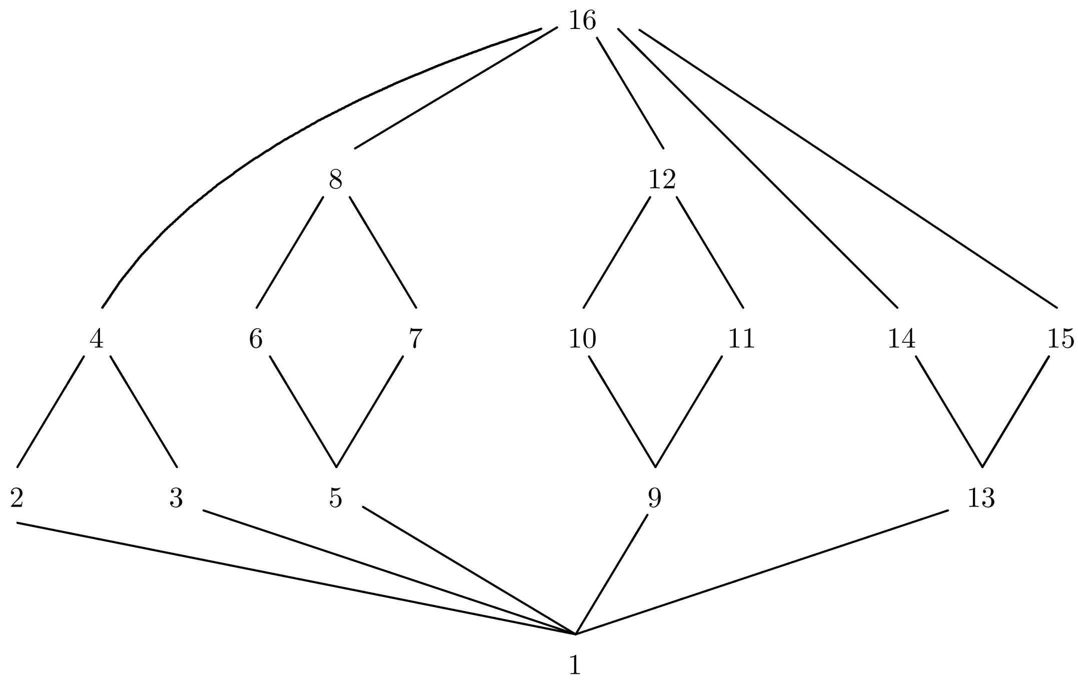


Figure 4.4.: Structure of $\Phi(\mathcal{G}^{[u_1\{v_1v_2\}]})$.

□

Example 4.4 The adjacency matrix of the component $\mathcal{G}^{[u_2\{v_2v_3\}]}$ displayed in Figure 4.2(c) is

$$T^{[u_2\{v_2v_3\}]} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

which, omitting the superscript $[u_2\{v_2v_3\}]$ for the sake of brevity, has spectrum and eigenspaces

$$\Lambda = \{0, \pm\sqrt{2+\sqrt{2}}, \pm\sqrt{2-\sqrt{2}}\},$$

$$E_0 = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle,\ E_{\sqrt{2+\sqrt{2}}} = \langle \begin{pmatrix} 2-\sqrt{2} \\ 2-\sqrt{2} \\ -2+2\sqrt{2} \\ (2-\sqrt{2})\sqrt{2+\sqrt{2}} \\ (-4+3\sqrt{2})\sqrt{2+\sqrt{2}} \end{pmatrix} \rangle,\ E_{-\sqrt{2+\sqrt{2}}} = \langle \begin{pmatrix} -(2-\sqrt{2}) \\ -(2-\sqrt{2}) \\ 2-2\sqrt{2} \\ (2-\sqrt{2})\sqrt{2+\sqrt{2}} \\ (-4+3\sqrt{2})\sqrt{2+\sqrt{2}} \end{pmatrix} \rangle$$

$$E_{\sqrt{2-\sqrt{2}}} = \langle \begin{pmatrix} 2+\sqrt{2} \\ 2+\sqrt{2} \\ -(2+2\sqrt{2}) \\ (2+\sqrt{2})\sqrt{2-\sqrt{2}} \\ (-4-3\sqrt{2})\sqrt{2-\sqrt{2}} \end{pmatrix} \rangle,\ \ E_{-\sqrt{2-\sqrt{2}}} = \langle \begin{pmatrix} -(2+\sqrt{2}) \\ -(2+\sqrt{2}) \\ 2+2\sqrt{2} \\ (2+\sqrt{2})\sqrt{2-\sqrt{2}} \\ (-4-3\sqrt{2})\sqrt{2-\sqrt{2}} \end{pmatrix} \rangle.$$

By Proposition 4.8, $\Omega_{v_2\cup v_3}$ is invariant under $T^{[u_2\{v_2v_3\}]}$ if and only if

$$\begin{aligned} \Omega_{v_2} &= \alpha D_0 \oplus \beta D_{\sqrt{2+\sqrt{2}}} \oplus \gamma D_{\sqrt{2-\sqrt{2}}} \\ \Omega_{v_3} &= \qquad\ \ \beta F_{\sqrt{2+\sqrt{2}}} \oplus \gamma F_{\sqrt{2-\sqrt{2}}} \end{aligned}$$

with $\alpha, \beta, \gamma \in \{0, 1\}$ and

$$D_0 = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle, \quad D_{\sqrt{2+\sqrt{2}}} = \langle \begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \end{pmatrix} \rangle, \quad D_{\sqrt{2-\sqrt{2}}} = \langle \begin{pmatrix} 1 \\ 1 \\ -\sqrt{2} \end{pmatrix} \rangle,$$

$$F_{\sqrt{2+\sqrt{2}}} = \langle \begin{pmatrix} 1 \\ -1+\sqrt{2} \end{pmatrix} \rangle, \quad F_{\sqrt{2-\sqrt{2}}} = \langle \begin{pmatrix} 1 \\ -1-\sqrt{2} \end{pmatrix} \rangle.$$

All 8 possible pairs of Ω_{v_2} and Ω_{v_3} are given in Table A.4 in Appendix A. Table A.5 in Appendix A gives the relations between $\mu_4, \ldots, \mu_8$ corresponding to the spaces in Table A.4, together with matrices X^{v_2} and X^{v_3}. If $\mathcal{G}^{[u_2\{v_2v_3\}]}$ is regarded as a vertex and edge coloured graph with one component only, then we have found all spaces in $\Phi(\mathcal{G}^{[u_2\{v_2v_3\}]})$. The structure of $\Phi(\mathcal{G}^{[u_2\{v_2v_3\}]})$ the same as the structure of $\Phi(\mathcal{G}^{[u_1\{v_1v_1\}]})$, which is represented by the diagram in Figure 4.3.

$\square$

We now combine the constraints of $\Phi(\mathcal{G})$ on each of the individual components to obtain all spaces Ω in $\Phi(\mathcal{G})$.

Example 4.5 For $\mathcal{G}$ displayed in Figure 4.1(a), by Theorem 4.2, $\Omega \leq \mathbb{R}^V$ lies in $\Phi(\mathcal{G})$ if and only if

$$\Omega = \tilde{\Omega}_{v_1} \oplus \tilde{\Omega}_{v_2} \oplus \tilde{\Omega}_{v_3}$$

and

$$\begin{aligned}
\Omega_{v_1} &= \bigoplus_{\lambda \in \Lambda^{[u_1\{v_1v_1\}]}} A_\lambda^{[u_1\{v_1v_1\}]} = \mathcal{I}_{v_1}^{v_1 \cup v_2}\left(\bigoplus_{\lambda \in \Lambda_{\geq}^{[u_1\{v_1v_2\}]}} A_\lambda^{[u_1\{v_1v_2\}]}\right) && (4.23)\\
\Omega_{v_2} &= \mathcal{I}_{v_2}^{v_1 \cup v_2}\left(\bigoplus_{\lambda \in \Lambda_{\geq}^{[u_1\{v_1v_2\}]}} A_\lambda^{[u_1\{v_1v_2\}]}\right) \\
&= \mathcal{I}_{v_2}^{v_2 \cup v_3}\left(\bigoplus_{\lambda \in \Lambda_{\geq}^{[u_2\{v_2v_3\}]}} A_\lambda^{[u_2\{v_2v_3\}]}\right) && (4.24)\\
\Omega_{v_3} &= \mathcal{I}_{v_3}^{v_2 \cup v_3}\left(\bigoplus_{\lambda \in \Lambda_{\geq}^{[u_2\{v_2v_3\}]}} A_\lambda^{[u_2\{v_2v_3\}]}\right) && (4.25)
\end{aligned}$$

for some $A_\lambda^{[u_i\{v_jv_k\}]}$ as in Theorem 4.2 for $1 \leq i, j \leq 2$ and $1 \leq k \leq 3$. Note how a choice for $A_\lambda^{[u_1\{v_1v_1\}]}$ limits the possibilities for $A_\lambda^{[u_1\{v_1v_2\}]}$ in order for (4.23) to be satisfied. Once $A_\lambda^{[u_1\{v_1v_2\}]}$ are chosen, in order for (4.24) to hold, $A_\lambda^{[u_2\{v_2v_3\}]}$ must be chosen accordingly, which, by (4.25), uniquely determines Ω_{v_3}.

We begin by considering the constraints imposed on Ω_{v_1} by the component pair $\mathcal{G}^{[u_1\{v_1v_1\}]}$ and $\mathcal{G}^{[u_1\{v_1v_2\}]}$. Consulting Table A.1 for $\mathcal{G}^{[u_1\{v_1v_1\}]}$, and Tables A.2 and A.3 for $\mathcal{G}^{[u_1\{v_1v_2\}]}$, we obtain the only compatible possibilities to be

1. $\Omega_{v_1} = 0, \qquad \Omega_{v_2} = 0$
2. $\Omega_{v_1} = 0, \qquad \Omega_{v_2} = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle$
3. $\Omega_{v_1} = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle, \qquad \Omega_{v_2} = 0$
4. $\Omega_{v_1} = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle, \qquad \Omega_{v_2} = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle$
5. $\Omega_{v_1} = \left(\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_2} = \left(\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle \right)^{\perp}$
6. $\Omega_{v_1} = \left(\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_2} = \mathbb{R}^{v_2}$
7. $\Omega_{v_1} = \mathbb{R}^{v_1}, \qquad \Omega_{v_2} = \left(\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle \right)^{\perp}$
8. $\Omega_{v_1} = \mathbb{R}^{v_1}, \qquad \Omega_{v_2} = \mathbb{R}^{v_2}.$

Considering the component pair $\mathcal{G}^{[u_1\{v_1v_2\}]}$ and $\mathcal{G}^{[u_2\{v_2v_3\}]}$ gives the following four compatible combinations for Ω_{v_2} and Ω_{v_3}

1. $\Omega_{v_2} = 0, \qquad \Omega_{v_3} = 0$

2. $\Omega_{v_2} = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle, \qquad \Omega_{v_3} = 0$

3. $\Omega_{v_2} = \left(\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_3} = \mathbb{R}^{v_3}$

4. $\Omega_{v_2} = \mathbb{R}^{v_2}, \qquad \Omega_{v_3} = \mathbb{R}^{v_3}.$

Forming the intersection, $\Omega = \tilde{\Omega}_{v_1} \oplus \tilde{\Omega}_{v_2} \oplus \tilde{\Omega}_{v_3} \in \Phi(\mathcal{G})$ if and only if Ω_{v_1}, Ω_{v_2} and Ω_{v_3} fall into one of the following cases

1. $\Omega_{v_1} = 0, \qquad \Omega_{v_2} = 0, \qquad \Omega_{v_3} = 0$

2. $\Omega_{v_1} = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle, \qquad \Omega_{v_2} = 0, \qquad \Omega_{v_3} = 0$

3. $\Omega_{v_1} = 0, \qquad \Omega_{v_2} = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle, \qquad \Omega_{v_3} = 0$

4. $\Omega_{v_1} = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle, \qquad \Omega_{v_2} = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle, \qquad \Omega_{v_3} = 0$

5. $\Omega_{v_1} = \left(\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_2} = \left(\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_3} = \mathbb{R}^{v_3}$

6. $\Omega_{v_1} = \mathbb{R}^{v_1}, \qquad \Omega_{v_2} = \left(\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_3} = \mathbb{R}^{v_3}$

7. $\Omega_{v_1} = \left(\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle \right)^{\perp}, \qquad \Omega_{v_2} = \mathbb{R}^{v_2}, \qquad \Omega_{v_3} = \mathbb{R}^{v_3}$

8. $\Omega_{v_1} = \mathbb{R}^{v_1}, \qquad \Omega_{v_2} = \mathbb{R}^{v_2}, \qquad \Omega_{v_3} = \mathbb{R}^{v_3}.$

The corresponding design matrices X are given by

1. X not defined.
2. $X = X^{v_1} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$; X^{v_2} and X^{v_3} not defined.
3. $X = X^{v_2} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$; X^{v_1} and X^{v_3} not defined.
4. $X = \text{diag}(X^{v_1}, X^{v_2}) = \text{diag}\left(\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right)$; X^{v_3} not defined.
5. $X = \text{diag}\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}, I_2\right)$
6. $X = \text{diag}\left(I_3, \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}, I_2\right)$
7. $X = \text{diag}\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}, I_3, I_2\right)$
8. $X = I_8$

giving the corresponding maximum likelihood estimates, based on a sample $Y^1 = y^1, \ldots, Y^n = y^n$, to be

1.

$$\hat{\mu}_i = 0 \quad \text{for } 1 \leq i \leq 8$$

2.
$$\hat{\mu}_i = \begin{cases} 0 & \text{for } i = 1,4,5,6,7,8 \\ \sum_{j=1}^{n}(y_2^j - y_3^j)/2n & \text{for } i = 2 \\ \sum_{j=1}^{n}(y_3^j - y_2^j)/2n & \text{for } i = 3 \end{cases}$$

3.
$$\hat{\mu}_i = \begin{cases} 0 & \text{for } i = 1,2,3,6,7,8 \\ \sum_{j=1}^{n}(y_4^j - y_5^j)/2n & \text{for } i = 4 \\ \sum_{j=1}^{n}(y_5^j - y_4^j)/2n & \text{for } i = 5 \end{cases}$$

4.
$$\hat{\mu}_i = \begin{cases} 0 & \text{for } i = 1,6,7,8 \\ \sum_{j=1}^{n}(y_i^j - y_{i+1}^j)/2n & \text{for } i = 2,4 \\ \sum_{j=1}^{n}(y_i^j - y_{i-1}^j)/2n & \text{for } i = 3,5 \end{cases}$$

5.
$$\hat{\mu}_i = \begin{cases} \sum_{j=1}^{n} y_i^j/n & \text{for } i = 1,6,7,8 \\ \sum_{j=1}^{n}(y_i^j + y_{i+1}^j)/2n & \text{for } i = 2,4 \\ \sum_{j=1}^{n}(y_i^j + y_{i-1}^j)/2n & \text{for } i = 3,5 \end{cases}$$

6.
$$\hat{\mu}_i = \begin{cases} \sum_{j=1}^{n} y_i^j/n & \text{for } i = 1,2,3,6,7,8 \\ \sum_{j=1}^{n}(y_4^j + y_5^j)/2n & \text{for } i = 4,5 \end{cases}$$

7.
$$\hat{\mu}_i = \begin{cases} \sum_{j=1}^{n} y_i^j/n & \text{for } i = 1,4,5,6,7,8 \\ \sum_{j=1}^{n}(y_2^j + y_3^j)/2n & \text{for } i = 2,3 \end{cases}$$

8.
$$\hat{\mu}_i = \sum_{j=1}^{n} y_i^j / n \quad \text{for } 1 \leq i \leq 8.$$

Note how all mean spaces Ω in $\Phi(\mathcal{G})$ can be grouped into orthogonal complements in $\mathbb{R}^{v_1} \oplus \mathbb{R}^{v_2} \oplus \mathbb{R}^{v_3}$. The structure of $\Phi(\mathcal{G})$ is given in Figure 4.5.

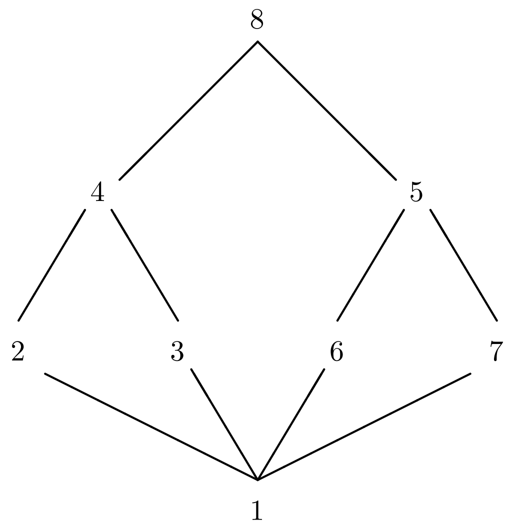


Figure 4.5.: Structure of $\Phi(\mathcal{G})$.

□

4.4. Equality Constraints on the Mean

Theorem 4.2 gave a necessary and sufficient condition for $\Omega \leq \mathbb{R}^V$ to lie inside $\Phi(\mathcal{G})$ for a coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. In this section we consider the special case of spaces Ω which are constant on given subsets of V, corresponding to equality constraints on the entries of the mean vector μ. More precisely, we consider spaces of the form

$$\Omega(\mathcal{M}) = \{\mu \in \mathbb{R}^V \mid \mu_\alpha = \mu_\beta \text{ whenever } \alpha \equiv \beta \ (\mathcal{M})\}$$

or equivalently,

$$\Omega(\mathcal{M}) = \bigoplus_{\eta \in \mathcal{M}} \tilde{\Omega}_\eta \qquad \text{with} \qquad \Omega_\eta = \langle e_\eta \rangle \quad \forall\, \eta \in \mathcal{M} \tag{4.26}$$

for partitions $\mathcal{M}$ of V, where $e_\eta \in \mathbb{R}^\eta$ denotes the constant vector whose entries are all equal to one. We will show that for such Ω, conditions (4.21) and (4.22) in Theorem 4.2 translate into regularity conditions on $\mathcal{M}$ and $(\mathcal{V}, \mathcal{E})$.

4.4.1. Equitable Partitions

Equitable partitions were first introduced by Sachs (1966):

Definition 4.2 (Sachs) Let $G = (V, E)$ be an undirected graph. Then a partition $\mathcal{V}$ of V is called *equitable with respect to* G if for all $v \in \mathcal{V}$, $\alpha, \beta \in v$,

$$|\mathrm{ne}(\alpha) \cap w| = |\mathrm{ne}(\beta) \cap w| \quad \forall\ w \in \mathcal{V}.$$

□

The following directly follows from the definition of equitable partitions:

Lemma 4.6 *Let $G = (V, E)$ be an undirected graph. Then $\mathcal{V}$ is equitable with respect to G if and only if for all $\mathcal{W} \subseteq \mathcal{V}$, $\mathcal{W}$ is equitable with respect to $G^W = (W, E_W)$ where $W = \cup_{w \in \mathcal{W}} w$.*

□

Lemma 4.7 *Let $G = (V, E)$ be an undirected graph with adjacency matrix T and let $\mathcal{V}$ be a partition of V. Then $T\Omega(\mathcal{V}) \subseteq \Omega(\mathcal{V})$ if and only if $\mathcal{V}$ is equitable with respect to G.*

□

Proof: Let $x \in \Omega(\mathcal{V})$ and consider the action of T on x. For $\alpha \in V$,

$$(Tx)_\alpha = \sum_{\beta \in \mathrm{ne}(\alpha)} x_\beta = \sum_{v \in \mathcal{V}} \sum_{\beta \in \mathrm{ne}(\alpha) \cap v} x_\beta.$$

As, by assumption, x is constant on each set $v \in \mathcal{V}$, taking the value x_v say, the above simplifies to

$$(Tx)_\alpha = \sum_{v \in \mathcal{V}} |\mathrm{ne}(\alpha) \cap v| x_v. \tag{4.27}$$

Now $(Tx) \in \Omega(\mathcal{V})$ if and only if $(Tx)_\alpha = (Tx)_\beta$ whenever $\alpha \equiv \beta \ (\mathcal{V})$. By (4.27) this happens if and only if

$$|\mathrm{ne}(\alpha) \cap v| = |\mathrm{ne}(\beta) \cap v| \quad \forall \ v \in \mathcal{V},$$

i.e. if and only if $\mathcal{V}$ is equitable with respect to G as claimed.

■

For an alternative, and essentially equivalent, proof in the framework of Algebraic Graph Theory we refer to Chan and Godsil (1997, Lemma 2.2). We deduce:

Lemma 4.8 *Let $G = (\mathcal{V}, E)$ be a vertex coloured graph. Let T be the adjacency matrix of $G = (V, E)$, let Λ denote the spectrum of T and for $\lambda \in \Lambda$ let E_λ be the eigenspace of λ. Then there exist $A_\lambda \leq E_\lambda$ for $\lambda \in \Lambda$ such that*

$$\Omega(\mathcal{V}) = \bigoplus_{\lambda \in \Lambda} A_\lambda$$

if and only if $\mathcal{V}$ is equitable with respect to G.

□

Proof: By Lemma 4.7, $\mathcal{V}$ is equitable with respect to G if and only if $T\Omega(\mathcal{V}) \subseteq \Omega(\mathcal{V})$. The claim then follows directly from Lemma 4.2.

■

4.4.2. Vertex Regular Colourings

The analogue of equitable partitions for vertex coloured graphs are *vertex regular* graph colourings for vertex and edge coloured graphs, which we define below.

Definition 4.3 For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ we say that $(\mathcal{V}, \mathcal{E})$ is *vertex regular* if $\mathcal{V}$ is equitable with respect to $G^u = (V, u)$ for all $u \in \mathcal{E}$.

□

Vertex regular colourings can be characterised by the following two properties:

Lemma 4.9 *A graph colouring $(\mathcal{V}, \mathcal{E})$ is vertex regular if and only if for all $u \in \mathcal{E}$ and $v, w \in \mathcal{V}$, all $u\{vv\}$-components are regular and all $u\{vw\}$-components are bi-regular.*

□

Proof: For $\alpha \in V$ and $u \in \mathcal{E}$ let

$$\mathrm{ne}_u(\alpha) = \{\beta \mid \alpha\beta \in u\}. \tag{4.28}$$

Then, by Definition 4.3, $(\mathcal{V}, \mathcal{E})$ is vertex regular if and only if for all $u \in \mathcal{E}$, $v, w \in \mathcal{V}$,

$$|\mathrm{ne}_u(\alpha) \cap w| = |\mathrm{ne}_u(\beta) \cap w| \quad \forall\ \alpha, \beta \in v. \tag{4.29}$$

For $v = w$ condition (4.29) is equivalent to all $u\{vv\}$-components being regular, for $v \neq w$ (4.29) is equivalent to all $u\{vw\}$-components being bi-regular.

■

Proposition 4.10 *A graph colouring $(\mathcal{V}, \mathcal{E})$ is vertex regular if and only if for all $\mathcal{W} \subseteq \mathcal{V}$, $(\mathcal{W}, \mathcal{E}_\mathcal{W})$ is vertex regular, where $\mathcal{E}_\mathcal{W} = \{u \cap E_W\}_{u \in \mathcal{E}}$ and $W = \cup_{w \in \mathcal{W}} w$.*

□

Put in words, $(\mathcal{V}, \mathcal{E})$ is vertex regular if and only if any collection of vertex colour classes induces a vertex regular colouring.

Proof of Proposition 4.10: If $(\mathcal{W}, \mathcal{E}_{\mathcal{W}})$ is vertex regular for all $\mathcal{W} \subseteq \mathcal{V}$, then clearly $(\mathcal{V}, \mathcal{E})$ is vertex regular as $\mathcal{E}_{\mathcal{V}} = \mathcal{E}$. So suppose that $(\mathcal{V}, \mathcal{E})$ is vertex regular. By Lemma 4.9, for all $u \in \mathcal{E}$ and $v, w \in \mathcal{V}$, all $u\{vv\}$-components are regular and all $u\{vw\}$-components are bi-regular. Since $(\mathcal{V}, \mathcal{E})$ and $(\mathcal{W}, \mathcal{E}_{\mathcal{W}})$ agree in their components for all $\mathcal{W} \subseteq \mathcal{V}$, $(\mathcal{W}, \mathcal{E}_{\mathcal{W}})$ must be vertex regular colouring for all $\mathcal{W} \subseteq \mathcal{V}$. ■

4.4.3. $\mathcal{G}$-Invariance

Theorem 4.3 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a vertex and edge coloured graph and let $\mathcal{M}$ be a partition of V. Then $\Omega(\mathcal{M}) \in \Phi(\mathcal{G})$ if and only if*

(i) $\mathcal{M} \leq \mathcal{V}$

(ii) $(\mathcal{M}, \mathcal{E})$ *is vertex regular.*

□

Proof: By Theorem 4.2, $\Omega \in \Phi(\mathcal{G})$ for $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ if and only if

$$\Omega = \bigoplus_{v \in \mathcal{V}} \tilde{\Omega}_v \tag{4.30}$$

and for all $u \in \mathcal{E}$ and all $v, w \in \mathcal{V}$ there exist $A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda^{[u\{vw\}]}$ such that

$$\Omega_v = \mathcal{I}_v^{v \cup w} \left(\bigoplus_{\lambda \in \Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]} \right) = \bigoplus_{\lambda \in \Lambda^{[u\{vv\}]}} A_\lambda^{[u\{vv\}]}. \tag{4.31}$$

By definition of $\Omega(\mathcal{M})$ in (4.26),

$$\Omega(\mathcal{M}) = \bigoplus_{v \in \mathcal{V}} \tilde{\Omega}_v$$

if and only if $\mathcal{M} \leq \mathcal{V}$, which is precisely condition (i) in the claim. So to prove the result we need to show that for $\mathcal{M} \leq \mathcal{V}$, $(\mathcal{M}, \mathcal{E})$ is vertex regular if and only if there exist $A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda^{[u\{vw\}]}$ such that (4.31) becomes

$$\Omega_v = \bigoplus_{\eta \subseteq v} \tilde{\Omega}_\eta^v, \qquad \Omega_\eta = \langle e_\eta \rangle \qquad \forall\ v \in \mathcal{V}.$$

So assume $\mathcal{M} \leq \mathcal{V}$, or equivalently that every set in $\mathcal{V}$ is a union of sets in $\mathcal{M}$. For $v, w \in \mathcal{V}$ let $\mathcal{M}_v$ denote the partition of v induced by $\mathcal{M}$ and similarly for $v \cup w$ and $\mathcal{M}_{v \cup w}$. Then, by definition of vertex regularity and Proposition 4.10, $(\mathcal{M}, \mathcal{E})$ is vertex regular if and only if for all $v, w \in \mathcal{V}$, $\mathcal{M}_v$ is equitable with respect to all $u\{vv\}$-components of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\mathcal{M}_{v \cup w}$ is equitable with respect to all $u\{vw\}$-components of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.

Thus, by Lemma 4.8 applied to each component individually, $(\mathcal{M}, \mathcal{E})$ is vertex regular if and only if for all $u \in \mathcal{E}$ and all $v \in \mathcal{V}$ there exist $A_\lambda^{[u\{vv\}]} \leq E_\lambda^{[u\{vv\}]}$ for $\lambda \in \Lambda^{[u\{vv\}]}$ such that

$$\Omega(\mathcal{M}_v) = \bigoplus_{\lambda \in \Lambda^{[u\{vv\}]}} A_\lambda^{[u\{vv\}]} \tag{4.32}$$

Together Lemma 4.3 and Lemma 4.8 give that for all $u \in \mathcal{E}$ and all $v, w \in \mathcal{V}$ there exist $A_\lambda^{[u\{vw\}]} \leq E_\lambda^{[u\{vw\}]}$ for $\lambda \in \Lambda^{[u\{vw\}]}$ such that

$$\Omega(\mathcal{M}_{v \cup w}) = \bigoplus_{\lambda \in \Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]} \tag{4.33}$$

if and only if $(\mathcal{M}, \mathcal{E})$ is vertex regular. Combining (4.32) and (4.33) we obtain that there exist $A_\lambda^{[u\{vv\}]} \leq E_\lambda^{[u\{vv\}]}$ for $\lambda \in \Lambda^{[u\{vv\}]}$ such that for all $v \in \mathcal{V}$,

$$\Omega_v = \mathcal{I}_v^{v \cup w} \left(\bigoplus_{\lambda \in \Lambda_{\geq 0}^{[u\{vw\}]}} A_\lambda^{[u\{vw\}]} \right) = \bigoplus_{\lambda \in \Lambda^{[u\{vv\}]}} A_\lambda^{[u\{vv\}]} = \bigoplus_{\eta \subseteq v} \tilde{\Omega}_\eta^v$$

for $\Omega_\eta = \langle e_\eta \rangle$ if and only if $(\mathcal{M}, \mathcal{E})$ is vertex regular.

■

Note that if $\mathcal{M}$ is the partition $\{\{\alpha\}\}_{\alpha \in V}$ of V into singletons, corresponding to $\Omega(\mathcal{M}) = \mathbb{R}^V$, and thus an unrestricted mean μ, then $\mathcal{M}$ is finer than any vertex colouring $\mathcal{V}$ and $(\mathcal{M}, \mathcal{E})$ is vertex regular for any edge colouring $\mathcal{E}$. Thus $\mathcal{M}$ always satisfies the two conditions in Theorem 4.3, which conforms with the fact that we always have equality of estimators $\hat{\mu} = \mu^*$ for an unrestricted mean μ.

Further, by Proposition 4.6, if $\Omega(\mathcal{M}) \in \Phi(\mathcal{G})$, then $\Omega(\mathcal{M})^\perp \in \Phi(\mathcal{G})$ with

$$\Omega(\mathcal{M})^\perp = \{\mu \in \mathbb{R}^V \mid \sum_{\alpha \in \eta} \mu_\alpha = 0, \ \forall\ \eta \in \mathcal{M}\}.$$

Thus all hypotheses in Section 2.3.1 which restricted the mean, namely all except H_{vc}, either equal $\mu \in \Omega(\mathcal{M})$ or $\mu \in \Omega(\mathcal{M})^\perp$ for a partition $\mathcal{M}$ of V which satisfies the conditions in Theorem 4.3.

A well-known example of when condition (i) in Theorem 4.3 is violated is the *Behrens-Fisher problem* (Behrens, 1929; Fisher, 1939). The problem, in its original univariate form, concerns testing the equality of the means μ_1 and μ_2 based on two samples $Y_1^1, \ldots, Y_1^m$ and $Y_2^1, \ldots, Y_2^n$ with different variances σ_1^2 and σ_2^2. If it is assumed that $\mu_1 = \mu_2 = \mu$, the likelihood function may have multiple local maxima in μ, so that standard statistical inference results cease to hold (e.g. Scheffé, 1944).

(Many alternative solutions have been developed over time, see e.g. Scheffé (1943) for the univariate case and Bennett (1951) for the multivariate generalisation, also summarised in Anderson (1984, Section 5.5). For more recent, algebraic, treatments, see Drton (2008) for the univariate and Buot et al. (2007) for the multivariate case.)

The Behrens-Fisher problem can be formulated in our framework, as, in our notation, the goal is to maximise the likelihood of a sample drawn from

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim \mathcal{N}_2(\mu, \Sigma), \quad \mu \in \Omega(\mathcal{M}), \quad \Sigma^{-1} \in \mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E}) \tag{4.34}$$

with $\mathcal{M} = \{\{1, 2\}\}$, $\mathcal{V} = \{\{1\}, \{2\}\}$ and $\mathcal{E} = \emptyset$. We give the corresponding graph colourings in Figure 4.6.

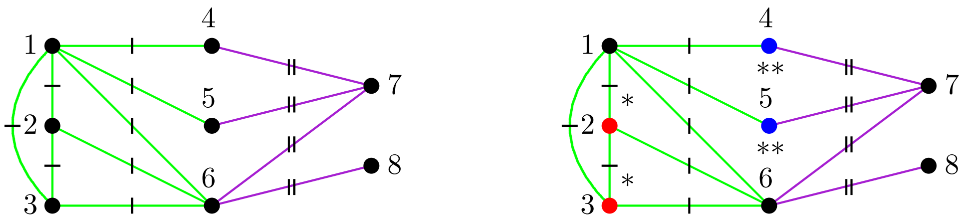
(a) $(\mathcal{V}, \mathcal{E})$ (b) $(\mathcal{M}, \mathcal{E})$

Figure 4.6.: The Behrens-Fisher problem.

Clearly $\mathcal{M} \not\leq \mathcal{V}$, so that, by Theorem 4.3, we do not have equality of estimators $\mu^* = \hat{\mu}$. Indeed, this is the simplest example of a graphical Gaussian model with symmetry constraints violating condition (i) in Theorem 4.3.

Looking back at Example 4.5 with coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ displayed in Figure 4.1(a), we observe that the $\mathcal{G}$-stable spaces Ω which are determined by equality constraints precisely correspond to refinements $\mathcal{M}$ of the vertex colouring $\mathcal{V}$ which give rise to vertex regular graph colourings $(\mathcal{M}, \mathcal{E})$.

These are the singleton partition $\mathcal{M}_1 = \{\{1\}, \{2\}, \ldots, \{8\}\}$, giving $\mu \in \mathbb{R}^V$, and $\mathcal{M}_2 = \{\{1\}, \{2, 3\}, \{4, 5\}, \{6\}, \{7\}, \{8\}\}$ corresponding to $\mu \in \mathbb{R}^V$ subject to $\mu_2 = \mu_3$ and $\mu_4 = \mu_5$. The graph colourings $(\mathcal{M}_1, \mathcal{E})$ and $(\mathcal{M}_2, \mathcal{E})$ are given in Figure 4.7.



(a) $(\mathcal{M}_1, \mathcal{E})$ (b) $(\mathcal{M}_2, \mathcal{E})$

Figure 4.7.: Vertex regular colourings of graph $\mathcal{G}$ displayed in Figure 4.1(a).

Theorem 4.3 further implies the following for RCOP models:

Corollary 4.3 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the coloured graph of an RCOP model generated by permutation group $\Gamma \leq \mathrm{Aut}(G)$. Then $\Omega(\mathcal{V}) \in \Phi(\mathcal{G})$.*

□

For the proof we require the following preliminary result.

Lemma 4.10 *Let $G = (V, E)$ be a graph and let $\Gamma \leq \mathrm{Aut}(G)$ have vertex orbits $\mathcal{V}$ and edge orbits $\mathcal{E}$. Then $(\mathcal{V}, \mathcal{E})$ is vertex regular.*

□

Proof: We showed in the proof of Proposition 3.2 in Section 2.3.1 that $(\mathcal{V}, \mathcal{E})$ is edge regular, so that whenever two edges are of equal colour in $\mathcal{E}$, they connect the same vertex colour classes in $\mathcal{V}$.

If $\alpha, \beta \in V$ are two equally coloured vertices in $\mathcal{V}$, then, by assumption, there exists a permutation $\sigma \in \Gamma$ which maps α to β leaving $(\mathcal{V}, \mathcal{E})$ invariant, so that α and β must have the same degree in every edge colour class. For $\mathrm{ne}_u(\alpha)$ as in (4.28), the above imply $|\mathrm{ne}_u(\alpha) \cap v| = |\mathrm{ne}_u(\beta) \cap v|$ for all $v \in \mathcal{V}$ and all pairs $\alpha, \beta \in V$ with $\alpha \equiv \beta$ $(\mathcal{V})$, which is precisely the definition of vertex regularity for $(\mathcal{V}, \mathcal{E})$.

■

Proof of Corollary 4.3: It suffices to show that $\mathcal{M} = \mathcal{V}$ satisfies conditions (i) and (ii) in Theorem 4.3. (i) trivially holds, and (ii) has been shown in Lemma 4.10.

∎

In the next chapter we will show that for every coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, there exists a unique coarsest partition $\mathcal{M}$ of V such that $\Omega(\mathcal{M}) \in \Phi(\mathcal{G})$.

4.5. Comparative Experiments

We illustrate the result in Theorem 4.3 by applying it to the design of comparative experiments which ensure mean treatment effects to be estimable by their least squares estimator if the response is assumed to follow a Gaussian distribution.

Suppose we wish to estimate mean treatment effects $(\tau_t)_{t\in T}$ of a set of treatments T with $|T| > 1$ by applying them to experimental sites labelled by V which are arranged in a grid with row set R and column set C. After treatments are applied to the sites, the response $Y = (Y_\alpha)_{\alpha\in V}$ is measured, assuming that

$$Y \sim \mathcal{N}_V(\mu, \Sigma) \quad \text{with} \quad \mathbb{E}(Y_\alpha) = \mu_\alpha = \tau_{T(\alpha)}, \qquad \alpha \in V \tag{4.35}$$

where $T(\alpha)$ denotes the treatment which has been applied to site α. If $\mathcal{T}$ is the partition of V, or *treatment allocation*, given by

$$\alpha \equiv \beta \ (\mathcal{T}) \quad \Longleftrightarrow \quad T(\alpha) = T(\beta), \quad \alpha, \beta \in V.$$

then (4.35) may be equivalently expressed as

$$Y \sim \mathcal{N}_V(\mu, \Sigma) \quad \text{with} \quad \mu \in \Omega(\mathcal{T}).$$

Suppose further that while the value of $\Sigma \in \mathcal{S}^+$ is unknown, the nature of the experiment, for instance the physical layout of the experimental sites, suggests (i) $\mathrm{Var}(Y_\alpha|Y_{V\setminus\{\alpha\}})$ to be constant for all $\alpha \in V$, (ii) the response Y_α at any given site α to be conditionally independent of the remaining variables given Y_β for β lying next to α, and (iii) the non-zero concentrations to be constant within rows and constant within columns.

If the response at the first and last site in each row are assumed to be uncorrelated, we are in the setting of a *homogeneous nearest-neighbour lattice scheme*, as it was first defined by Besag (1974) as a special case of the nearest-neighbour system introduced by Whittle (1963) and Bartlett (1955, Section 2.2). If the response at the first and last site in each row, and similarly column, are assumed to be subject to the same dependence as all nearest neighbours within rows, and similarly columns, following Besag and Moran (1975), we refer to the dependence structure as a *homogeneous nearest-neighbour torus scheme*.

Such assumptions are common in spatial statistics based on random Markov fields (Whittle, 1954; Ripley, 1981, 1988), with a classic application lying in modelling the yield of agricultural plots (Besag, 1974; Besag and Higdon, 1999).

In order to allow for row and column effects on the mean response, it is sensible to only consider treatment allocations $\mathcal{T}$ in which each treatment is applied in each row and column equally often. The simplest way in which this can be satisfied is if $|R| = |C| = |T|$ and each treatment appears in each row and each column precisely once, in which case treatments are allocated according to a *Latin square design* (Bailey, 2008, Section 6.2).

The question we address in this section is how treatments should be allocated in the above settings in order for the maximum likelihood estimator $\hat{\tau}$ of the mean treatment effects τ is to equal its least squares estimator τ^*. Applying Theorem 4.3, we show that while suitable Latin square designs exist for the lattice scheme only for $|T| = 2$, for a torus scheme suitable designs can be constructed for T of all sizes.

4.5.1. Homogeneous Nearest-Neighbour Lattice Scheme

For $\alpha \in V = R \times C$, let $\mathrm{ne}_R(\alpha)$ denote the immediate neighbourhood of the site labelled α in its row and similarly for $\mathrm{ne}_C(\alpha)$. If we let $N(\alpha) = \mathrm{ne}_R(\alpha) \cup \mathrm{ne}_C(\alpha)$, then a nearest-neighbour lattice scheme assumes (i) $\mathrm{Var}(Y_\alpha | Y_{V\setminus\{\alpha\}})$ to be constant for all $\alpha \in V$, and (ii) Y_α to be conditionally independent of $Y_{V\setminus(N(\alpha)\cup\{\alpha\})}$ given $Y_{N(\alpha)}$. A *homogeneous* nearest-neighbour lattice scheme assumes the dependence structure to be independent of the position in the grid, up to boundary effects.

Thus, a homogeneous nearest-neighbour lattice scheme restricts $\Sigma^{-1} \in \mathcal{S}^+$ to be of the form

$$\left[(\Sigma^l)^{-1}\right]_{\alpha\beta} = \begin{cases} \kappa^2 & \text{if } \alpha = \beta \\ \rho_1^l \kappa^2 & \text{if } \beta \in \mathrm{ne}_R(\alpha) \\ \rho_2^l \kappa^2 & \text{if } \beta \in \mathrm{ne}_C(\alpha) \\ 0 & \text{otherwise} \end{cases}$$

(where we add the superscript l for 'lattice'). If we let

$$E_R^l = \{\alpha\beta \mid \alpha, \beta \in V,\ \beta \in \mathrm{ne}_R(\alpha)\} \qquad E_C^l = \{\alpha\beta \mid \alpha, \beta \in V, \beta \in \mathrm{ne}_C(\alpha)\}$$

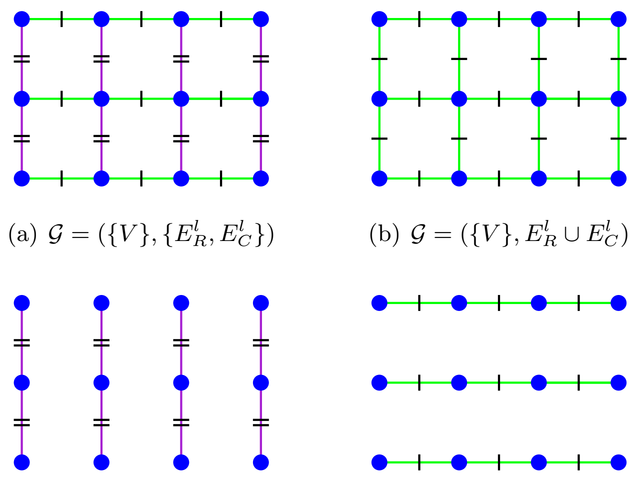
then the dependence structure can be represented by a coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

with $\mathcal{V} = \{V\}$ and

$$\mathcal{E} = \begin{cases} \{E_R^l, E_C^l\} & \text{for } \rho_1^l \neq \rho_2^l, \;\; \rho_1^l, \rho_2^l \neq 0 \\ \{E_R^l \cup E_C^l\} & \text{for } \rho_1^l = \rho_2^l \neq 0 \\ \{E_C^l\} & \text{for } \rho_1^l = 0, \;\; \rho_2^l \neq 0 \\ \{E_R^l\} & \text{for } \rho_1^l \neq 0, \;\; \rho_2^l = 0 \end{cases}$$

where we discard the possibility of an empty edge set as it corresponds to $\Sigma = \sigma^2 I$, as then we always have $\hat{\tau} = \tau^*$ by default.

We give $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for $|R| = 3$ and $|C| = 4$ in Figure 4.8. As all vertices are of equal colour, we omit asterisks for ease of exposition. Note that because all vertices are collected in one colour class, we have $\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E})$.



(a) $\mathcal{G} = (\{V\}, \{E_R^l, E_C^l\})$ (b) $\mathcal{G} = (\{V\}, E_R^l \cup E_C^l)$

(c) $\mathcal{G} = (\{V\}, \{E_C^l\})$ (d) $\mathcal{G} = (\{V\}, \{E_R^l\})$

Figure 4.8.: Homogeneous nearest-neighbour lattice schemes.

Proposition 4.11 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the coloured graph of a homogeneous nearest-*

neighbour lattice scheme. Then there exists a treatment allocation $\mathcal{T}$ according to a $|T| \times |T|$ Latin square such that $\Omega(\mathcal{T}) \in \Phi(\mathcal{G})$ if and only if $|T| = 2$.

□

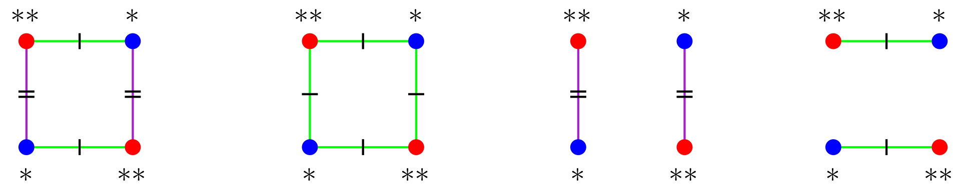
Proof: In order to prove the claim, by Theorem 4.3, it suffices to show that for $\mathcal{T}$ giving rise to a Latin square design and $\mathcal{E}$ being equal to any of $\{E_R^l, E_C^l\}$, $\{E_R^l \cup E_C^l\}$, $\{E_R^l\}$, $\{E_C^l\}$, $(\mathcal{T}, \mathcal{E})$ is vertex regular if only if $|T| = 2$.

A lattice has three kinds of vertices: firstly, the corner vertices of degree two; secondly, the boundary vertices which are not corner vertices, which have degree three; and finally, the off-boundary vertices of degree four. Let $\mathcal{E}$ be any of $\{E_R^l, E_C^l\}$, $\{E_C^l\}$, and $\{E_R^l\}$. If $\alpha \in V$ is a corner vertex, then α has degree one in every edge colour class in $\mathcal{E}$. In order for $(\mathcal{T}, \mathcal{E})$ to be vertex regular, this must be true for each vertex inside the same set as α in $\mathcal{T}$. Thus if α is not a singleton in $\mathcal{T}$ (in which case we have $|T| = 1$), then any vertex β in the same set as α in $\mathcal{T}$ must be a corner vertex as well. However as every treatment appears exactly once in every row and column, β must be the corner vertex lying diagonally opposite α in $\mathcal{G}$, giving $|T| = 2$.

For $\mathcal{E} = \{E_R^l \cup E_C^l\}$, as before, let α be in a corner vertex in V. Then α has degree two in the only edge colour class in the graph. This must be true for all other vertices β inside the same set as α in $\mathcal{T}$, and as $\mathcal{T}$ describes a Latin square design, β must be the corner vertex lying diagonally opposite α in V. Again we obtain $|T| = 2$.

■

The corresponding coloured graphs $\mathcal{G} = (\mathcal{T}, \mathcal{E})$ are given in Figure 4.9. The edge colouring $\mathcal{E}$ represents the restrictions on Σ^{-1} while the vertex colouring $\mathcal{T}$ represents the allocation of treatments.



(a) $\mathcal{G} = (\mathcal{T}, \{E_R^l, E_C^l\})$ (b) $\mathcal{G} = (\mathcal{T}, \{E_R^l \cup E_C^l\})$ (c) $\mathcal{G} = (\mathcal{T}, \{E_C^l\})$ (d) $\mathcal{G} = (\mathcal{T}, \{E_R^l\})$

Figure 4.9.: Latin square design for homogeneous nearest-neighbour lattice scheme.

4.5.2. Homogenous Nearest-Neighbour Torus Scheme

If additionally there is a carryover effect from the end of a row to its beginning and similarly for the columns, for example if the sites are physically arranged on a torus, Σ^{-1} will be restricted according to what, following Besag and Moran (1975), we define to be a *homogeneous nearest-neighbour torus scheme*. It is given by assuming

$$\left[(\Sigma^t)^{-1}\right]_{\alpha\beta} = \begin{cases} \kappa^2 & \text{if } \alpha = \beta \\ \rho_1^t \kappa^2 & \text{if } \beta \in \mathrm{Ne}_R(\alpha) \\ \rho_2^t \kappa^2 & \text{if } \beta \in \mathrm{Ne}_C(\alpha) \\ 0 & \text{otherwise} \end{cases}$$

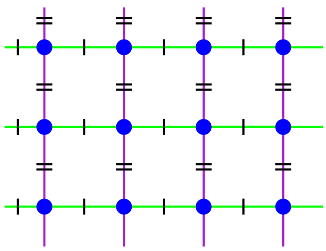
(where we add the superscript t for 'torus' and) where $\mathrm{Ne}_R(\alpha)$ and $\mathrm{Ne}_C(\alpha)$ denote the immediate neighbourhoods of $\alpha \in V$ within its row and column respectively if the first and last vertex in each row are connected, as are the first and last vertex in each column. If

$$E_R^t = \{\alpha\beta \mid \alpha, \beta \in V,\ \beta \in \mathrm{Ne}_R(\alpha)\} \qquad E_C^t = \{\alpha\beta \mid \alpha, \beta \in V,\ \beta \in \mathrm{Ne}_C(\alpha)\}$$

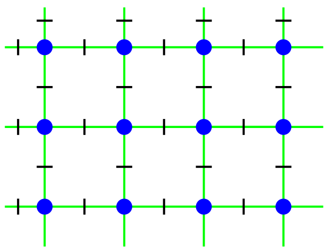
the corresponding coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is given by $\mathcal{V} = \{V\}$ and

$$\mathcal{E} = \begin{cases} \{E_R^t, E_C^t\} & \text{for } \rho_1^t \neq \rho_2^t,\ \ \rho_1^t, \rho_2^t \neq 0 \\ \{E_R^t \cup E_C^t\} & \text{for } \rho_1^t = \rho_2^t \neq 0 \\ \{E_C^t\} & \text{for } \rho_1^t = 0,\ \ \rho_2^t \neq 0 \\ \{E_R^t\} & \text{for } \rho_1^t \neq 0,\ \ \rho_2^t = 0 \end{cases}$$

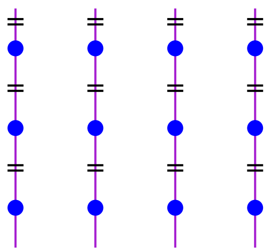
The graphs for $|R| = 3$ and $|C| = 4$ are given in Figure 4.10. For ease of exposition, in our display we indicate edges between first and last vertices in rows and columns by half edges which point outside the plot. The halves are then assumed to join up to a full edge. Just as for lattice schemes, we have $\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E})$.



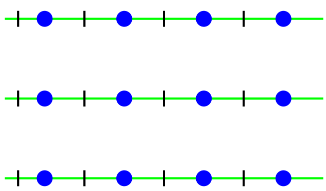
(a) $\mathcal{G} = (\{V\}, \{E_R^t, E_C^t\})$



(b) $\mathcal{G} = (\{V\}, \{E_R^t \cup E_C^t\})$



(c) $\mathcal{G} = (\{V\}, \{E_C^t\})$



(d) $\mathcal{G} = (\{V\}, \{E_R^t\})$

Figure 4.10.: Homogeneous nearest-neighbour torus schemes.

Section 6.2 in Bailey (2008) describes various construction algorithms for Latin squares. One of them is the cyclic method, in which the consecutiveness of treatments is fixed and the treatments are applied in this order to each row with a constant shift of one. To give an example, let $T = \{t_1, t_2, t_3\}$ and let V be given by a 3×3 grid. Then a Latin square can be constructed by the cyclic method by applying treatments (t_1, t_2, t_3) to the first row, treatments (t_3, t_1, t_2) to the second row and treatments (t_2, t_3, t_1) to the third row in that order. Clearly, each treatment appears precisely once in each row and column.

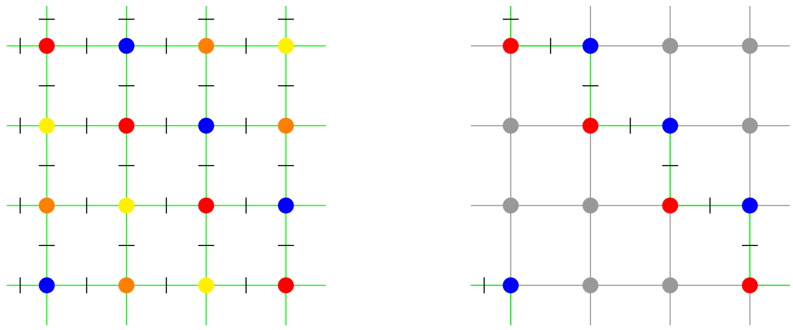
Proposition 4.12 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the coloured graph of a homogeneous nearest-neighbour torus scheme. Then if $\mathcal{T}$ is a treatment allocation according to a $|T| \times |T|$ cyclic Latin square, then $\Omega(\mathcal{T}) \in \Phi(\mathcal{G})$ for all $|T| > 1$.*

$\square$

Proof: In order to prove the claim, by Theorem 4.3, it suffices to show that for $\mathcal{T}$ corresponding to a cyclic Latin square design and $\mathcal{E}$ being equal to any of $\{E^t_R, E^t_C\}$, $\{E^t_R \cup E^t_C\}$, $\{E^t_R\}$, $\{E^t_C\}$, $(\mathcal{T}, \mathcal{E})$ is vertex regular for all $|T| > 1$.

If $\mathcal{E}$ is any of $\{E^l_R, E^l_C\}$, $\{E^l_C\}$, and $\{E^l_R\}$, then clearly $(\mathcal{T}, \mathcal{E})$ is vertex regular, as the adjacencies between the treatments are identical in each row and column by construction. For $\mathcal{E} = \{E^l_R \cup E^l_C\}$, it suffices to consider the example displayed in Figure 4.11. Figure 4.11(a) displays a cyclic Latin square design for $|T| = 4$. The labelling of vertices by asterisks to indicate their colour class is omitted for the sake of legibility.

To show that $(\mathcal{T}, \mathcal{E})$ is vertex regular in this case, we need to show that each graph component $\mathcal{G}^{[\mathcal{E}\{vw\}]}$ has constant degree in v and in w respectively, for all $v, w \in \mathcal{T}$. This holds, as either $\mathcal{G}^{[\mathcal{E}\{vw\}]}$ is isomorphic to the graph in Figure 4.11(b), or it is empty.



(a) $(\mathcal{T},\mathcal{E})$ (b) $\mathcal{G}^{[\mathcal{E}\{vw\}]}$

Figure 4.11.: Latin square design for homogeneous nearest-neighbour torus scheme.

■

Note that Propositions 4.11 and 4.12 remain unchanged if instead of $(\Sigma^{-1})_{\alpha\alpha} = (\Sigma^{-1})_{\beta\beta}$ for all $\alpha, \beta \in V$ we assume that $(\Sigma^{-1})_{\alpha\alpha} = (\Sigma^{-1})_{\beta\beta}$ whenever $T(\alpha) = T(\beta)$. This is because for $\mathcal{V} = \{V\}$, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\mathcal{T}$ satisfy conditions (i) and (ii) in Theorem 4.3 if and only if $\mathcal{G}' = (\mathcal{T}, \mathcal{E})$ and $\mathcal{T}$ do, for any partition $\mathcal{T}$ of V.

5. Structures of Model Classes

In the previous chapters we identified three colouring regularities which correspond to statistically desirable properties in the represented RCON and RCOR models. Edge regular colourings (see Definition 3.4 in Section 3.4) ensure $\mathcal{S}^+(\mathcal{V},\mathcal{E}) = \mathcal{R}^+(\mathcal{V},\mathcal{E})$, while vertex regular colourings (see Definition 4.3 in Section 4.4.2) guarantee a model mean subject to equality constraints to be estimable by its least squares estimate. Permutation generated colourings (Definition 3.6 in Section 3.4.1), representing RCOP models in the terminology of Højsgaard and Lauritzen (2008), are equipped with both properties and additionally allow a particularly simple maximisation of the likelihood function. A fourth statistically meaningful colouring regularity is given by the intersection of vertex and edge regular colourings, which, following Siemons (1983), we term to be regular.

Motivated by the need for model selection methods for RCON and RCOR models this chapter is devoted to the study of the structure of RCON and RCOR models, and the model classes corresponding to the above colouring regularities. The motivation to study the latter is threefold: firstly, we demonstrate that the number of RCON and RCOR models grows dramatically with the number of variables, so that smaller model search spaces are desirable. Secondly, generic equality constraints of RCON and RCOR models are generally not readily interpretable and, lastly, do not guarantee the corresponding model to have any particular statistical properties.

The main results presented in this chapter are that each of the model classes forms a non-distributive lattice of models and the identification of their meet and join operations (Theorems 5.2, 5.3, 5.4, 5.5). They are established by showing that each model class is stable under model intersection, which gives the shared meet operation, and by demonstrating that whenever a model does not fall inside a given class there is a unique smallest larger model which does, giving the distinct join operations.

Most of the presented results, partially also to be found in Gehrmann (2011), are of combinatorial or graph theoretic nature. Their interpretation and practical implications for model selection are deferred to the next chapter.

5.1. Preliminaries

5.1.1. Notation

We let U_V denote the set of undirected graphs with vertex set V, and let $\mathcal{S}_V^+$ be the set of symmetric positive definite $V \times V$ matrices. $\mathcal{S}_V^G = \{\mathcal{S}^+(G) \mid G \in U_V\}$ is to denote the family of mean-zero graphical Gaussian models with variable set V, and F_V the set of their factor graphs. For a discrete set S we let $P(S)$ denote the power set of S and $\mathcal{P}(S)$ the set of partitions of S.

In order to keep our exposition readable, in this chapter we identify the set of coloured graphs on vertex set V with the set of graph colourings of U_V and denote both by $\mathcal{C}_V$. The two sets are clearly isomorphic and whether we are referring to a graph colouring or coloured graph shall be determined by context. For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, $G = (V, E) \in U_V$ denotes the underlying uncoloured graph.

We let $\mathcal{S}_V^{\mathcal{G}} = \{\mathcal{S}^+(\mathcal{V}, \mathcal{E}) \mid (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V\}$ and $\mathcal{R}_V^{\mathcal{G}} = \{\mathcal{R}^+(\mathcal{V}, \mathcal{E}) \mid (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V\}$ denote the sets of corresponding sets of RCON and RCOR models with variable set V. $\mathcal{F}_V$ is to denote the set of coloured factor graphs of RCON and RCOR models, as well as the set of colourings of F_V. For $\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F) \in \mathcal{F}_V$, $G_F = (V \cup N, E_F) \in F_V$ denotes the underlying uncoloured factor graph.

5.1.2. Partially Ordered Sets and Lattices

A *binary relation* ρ on a set A is a subset of $A \times A$ with two elements $a, b \in A$ being *in relation* with respect to ρ if $(a, b) \in \rho$, denoted by $a \ \rho \ b$. If a binary relation ρ satisfies (P1)-(P3) below for all $a, b, c \in A$, it is a *partial ordering relation* and A is a *partially ordered set* or *poset*, which we also denote by $\langle A; \leq \rangle$.

(P1) Reflexivity: $a \ \rho \ a$

(P2) Antisymmetry: $a \ \rho \ b$ and $b \ \rho \ a \ \Rightarrow \ a = b$

(P3) Transitivity: $a \ \rho \ b$ and $b \ \rho \ c \ \Rightarrow \ a \ \rho \ c$

If $\langle A; \leq \rangle$ is a poset and B is a nonempty subset of A, then '$\leq$' is also a partial ordering on B and we call $\langle B; \leq \rangle$ a *subposet* of $\langle A; \leq \rangle$. '$\leq$' is our generic symbol for a partial ordering. Which ordering is meant will be determined by context.

For $H \subseteq A$ and $a \in A$, a is an *upper bound* of H if $h \leq a$ for all $h \in H$. An upper bound a of H is the *least upper bound*, or *supremum*, of H, denoted $a = \sup H$ or $a = \bigvee H$, if $a \leq b$ for any upper bound b of H. The *lower bound* and *greatest lower bound*, or *infimum*, denoted $\inf H$ or $\bigwedge H$, are defined similarly. For $H = \emptyset$, $\bigvee H$ is the smallest element in A, called *zero*, if it exists, and $\bigwedge H$ is the largest element in A, called *unit*, if it exists. For $H = \{a, b\}$, we write $a \wedge b$ for $\bigwedge H$ and $a \vee b$ for $\bigvee H$, and call $\wedge$ the *meet* and $\vee$ the *join* operation respectively.

An elegant principle which applies to posets is the *duality principle*: rather than writing $a \leq b$ we may equivalently write $b \geq a$, which also is a binary relation on A satisfying properties (P1)-(P3). Thus $\langle A; \geq \rangle$ is a poset, called the *dual* of $\langle A; \leq \rangle$.

$\langle A; \leq \rangle$ is a *lattice* if $a \wedge b$ and $a \vee b$ exists for all $a, b \in A$. Thus in lattices, $\wedge$ and $\vee$ are binary operations, which uniquely define the lattice. We therefore sometimes denote a lattice $\langle A; \leq \rangle$ as $\langle A; \wedge, \vee \rangle$. An equivalent necessary and sufficient condition for $\langle A; \leq \rangle$ to be a lattice is that $\bigwedge H$ and $\bigvee H$ exist for all finite nonempty $H \subseteq A$.

A lattice $\langle L; \leq \rangle$ is *bounded* if it has both a zero and a unit. In a bounded lattice $\langle L; \wedge, \vee \rangle$, $a \in L$ is a *complement* of $b \in L$ if $a \vee b = 1$ and $a \wedge b = 0$. If every element in L has a complement it is a *complemented* lattice. L is *distributive* if for all $a, b, c \in L$,

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c). \tag{5.1}$$

If L is both complemented and distributive we call it a *Boolean algebra*. Bounded distributive lattices are always Boolean algebras. If $\langle L; \leq \rangle$ is a lattice and $B \subseteq L$ is non-empty, then $\langle B; \leq \rangle$ is a *sublattice* $\langle B; \wedge, \vee \rangle$ of L if it is stable under $\wedge$ and $\vee$. If B is stable under $\wedge$, it is a *meet-sublattice*.

Lemma 5.1 *Let $\langle L; \leq \rangle = \langle L; \wedge, \vee \rangle$ be a finite lattice and let $\langle B; \leq \rangle$ be a meet-sublattice of L. Then $\langle B; \leq \rangle$ is a lattice with meet operation $\wedge$ inherited from L, and join operation $\vee_B$ which for $A \subseteq B$ is given by*

$$\bigvee_B A = \bigwedge \{s \in B \mid a \leq s, \ \textit{all } a \in A\}$$

and may be distinct from $\vee$. $\square$

We omit the proof as it is immediate. The structure of a lattice $\langle L; \leq \rangle$ may be visualised by a *Hasse diagram*, in which each element pair $a, b, \in L$ is joined by an edge whenever $a \leq b$ and there is no $x \in L \setminus \{a, b\}$ such that $a \leq x \leq b$.

Lemma 5.2 *$\langle P(S); \subseteq \rangle$ is a Boolean algebra for all discrete sets S.*

□

Proof: It suffices to observe that $\langle P(S); \subseteq \rangle$ is a bounded lattice with zero $\emptyset$ and unit S. The meet and join operations are given by set intersection and set union, which are distributive operations.

■

Lemma 5.3 *$\langle \mathcal{P}(S); \leq \rangle$ is a lattice for all discrete sets S. However, $\langle \mathcal{P}(S); \leq \rangle$ is generally not a Boolean algebra.*

□

Proof: To prove the claim, we need to show that $\bigwedge H$ exists for all $H \subseteq \mathcal{P}(S)$. So let $H = \{P_i\}_{i \in I} \subseteq \mathcal{P}(S)$, and let $Q \in \mathcal{P}(S)$ be given by $Q = \{\{\alpha\}\}_{\alpha \in S}$ if $I = H = \emptyset$, and otherwise let Q be determined by

$$\alpha \equiv \beta \ (Q) \qquad \Longleftrightarrow \qquad \alpha \equiv \beta \ (P_i) \quad \forall \ i \in I, \quad \alpha, \beta \in S.$$

Then $Q \leq P_i$ for all $i \in I$ while being the coarsest such partition. Thus, $Q = \bigwedge H$, which therefore exists for all $H \subseteq \mathcal{P}(S)$, establishing the claim. The meet and join operations of $\langle \mathcal{P}(S); \leq \rangle$ are given by

$$\alpha \equiv \beta \ (P_1 \wedge P_2) \qquad \Longleftrightarrow \qquad \alpha \equiv \beta \ (P_1) \quad \text{and} \quad \alpha \equiv \beta \ (P_2) \tag{5.2}$$

$$\begin{aligned} \alpha \equiv \beta \ (P_1 \vee P_2) \qquad \Longleftrightarrow \qquad & \exists \ \gamma_0, \ldots, \gamma_k \in S \text{ such that } \gamma_0 = \alpha, \gamma_k = \beta, \\ & \gamma_{i-1} \equiv \gamma_i \ (P_{j_i}), j_i \in \{1, 2\} \text{ for } 1 \leq i \leq k \end{aligned} \tag{5.3}$$

for $P_1, P_2 \in \mathcal{P}(S)$ and $\alpha, \beta \in S$. The zero of $\langle \mathcal{P}(S); \leq \rangle$ is the singleton partition $\{\{\alpha\}\}_{\alpha \in S}$ and its unit is $\{S\}$. $\langle \mathcal{P}(S); \leq \rangle$ is generally not a Boolean algebra as it is not distributive. For instance, relation (5.1) is violated for $S = [5]$ by $a = \{\{1, 2\}, \{3, 4\}, \{5\}\}$, $b = \{\{1, 4\}, \{2, 3, 5\}\}$ and $c = \{\{1, 3\}, \{2, 4, 5\}\}$ as

$$a \vee (b \wedge c) = \{\{1,2,5\},\{3,4\}\} \neq \{\{1,2,3,4,5\}\} = (a \vee b) \wedge (a \vee c).$$

$\blacksquare$

Most concepts presented in this section can be found in or trivially deduced from facts in Chapter I of Grätzer (1998).

5.2. Unrestricted Graphical Gaussian Models

It is well known that U_V is a partially ordered set with partial ordering '$\leq$' given by edge set inclusion, i.e. for $G_1 = (V, E_1), G_2 = (V, E_2) \in U_V$,

$$G_1 = (V, E_1) \leq (V, E_2) = G_2 \qquad \Longleftrightarrow \qquad E_1 \subseteq E_2. \tag{5.4}$$

Proposition 5.1 *$\langle U_V; \leq \rangle$ is a Boolean algebra with*

$$G_1 \wedge G_2 = (V, E_1 \cap E_2) \qquad \text{and} \qquad G_1 \vee G_2 = (V, E_1 \cup E_2) \tag{5.5}$$

for $G_1 = (V, E_1), G_2 = (V, E_2) \in U_V$.

$\square$

Proof: The claim follows from the fact that $\langle U_V; \leq \rangle = \langle P(V_{\neq}^{(2)}); \subseteq \rangle$ (with '$\subseteq$' only applying to the edges) and Lemma 5.2. The smallest and largest graph in U_V with respect to '$\leq$' defined in (5.4) are the empty graph and complete graph respectively.

$\blacksquare$

The structure of $\langle U_V, \leq \rangle$ directly translates to $\mathcal{S}_V^+$:

Proposition 5.2 *For $G_1, G_2 \in U_V$,*

$$\mathcal{S}^+(G_1) \subseteq \mathcal{S}^+(G_2) \qquad \Longleftrightarrow \qquad G_1 \leq G_2.$$

$\square$

Proof: $\mathcal{S}^+(G_1) \subseteq \mathcal{S}^+(G_2)$ if and only if whenever an entry is zero in all matrices in $\mathcal{S}^+(G_2)$ it is zero in all matrices in $\mathcal{S}^+(G_1)$. This happens if and only if the set of missing edges in G_2 is a subset of the missing edges in G_1, i.e. if and only if $G_1 \leq G_2$.

■

Combining Proposition 5.1 and Proposition 5.2 we obtain:

Proposition 5.3 *$\langle \mathcal{S}_V^G; \subseteq \rangle$ is a Boolean algebra with partial ordering induced by '$\leq$' on U_V as defined in (5.5). For $G_1, G_2 \in U_V$ the meet and join operation are*

$$\mathcal{S}^+(G_1) \wedge \mathcal{S}^+(G_2) = \mathcal{S}^+(G_1 \wedge G_2) \qquad \mathcal{S}^+(G_1) \vee \mathcal{S}^+(G_2) = \mathcal{S}^+(G_1 \vee G_2).$$

□

Corollary 5.1

$$|\mathcal{S}_V^G| = |U_V| = 2^{\binom{|V|}{2}}$$

□

Proof: This follows directly from the fact that each graph in U_V has at most $\binom{|V|}{2}$ edges, each of which is either present or not.

■

5.3. RCON and RCOR Models

Definition 5.1 *For $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G})$, $\mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \in \mathcal{C}_V$, we write $\mathcal{G} \preceq \mathcal{H}$ if*

(i) $G \leq H$

(ii) $\mathcal{V}_\mathcal{G} \geq \mathcal{V}_\mathcal{H}$

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

□

Put in words, $\mathcal{G} \preceq \mathcal{H}$ if $\mathcal{G}$ can be obtained from $\mathcal{H}$ through combinations of

- dropping existing edge colour classes
- merging existing edge colour classes
- merging vertex colour classes

or equivalently, if $\mathcal{H}$ can be obtained from $\mathcal{G}$ by combinations of

- adding new edge colour classes
- splitting existing edge colour classes
- splitting vertex colour classes.

Clearly '$\preceq$' defines a partial ordering on $\mathcal{C}_V$. For example, for the graphs $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$ for $i = 1, 2, 3$ in Figure 5.1, $\mathcal{G}_1 \preceq \mathcal{G}_2$ as conditions (i)-(iii) in Definition 5.1 above are satisfied, whereas $\mathcal{G}_1 \npreceq \mathcal{G}_3$ because conditions (ii) and (iii) are violated.

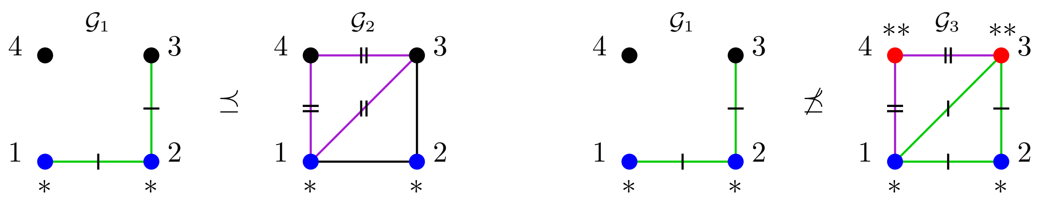


Figure 5.1.: Partial ordering in $\mathcal{C}_{[4]}$.

Proposition 5.4 *Let $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}), \mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \in \mathcal{C}_V$. Then*

$$\mathcal{S}^+(\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}) \subseteq \mathcal{S}^+(\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \quad \Longleftrightarrow \quad \mathcal{G} \preceq \mathcal{H} \quad \Longleftrightarrow \quad \mathcal{R}^+(\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}) \subseteq \mathcal{R}^+(\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}).$$

□

Proof: $\mathcal{S}^+(\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}) \subseteq \mathcal{S}^+(\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H})$ if and only if whenever an entry is zero for all matrices in $\mathcal{S}^+(\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H})$ it is zero for all matrices in $\mathcal{S}^+(\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G})$, and whenever two entries in all matrices in $\mathcal{S}^+(\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H})$ are equal they are also equal in all matrices in $\mathcal{S}^+(\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G})$. This is equivalent to $\mathcal{G}$ and $\mathcal{H}$ satisfying the conditions in Definition 5.1. The same arguments give $\mathcal{R}^+(\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}) \subseteq \mathcal{R}^+(\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H})$ if and only if $\mathcal{G} \preceq \mathcal{H}$. ■

Proposition 5.5 *For $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}), \mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \in \mathcal{C}_V$, let*

$$\mathcal{G} \sqcap \mathcal{H} = (\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}) \qquad \textit{and} \qquad \mathcal{G} \sqcup \mathcal{H} = (\mathcal{V}_\mathcal{G} \wedge \mathcal{V}_\mathcal{H}, \mathcal{E}^{**}_\mathcal{G} \wedge \mathcal{E}^{**}_\mathcal{H}) \tag{5.6}$$

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 edge set in $E_\mathcal{G} \cap E_\mathcal{H}$, $\mathcal{E}^{**}_\mathcal{G} = \mathcal{E}_\mathcal{G} \cup \{\{E_\mathcal{H} \backslash E_\mathcal{G}\}\}$ and $\mathcal{E}^{**}_\mathcal{H} = \mathcal{E}_\mathcal{H} \cup \{\{E_\mathcal{G} \backslash E_\mathcal{H}\}\}$. $\langle \mathcal{C}_V; \preceq \rangle$ is a non-distributive lattice with meet and join operations given by

$$\mathcal{G} \wedge \mathcal{H} = \mathcal{G} \sqcap \mathcal{H} \qquad \textit{and} \qquad \mathcal{G} \vee \mathcal{H} = \mathcal{G} \sqcup \mathcal{H}.$$

□

Proof: Let $\mathcal{G}_\sqcap = (\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}) = (\mathcal{V}_\sqcap, \mathcal{E}_\sqcap)$, $\mathcal{G}^* = (\mathcal{V}_\mathcal{G}, \mathcal{E}^*_\mathcal{G})$ and $\mathcal{H}^* = (\mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{H})$. Then, by (5.6),

$$\mathcal{G} \sqcap \mathcal{H} = \mathcal{G}^* \sqcap \mathcal{H}^*,$$

so that $\mathcal{G}_\sqcap$ is the infimum of $\{\mathcal{G}, \mathcal{H}\}$ in $\mathcal{C}_V$ if and only if it is the infimum of $\{\mathcal{G}^*, \mathcal{H}^*\}$.

To show the latter, we need to prove that $\mathcal{G}_\sqcap \preceq \mathcal{G}^*$, $\mathcal{G}_\sqcap \preceq \mathcal{H}^*$ and that $\mathcal{G}_\sqcap$ is the largest such graph. We consider conditions (i), (ii) and (iii) in Definition 5.1.

(i) As $E_\sqcap = E^*_\mathcal{G} = E^*_\mathcal{H}$, we clearly have $G_\sqcap \leq G^*$ and $G_\sqcap \leq H^*$.

(ii) As $\mathcal{V}_\sqcap = \mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}$, by construction $\mathcal{V}_\sqcap \geq \mathcal{V}_\mathcal{G}$ and $\mathcal{V}_\sqcap \geq \mathcal{V}_\mathcal{H}$.

(iii) $\mathcal{E}_\sqcap = \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}$ which directly implies $\mathcal{E}_\sqcap \geq \mathcal{E}^*_\mathcal{G}$ and $\mathcal{E}_\sqcap \geq \mathcal{E}^*_\mathcal{H}$.

$\mathcal{G}_\sqcap$ is the largest graph which is smaller than both $\mathcal{G}^*$ and $\mathcal{H}^*$ as $E_\sqcap$ is chosen to be of maximal order, and $\mathcal{V}_\sqcap$ and $\mathcal{E}_\sqcap$ are chosen to be as fine as possible.

For the join operation, let $\mathcal{G}_\sqcup = (\mathcal{V}_\mathcal{G} \wedge \mathcal{V}_\mathcal{H}, \mathcal{E}^{**}_\mathcal{G} \wedge \mathcal{E}^{**}_\mathcal{H}) = (\mathcal{V}_\sqcup, \mathcal{E}_\sqcup)$ and similarly let $\mathcal{G}^{**} = (\mathcal{V}_\mathcal{G}, \mathcal{E}^{**}_\mathcal{G})$ and $\mathcal{H}^{**} = (\mathcal{V}_\mathcal{H}, \mathcal{E}^{**}_\mathcal{H})$. Then

$$\mathcal{G} \sqcup \mathcal{H} = \mathcal{G}^{**} \sqcup \mathcal{H}^{**}.$$

We need to show that $\mathcal{G}_\sqcup \succeq \mathcal{G}^{**}$, $\mathcal{G}_\sqcup \succeq \mathcal{H}^{**}$ and that $\mathcal{G}_\sqcup$ is the smallest such graph. Verifying conditions (i), (ii) and (iii) in Definition 5.1 gives the following.

(i) As $E_\sqcup = E^{**}_\mathcal{G} = E^{**}_\mathcal{H} = E_\mathcal{G} \cup E_\mathcal{H}$, we clearly have $G_\sqcup \geq G^{**}$ and $G_\sqcup \geq H^{**}$.

(ii) As $\mathcal{V}_\sqcup = \mathcal{V}_\mathcal{G} \wedge \mathcal{V}_\mathcal{H}$, by construction $\mathcal{V}_\mathcal{G} \geq \mathcal{V}_\sqcup$ and $\mathcal{V}_\mathcal{H} \geq \mathcal{V}_\sqcup$.

(iii) $\mathcal{E}_\sqcup = \mathcal{E}^{**}_\mathcal{G} \wedge \mathcal{E}^{**}_\mathcal{H}$ which directly implies $\mathcal{E}^{**}_\mathcal{G} \geq \mathcal{E}_\sqcup$ and $\mathcal{E}^{**}_\mathcal{H} \geq \mathcal{E}_\sqcup$.

$\mathcal{G}_\sqcup$ is the smallest graph which is larger than both $\mathcal{G}^{**}$ and $\mathcal{H}^{**}$ as $E_\sqcup$ is chosen to be of minimal order, and $\mathcal{V}_\sqcup$ and $\mathcal{E}_\sqcup$ are chosen to be as coarse as possible.

The zero of $\langle \mathcal{C}_V; \preceq \rangle$ is given by the empty graph with all vertices in one colour class $\mathcal{G} = (\{V\}, \emptyset)$ and the unit is the complete graph in which all vertices and edges form singleton colour classes $\mathcal{G} = (\{\{\alpha\}\}_{\alpha \in V}, \{\{\alpha\beta\}\}_{\alpha\beta \in V^{(2)}_{\neq}})$.

Non-distributivity of $\langle \mathcal{C}_V; \preceq \rangle$ is inherited from non-distributivity of the sets of possible vertex and edge colourings $\mathcal{P}(V)$ and $\mathcal{P}(V^{(2)}_{\neq})$ respectively. Let $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}), \mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}), \mathcal{K} = (\mathcal{V}_\mathcal{K}, \mathcal{E}) \in \mathcal{C}_V$, all agreeing in their edge colouring $\mathcal{E}$. If $\langle \mathcal{C}_V; \preceq \rangle$ is distributive then

$$\mathcal{G} \vee (\mathcal{H} \wedge \mathcal{K}) = (\mathcal{G} \vee \mathcal{H}) \wedge (\mathcal{G} \vee \mathcal{K}). \tag{5.7}$$

for all $\mathcal{V}_\mathcal{G}, \mathcal{V}_\mathcal{H}, \mathcal{V}_\mathcal{K} \in \mathcal{P}(V),\ \mathcal{E} \in \mathcal{P}(E)$. As

$$\begin{aligned} \mathcal{G} \vee (\mathcal{H} \wedge \mathcal{K}) &= (\mathcal{V}_\mathcal{G} \wedge (\mathcal{V}_\mathcal{H} \vee \mathcal{V}_\mathcal{K}), \mathcal{E}) \\ (\mathcal{G} \vee \mathcal{H}) \wedge (\mathcal{G} \vee \mathcal{K}) &= ((\mathcal{V}_\mathcal{G} \wedge \mathcal{V}_\mathcal{H}) \vee (\mathcal{V}_\mathcal{G} \wedge \mathcal{V}_\mathcal{K}), \mathcal{E}), \end{aligned}$$

by the duality principle, (5.7) holds if and only if

$$\mathcal{V}_{\mathcal{G}} \vee (\mathcal{V}_{\mathcal{H}} \wedge \mathcal{V}_{\mathcal{K}}) = (\mathcal{V}_{\mathcal{G}} \vee \mathcal{V}_{\mathcal{H}}) \wedge (\mathcal{V}_{\mathcal{G}} \vee \mathcal{V}_{\mathcal{K}})$$

for all $\mathcal{V}_{\mathcal{G}}, \mathcal{V}_{\mathcal{H}}, \mathcal{V}_{\mathcal{K}} \in \mathcal{P}(V)$, i.e. if and only if $\mathcal{P}(V)$ is distributive, which it isn't by Lemma 5.3. Thus $\langle \mathcal{C}_V; \preceq \rangle$ is non-distributive as well.

■

We illustrate the meet and join operations in $\langle \mathcal{C}_V; \preceq \rangle$ given in (5.6) in Figure 5.2.

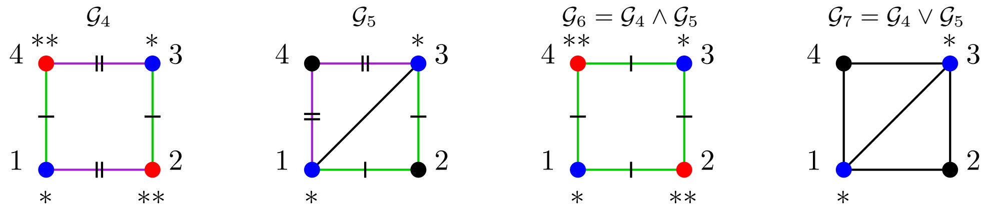


Figure 5.2.: Meet and join operations in $\mathcal{C}_{[4]}$.

Proposition 5.4 and Proposition 5.5 together give the following result:

Theorem 5.1 *$\langle \mathcal{S}_V^{\mathcal{G}}; \subseteq \rangle$ and $\langle \mathcal{R}_V^{\mathcal{G}}; \subseteq \rangle$ are non-distributive lattices with partial ordering induced by '$\preceq$' on $\mathcal{C}_V$ as defined in Definition 5.1. For $\mathcal{G} = (\mathcal{V}_{\mathcal{G}}, \mathcal{E}_{\mathcal{G}})$, $\mathcal{H} = (\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}}) \in \mathcal{C}_V$, we let $\mathcal{G} \wedge \mathcal{H} = (\mathcal{V}_\wedge, \mathcal{E}_\wedge)$ and $\mathcal{G} \vee \mathcal{H} = (\mathcal{V}_\vee, \mathcal{E}_\vee)$, then the corresponding meet and join operations are*

$$\mathcal{S}^+(\mathcal{V}_{\mathcal{G}}, \mathcal{E}_{\mathcal{G}}) \wedge \mathcal{S}^+(\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}}) = \mathcal{S}^+(\mathcal{V}_\wedge, \mathcal{E}_\wedge) \qquad \mathcal{S}^+(\mathcal{V}_{\mathcal{G}}, \mathcal{E}_{\mathcal{G}}) \vee \mathcal{S}^+(\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}}) = \mathcal{S}^+(\mathcal{V}_\vee, \mathcal{E}_\vee),$$

and similarly for $\langle \mathcal{R}_V^{\mathcal{G}}; \subseteq \rangle$.

□

Proposition 5.6 *The sets $\mathcal{C}_V$ and $\mathcal{F}_V$ are isomorphic.*

□

Proof: For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, the corresponding factor graph $\mathcal{G}_F$ is constructed by

1. replacing each edge $\alpha\beta \in E$ by a node and connecting it to α and β in V,
2. colouring the nodes with colours not appearing in $\mathcal{V}$, giving two nodes the same colour if and only if the edges they replace were of equal colour.

Let the corresponding map be denoted by $\phi_V : \mathcal{C}_V \to \mathcal{F}_V$. Then ϕ_V is a bijective homomorphism as is ϕ_V^{-1}, as adjacency of two vertices in $\mathcal{G}$ directly translates into a node being connected to the two end vertices in $\mathcal{G}_F$ and vice versa. Thus $\mathcal{C}_V$ and $\mathcal{F}_V$ are isomorphic.

■

5.3.1. Number of RCON and RCOR Models

As, by (3.1) and (3.3), for both model types there is one model parameter for each vertex colour class in $\mathcal{V}$ and one for each edge colour class in $\mathcal{E}$ in the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, there are as many RCON and RCOR models with variables V as there are coloured graphs with vertex set V, i.e. $|\mathcal{S}_V^+| = |\mathcal{R}_V^+| = |\mathcal{C}_V|$. Given that the number of graph colourings of a given graph $G = (V, E)$ is given by the product $|P(V)||P(E)|$ of the number of partitions of V multiplied by the number of partitions of E, we obtain

$$|\mathcal{C}_V| = \sum_{E \subseteq V_{\neq}^{(2)}} |\mathcal{P}(V)||\mathcal{P}(E)| = |\mathcal{P}(V)| \sum_{E \subseteq V_{\neq}^{(2)}} |\mathcal{P}(E)|. \tag{5.8}$$

For a discrete set S of size d, $|\mathcal{P}(S)|$ is given by the d^{th} *Bell number* B_d (Bell, 1934; Pitman, 1997) which satisfy the recursive relationship

$$B_{d+1} = \sum_{k=0}^{d} \binom{d}{k} B_k \tag{5.9}$$

so that

$$|\mathcal{C}_V| = B_{|V|} \sum_{E \subseteq V^{(2)}_{\neq}} B_{|E|} = B_{|V|} \sum_{k=0}^{\binom{|V|}{2}} \binom{\binom{|V|}{2}}{k} B_k = B_{|V|} B_{\binom{|V|}{2}+1}.$$

We give the first 11 Bell numbers in Table 5.1 together with $|\mathcal{C}_{[4]}|$ and $|\mathcal{C}_{[5]}|$.

d	0	1	2	3	4	5	6	7	8	9	10	11
B_d	1	1	2	5	15	52	203	877	4,140	21,147	115,975	678,570

Table 5.1.: Bell numbers B_d for $0 \leq d \leq 11$.

$$|\mathcal{C}_{[4]}| = B_4 B_7 = 13,155 \qquad |\mathcal{C}_{[5]}| = B_5 B_{11} = 35,285,640$$

For comparison,

$$|U_{[4]}| = 2^{\binom{4}{2}} = 2^6 = 64 \qquad |U_{[5]}| = 2^{\binom{5}{2}} = 2^{10} = 1,024.$$

Berend and Tassa (2010) show that for all $\epsilon > 0$ there exists a d_ϵ such that for $d > d_\epsilon$,

$$\left(\frac{d}{e \log d}\right)^d < B_d < \left(\frac{d}{e^{1-\epsilon} \log d}\right)^d$$

giving that asymptotically $|\mathcal{C}_V|$ grows faster than $\left(\frac{|V|}{e \log |V|}\right)^{|V|} \left(\frac{\binom{|V|}{2}}{e \log \binom{|V|}{2}}\right)^{\binom{|V|}{2}}$.

5.4. Model Classes within RCON and RCOR Models

In the previous chapters three graph colouring regularities have been identified which correspond to desirable statistical properties in the represented RCON and RCOR models. This section is devoted to the identification of structural relations between the model classes and, most importantly, their structures. We begin by (re)stating their definitions.

5.4.1. Definitions

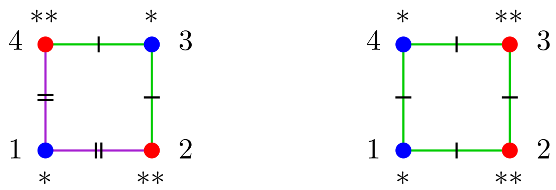
Models Represented by Edge Regular Colourings

Edge regular colourings were defined in Definition 3.4 in Section 3.4 as colourings in which every pair of equally coloured edges connects the same vertex colour classes. Employing the notion of a graph component, subsequently introduced in Section 4.3.2, we give an equivalent definition below.

Definition 5.2 For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, $(\mathcal{V}, \mathcal{E})$ is edge regular if every $u \in \mathcal{E}$ appears in precisely one $u\{vw\}$-component of $\mathcal{G}$ for $v, w \in \mathcal{V}$. The set of edge regular colourings in $\mathcal{C}_V$ is denoted $\mathcal{B}_V$.

$\square$

While the colouring in Figure 5.3(a) is edge regular (both green edges and both purple edges connect a red vertex to a blue one), the colouring in Figure 5.3(b) is not, as the green edges appear between different pairs of vertex colours.



(a) $(\mathcal{V}, \mathcal{E}) \in \mathcal{B}_V$ (b) $(\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V \setminus \mathcal{B}_V$

Figure 5.3.: An edge regular colouring and one which is not edge regular.

In the above notation, Proposition 3.1 in Section 3.4 states that for $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, $\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ if and only if $(\mathcal{V}, \mathcal{E}) \in \mathcal{B}_V$.

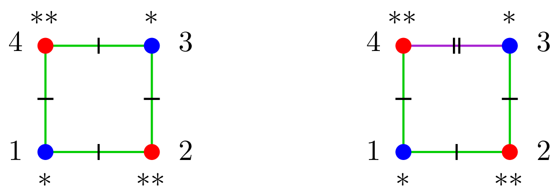
Models Represented by Vertex Regular Colourings

Vertex regular colourings were defined in Definition 4.3 in Section 4.4.2.

Definition 5.3 For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, we say that $(\mathcal{V}, \mathcal{E})$ is *vertex regular* if $\mathcal{V}$ is an equitable partition with respect to $G^u = (V, u)$ for all $u \in \mathcal{E}$. The set of vertex regular colourings in $\mathcal{C}_V$ is denoted $\mathcal{P}_V$.

□

While the colouring in Figure 5.4(a) is vertex regular, the colouring in Figure 5.4(b) is not. The former has only one edge colour class, so that it is vertex regular if and only if its vertex colouring is equitable with respect to G, which it is. The colouring in Figure 5.4(b) cannot be vertex regular as while vertex 4 is incident to a purple edge, vertex 2 isn't even though they are of the same colour.



(a) $(\mathcal{V}, \mathcal{E}) \in \mathcal{P}_V$ (b) $(\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V \setminus \mathcal{P}_V$

Figure 5.4.: A vertex regular colouring and one which is not vertex regular.

By Theorem 4.3 in Section 4.4.3, for $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, $\Omega(\mathcal{V}) \in \Phi(\mathcal{G})$ if and only if $(\mathcal{V}, \mathcal{E}) \in \mathcal{P}_V$.

Models Represented by Regular Colourings

Following Siemons (1983), we define colourings which are both vertex regular and edge regular to be *regular*.

Definition 5.4 For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ we say that $(\mathcal{V}, \mathcal{E})$ is *regular* if $(\mathcal{V}, \mathcal{E}) \in \mathcal{R}_V = \mathcal{B}_V \cap \mathcal{P}_V$.

□

Thus regular colourings $(\mathcal{V}, \mathcal{E})$ are the only colourings which represent models which satisfy $\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ and $\Omega(\mathcal{V}) \in \Phi(\mathcal{G})$.

Proposition 5.7 *For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, $(\mathcal{V}, \mathcal{E})$ lies inside $\mathcal{R}_V$ if and only if*

(i) all equally coloured edges connect the same vertex colour classes

(ii) all equally coloured vertices have the same degree in every edge colour class.

□

Proof: By definition of $\mathcal{B}_V$, $(\mathcal{V}, \mathcal{E}) \in \mathcal{B}_V$ is equivalent to $(\mathcal{V}, \mathcal{E})$ satisfying condition (i) above. For condition (ii), by definition of $\mathcal{P}_V$, we have $(\mathcal{V}, \mathcal{E}) \in \mathcal{P}_V$ if and only if $\mathcal{V}$ is equitable with respect to $G^u = (V, u)$ for all $u \in \mathcal{E}$. For graph colourings in $\mathcal{B}_V$, G^u only contains edges between one pair of vertex colours or within one vertex colour class, so that $\mathcal{V}$ is equitable with respect to G^u for all $u \in \mathcal{E}$ if and only if every pair of equally coloured vertices has the same degree in every edge colour class, which is precisely condition (ii).

■

By the above, the colourings in Figure 5.3(b) and Figure 5.4(b) cannot be regular. While the colouring in Figure 5.4(a) is regular, the colouring in Figure 5.3(a) is not.

Models Represented by Permutation Generated Colourings

Definition 3.6 in Section 3.4.1 defined permutation generated colourings.

Definition 5.5 Let $\Gamma \leq S(V)$ and let $\mathcal{V}_\Gamma$ and $\mathcal{E}_\Gamma$ denote the vertex orbit partition of V and edge orbit partition of $V_{\neq}^{(2)}$ generated by $\tilde{\Gamma}$ respectively. A graph colouring $(\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ is *permutation generated with generating group* Γ if $\mathcal{V} = \mathcal{V}_\Gamma$ and $\mathcal{E} \subseteq \mathcal{E}_\Gamma$. The set of permutation generated colourings in $\mathcal{C}_V$ is denoted Π_V.

□

Among the graph colourings in Figure 5.3 and 5.4, only the colouring in Figure 5.4(a) is permutation generated, with generating group $\Gamma = \langle(13)(24)\rangle$.

Permutation generated colourings represent RCOP models and possess all desirable properties discussed thus far.

5.4.2. Structural Relations

Relations between the colouring classes $\mathcal{B}_V$, $\mathcal{P}_V$, $\mathcal{R}_V$ and Π_V are summarised in the diagram in Figure 5.5 below.

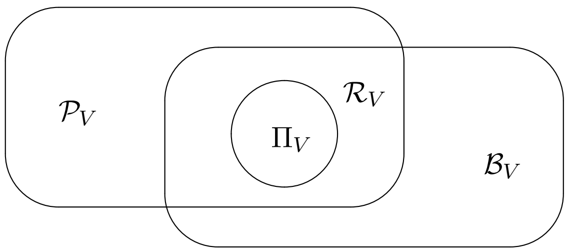


Figure 5.5.: Structural relations between colouring classes.

As $\mathcal{R}_V = \mathcal{B}_V \cap \mathcal{P}_V$ by definition, trivially $\mathcal{R}_V \subseteq \mathcal{B}_V$ and $\mathcal{R}_V \subseteq \mathcal{P}_V$. The colouring displayed in Figure 5.3(b) lies in $\mathcal{P}_{[4]} \setminus \mathcal{R}_{[4]}$ and the colouring in Figure 5.4(b) lies in $\mathcal{B}_{[4]} \setminus \mathcal{R}_{[4]}$, so both inclusions may be strict. That $\Pi_V \subseteq \mathcal{R}_V$ was established in the proof of Lemma 4.10 in Section 4.4.3.

The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ in Figure 5.6 establishes that $\mathcal{R}_V \setminus \Pi_V$ may be non-empty. $\mathcal{G}$ clearly has a regular colouring $(\mathcal{V}, \mathcal{E}) \in \mathcal{R}_{[11]}$. Suppose we also have $(\mathcal{V}, \mathcal{E}) \in \Pi_{[11]}$.

Then there exists a permutation group $\Gamma \leq \mathrm{Aut}(G)$ with vertex and edge orbits given by $\mathcal{V}$ and $\mathcal{E}$ respectively. Thus there exists a permutation $\sigma \in \Gamma$ preserving $(\mathcal{V}, \mathcal{E})$ which maps vertex 1 to vertex 2 as they have the same colour in $\mathcal{V}$. If 1 is mapped to 2, the set $\{4, 5\}$ must be mapped to $\{6, 7\}$. As 4 and 5 are both incident with vertex 10, 10 must be mapped to a vertex incident to both 6 and 7. However such a vertex does not exist, so $(\mathcal{V}, \mathcal{E}) \notin \Pi_{[11]}$.

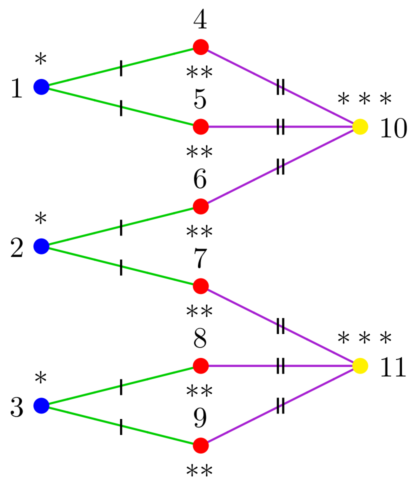


Figure 5.6.: A graph with colouring in $\mathcal{R}_{[11]} \setminus \Pi_{[11]}$.

For illustration we give the sizes of $U_{[4]}$, $\mathcal{C}_{[4]}$, $\mathcal{B}_{[4]}$, $\mathcal{R}_{[4]}$, $\mathcal{P}_{[4]}$ and $\Pi_{[4]}$ in Table 5.2. How the four last mentioned sizes can be obtained is described in Appendix B.

Set	$U_{[4]}$	$\mathcal{C}_{[4]}$	$\mathcal{B}_{[4]}$	$\mathcal{R}_{[4]}$	$\mathcal{P}_{[4]}$	$\Pi_{[4]}$
Size	64	13,155	3,065	251	1,380	251

Table 5.2.: Model set sizes for $V = [4]$.

The relative sizes in Table 5.2 are representative of the general case: $\mathcal{B}_V$ is the largest model class, followed by $\mathcal{P}_V$ and $\mathcal{R}_V$. $|\mathcal{P}_V|$ will generally be considerably smaller than $|\mathcal{B}_V|$ as the defining conditions of $\mathcal{P}_V$ are far more restrictive than those of $\mathcal{B}_V$. Π_V is the smallest class of the four for all V, however may equal $\mathcal{R}_V$ for some V, as for example for $V = [4]$.

5.4.3. Structures

While the structure of each colouring class requires individual treatment, some properties are shared by the classes and can be obtained simultaneously for all of them. Below we provide a brief overview of the results we will show in this section and obtain as many of them as we can in a unified manner. We will show that

(a) each colouring class $\mathcal{B}_V$, $\mathcal{P}_V$, $\mathcal{R}_V$ and Π_V is a non-distributive meet-sublattice of $\langle \mathcal{C}_V; \preceq \rangle$;

(b) the respective join operations $\vee_{\mathcal{X}}$ for $\mathcal{G}, \mathcal{H} \in \mathcal{X}_V$ with $\mathcal{X} \in \{\mathcal{B}, \mathcal{P}, \mathcal{R}, \Pi\}$ are given by $\mathcal{G} \vee_{\mathcal{X}} \mathcal{H} = s_{\mathcal{X}}(\mathcal{G} \vee \mathcal{H})$ where $\vee$ is the join operation in $\mathcal{C}_V$ and the functions $s_{\mathcal{X}} : \mathcal{C}_V \to \mathcal{C}_V$ return the supremum of a given graph inside $\mathcal{X}_V$, which we specify explicitly.

Clearly each colouring class is a subposet of $\langle \mathcal{C}_V; \preceq \rangle$ with partial ordering '$\preceq$'. In order to establish (a) one needs to show that each colouring class is stable under the meet operation $\wedge$ in $\langle \mathcal{C}_V; \preceq \rangle$ and that each colouring class is non-distributive. The first claim will be shown for each colouring class individually in the following sections, while we demonstrate non-stability under $\vee$ and non-distributivity through suitable examples below.

For $\mathcal{P}_V$, $\mathcal{R}_V$ and Π_V, non-stability under $\vee$ is established by observing that while graphs $\mathcal{G}_4$ and $\mathcal{G}_5$ in Figure 5.2 lie inside Π_V, with generating groups $\Gamma_4 = \langle (13)(24) \rangle$, $\Gamma_5 = \langle (13) \rangle$ respectively, and therefore also in $\mathcal{P}_V$ and $\mathcal{R}_V$, their join $\mathcal{G}_7 = \mathcal{G}_4 \vee \mathcal{G}_5$ does not. For $\mathcal{B}_V$, the same is established by the graphs in in Figure 5.7.

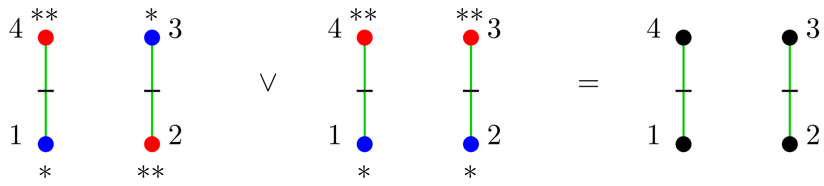


Figure 5.7.: $B_{[4]}$ is not stable under $\vee$.

Combined with the fact that each colouring class is stable under $\wedge$, the above establishes that each is a meet-sublattice of $\langle \mathcal{C}_V; \preceq \rangle$, so that by Lemma 5.1 each is a meet-sublattice of $\langle \mathcal{C}_V; \preceq \rangle$. As the zero and unit in $\langle \mathcal{C}_V; \preceq \rangle$, given by the empty graph with all vertices in one colouring class and the complete graph with singleton colour classes, lie in Π_V, with generating groups $\Gamma = S(V)$ and $\Gamma = \langle Id \rangle$ respectively, they are also the zero and unit in all four colouring classes. Note that graphs $\mathcal{G}_4$, $\mathcal{G}_5$ and $\mathcal{G}_6 = \mathcal{G}_4 \wedge \mathcal{G}_5$ in Figure 5.2 illustrate the stability of all four colouring classes under $\wedge$, with the colouring of $\mathcal{G}_6$ being generated by $\Gamma_6 = \langle (13), (24) \rangle$.

Even before identifying the individual join operations $\vee_{\mathcal{B}}$, $\vee_{\mathcal{P}}$, $\vee_{\mathcal{R}}$ and $\vee_{\Pi}$ we can establish non-distributivity of each lattice by a simple example. All graphs $\mathcal{G}_8$, $\mathcal{G}_9$, $\mathcal{G}_{10}$, $\mathcal{G}_9 \wedge \mathcal{G}_{10}$ and $\mathcal{G}_8 \vee \mathcal{G}_9 = \mathcal{G}_8 \vee \mathcal{G}_{10}$ displayed in Figure 5.8 lie in Π_V, with generating groups $\Gamma_8 = \langle (124) \rangle$, $\Gamma_9 = \langle (234) \rangle$, $\Gamma_{10} = \langle (13)(24) \rangle$, $\Gamma_{9 \wedge 10} = S([4])$ and $\Gamma_{8 \vee 9} = \Gamma_{8 \vee 10} = \langle (24) \rangle$ respectively, and therefore also in $\mathcal{B}_V$, $\mathcal{P}_V$ and $\mathcal{R}_V$.

While the individual join operations are not known at this stage, clearly if for $\mathcal{G}, \mathcal{H} \in \mathcal{B}_V$, their join $\mathcal{G} \vee \mathcal{H}$ also lies in $\mathcal{B}_V$, then $\vee = \vee_{\mathcal{B}}$ and similarly for the other colouring classes. This establishes that the $\vee$-joins of the graphs displayed in Figure 5.8 must equal the $\vee_{\mathcal{B}}$-, $\vee_{\mathcal{P}}$-, $\vee_{\mathcal{R}}$- and $\vee_{\Pi}$-joins. Now as

$$\mathcal{G}_8 = \mathcal{G}_8 \vee (\mathcal{G}_9 \wedge \mathcal{G}_{10}) \neq (\mathcal{G}_8 \vee \mathcal{G}_9) \wedge (\mathcal{G}_8 \vee \mathcal{G}_{10}) = \mathcal{G}_8 \vee \mathcal{G}_9$$

condition (5.1) for $a = \mathcal{G}_8$, $b = \mathcal{G}_9$ and $c = \mathcal{G}_{10}$ is violated, so that each colouring class is non-distributive.

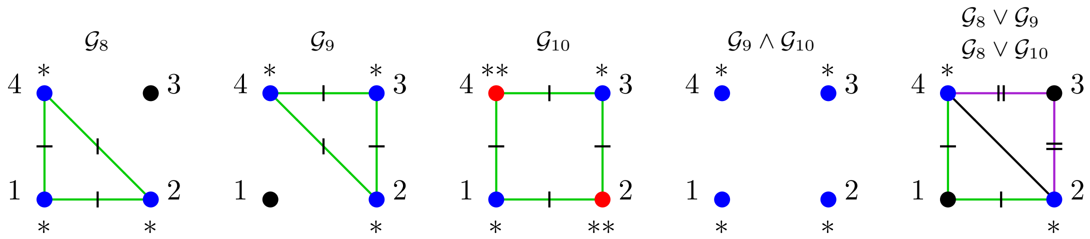


Figure 5.8.: Non-distributivity of $\langle \mathcal{B}_V; \preceq \rangle$, $\langle \mathcal{P}_V; \preceq \rangle$, $\langle \mathcal{R}_V; \preceq \rangle$ and $\langle \Pi_V; \preceq \rangle$.

For (b), let $\mathcal{X} \in \{\mathcal{B}, \mathcal{P}, \mathcal{R}, \Pi\}$ and let $\mathcal{G}, \mathcal{H} \in \mathcal{X}_V$. By the above we know $\mathcal{X}_V$ to be a lattice so that the join operation $\vee_{\mathcal{X}}$ is certain to exist. By definition, $\mathcal{K} = \mathcal{G} \vee_{\mathcal{X}} \mathcal{H}$ is the smallest graph which satisfies $\mathcal{K} \succeq \mathcal{G}$, $\mathcal{K} \succeq \mathcal{H}$ and $\mathcal{K} \in \mathcal{X}_V$. The supremum of $\mathcal{G}$ and $\mathcal{H}$ in $\langle \mathcal{C}_V; \preceq \rangle$ is the smallest graph satisfying the first two relations so that $\mathcal{K}$ is in fact the smallest graph satisfying $\mathcal{K} \succeq (\mathcal{G} \vee \mathcal{H})$ and $\mathcal{K} \in \mathcal{X}_V$. Thus the join operation in $\mathcal{X}_V$ is given by

$$\mathcal{G} \vee_{\mathcal{X}} \mathcal{H} = s_{\mathcal{X}}(\mathcal{G} \vee \mathcal{H})$$

where $s_{\mathcal{X}} : \mathcal{C}_V \to \mathcal{C}_V$ for $\mathcal{G} \in \mathcal{C}_V$ returns the smallest graph which lies in $\mathcal{X}_V$ and is at least as large as $\mathcal{G}$ with respect to $\preceq$.

A vertex and edge coloured graph is enlarged with respect to the partial ordering $\preceq$ by adding edge colour classes and splitting vertex and edge colour classes. As no colouring criterion can be satisfied through adding new edges, $s_{\mathcal{B}}$, $s_{\mathcal{P}}$, $s_{\mathcal{R}}$ and s_{Π} only operate on the colouring of a graph through splitting colour classes.

While the main result we establish is that each colouring class forms a lattice, the functions $s_{\mathcal{B}}$, $s_{\mathcal{P}}$, $s_{\mathcal{R}}$ and s_{Π} are also of interest on their own. They can be used to test whether a given graph colouring falls into a desired colouring class, which it does whenever its supremum in the class is given by the colouring itself. We have implemented these functions in our computer program `GraphCheck`, which exploits Brendan McKay's computer program `nauty` (McKay, 2010). We briefly describe it at the end of the section and provide sample output in Appendix C.

Interestingly, while every graph falling outside a given colouring class has a supremum inside the class, this is not true for the infimum which we demonstrate by examples for each colouring class.

Edge Regular Colourings $\mathcal{B}_V$

Proposition 5.8 *$\mathcal{B}_V$ is stable under the meet operation $\wedge$ in $\langle \mathcal{C}_V; \preceq \rangle$ in (5.6).*

□

Proof: Let $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}), \mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \in \mathcal{B}_V$. $\mathcal{G} \wedge \mathcal{H} = (\mathcal{V}_{\mathcal{G}\wedge\mathcal{H}}, \mathcal{E}_{\mathcal{G}\wedge\mathcal{H}})$ is obtained from $\mathcal{G}$ and $\mathcal{H}$ by dropping of edge colour classes and merging of colour classes. The only operation potentially leading to $\mathcal{G} \wedge \mathcal{H}$ lying outside of $\mathcal{B}_V$ is merging of edge colour classes.

So let $\alpha\beta$ and $\gamma\delta$ be two edges in $\mathcal{G} \wedge \mathcal{H}$ of equal colour. Then, by definition of $\wedge$ in $\langle \mathcal{C}_V; \preceq \rangle$ in (5.6) and of $\vee$ in partition spaces in (5.3), there exists a sequence $\alpha_0\beta_0, \ldots, \alpha_k\beta_k$ in $E_{\mathcal{G}\wedge\mathcal{H}}$ such that $\alpha_0\beta_0 = \alpha\beta$, $\alpha_k\beta_k = \gamma\delta$ and $\alpha_{i-1}\beta_{i-1} \equiv \alpha_i\beta_i$ $(\mathcal{E}_\mathcal{G})$ or $\alpha_{i-1}\beta_{i-1} \equiv \alpha_i\beta_i$ $(\mathcal{E}_\mathcal{H})$ for $1 \leq i \leq k$. As both $\mathcal{G}$ and $\mathcal{H}$ have edge regular colourings, $\alpha_{i-1}\beta_{i-1}$ and $\alpha_i\beta_i$ connect the same vertex colour classes in the graph in which they are of equal colour, which we denote by $\{\alpha_{i-1}, \beta_{i-1}\} \equiv \{\alpha_i, \beta_i\}$ $(\mathcal{V}_\mathcal{X})$ with $\mathcal{X} \in \{\mathcal{G}, \mathcal{H}\}$. By (5.3) this gives that $\{\alpha, \beta\} = \{\alpha_0, \beta_0\} \equiv \{\alpha_k, \beta_k\} = \{\gamma, \delta\}$ $(\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H})$. As $\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H} = \mathcal{V}_{\mathcal{G}\wedge\mathcal{H}}$, $\alpha\beta$ and $\gamma\delta$ connect the same vertex colour classes in $\mathcal{G} \wedge \mathcal{H}$.

■

Proposition 5.9 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ and let $\mathcal{E}_\mathcal{B}$ be the partition of E given by $\alpha\beta \equiv \gamma\delta$ $(\mathcal{E}_\mathcal{B})$ if and only if $\alpha\beta$ and $\gamma\delta$ lie in the same graph component in $\mathcal{G}$ for $\alpha\beta, \gamma\delta \in E$. Then $\mathcal{G}$ has a supremum in $\mathcal{B}_V$ with colouring*

$$s_\mathcal{B}(\mathcal{G}) = (\mathcal{V}, \mathcal{E} \wedge \mathcal{E}_\mathcal{B}). \qquad \square$$

Proof: The claim is trivially true for $(\mathcal{V}, \mathcal{E}) \in \mathcal{B}_V$. If $(\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V \backslash \mathcal{B}_V$, an edge regular colouring cannot be achieved through splitting vertex colour classes or adding edge colour classes. The only effective manipulation is therefore the splitting of edge colour classes. The coarsest partition which is finer than $\mathcal{E}$ and gives an edge regular colouring is $\mathcal{E} \wedge \mathcal{E}_\mathcal{B}$, as it splits edge colour classes across graph components of $\mathcal{G}$. ■

Note that $s_\mathcal{B}(\mathcal{G})$ gives precisely the colouring denoted by $(\mathcal{V}, \mathcal{E}^*)$ in Corollary 4.2 in Section 4.3.3. Combining Proposition 5.8 and Proposition 5.9 with the introduction:

Theorem 5.2 *$\langle \mathcal{B}_V; \preceq \rangle$ is a non-distributive lattice with meet operation $\wedge$ inherited from $\langle \mathcal{C}_V; \preceq \rangle$, and join operation $\vee_\mathcal{B}$ given by*

$$\mathcal{G} \vee_\mathcal{B} \mathcal{H} = s_\mathcal{B}(\mathcal{G} \vee \mathcal{H}), \qquad \mathcal{G}, \mathcal{H} \in \mathcal{B}_V.$$

$\square$

It may be interesting to note that while by Proposition 5.9 every graph colouring inside $\mathcal{C}_V \setminus \mathcal{B}_V$ has a supremum in $\mathcal{B}_V$, this is not true for the infimum. The graph $\mathcal{G}$ in Figure 5.9 does not have an edge regular colouring (as black vertices stand for singleton colour classes), while both $\mathcal{H} \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{G}$ do. They are largest smaller graphs as there exist no $\mathcal{H}', \mathcal{K}' \in \mathcal{B}_V$ with $\mathcal{H} \prec \mathcal{H}' \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{K}' \prec \mathcal{G}$.

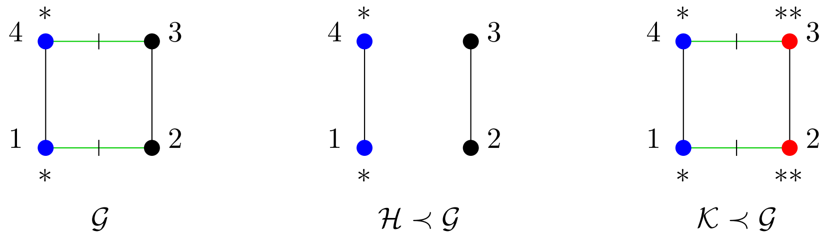


Figure 5.9.: Not every graph in $\mathcal{C}_{[4]} \setminus \mathcal{B}_{[4]}$ has an infimum in $\mathcal{B}_{[4]}$.

Regular Colourings $\mathcal{R}_V$

We treat regular colourings before vertex regular colourings for reasons to become apparent in the text. The following result allows us to exploit the isomorphism between $\mathcal{C}_V$ and $\mathcal{F}_V$ established in Proposition 5.6.

Lemma 5.4 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ and let $\phi_V(\mathcal{G}) = \mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, \mathcal{E}_F) \in \mathcal{F}_V$ be the corresponding coloured factor graph. Then*

$$(\mathcal{V}, \mathcal{E}) \in \mathcal{R}_V \qquad \Longleftrightarrow \qquad (\mathcal{V} \cup \mathcal{N}) \text{ is equitable with respect to } G_F. \tag{5.10}$$

□

Proof: First note that factor graphs by construction never contain edges between two nodes or between two vertices, and that a vertex never has the same colour as a node. This implies that $\mathcal{V} \cup \mathcal{N}$ is equitable with respect to G_F if and only if

(i) nodes of the same colour in $\mathcal{N}$ always have equally many neighbours in every vertex colour class in $\mathcal{V}$, and

(ii) vertices of the same colour in $\mathcal{V}$ always have equally many neighbours in every node colour class in $\mathcal{N}$.

Observing that conditions (i) and (ii) above are equivalent to conditions (i) and (ii) in Proposition 5.7 for $(\mathcal{V}, \mathcal{E})$ completes the proof. ■

We will further require the following results (McKay, 1976):

Lemma 5.5 (McKay) *Let $G = (V, E)$ be an undirected graph and let $\mathcal{V}$ be a partition of V. There is a unique coarsest partition that is finer than $\mathcal{V}$ and equitable with respect to G. We denote this partition by $r_G(\mathcal{V})$ and call it* coarsest equitable refinement of $\mathcal{V}$ with respect to G.

□

Lemma 5.6 (McKay) *Let $G = (V, E)$ be an undirected graph and let $\mathcal{V}_1$ and $\mathcal{V}_2$ be two partitions of V both equitable with respect to G. Then the join of the two partitions $\mathcal{V}_1 \vee \mathcal{V}_2$ is also equitable with respect to G.*

□

This allows us to prove the following:

Proposition 5.10 *$\mathcal{R}_V$ is stable under the meet operation $\wedge$ in $\langle \mathcal{C}_V; \preceq \rangle$ in (5.6).*

□

Proof: Let $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}), \mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \in \mathcal{R}_V$. $\mathcal{G} \wedge \mathcal{H}$ is obtained from $\mathcal{G}$ and $\mathcal{H}$ by dropping edge colour classes and merging colour classes. Dropping edge colour classes cannot lead to $\mathcal{G} \wedge \mathcal{H}$ lying outside of $\mathcal{R}_V$, so we only need to show that colour classes are merged in a permissible way.

So let E^* be the edge set of $\mathcal{G} \wedge \mathcal{H}$, $\mathcal{E}^*_\mathcal{G}$ and $\mathcal{E}^*_\mathcal{H}$ its colourings in $\mathcal{E}_\mathcal{G}$ and $\mathcal{E}_\mathcal{H}$, and let $\mathcal{G}^* = (\mathcal{V}_\mathcal{G}, \mathcal{E}^*_\mathcal{G})$ and $\mathcal{H}^* = (\mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{H})$. Then $\mathcal{G}^*, \mathcal{H}^* \in \mathcal{R}_V$ and $\mathcal{G} \wedge \mathcal{H} = \mathcal{G}^* \wedge \mathcal{H}^* = (\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H})$, so that in order to prove the claim we need to show that $(\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}) \in \mathcal{R}_V$.

As $\mathcal{G}^*, \mathcal{H}^* \in \mathcal{R}_V$, by relation (5.10), $\phi_V(\mathcal{G}^*) = \mathcal{G}^*_F = (\mathcal{V}_\mathcal{G} \cup \mathcal{N}^*_\mathcal{G}, E^*_F) \in \mathcal{F}_V$ and $\phi_V(\mathcal{H}^*) = \mathcal{H}^*_F = (\mathcal{V}_\mathcal{H} \cup \mathcal{N}^*_\mathcal{H}, E^*_F) \in \mathcal{F}_V$ both have equitable vertex colourings with respect to $G^*_F = H^*_F$. The join of the two colourings satisfies

$$(\mathcal{V}_{\mathcal{G}} \cup \mathcal{N}^*_{\mathcal{G}}) \vee (\mathcal{V}_{\mathcal{H}} \cup \mathcal{N}^*_{\mathcal{H}}) = (\mathcal{V}_{\mathcal{G}} \vee \mathcal{V}_{\mathcal{H}}) \cup (\mathcal{N}^*_{\mathcal{G}} \vee \mathcal{N}^*_{\mathcal{H}}) \tag{5.11}$$

as elements in $\mathcal{V}$ and $\mathcal{N}$ are never of the same colour. Combining relations (5.10) and (5.11) with Lemma 5.6, we obtain

$$\phi_V^{-1}(((\mathcal{V}_{\mathcal{G}} \cup \mathcal{N}^*_{\mathcal{G}}) \vee (\mathcal{V}_{\mathcal{H}} \cup \mathcal{N}^*_{\mathcal{H}}), E^*_F)) = (\mathcal{V}_{\mathcal{G}} \vee \mathcal{V}_{\mathcal{H}}, \mathcal{E}^*_{\mathcal{G}} \vee \mathcal{E}^*_{\mathcal{H}}) = \mathcal{G}^* \wedge \mathcal{H}^* = \mathcal{G} \wedge \mathcal{H} \in \mathcal{R}_V.$$

■

Proposition 5.11 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ and let $\phi_V(\mathcal{G}) = \mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, \mathcal{E}_F)$ be the corresponding coloured factor graph. Then $\mathcal{G}$ has a supremum in $\mathcal{R}_V$ with colouring*

$$s_{\mathcal{R}}(\mathcal{G}) = \phi_V^{-1}((r_{G_F}(\mathcal{V} \cup \mathcal{N}), E_F))$$

where $\phi_V : \mathcal{C}_V \longrightarrow \mathcal{F}_V$ is the isomorphism defined in the proof of Proposition 5.6.

□

Proof: The claim is trivially true if $(\mathcal{V}, \mathcal{E}) \in \mathcal{R}_V$, as then $r_{G_F}(\mathcal{V} \cup \mathcal{N}) = (\mathcal{V} \cup \mathcal{N})$ and $\phi_V^{-1}((r_{G_F}(\mathcal{V} \cup \mathcal{N}), E_F)) = (\mathcal{V}, \mathcal{E})$. For $(\mathcal{V}, \mathcal{E}) \notin \mathcal{R}_V$, the operations which give a bigger graph are adding edge colour classes and splitting colour classes. If $(\mathcal{V}, \mathcal{E})$ is not regular, it doesn't satisfy at least one of the conditions (i) and (ii) in Proposition 5.7. As adding edge colour classes does not help satisfy either, the supremum of $\mathcal{G}$ in $\mathcal{R}_V$, if it exists, must have the same edge set as $\mathcal{G}$ and a finer colouring.

We consider the graph $\phi_V^{-1}((r_{G_F}(\mathcal{V} \cup \mathcal{N}), E_F))$ given in the claim. Firstly, by definition of r_{G_F}, the coloured factor graph $(r_{G_F}(\mathcal{V} \cup \mathcal{N}), E_F)$ has an equitable vertex and node colouring, which by Lemma 5.4 implies that $\phi^{-1}((r_{G_F}(\mathcal{V} \cup \mathcal{N}), E_F))$ is indeed regular. Further, it certainly is larger than $\mathcal{G}$ as the colouring of $\mathcal{G}_F$ is refined by r_{G_F}, thus leading to a larger graph than the original. As $r_{G_F}(\mathcal{V} \cup \mathcal{N})$ is the coarsest refinement of $(\mathcal{V} \cup \mathcal{N})$ which is equitable, $\phi_V^{-1}((r_{G_F}(\mathcal{V} \cup \mathcal{N}), E_F))$ must be the smallest graph in $\mathcal{R}_V$ which is larger than $\mathcal{G}$.

■

Combining Proposition 5.10 and Proposition 5.11 with the introduction:

Theorem 5.3 *$\langle \mathcal{R}_V; \preceq \rangle$ is a non-distributive lattice with meet operation $\wedge$ inherited from $\langle \mathcal{C}_V; \preceq \rangle$, and join operation $\vee_{\mathcal{R}}$ given by*

$$\mathcal{G} \vee_{\mathcal{R}} \mathcal{H} = s_{\mathcal{R}}(\mathcal{G} \vee \mathcal{H}), \qquad \mathcal{G}, \mathcal{H} \in \mathcal{R}_V.$$

□

Just as in $\langle \mathcal{B}_V; \preceq \rangle$, graph colourings outside $\mathcal{R}_V$ do not necessarily have an infimum in $\mathcal{R}_V$, as illustrated in Figure 5.10. While $\mathcal{G}$ does not have a regular colouring (as green edges appear between blue and red vertices, and between red vertices and a black vertex), both $\mathcal{H} \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{G}$ do. There exist no $\mathcal{H}', \mathcal{K}' \in \mathcal{R}_V$ with $\mathcal{H} \prec \mathcal{H}' \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{K}' \prec \mathcal{G}$.

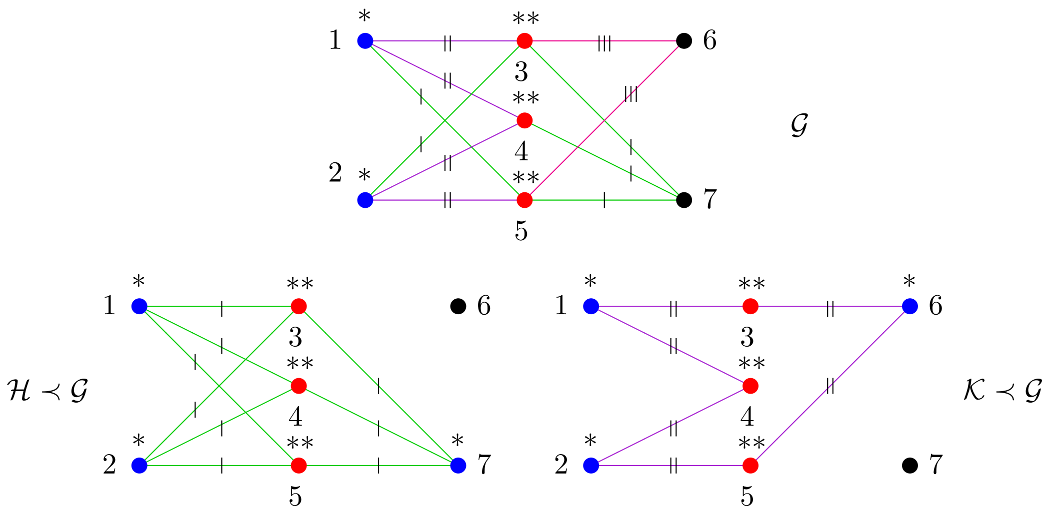


Figure 5.10.: Not every graph in $\mathcal{C}_{[7]} \setminus \mathcal{R}_{[7]}$ has an infimum in $\mathcal{R}_{[7]}$.

Vertex Regular Colourings $\mathcal{P}_V$

Proposition 5.12 *$\mathcal{P}_V$ is stable under the meet operation $\wedge$ in $\langle \mathcal{C}_V; \preceq \rangle$ in (5.6).*

□

Proof: Let $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G}), \mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \in \mathcal{P}_V$. Let $E_{\mathcal{G}\wedge\mathcal{H}}$ denote the edge set of $\mathcal{G}\wedge\mathcal{H}$ and let $\mathcal{E}^*_\mathcal{G}$ and $\mathcal{E}^*_\mathcal{H}$ be its colourings in $\mathcal{E}_\mathcal{G}$ and $\mathcal{E}_\mathcal{H}$ respectively. By definition of the meet operation in $\langle \mathcal{C}_V; \preceq\rangle$, $\mathcal{G} \wedge \mathcal{H} = (\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}, \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H})$ so that the edge colouring of $\mathcal{G} \wedge \mathcal{H}$ is obtained from $\mathcal{E}_\mathcal{G}$ and $\mathcal{E}_\mathcal{H}$ by dropping and merging of edge colour classes.

Since $\mathcal{V}_\mathcal{G}$ is equitable with respect to $G^u = (V, u)$ for all $u \in \mathcal{E}_\mathcal{G}$, it must also be equitable with respect to $G^{u^*} = (V, u^*)$ for all $u^* \in \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}$. Similarly, $\mathcal{V}_\mathcal{H}$ is equitable with respect to $G^{u^*} = (V, u^*)$ for all $u^* \in \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}$. Applying Lemma 5.6 gives that then also $\mathcal{V}_\mathcal{G} \vee \mathcal{V}_\mathcal{H}$ is equitable with respect to $G^{u^*} = (V, u^*)$ for all $u^* \in \mathcal{E}^*_\mathcal{G} \vee \mathcal{E}^*_\mathcal{H}$, giving $\mathcal{G} \wedge \mathcal{H} \in \mathcal{P}_V$ as claimed.

■

Lemma 5.7 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ and let $s_\mathcal{R}(\mathcal{G}) = (\mathcal{V}_\mathcal{R}, \mathcal{E}_\mathcal{R})$. Then for all $\mathcal{G}' = (\mathcal{V}', \mathcal{E}) \in \mathcal{C}_V$ with $\mathcal{V}_\mathcal{R} \le \mathcal{V}' \le \mathcal{V}$, $s_\mathcal{R}(\mathcal{G}') = (\mathcal{V}_\mathcal{R}, \mathcal{E}_\mathcal{R})$.*

□

Proof: For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, $s_\mathcal{R}(\mathcal{G})$ finds the unique coarsest refinement $(\mathcal{V}_\mathcal{R} \cup \mathcal{N}_\mathcal{R})$ of $(\mathcal{V} \cup \mathcal{N})$ which is equitable with respect to the factor graph $G_F = (V \cup N, E_F)$ of $\mathcal{G}$ and returns the corresponding graph colouring $\phi_V^{-1}((\mathcal{V}_\mathcal{R} \cup \mathcal{N}_\mathcal{R}, E_F))$. By the isomorphism between $\mathcal{C}_V$ and $\mathcal{F}_V$,

$$\begin{aligned}
\mathcal{V}_\mathcal{R} \le \mathcal{V}' \le \mathcal{V} \qquad &\iff \qquad (\mathcal{V}_\mathcal{R}, \mathcal{E}_\mathcal{R}) \succeq (\mathcal{V}_\mathcal{R}, \mathcal{E}) \succeq (\mathcal{V}', \mathcal{E}) \succeq (\mathcal{V}, \mathcal{E}) \\
&\iff \qquad (\mathcal{V}_\mathcal{R} \cup \mathcal{N}_\mathcal{R}) \le (\mathcal{V}_\mathcal{R} \cup \mathcal{N}) \le (\mathcal{V}' \cup \mathcal{N}) \le (\mathcal{V} \cup \mathcal{N}).
\end{aligned}$$

As $(\mathcal{V}_\mathcal{R} \cup \mathcal{N}_\mathcal{R})$ is the unique coarsest refinement of $(\mathcal{V} \cup \mathcal{N})$ which is equitable with respect to $G_F = (V \cup N, E_F)$, there exists no coarser partition of $V \cup N$ which is finer than $(\mathcal{V} \cup \mathcal{N})$ with the property. Thus, $(\mathcal{V}_\mathcal{R} \cup \mathcal{N}_\mathcal{R})$ is also the unique coarsest refinement of $(\mathcal{V}' \cup \mathcal{N})$ equitable with respect to G_F, giving $s_\mathcal{R}(\mathcal{G}') = (\mathcal{V}_\mathcal{R}, \mathcal{E}_\mathcal{R})$.

■

Proposition 5.13 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$ and let $s_{\mathcal{R}}(\mathcal{G}) = (\mathcal{V}_{\mathcal{R}}, \mathcal{E}_{\mathcal{R}})$. Then $\mathcal{G}$ has a unique supremum in $\mathcal{P}_V$ with colouring*

$$s_{\mathcal{P}}(\mathcal{G}) = (\mathcal{V}_{\mathcal{R}}, \mathcal{E}).$$

□

Proof: We first show that $(\mathcal{V}_{\mathcal{R}}, \mathcal{E}) \in \mathcal{P}_V$ and that $(\mathcal{V}, \mathcal{E}) \preceq (\mathcal{V}_{\mathcal{R}}, \mathcal{E})$. As $\mathcal{R}_V \subseteq \mathcal{P}_V$, by definition of $s_{\mathcal{R}}(\cdot)$, $s_{\mathcal{R}}(\mathcal{G}) = (\mathcal{V}_{\mathcal{R}}, \mathcal{E}_{\mathcal{R}}) \in \mathcal{P}_V$. As $\mathcal{E}_{\mathcal{R}} \leq \mathcal{E}$ and $\mathcal{P}_V$ is stable under the operation of merging edge colour classes, we obtain $(\mathcal{V}_{\mathcal{R}}, \mathcal{E}) \in \mathcal{P}_V$. Further, as $\mathcal{V}_{\mathcal{R}} \leq \mathcal{V}$, clearly $(\mathcal{V}, \mathcal{E}) \preceq (\mathcal{V}_{\mathcal{R}}, \mathcal{E})$.

It remains to be shown that $(\mathcal{V}_{\mathcal{R}}, \mathcal{E})$ is the smallest graph with these properties. So suppose this was false and there exists $\mathcal{G}' = (\mathcal{V}', \mathcal{E}') \in \mathcal{P}_V$ with $(\mathcal{V}, \mathcal{E}) \prec (\mathcal{V}', \mathcal{E}') \prec (\mathcal{V}_{\mathcal{R}}, \mathcal{E})$. Then clearly $\mathcal{E}' = \mathcal{E}$ and $\mathcal{G}' = (\mathcal{V}', \mathcal{E})$ with $\mathcal{V}_{\mathcal{R}} < \mathcal{V}' < \mathcal{V}$. Then, as $(\mathcal{V}', \mathcal{E}) \in \mathcal{P}_V$ and by definition of $\mathcal{B}_V$, $\mathcal{P}_V$ and $\mathcal{R}_V$, we have

$$s_{\mathcal{R}}(\mathcal{G}') = s_{\mathcal{B}}(\mathcal{G}') = (\mathcal{V}', \mathcal{E} \wedge \mathcal{E}'_{\mathcal{B}}) \tag{5.12}$$

where $\mathcal{E}'_{\mathcal{B}}$ is a partition of E in which two edges are in the same set if and only if they lie in the same graph component in $\mathcal{G}'$. However by Lemma 5.7,

$$s_{\mathcal{R}}(\mathcal{G}') = (\mathcal{V}_{\mathcal{R}}, \mathcal{E}_{\mathcal{R}})$$

which by uniqueness of $s_{\mathcal{R}}(\mathcal{G}')$ together with (5.12) gives $\mathcal{V}' = \mathcal{V}_{\mathcal{R}}$. Thus indeed $s_{\mathcal{P}}(\mathcal{G}) = (\mathcal{V}_{\mathcal{R}}, \mathcal{E})$, as claimed.

■

Note that this proves our claim at the very end of Section 4.4.3 that for every coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, there exists a unique coarsest partition $\mathcal{M}$ of V such that $\Omega(\mathcal{M}) \in \Phi(\mathcal{G})$. This is because, by Theorem 4.3, $\Omega(\mathcal{M}) \in \Phi(\mathcal{G})$ if and only if $\mathcal{M} \leq \mathcal{V}$ and $(\mathcal{M}, \mathcal{E}) \in \mathcal{P}_V$. $\mathcal{M}$ is precisely given by $\mathcal{V}_{\mathcal{R}}$ in Proposition 5.13.

Combining Proposition 5.12 and Proposition 5.13 with the introduction:

Theorem 5.4 *$\langle \mathcal{P}_V; \preceq \rangle$ is a non-distributive lattice with meet operation $\wedge$ inherited from $\langle \mathcal{C}_V; \preceq \rangle$, and join operation $\vee_{\mathcal{P}}$ given by*

$$\mathcal{G} \vee_{\mathcal{P}} \mathcal{H} = s_{\mathcal{P}}(\mathcal{G} \vee \mathcal{H}), \qquad \mathcal{G}, \mathcal{H} \in \mathcal{P}_V.$$

$\square$

Just as in $\langle \mathcal{B}_V; \preceq \rangle$ and $\langle \mathcal{R}_V; \preceq \rangle$, graph colourings outside $\mathcal{P}_V$ do not necessarily have an infimum in $\mathcal{P}_V$, as illustrated in Figure 5.11. While $\mathcal{G}$ does not have a vertex regular colouring (as for example only one green edge is incident with vertex 4, while two green edges are incident with vertices 3 and 5), both $\mathcal{H} \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{G}$ do. There exist no $\mathcal{H}', \mathcal{K}' \in \mathcal{R}_V$ with $\mathcal{H} \prec \mathcal{H}' \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{K}' \prec \mathcal{G}$.

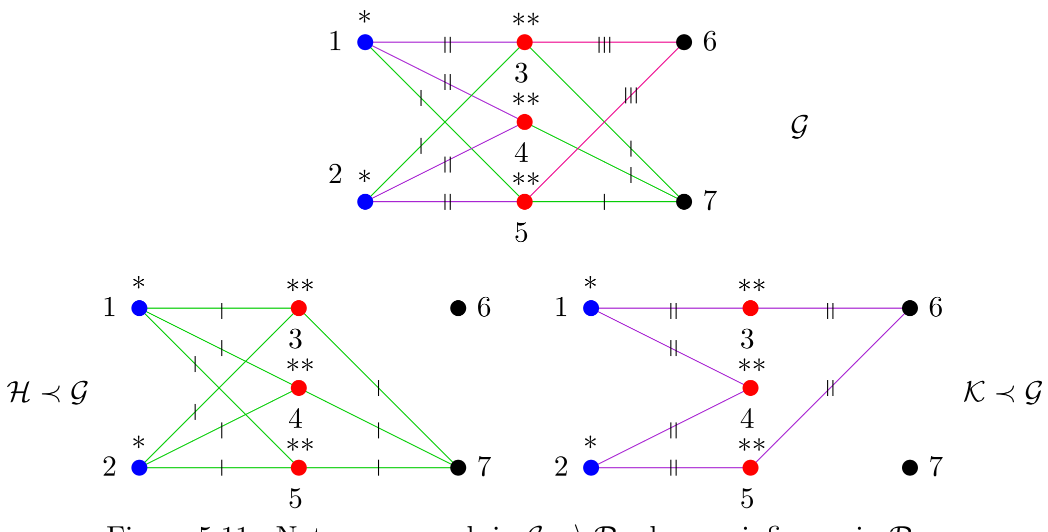


Figure 5.11.: Not every graph in $\mathcal{C}_{[7]} \setminus \mathcal{P}_{[7]}$ has an infimum in $\mathcal{P}_{[7]}$.

Permutation Generated Colourings Π_V

Understanding the structure of Π_V requires a study of the structure of $S(V)$ and the relation between a group $\Gamma \leq S(V)$ and the colouring $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \in \Pi_V$ it generates.

It is a well-established fact that for all $\Gamma \leq S(V)$ the set of subgroups of Γ forms a lattice with respect to set inclusion (Schmidt, 1994). The lattice is distributive if and only if Γ is cyclic, i.e. if and only if Γ is generated by one element only (Corollary 1.2.4 in Schmidt, 1994). For $\Gamma = S(V)$, this in particular implies:

Proposition 5.14 $\langle S(V); \leq \rangle$ *is a non-distributive lattice, with meet and join operations*

$$\Gamma_1 \wedge \Gamma_2 = \Gamma_1 \cap \Gamma_2 \qquad \Gamma_1 \vee \Gamma_2 = \langle \Gamma_1 \cup \Gamma_2 \rangle, \qquad \Gamma_1, \Gamma_2 \in S(V).$$

□

Proof: The fact that $S(V) \leq S(V)$ directly establishes that $\langle S(V); \leq \rangle$ is a lattice. It is non-distributive as there exists no $\sigma \in S(V)$ such that $S(V) = \langle \sigma \rangle$.

For $\Gamma_1, \Gamma_2 \in S(V)$, $\Gamma_1 \wedge \Gamma_2$ is the largest group in $S(V)$ satisfying $\Gamma_1 \wedge \Gamma_2 \leq \Gamma_1$ and $\Gamma_1 \wedge \Gamma_2 \leq \Gamma_2$. This implies that $\Gamma_1 \wedge \Gamma_2 \leq \Gamma_1 \cap \Gamma_2$. As $\Gamma_1 \cap \Gamma_2$ is a group itself, by maximality, $\Gamma_1 \wedge \Gamma_2 = \Gamma_1 \cap \Gamma_2$.

For the join, $\Gamma_1 \vee \Gamma_2$ is the smallest group satisfying $\Gamma_1 \leq \Gamma_1 \vee \Gamma_2$ and $\Gamma_2 \leq \Gamma_1 \vee \Gamma_2$, which implies $\Gamma_1 \cup \Gamma_2 \leq \Gamma_1 \vee \Gamma_2$. As $\Gamma_1 \vee \Gamma_2$ is a group, $\langle \Gamma_1 \cup \Gamma_2 \rangle \leq \Gamma_1 \vee \Gamma_2$. Since $\langle \Gamma_1 \cup \Gamma_2 \rangle$ is itself a group, by minimality, $\Gamma_1 \vee \Gamma_2 = \langle \Gamma_1 \cup \Gamma_2 \rangle$.

■

Unfortunately the lattice structure of $\langle S(V); \leq \rangle$ does not directly translate to Π_V. Firstly, while every group $\Gamma \leq S(V)$ generates a unique graph colouring $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$, the converse is generally false. For example, for $V = [4]$, $\Gamma_1 = \langle (1234), (12) \rangle$ and $\Gamma_2 = \langle (1234), (13) \rangle$, we have $\Gamma_1 \neq \Gamma_2$, but $\mathcal{V}_{\Gamma_1} = \mathcal{V}_{\Gamma_2} = \{V\}$ and $\mathcal{E}_{\Gamma_1} = \mathcal{E}_{\Gamma_2} = \{V^{(2)}_{\neq}\}$.

The non-trivial relationship between permutation groups and the colourings they generate has been addressed in group theory, however is still poorly understood in general. The most relevant results available to us have been established by Wielandt (1969), who found a sufficient condition for a group $\Gamma \in S(V)$ not to generate the same vertex and edge orbits as any other distinct group $\Gamma' \in S(V)$, and further showed that for all groups, edge orbits always determine vertex orbits, both given below. (More results on permutation groups and their orbits are available. However we abstain from including them, as they lack the generality we require by only applying to particular classes of groups.)

Proposition 5.15 (Wielandt) *Let $\Gamma \leq S(V)$ be such that there exists a vertex $\alpha \in V$ such that the only element in Γ which fixes α is the identity. Then there is no other group $\Gamma' \leq S(V)$ such that $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) = (\mathcal{V}_{\Gamma'}, \mathcal{E}_{\Gamma'})$.*

□

Proposition 5.16 (Wielandt) *For $\Gamma \leq S(V)$, and $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$ as above, $\mathcal{E}_\Gamma$ uniquely determines $\mathcal{V}_\Gamma$, i.e. for $\Gamma, \Gamma' \leq S(V)$, if $\mathcal{E}_\Gamma = \mathcal{E}_{\Gamma'}$ then $\mathcal{V}_\Gamma = \mathcal{V}_{\Gamma'}$.*

□

A further fact which makes the relation between $S(V)$ and Π_V non-trivial is that their orderings are not 'synchronised', not even for colourings of the complete graph which lie in Π_V. What we mean by this is made precise in Proposition 5.17.

Lemma 5.8 *If $\Gamma \leq S(V)$ and $J_\Gamma = \{g_1, \ldots, g_k\}$ is a generating set of Γ, then*

$$(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) = \bigwedge_{g \in J_\Gamma} (\mathcal{V}_{\langle g \rangle}, \mathcal{E}_{\langle g \rangle}) = \left(\bigvee_{g \in J_\Gamma} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_\Gamma} \mathcal{E}_{\langle g \rangle} \right).$$

□

Proof: Let $J_\Gamma = \{g_1, \ldots, g_k\}$. For $\alpha, \beta \in V$,

$$\begin{aligned}
\alpha \equiv \beta \ (\mathcal{V}_\Gamma) \quad &\iff \quad \exists\ \sigma \in \Gamma \text{ such that } \alpha = \sigma(\beta) \\
&\iff \quad \exists\ h_i \in \langle g_i \rangle \text{ for } 1 \le i \le k \text{ such that} \\
&\qquad\quad \alpha = \sigma(\beta) \text{ for } \sigma = h_1 h_2 \ldots h_k \\
&\iff \quad \exists\ \gamma_0, \gamma_1, \ldots, \gamma_k \in V \text{ with } \gamma_0 = \alpha,\ \gamma_k = \beta \\
&\qquad\quad \text{such that } \gamma_{i-1} \equiv \gamma_i \ (\mathcal{V}_{\langle g_i \rangle}) \text{ for } 1 \le i \le k \\
&\iff \quad \alpha \equiv \beta \ (\bigvee_{g \in J_\Gamma} \mathcal{V}_{\langle g \rangle}).
\end{aligned}$$

Similarly for $\alpha\beta, \gamma\delta \in V^{(2)}_{\neq}$,

$$\alpha\beta \equiv \gamma\delta \ (\mathcal{E}_\Gamma) \quad \iff \quad \alpha\beta \equiv \gamma\delta \ (\bigvee_{g \in J_\Gamma} \mathcal{E}_{\langle g \rangle}).$$

■

Proposition 5.17 *For $\Gamma_1, \Gamma_2 \le S(V)$,*

$$\Gamma_1 \le \Gamma_2 \quad \Longrightarrow \quad (\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) \preceq (\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}),$$

with the converse being generally false.

□

Proof: Let J_{Γ_1} and J_{Γ_2} be generating sets of Γ_1 and Γ_2 respectively. Then, by Lemma 5.8,

$$(\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}) = \left(\bigvee_{g \in J_{\Gamma_1}} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_{\Gamma_1}} \mathcal{E}_{\langle g \rangle} \right) \qquad (\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) = \left(\bigvee_{g \in J_{\Gamma_2}} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_{\Gamma_2}} \mathcal{E}_{\langle g \rangle} \right).$$

Clearly, $J_{\Gamma_1} = \Gamma_1$ and $J_{\Gamma_2} = \Gamma_2$ are valid generating sets, so that

$$\begin{aligned}
(\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) &= \left(\bigvee_{g\in\Gamma_2} \mathcal{V}_{\langle g\rangle}, \bigvee_{g\in\Gamma_2} \mathcal{E}_{\langle g\rangle} \right) \\
&= \left(\left(\bigvee_{g\in\Gamma_1} \mathcal{V}_{\langle g\rangle} \right) \vee \left(\bigvee_{g\in\Gamma_2\setminus\Gamma_1} \mathcal{V}_{\langle g\rangle} \right), \left(\bigvee_{g\in\Gamma_1} \mathcal{E}_{\langle g\rangle} \right) \vee \left(\bigvee_{g\in\Gamma_2\setminus\Gamma_1} \mathcal{E}_{\langle g\rangle} \right) \right) \\
&= (\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}) \wedge \left(\bigvee_{g\in\Gamma_2\setminus\Gamma_1} \mathcal{V}_{\langle g\rangle}, \bigvee_{g\in\Gamma_2\setminus\Gamma_1} \mathcal{E}_{\langle g\rangle} \right),
\end{aligned}$$

directly implying the result.

That the converse is generally false can be established by a simple example. Let $V = [4]$, $\Gamma_1 = \langle(12)\rangle$ and $\Gamma_2 = \langle(123)\rangle$. Then $\Gamma_1 \not\leq \Gamma_2$, however $(\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) \preceq (\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1})$, as shown in Figure 5.12.

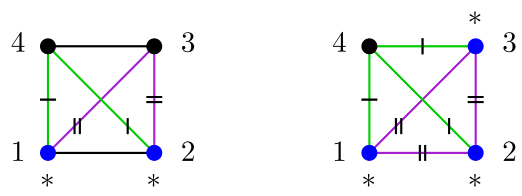


Figure 5.12.: Two permutation generated colourings.

■

Despite the non-trivial relationship between $\langle S(V); \leq\rangle$ and $\langle \Pi_V; \preceq\rangle$ we still obtain that Π_V is a non-distributive lattice. We begin with a first Lemma.

Lemma 5.9 $\mathcal{K}_V = \{(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \mid \Gamma \leq S(V)\} \subset \Pi_V$ *is stable under the meet operation* $\wedge$ *in* $\langle \mathcal{C}_V; \preceq\rangle$ *given in (5.6). In particular,*

$$(\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}) \wedge (\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) = (\mathcal{V}_{\Gamma_1\vee\Gamma_2}, \mathcal{E}_{\Gamma_1\vee\Gamma_2}).$$

□

Proof: Let $\Gamma_1, \Gamma_2 \leq S(V)$ and let J_{Γ_1} and J_{Γ_2} be generating sets of Γ_1 and Γ_2 respectively. Then, by Lemma 5.8,

$$(\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}) = \left(\bigvee_{g \in J_{\Gamma_1}} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_{\Gamma_1}} \mathcal{E}_{\langle g \rangle} \right) \qquad (\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) = \left(\bigvee_{g \in J_{\Gamma_2}} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_{\Gamma_2}} \mathcal{E}_{\langle g \rangle} \right)$$

so that

$$(\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}) \wedge (\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) = \left(\bigvee_{g \in J_{\Gamma_1} \cup J_{\Gamma_2}} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_{\Gamma_1} \cup J_{\Gamma_2}} \mathcal{E}_{\langle g \rangle} \right).$$

By Proposition 5.14, the join of Γ_1 and Γ_2 in $S(V)$ is $\langle \Gamma_1 \cup \Gamma_2 \rangle$, which contains all elements which can be written as a combination of elements from Γ_1 and elements from Γ_2. Now each element in Γ_1 can be written as a combination of elements in J_{Γ_1} and their inverses, and similarly for Γ_2 and J_{Γ_2}, so that every element in $\langle \Gamma_1 \cup \Gamma_2 \rangle$ can be written as a combination of elements in $J_{\Gamma_1} \cup J_{\Gamma_2}$ and their inverses. Thus $J_{\Gamma_1} \cup J_{\Gamma_2}$ is a generating set of $\langle \Gamma_1 \cup \Gamma_2 \rangle$, so that by Lemma 5.8,

$$(\mathcal{V}_{\Gamma_1}, \mathcal{E}_{\Gamma_1}) \wedge (\mathcal{V}_{\Gamma_2}, \mathcal{E}_{\Gamma_2}) = \left(\bigvee_{g \in J_{\Gamma_1} \cup J_{\Gamma_2}} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_{\Gamma_1} \cup J_{\Gamma_2}} \mathcal{E}_{\langle g \rangle} \right) = (\mathcal{V}_{\Gamma_1 \vee \Gamma_2}, \mathcal{E}_{\Gamma_1 \vee \Gamma_2}).$$

■

Proposition 5.18 Π_V *is stable under the meet operation* $\wedge$ *in* $\langle \mathcal{C}_V; \preceq \rangle$ *given in (5.6). In particular, for* $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$, $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2) \in \Pi_V$ *with colourings generated by* $\Gamma_1, \Gamma_2 \leq S(V)$ *respectively, the colouring of* $\mathcal{G}_1 \wedge \mathcal{G}_2$ *is generated by* $\Gamma_1 \vee \Gamma_2$.

□

Proof: Let $\mathcal{G}_3 = \mathcal{G}_1 \wedge \mathcal{G}_2 = (\mathcal{V}_3, \mathcal{E}_3)$. By definition of Π_V, $\mathcal{V}_i = \mathcal{V}_{\Gamma_i}$ and $\mathcal{E}_i \subseteq \mathcal{E}_{\Gamma_i}$ for $i = 1, 2$. By definition of the meet operation in $\langle \mathcal{C}_V; \preceq \rangle$ and Lemma 5.9,

$$\begin{aligned} \mathcal{V}_3 &= \mathcal{V}_1 \vee \mathcal{V}_2 = \mathcal{V}_{\Gamma_1} \vee \mathcal{V}_{\Gamma_2} = \mathcal{V}_{\Gamma_1 \vee \Gamma_2} \\ \mathcal{E}_3 &\subseteq \mathcal{E}_1^* \vee \mathcal{E}_2^* \subseteq \mathcal{E}_{\Gamma_1} \vee \mathcal{E}_{\Gamma_2} = \mathcal{E}_{\Gamma_1 \vee \Gamma_2} \end{aligned}$$

so that $\mathcal{G}_1 \wedge \mathcal{G}_2 \in \Pi_V$ with colouring generated by $\Gamma_1 \vee \Gamma_2$ as claimed. ■

Proposition 5.19 *For $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{C}_V$, let* $\mathrm{Aut}(\mathcal{V}, \mathcal{E}) \leq S(V)$ *denote the largest group leaving $(\mathcal{V}, \mathcal{E})$ invariant and let $(\mathcal{V}_{\mathrm{Aut}}, \mathcal{E}_{\mathrm{Aut}})$ denote the graph colouring of $G = (V, E)$ given by the orbits of* $\mathrm{Aut}(\mathcal{V}, \mathcal{E})$ *in V and E respectively. Then $\mathcal{G}$ has a supremum in Π_V with colouring*

$$s_\Pi(\mathcal{G}) = (\mathcal{V}_{\mathrm{Aut}}, \mathcal{E}_{\mathrm{Aut}}).$$

□

Proof: The claim is trivially true for $\mathcal{G} \in \Pi_V$. So suppose $\mathcal{G} \in \mathcal{C}_V \setminus \Pi_V$. $\mathcal{G}$ is modified to a larger graph by adding edge colour classes and splitting colour classes. As the former will not enforce a permutation generated colouring, in order to prove the claim we need to show that $(\mathcal{V}_{\mathrm{Aut}}, \mathcal{E}_{\mathrm{Aut}})$ is the coarsest refinement of $(\mathcal{V}, \mathcal{E})$ which lies in Π_V.

By definition of $\mathrm{Aut}(\mathcal{V}, \mathcal{E})$, it contains all permutations which leave $\mathcal{V}$ and $\mathcal{E}$ invariant and therefore is the largest group $\Gamma \leq S(V)$ satisfying $\mathcal{V}_\Gamma \leq \mathcal{V}$ and that $\mathcal{E}_\Gamma$ restricted to E is finer than $\mathcal{E}$. It therefore, by Proposition 5.17, generates the coarsest such colouring.

■

Combining Proposition 5.18 and Proposition 5.19 with the introduction:

Theorem 5.5 *$\langle \Pi_V; \preceq \rangle$ is a non-distributive lattice with meet operation $\wedge$ inherited from $\langle \mathcal{C}_V; \preceq \rangle$, and join operation $\vee_\Pi$ given by*

$$\mathcal{G} \vee_\Pi \mathcal{H} = s_\Pi(\mathcal{G} \vee \mathcal{H}), \qquad \mathcal{G}, \mathcal{H} \in \Pi_V.$$

□

Not all coloured graphs in $\mathcal{C}_V \setminus \Pi_V$ have an infimum in Π_V, as shown in Figure 5.13. While the colouring of $\mathcal{G}$ does not lie in Π_V, the colourings of both $\mathcal{H}$, generated by by $\Gamma_{\mathcal{H}} = \langle(234)\rangle$, and $\mathcal{K}$, generated by $\Gamma_{\mathcal{K}} = \langle(1234)\rangle$, do. Clearly, $\mathcal{H} \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{G}$ while no $\mathcal{H}', \mathcal{K}' \in \Pi_V$ exist such that $\mathcal{H} \prec \mathcal{H}' \prec \mathcal{G}$ and $\mathcal{K} \prec \mathcal{K}' \prec \mathcal{G}$.

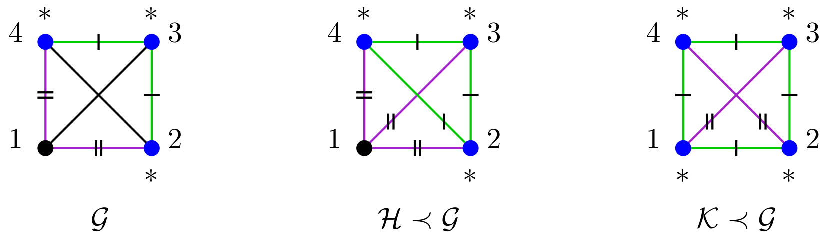


Figure 5.13.: Not every graph in $\mathcal{C}_{[4]} \setminus \Pi_{[4]}$ has an infimum in $\Pi_{[4]}$.

5.4.4. `GraphCheck` Algorithms

`GraphCheck` is a program we have developed for the identification of the colouring classes we have described in this chapter. Given a graph colouring, `GraphCheck` determines whether it falls into the desired class, and if not returns the supremum of the given graph in the class. Hypothesised or inferred RCON and RCOR models may then be tested for membership in the model classes treated in this chapter.

The corresponding identification algorithms are directly based on Proposition 5.9 for edge regular colourings, Proposition 5.11 for regular colourings, Proposition 5.13 for vertex regular colourings and on Proposition 5.19 for permutation generated colourings. Outlines of the algorithms together with output for an example graph are given in Appendix C.

6. Model Selection

Currently no model selection procedures for RCON and RCOR models are available, however they are vital for the models to become widely applicable in practice. In the previous chapter we argued that searching within the four model classes represented by edge regular, vertex regular, regular and permutation generated colourings will generally be more feasible than searching in the full sets of RCON or RCOR models.

One way to develop appropriate model selection procedures is by adapting existing model search algorithms for unconstrained graphical Gaussian models. As each model class is a complete lattice, it is natural to consider methods exploiting this structural property. Prominent such methods include stepwise procedures (Dempster, 1972; Whittaker, 1990; Edwards, 2000), the Edwards-Havránek model selection procedure (Edwards and Havránek, 1987), and, more recently, neighbourhood selection with the lasso (Meinshausen and Bühlmann, 2006), stability selection (Meinshausen and Bühlmann, 2010), and the SINful approach (Drton and Perlman, 2008).

A crucial difference between the search spaces of unconstrained graphical Gaussian models and the model classes studied here is that while the former constitute a distributive lattice (Proposition 5.3 in Section 5.2), the latter are all non-distributive (Theorems 5.2, 5.3, 5.4, 5.5 in Section 5.4.3). This directly disqualifies neighbourhood selection with the lasso, stability selection and the SINful approach, as they all require distributivity. By contrast, stepwise procedures and the Edwards-Havránek algorithm can operate on any lattice of models. We concentrate on the latter.

In the most general form, an Edwards-Havránek model search at each stage may either move up or down the Hasse diagram of the search lattice. It turns out that for models with edge regular colourings upward moves are computationally significantly more expensive than downward moves, so that a downwards only procedure, searching within the set of submodels of models initially assumed acceptable, seems most feasible. Motivated by the very encouraging performance on the data described in Example 3.1 in Section 3.1, for which the corresponding algorithm fitted 232 models out of $1.3 \cdot 10^6$ before arriving at 4 smallest acceptable models, we strongly believe it may perform well in general.

While the other three model classes all equally qualify for an Edwards-Havránek search procedure, we turn our attention to models with a permutation generated colouring, as an Edwards-Havránek search for vertex regular and regular colourings would require repeated non-trivial combinatorial constructions. Further, as permutation generated colourings are special cases of vertex regular and regular colourings, a search strictly within the model classes represented by the latter may be achieved by combining a search in permutation generated colourings with a subsequent local search. One may for instance investigate whether a better suited model with a vertex regular colouring can be obtained from a minimally accepted model with permutation generated constraints through merging of edge colour classes. The required functionality is already available in the `R` package `gRc` (Højsgaard and Lauritzen, 2007).

The intricate relationship between permutation groups acting on the vertex set of an undirected graph and their vertex and edge orbits makes it non-trivial to develop an explicit Edwards-Havránek model selection procedure for models represented by permutation generated colourings. However, a construction of a model search algorithm obeying the Edwards-Havránek principles is still possible and we develop a corresponding procedure. Applied to the data in Example 3.5 in Section 3.4.1, after fitting 48 out of 251 models our algorithm arrived at 9 smallest acceptable models, collectively in very good agreement with previous analyses (Whittaker, 1990).

Most of the results presented in this chapter can be found in Gehrmann (2011).

6.1. The Edwards-Havránek Model Selection Procedure

The Edwards-Havránek model selection procedure operates on model search spaces which are lattices and is closely related to the all possible models approach, but is considerably faster. It is based on the following two principles: (i) if a model is accepted then all models that include it are (weakly) accepted, and (ii) if a model is rejected then all of its submodels are considered to be (weakly) rejected.

The procedure starts by initially testing a set of models and assigns the accepted models to a set $\mathcal{A}$ and the rejected models to set $\mathcal{R}$. By assumption, all models larger than $\mathcal{A}$ are (weakly) accepted and the ones smaller than $\mathcal{R}$ are (weakly) rejected, so that only $\min \mathcal{A}$, the smallest models in $\mathcal{A}$, and $\max \mathcal{R}$, the largest models in $\mathcal{R}$, are of interest. The procedure repeatedly tests model lying immediately below $\min \mathcal{A}$ and immediately above $\max \mathcal{R}$, and updates the two sets accordingly. Once the set to be updated remains unchanged, the algorithm returns $\min \mathcal{A}$ and terminates. The method to determine whether a model is to be rejected can be any suitable statistical test in accordance to the principle of *coherence* (Gabriel, 1969), stating that a test should not accept a model while rejecting a larger one.

Following Edwards and Havránek (1987), for a set of models S let $D_a(S)$ denote the set of models in the search lattice $\mathcal{F}$ say which are smallest with the property that they are not contained in any model in S,

$$D_a(S) = \min\{d \in \mathcal{F} \mid d \not\subseteq s \text{ for all } s \in S\},$$

and let $D_r(S)$ be the set of largest models that do not contain any model in S,

$$D_r(S) = \max\{d \in \mathcal{F} \mid s \not\subseteq d \text{ for all } s \in S\}.$$

$D_a(S)$ is referred to as the *acceptance dual* of S, and $D_r(S)$ as the *rejection dual* of S. The procedure may then be summarised as:

1. Test an initial set of models and assign the accepted models to the set $\mathcal{A}$ and the rejected models to the set $\mathcal{R}$.
2. Choose between 3 and 4.
3. Test the models in $D_r(\mathcal{A}) \setminus \mathcal{R}$. If all are rejected, stop; otherwise, update $\mathcal{A}$ and $\mathcal{R}$ and go to 2.
4. Test the models in $D_a(\mathcal{R}) \setminus \mathcal{A}$. If all are accepted, stop; otherwise, update $\mathcal{A}$ and $\mathcal{R}$ and go to 2.

Acceptance and rejection duals can be computed in a recursive manner by using the following two relations. If S and T are two sets of models, then

$$\begin{aligned} D_a(S \cup T) &= \min\{s \vee t \mid s \in D_a(S), t \in D_a(T)\} \\ D_r(S \cup T) &= \max\{s \wedge t \mid s \in D_r(S), t \in D_r(T)\}. \end{aligned}$$

Thus describing the duals of a single model is enough.

6.2. Models Represented by Edge Regular Colourings

6.2.1. Acceptance Duals

Proposition 6.1 *For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the acceptance dual $D_a(\mathcal{S}^+(\mathcal{V}, \mathcal{E}))$ inside the set of models represented by graphs with edge regular colourings consists of all models represented by $\mathcal{G}_a = (\mathcal{V}_a, \mathcal{E}_a)$ satisfying one of the two conditions:*

(1i) $\mathcal{V}_a = \{V_1, V_2\}$ *such that* $\mathcal{V}_a \not\geq \mathcal{V}$ *and* $\mathcal{E}_a = \emptyset$

(1ii) $\mathcal{V}_a = \{V\}$ *and* $\mathcal{E}_a = \{E_a\}$ *with* $\mathcal{E}_a \neq \emptyset$ *and* $\mathcal{E}_a \not\geq \mathcal{E}$

□

Proof: Let $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{B}_V$ represent a model in $D_a(\mathcal{S}^+(\mathcal{V}, \mathcal{E}))$. Then, by definition of acceptance duals, $\mathcal{G}^* \not\preceq \mathcal{G}$ and there exists no $\mathcal{G}' \in \mathcal{B}_V$ such that $\mathcal{G}' \prec \mathcal{G}^*$ and $\mathcal{G}' \not\preceq \mathcal{G}$. By Definition 5.1 in Section 5.3, $\mathcal{G}^* \not\preceq \mathcal{G}$ if and only if at least one of the following holds:

(i) $\mathcal{V}^* \not\geq \mathcal{V}$

(ii) $\mathcal{E}^* \neq \emptyset$ and for all $\tilde{\mathcal{E}} \subseteq \mathcal{E}$, $\mathcal{E}^* \not\geq \mathcal{E} \setminus \tilde{\mathcal{E}}$

By minimality of models in acceptance duals, only one of the two condition is satisfied by models in the dual. For suppose both (i) and (ii) hold. Then in particular $\mathcal{E}^* \neq \emptyset$, so that for $\mathcal{G}' = (\mathcal{V}^*, \emptyset)$, $\mathcal{G}' \prec \mathcal{G}^*$ while also $\mathcal{G}' \not\preceq \mathcal{G}$, which is a contradiction.

So suppose first that $\mathcal{V}^* \not\geq \mathcal{V}$ and that (ii) does not hold. As $\mathcal{G}^*$ represents the simplest such model, it must have as few edges as possible, so that $\mathcal{E}^* = \emptyset$. The vertex partition giving the simplest model satisfying $\mathcal{V}^* \not\geq \mathcal{V}$ is one with as few vertex colour classes as possible, which is two. Thus $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*)$ is of type (1i).

Now suppose that $\mathcal{E}^* \neq \emptyset$, that for all sets of edge colour classes $\tilde{\mathcal{E}} \subseteq \mathcal{E}$ we have $\mathcal{E}^* \not\geq \mathcal{E} \setminus \tilde{\mathcal{E}}$ and condition (i) above does not hold. The vertex partition giving the simplest such model is $\mathcal{V}^* = \{V\}$, so that in particular all possible edge colourings $\mathcal{E}^*$ will give rise to an edge regular graph colouring.

Suppose that $\mathcal{E}^*$ has at least two edge colour classes. Then two cases are possible:

(a) E^* is a subset of the edges present in $\mathcal{G}$.

(b) E^* contains edges which are not present in $\mathcal{G}$, called *non-edges* below.

In case (a), in order for $\mathcal{G}^* = (\{V\}, \mathcal{E}^*) \not\preceq \mathcal{G}$ to hold, at least one edge colour class, E' say, must not be a union of edge colour classes in $\mathcal{E}$. Since there are at least two colour classes in $\mathcal{E}^*$, there exists $E'' \in \mathcal{E}^*$ with $E'' \neq E'$. Let $\mathcal{G}' = (\mathcal{V}, \mathcal{E}') = (\{V\}, \mathcal{E}^* \setminus E'')$. Then $\mathcal{G}' \prec \mathcal{G}^*$ while also $\mathcal{G}' \not\preceq \mathcal{G}$, so that $\mathcal{G}^*$ cannot represent a model in $D_a(\mathcal{S}^+(\mathcal{V}, \mathcal{E}))$, which is a contradiction.

In case (b), $\mathcal{E}^*$ must contain a class consisting of non-edges only or a class being a union of edges and non-edges. Let this class be E' and let $\mathcal{G}' = (\mathcal{V}, \mathcal{E}') = (\{V\}, \mathcal{E}^* \setminus E'')$ with $E'' \in \mathcal{E}^*$ and $E'' \neq E'$. Then $\mathcal{G}' \prec \mathcal{G}^*$ and $\mathcal{G}' \not\preceq \mathcal{G}$, again leading to a contradiction.

We have established that $\mathcal{E}^* = \{E^*\}$ for some $E^* \subseteq V_{\neq}^{(2)}$. For such edge colourings condition (ii) above becomes $\mathcal{E}^* \neq \emptyset$ and $\mathcal{E}^* \not\succeq \mathcal{E}$, completing the proof.

■

Put in words, models in the edge regular acceptance dual of $\mathcal{S}^+(\mathcal{V}, \mathcal{E})$ are either represented by the empty graph with two vertex colour classes which are not unions of colour classes in $\mathcal{V}$, or by graphs in which all vertices are of the same colour as are the edges, and the edge set is not a union of colour classes in $\mathcal{E}$.

For example, the coloured graphs in Figure 6.1 represent models which lie in the acceptance dual of the model represented by the graph with edge regular colouring in Figure 5.3(a) in Section 5.4.1.

(a) A graph of type (1i). (b) A graph of type (1ii).

Figure 6.1.: Acceptance dual corresponding to the graph in Figure 5.3(a).

By construction, acceptance duals are used to test whether there exist models immediately larger than $\max \mathcal{R}$ which can be rejected. Effectively, graphs of type (1i) in acceptance duals refine the colouring of the maximally rejected models, while graphs of type (1ii) add edge colour classes, which both give larger models. Thus the larger a graph $\mathcal{G}$, the fewer colourings there are of type (1i) and (1ii), giving that the size of acceptance duals decreases with increasing models. This is in full agreement with the definition of acceptance duals.

For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, there are $2^{|V|-1} - 2^{|\mathcal{V}|-1}$ colourings of type (1i). As each vertex inside V is either in one set of $\mathcal{V}_a$ or the other, there are $\frac{1}{2}(2^{|V|} - 2) = 2^{|V|-1} - 1$ partitions of V into two non-empty sets. Out of these precisely $2^{|\mathcal{V}|-1} - 1$ partitions do not qualify for $\mathcal{V}_a$ as their classes are unions of the colour classes in $\mathcal{V}$. The same argument gives there are $2^{\frac{|V|(|V|-1)}{2}-1} - 2^{|\mathcal{E}|-1}$ of models of type (1ii).

Thus the size of an edge regular acceptance dual of $\mathcal{S}^+(\mathcal{V}, \mathcal{E})$ is on the order of

$$2^{|V|-1} - 2^{|\mathcal{V}|-1} + 2^{\frac{|V|(|V|-1)}{2}-1} - 2^{|\mathcal{E}|-1} \sim 2^{|V|^2/2}$$

for large $|V|$. Hence the size of an edge regular acceptance dual $\mathcal{D}_a(S)$ of a set of models S has order up to $\binom{|S|}{2} 2^{|V|^2}$.

6.2.2. Rejection Duals

For a discrete set A we let $\operatorname{sing}(A)$ denote the partition of A into singleton colour classes, so that e.g. $\operatorname{sing}(\{1,2,3\}) = \{\{1\},\{2\},\{3\}\}$.

Proposition 6.2 *For $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the rejection dual $D_r(\mathcal{S}^+(\mathcal{V}, \mathcal{E}))$ inside the set of models represented by graphs with edge regular colourings consists of all models represented by $\mathcal{G}_r = (\mathcal{V}_r, \mathcal{E}_r)$ satisfying one of the three conditions:*

(2i) $\mathcal{V}_r = \{\{\alpha,\beta\}\} \cup \operatorname{sing}(V \setminus \{\alpha,\beta\})$ *such that* $\mathcal{V}_r \not\leq \mathcal{V}$ *and* $\mathcal{E}_r = \operatorname{sing}(V^{(2)}_{\neq})$

(2ii) $\mathcal{V}_r = \operatorname{sing}(V)$ *and* $\mathcal{E}_r = \operatorname{sing}(V^{(2)}_{\neq}) \setminus \{\alpha\beta\}$ *with* $\alpha\beta \in E$

(2iii) $\mathcal{V}_r = \{\{\alpha,\beta\},\{\gamma,\delta\}\} \cup \operatorname{sing}(V \setminus \{\alpha,\beta,\gamma,\delta\})$ *with* $\alpha \equiv \beta\ (\mathcal{V})$, $\gamma \equiv \delta\ (\mathcal{V})$ *and* $\mathcal{E}_r = \{\alpha\gamma, \beta\delta\} \cup \operatorname{sing}(V^{(2)}_{\neq} \setminus \{\alpha\gamma,\beta\delta\})$, *allowing* $\alpha = \beta$ *or* $\gamma = \delta$ *but not both, such that* $(\mathcal{V},\mathcal{E}) \not\preceq (\mathcal{V}_r,\mathcal{E}_r)$

□

Proof: Let $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*) \in \mathcal{B}_V$ represent a model in $D_r(\mathcal{S}^+(\mathcal{V},\mathcal{E}))$. Then, by definition of rejection duals, $\mathcal{G} \not\preceq \mathcal{G}^*$ and there exists no $\mathcal{G}' \in \mathcal{B}_V$ such that $\mathcal{G}^* \prec \mathcal{G}'$ and $\mathcal{G} \not\preceq \mathcal{G}'$. By Definition 5.1 in Section 5.3, $\mathcal{G} \not\preceq \mathcal{G}^*$ if and only if at least one of the following holds:

(i) $\mathcal{V}^* \not\leq \mathcal{V}$

(ii) For all $\tilde{\mathcal{E}} \subseteq \mathcal{E}^*$, $\mathcal{E}^* \setminus \tilde{\mathcal{E}} \not\leq \mathcal{E}$.

By maximality of models in rejection duals, only one of the condition above holds for any model in the dual. For suppose both (i) and (ii) are satisfied. Then in particular $\mathcal{E}^* \neq \operatorname{sing}(V^{(2)}_{\neq})$, so that for $\mathcal{G}' = (\mathcal{V}^*, \operatorname{sing}(V^{(2)}_{\neq}))$, $\mathcal{G}^* \prec \mathcal{G}'$ while also $\mathcal{G} \not\preceq \mathcal{G}'$, which is a contradiction.

Suppose first that $\mathcal{V}^* \not\leq \mathcal{V}$ and that (ii) does not hold. As $\mathcal{G}^*$ represents the largest such model, it must have as many edges as possible and the finest colouring possible. The first condition implies that $\mathcal{E}^* = \text{sing}(V^{(2)}_{\neq})$. Further, for a colouring $(\mathcal{V}^*, \mathcal{E}^*)$ of $G = (V, V^{(2)}_{\neq})$, if any two edges in $\mathcal{E}^*$ have the same colour, their end vertices must receive the same colouring in order to ensure edge regularity. The converse however does not hold, as giving any two vertices the same colour does not put any constraints on the edge colouring. Thus giving two vertices in V the same colour such that $\mathcal{V}^* \not\leq \mathcal{V}$ gives the largest possible colouring, establishing that if $\mathcal{V} \not\geq \mathcal{V}^*$, then $(\mathcal{V}^*, \mathcal{E}^*)$ must be of type (2i) in the claim.

So suppose that condition (i) does not hold, i.e. that $\mathcal{V}^* \leq \mathcal{V}$. There are two ways in which condition (ii) can fail to hold:

(a) There exists $\alpha\beta \in V^{(2)}_{\neq}$ such that $\alpha\beta \in E$ and $\alpha\beta \notin E^*$, or

(b) there exist $\alpha\beta, \gamma\delta \in V^{(2)}_{\neq}$ such that $\alpha\beta \equiv \gamma\delta \ (\mathcal{E}^*)$ and $\alpha\beta \not\equiv \gamma\delta \ (\mathcal{E})$.

By maximality of $\mathcal{G}^*$ only one of (a) and (b) can hold. For suppose not and consider $\mathcal{G}' = (\mathcal{V}^*, \mathcal{E}^* \cup \{\{\alpha\beta\}\})$ with $\alpha\beta \in E$ as in (a). Then $\mathcal{G}^* \prec \mathcal{G}'$ and $\mathcal{G} \not\preceq \mathcal{G}'$, which is a contradiction.

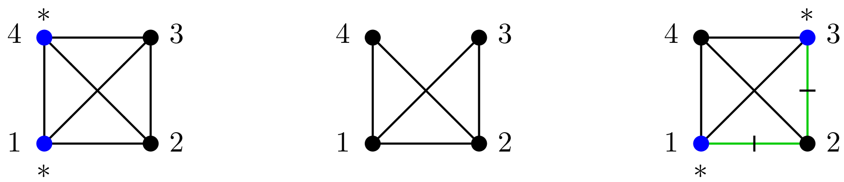
In case (a), the largest such model has vertex colouring $\mathcal{V}^* = \text{sing}(V)$, giving a colouring of type (2ii). In case of (b), in order to give an edge regular colouring the end vertices of $\alpha\beta$ and $\gamma\delta$ must be of the same colour in $\mathcal{V}^*$. By maximality of $(\mathcal{V}^*, \mathcal{E}^*)$ they must also be of the same colour in $\mathcal{V}$, giving that $(\mathcal{V}^*, \mathcal{E}^*)$ is precisely of type (2iii).

■

Graphs representing models in the rejection dual of a model represented by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ almost represent the unrestricted saturated model except for a minor modification: in graphs of type (2i) two vertices which are not of the same colour in $\mathcal{V}$

form the only composite colour class in $\mathcal{V}_r$, while graphs of type (2ii) are missing an edge present in $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Graphs of type (2iii) have a pair of equally coloured edges which are not of the same colour in $\mathcal{E}$, while all vertices which receive the same colour to give an edge regular colouring are also of the same colour in $\mathcal{V}$.

Examples of coloured graphs in the rejection dual of the model represented by the graph displayed in Figure 5.3(a) in Section 5.4.1 are given in Figure 6.2.



(a) A graph of type (2i). (b) A graph of type (2ii). (c) A graph of type (2iii).

Figure 6.2.: Rejection dual corresponding to the graph in Figure 5.3(a).

By construction, rejection duals contain models which lie immediately below $\min \mathcal{A}$. Graphs of type (2i) merge vertex colour classes, graphs of type (2ii) cause edge colour classes to be dropped and graphs of type (2iii) merge edge colour classes, as well as vertex colour classes to ensure the resulting model to be edge regular. All operations give graphs which represent smaller models, giving that the size of rejection duals grows with increasing models. This is in full agreement with the definition of rejection duals.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{V_1, \dots, V_k\}$ and $\mathcal{E} = \{E_1, \dots, E_l\}$. Then there are $\sum_{V_i, V_j \in \mathcal{V}, i<j} |V_i||V_j|$ graph colourings of type (2i), as there are this many ways of choosing two vertices in V which are of different colour in $\mathcal{V}$. There are $|E|$ graphs of type (2ii) and at most $\sum_{E_i, E_j \in \mathcal{E}, i<j} |E_i||E_j|$ models of type (2iii). (At most as not every pair of edges of different colours in $\mathcal{E}$ is suitable, as they need to connect the

same vertex colour classes in $\mathcal{V}$.) This gives that $|D_r(\mathcal{S}^+(\mathcal{V}, \mathcal{E}))|$ has order at most

$$\sum_{\substack{V_i, V_j \in \mathcal{V}, \\ i<j}} |V_i||V_j| \;+\; |E| \;+ \sum_{\substack{E_i, E_j \in \mathcal{E}, \\ i<j}} |E_i||E_j| \sim c|V|^4$$

for some constant c for large $|V|$. Hence the size of an edge regular rejection dual $D_a(S)$ of a set of models S has order at most $\binom{|S|}{2} c^2 |V|^8$.

6.2.3. An Edwards-Havránek Model Selection Algorithm

The orders of acceptance and rejection duals found in Sections 6.2.1 and 6.2.2 suggest that, from a computational point of view, working with rejection duals may be more efficient. This corresponds to the following algorithm.

1. Test an initial set of models and assign each to $\mathcal{R}$ if it is rejected and to $\mathcal{A}$ otherwise.
2. Test the models in $D_r(\mathcal{A}) \setminus \mathcal{R}$. If these are all rejected, stop. Otherwise update $\mathcal{A}$ and $\mathcal{R}$ and repeat.

We executed the above algorithm on the data set described in Example 3.1 in Section 3.1 with the saturated uncoloured model as our initial set of accepted models $\mathcal{A}$ and let $\mathcal{R} = \emptyset$ initially. Models were tested for acceptance by a likelihood ratio test relative to the saturated unconstrained model at significance level 5%, assuming an unconstrained mean μ, with functionality implemented in the `R` package `gRc` (Højsgaard and Lauritzen, 2007).

Out of the $1.3 \cdot 10^6$ models in the search space (Section B.1 in Appendix B), the algorithm fitted 232 models in 8 stages before arriving at 4 minimally accepted models whose graphs are displayed in Figure 6.3 together with their BIC values. A list of all fitted models, grouped by stages of the algorithm, is given in Appendix D.

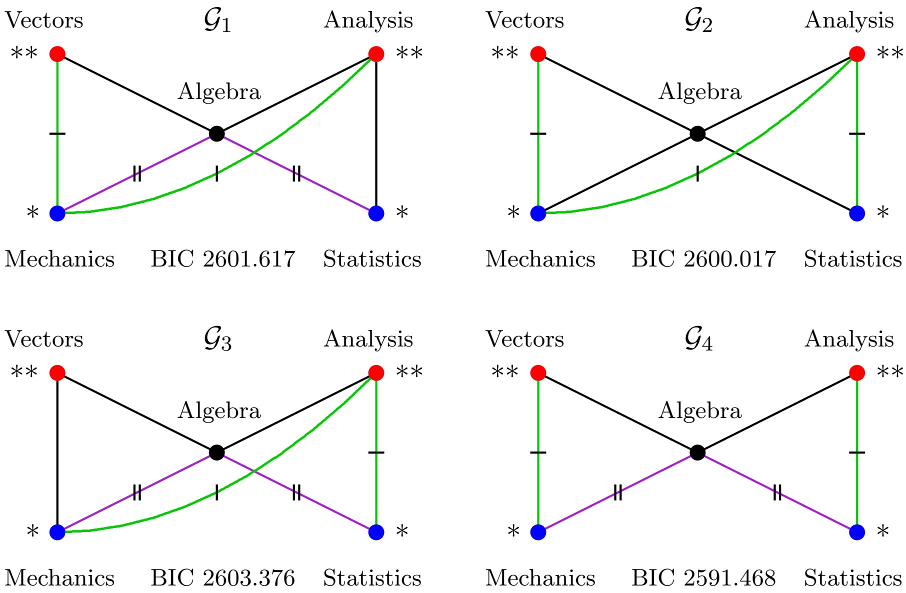


Figure 6.3.: Minimally accepted models by Edwards-Havránek procedure in $\mathcal{B}_{[5]}$.

The uncoloured graphs underlying the graphs in Figure 6.3 contain the graph in Figure 3.1(a) in Section 3.1, which represents the previously inferred unconstrained graphical Gaussian model for the data, as a subgraph. $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$ only differ by one edge and $\mathcal{G}_4$ has exactly the same edge set. Thus the algorithm picked up the previously established conditional independence structure correctly.

The minimally accepted model with the lowest BIC value is represented by $\mathcal{G}_4$. It is different from but not unsimilar to the RCON model fitted in Højsgaard and Lauritzen (2008), whose graph is displayed in Figure 3.1(b) in Section 3.1, which has a slightly lower BIC value of 2587.404 but no specific properties as it is chosen in an ad hoc manner. Note that the model fitted in Højsgaard and Lauritzen (2008) is not edge regular so that it could not have been considered by the algorithm.

Interestingly, the colouring of graph $\mathcal{G}_4$ becomes permutation generated after merging the two singleton edge colour classes into one. The corresponding model has a BIC value of 2589.148. We conducted local stepwise searches within the set of RCON models based on BIC value starting from the models represented by graphs $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$ and $\mathcal{G}_4$, and arrived at the models represented by $\mathcal{G}_5$, $\mathcal{G}_6$ and $\mathcal{G}_7$ displayed in Figure 6.4. ($\mathcal{G}_5$ was obtained from $\mathcal{G}_1$, $\mathcal{G}_6$ from $\mathcal{G}_2$, and $\mathcal{G}_7$ from both $\mathcal{G}_3$ and $\mathcal{G}_4$.)

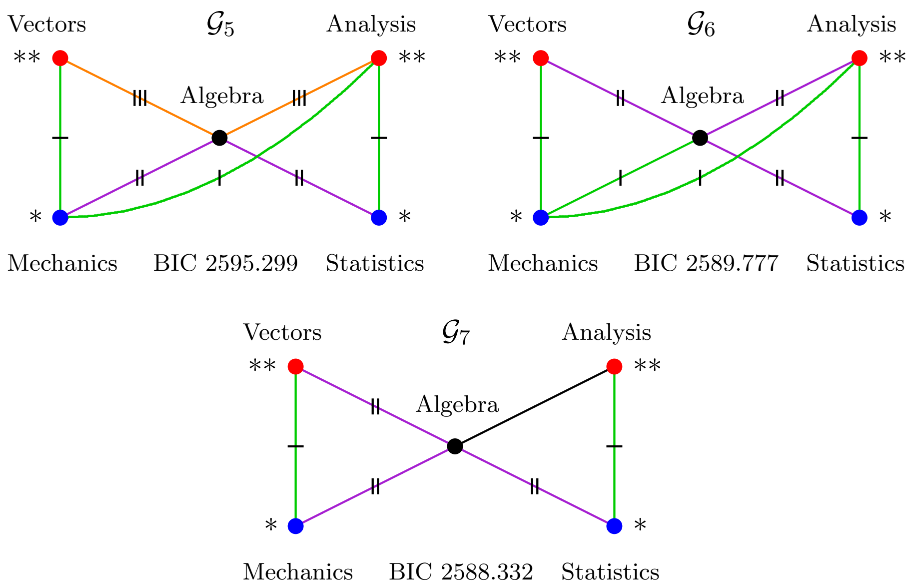


Figure 6.4.: Local RCON search starting from the models displayed in Figure 6.3.

Note the similarity between $\mathcal{G}_7$ and the graph of the model obtained by Højsgaard and Lauritzen (2008). The latter is obtained from the former by merging the two composite edge colour classes.

This example strongly suggests that an Edwards-Havránek model selection procedure for models with edge regular colourings may be feasible in general.

6.3. Models Represented by Permutation Generated Colourings

The intricate relationship between permutation groups and their vertex and edge orbits makes it non-trivial to develop an explicit description of the acceptance and rejection duals in Π_V. However, rejection and acceptance duals are nothing but constructions which allow an Edwards-Havránek search to move along the edges in the Hasse diagram of a lattice search space, so that if the diagram can be constructed, an Edwards-Havránek search is possible.

Below we describe the structure of the Hasse diagram of Π_V to subsequently exploit it for model selection.

6.3.1. The Hasse Diagram of Π_V

Let

$$\mathcal{K}_V = \{(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \mid \Gamma \leq S(V)\}$$

and

$$\pi_V : S(V) \to \mathcal{K}_V, \qquad \Gamma \mapsto (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma).$$

Then, as Proposition 5.15 in Section 5.4.3 suggests, π_V is not an injective function. In particular, if we let

$$\pi'_V : S(V) \to \mathcal{P}(V^{(2)}_{\neq}), \qquad \Gamma \mapsto \mathcal{E}_\Gamma,$$

by Proposition 5.16 in Section 5.4.3, for $\Gamma_1, \Gamma_2 \leq S(V)$,

$$\pi'_V(\Gamma_1) = \pi'_V(\Gamma_2) \qquad \Longrightarrow \qquad \pi_V(\Gamma_1) = \pi_V(\Gamma_2). \tag{6.1}$$

Further let

$$\mathcal{O}_V = \{(\mathcal{V}_\Gamma, \emptyset) \mid \Gamma \leq S(V)\}$$

and for $\Gamma \leq S(V)$,

$$D(\mathcal{E}_\Gamma) = \{(\mathcal{V}_\Gamma, \mathcal{E}) \mid \mathcal{E} \subseteq \mathcal{E}_\Gamma\}.$$

Then

$$\Pi_V = \mathcal{K}_V \cup \mathcal{O}_V \cup \left(\bigcup_{\Gamma \leq S(V)} D(\mathcal{E}_\Gamma) \right)$$

and we can describe the Hasse diagram of $\langle \Pi_V; \wedge, \vee_\Pi \rangle$ by describing the links within each of $\mathcal{K}_V$, $\mathcal{O}_V$, $D(\mathcal{E}_\Gamma)$ for $\Gamma \leq S(V)$ and the links between the sets.

We begin by describing the structure of $\mathcal{K}_V$. By Lemma 5.8 in Section 5.4.3, for $\Gamma \leq S(V)$ with generating set J_Γ,

$$(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) = \left(\bigvee_{g \in J_\Gamma} \mathcal{V}_{\langle g \rangle}, \bigvee_{g \in J_\Gamma} \mathcal{E}_{\langle g \rangle} \right). \tag{6.2}$$

Thus all colourings in $\mathcal{K}_V$ can be obtained from colourings of the form $(\mathcal{V}_{\langle \sigma \rangle}, \mathcal{E}_{\langle \sigma \rangle})$ for $\sigma \in S(V)$. The colourings generated by groups which have more than one generator can then be obtained by applying relation (6.2). While there are $|V|!$ permutation in $S(V)$, not all of them will generate distinct graph colourings as for example $\langle \sigma_1 \rangle = \langle \sigma_2 \rangle$ if $\sigma_2 \in \langle \sigma_1 \rangle$. Further, all vertex colourings $\mathcal{V} \in \mathcal{P}(V)$ are the vertex colouring of some graph in $\mathcal{K}_V$. For if we take $\mathcal{V} = \{V_1, \ldots, V_k\}$, then $\mathcal{V} = \mathcal{V}_\Gamma$ for $\Gamma = S(V_1) \times \ldots \times S(V_k)$.

Lemma 5.9 in Section 5.4.3 gives that $\mathcal{K}_V$ is a complete (non-distributive) sublattice $\langle \mathcal{K}_V; \wedge, \vee_\Pi \rangle$ of $\langle \Pi_V; \wedge, \vee_\Pi \rangle$. The links in the Hasse diagram of Π_V within $\mathcal{K}_V$ are thus given by the Hasse diagram of $\langle \mathcal{K}_V; \wedge, \vee_\Pi \rangle$.

For $D(\mathcal{E}_\Gamma)$, $\Gamma \leq S(V)$ observe first that

$$D(\mathcal{E}_\Gamma) \cap \mathcal{K}_V = (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \qquad D(\mathcal{E}_\Gamma) \cap \mathcal{O}_V = (\mathcal{V}_\Gamma, \emptyset). \tag{6.3}$$

Further, by Lemma 5.2 in Section 5.1.2, $D(\mathcal{V}_\Gamma)$ is a complete (distributive) sublattice of $\langle \Pi_V; \wedge, \vee_\Pi \rangle$ since it consists of all coloured graphs $(\mathcal{V}_\Gamma, \mathcal{E})$ with $\mathcal{E}$ lying in the powerset $P(\mathcal{E}_\Gamma)$ of $\mathcal{E}_\Gamma$. The unit of $D(\mathcal{V}_\Gamma)$ is $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \in \mathcal{K}_V$, the zero is $(\mathcal{V}_\Gamma, \emptyset) \in \mathcal{O}_V$, and $(\mathcal{V}_\Gamma, \mathcal{E}_1), (\mathcal{V}_\Gamma, \mathcal{E}_2) \in D(\mathcal{E}_\Gamma)$ are connected by a link in $D(\mathcal{E}_\Gamma)$ if and only if $\mathcal{E}_1$ and $\mathcal{E}_2$ differ by precisely one edge colour class. The links in the Hasse diagram of Π_V within $D(\mathcal{E}_\Gamma)$ are thus given by the Hasse diagram of $D(\mathcal{E}_\Gamma)$. By (6.3), the Hasse diagram of each $D(\mathcal{E}_\Gamma)$ can be viewed as a connection between $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \in \mathcal{K}_V$ and $(\mathcal{V}_\Gamma, \emptyset) \in \mathcal{O}_V$ in the Hasse diagram of Π_V.

While by (6.1), for $\Gamma \leq S(V)$, $\mathcal{E}_\Gamma$ uniquely determines $\mathcal{V}_\Gamma$, the converse is generally false. (Simply consider graphs $\mathcal{G}_{15}$ and $\mathcal{G}_{16}$ in Appendix B.) Thus if for $\Gamma_1 \neq \Gamma_2$, $\mathcal{V}_{\Gamma_1} = \mathcal{V}_{\Gamma_2}$ (while $\mathcal{E}_{\Gamma_1} \neq \mathcal{E}_{\Gamma_2}$) then $(\mathcal{V}_{\Gamma_1}, \emptyset) = (\mathcal{V}_{\Gamma_2}, \emptyset) \in D(\mathcal{E}_{\Gamma_1})$ and $(\mathcal{V}_{\Gamma_1}, \emptyset) = (\mathcal{V}_{\Gamma_2}, \emptyset) \in D(\mathcal{E}_{\Gamma_2})$ so that $D(\mathcal{E}_{\Gamma_1}) \cap D(\mathcal{E}_{\Gamma_2})$ is non-empty. More generally, in that case there exist maximal elements $(\mathcal{V}_1^*, \mathcal{E}_1^*), \ldots, (\mathcal{V}_k^*, \mathcal{E}_k^*)$ such that all elements smaller than or equal to some $(\mathcal{V}_i^*, \mathcal{E}_i^*)$ lie in $D(\mathcal{E}_{\Gamma_1}) \cap D(\mathcal{E}_{\Gamma_2})$ and all other elements in $D(\mathcal{E}_{\Gamma_1}) \cup D(\mathcal{E}_{\Gamma_2})$ lie in only one of $D(\mathcal{E}_{\Gamma_1})$ and $D(\mathcal{E}_{\Gamma_2})$. Further, for $\mathcal{V}_{\Gamma_1} \neq \mathcal{V}_{\Gamma_2}$ links may be present between elements of $D(\mathcal{E}_{\Gamma_1})$ and $D(\mathcal{E}_{\Gamma_2})$ as well.

For the structure of $\mathcal{O}_V$, first note that just as all possible vertex colourings each $\mathcal{V} \in \mathcal{P}(V)$ appear as some vertex colouring in $\mathcal{K}_V$ the same must be true for $\mathcal{O}_\Gamma$ which therefore is isomorphic to the partition lattice $\mathcal{P}(V)$. By Lemma 5.3 in Section 5.1.2, $\mathcal{O}_V$ is a complete (non-distributive) lattice and the links with the Hasse diagram of Π_V are given by the Hasse diagram of $\mathcal{O}_V$.

We illustrate the above description of the structure of Π_V by the sketch diagram in Figure 6.5. The two ellipses represent $\mathcal{K}_V$ and $\mathcal{O}_V$, where the links within $\mathcal{K}_V$ and $\mathcal{O}_V$ omitted for clarity of exposition. The vertices represent graph colourings $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$ and $(\mathcal{V}_\Gamma, \emptyset)$ for $\Gamma \leq S(V)$, and each path with solid edges between two such vertices stands for the links in the diagram of $D(\mathcal{E}_\Gamma)$. The dashed edges represent links between the $D(\mathcal{E}_\Gamma)$.

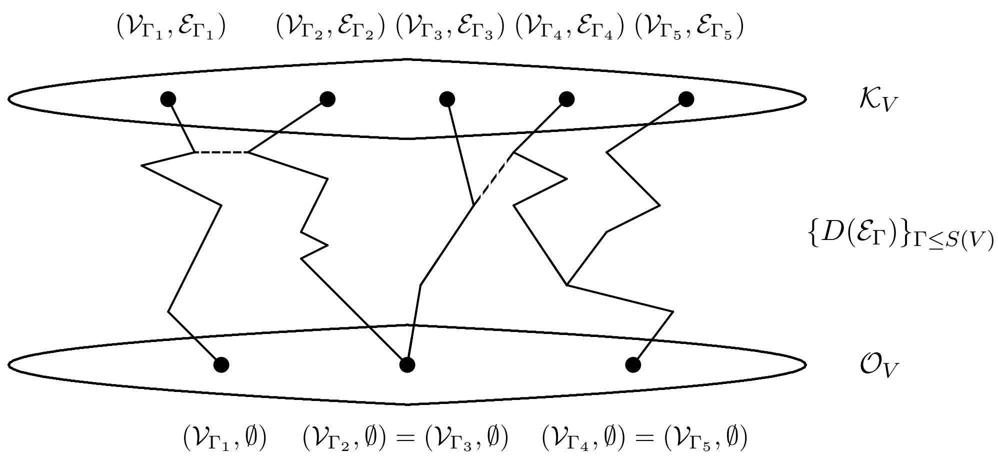


Figure 6.5.: Structure of Hasse diagram of $\langle \Pi_V; \wedge, \vee_\Pi \rangle$.

Three cases can occur for $(\mathcal{V}, \emptyset)$: either $(\mathcal{V}, \emptyset)$ is the zero in only one $D(\mathcal{E}_\Gamma)$, represented by $(\mathcal{V}_{\Gamma_1}, \emptyset)$; $(\mathcal{V}, \emptyset)$ is the zero in more than one $(\mathcal{V}_{\Gamma_1}, \emptyset)$ and they intersect in $(\mathcal{V}, \emptyset)$ only, see $(\mathcal{V}_{\Gamma_2}, \emptyset) = (\mathcal{V}_{\Gamma_3}, \emptyset)$; or $(\mathcal{V}, \emptyset)$ is the zero of more than one $D(\mathcal{E}_\Gamma)$ which share more elements than only $(\mathcal{V}, \emptyset)$, represented by $(\mathcal{V}_{\Gamma_4}, \emptyset) = (\mathcal{V}_{\Gamma_5}, \emptyset)$.

The only non-trivial connections we could not describe explicitly are those between $D(\mathcal{E}_{\Gamma_1})$ and $D(\mathcal{E}_{\Gamma_2})$ for $\mathcal{E}_{\Gamma_1} \neq \mathcal{E}_{\Gamma_2}$ and the links within $\mathcal{K}_V$. While this prevents us from giving an explicit description of the Hasse diagram of Π_V, for small sizes $|V|$ it can certainly be constructed computationally. For example by performing the following steps:

1. Find all $|V|!$ permutations in $S(V)$ and construct their vertex and edge orbits.

2. Find all graph colourings in $\mathcal{K}_V$ which can be obtained from the graphs in 1. by repeatedly applying the meet operation $\wedge$ in $\langle \mathcal{C}_V; \preceq \rangle$.
3. Obtain the Hasse diagram of $\mathcal{K}_V$.
4. Obtain the Hasse diagram of $D(\mathcal{E}_\Gamma)$ for all distinct $\mathcal{E}_\Gamma$ and the Hasse diagram of $\mathcal{O}_V$.
5. Check for additional links between all pairs $D(\mathcal{E}_\Gamma)$ and $D(\mathcal{E}_{\Gamma'})$ for $\Gamma \neq \Gamma'$.

Once the Hasse diagram of Π_V is obtained, a downwards search obeying the Edwards-Havránek principles may be carried out exploiting the diagram in the following way:

1. Let $\mathcal{A}$ contain the saturated model without symmetry constraints and let $\mathcal{R} = \emptyset$.
2. If all models one link below $\min \mathcal{A}$ lie in $\mathcal{R}$, stop and return $\min \mathcal{A}$. Otherwise test all models which lie one link below $\min \mathcal{A}$ and do not lie in $\mathcal{R}$, assign the accepted models to $\mathcal{A}$ and the rejected models to $\mathcal{R}$, and repeat.

6.3.2. An Example of an Edwards-Havránek Search

We conducted a downwards only Edwards-Havránek search as described above in the set of models represented by $\Pi_{[4]}$ for the Frets' heads data in Example 3.5 in Section 3.4.1. As before, models were tested for acceptance by performing a likelihood ratio test relative to the saturated unconstrained model at significance level 5%, assuming an unconstrained mean μ.

After fitting 48 models out of the 251 models in the search space in 4 stages, the algorithm arrived at 9 minimally accepted models, whose graphs are displayed in Figure 6.6 together with their BIC values. How the search navigated through the search space is described in full detail in Appendix E.

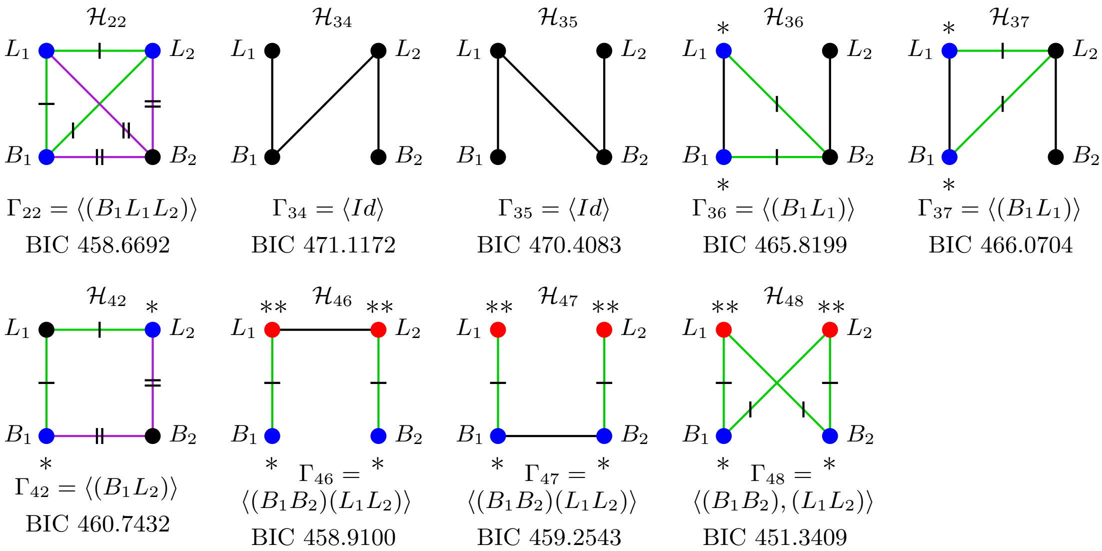


Figure 6.6.: Graphs of minimally accepted models in $\Pi_{[4]}$ for Frets' heads data.

Note that all found models have BIC value lower than the BIC value of 471.2982 of the model fitted in Højsgaard and Lauritzen (2008), whose graph is given in Figure 3.7(b) in Section 3.4.1. The lowest among them corresponds to the model represented by graph $\mathcal{H}_{48}$.

Further, the model selected in Højsgaard and Lauritzen (2008) is a supermodel, in fact the supremum, of the models represented by $\mathcal{H}_{46}$ and $\mathcal{H}_{47}$, with a further edge to complete the four cycle. Interestingly, $\mathcal{H}_{34}$ and $\mathcal{H}_{35}$ are two of the graphs found in Section 8.3 in Whittaker (1990), the other two being the underlying uncoloured graphs of $\mathcal{H}_{46}$ and $\mathcal{H}_{47}$. The graph $\mathcal{H}_{22}$ vividly stands out from the rest, with a remarkably small BIC value.

7. Discussion and Further Work

In this thesis we studied RCON and RCOR models as they were introduced by Højsgaard and Lauritzen (2008). We addressed maximum likelihood estimation of a non-zero model mean, we studied the structure of model classes within the sets of RCON and RCOR models and developed model selection techniques. In this section we discuss the obtained results and suggest direction for possible further work.

7.1. Estimation of Means

In Chapter 4 we obtained a characterisation of linear spaces Ω such that if Σ is restricted to satisfy the constraints of an RCON or RCOR model, the maximum likelihood estimator $\hat{\mu}$ and least squares estimator μ^* of μ coincide if μ is restricted to lie in Ω. We further demonstrated that for equality constraints on μ, the characterisation of Ω translates into a necessary and sufficient colouring regularity condition. The latter was illustrated with examples from comparative experiments according to Latin square designs with the covariance structure of the errors assumed to satisfy the constraints of a homogeneous nearest-neighbour lattice scheme or its torus equivalent.

Within comparative experiments, the above combination of constraints on Σ and μ is only one possibility of many and other combinations of treatment allocations and concentration structures with inherent symmetries may lead to interesting results. Further, for the torus concentration structure, we merely stated that Latin square designs constructed in a particular way guarantee estimability of treatment effects by the least squares estimator. In fact, others are possible and the characterisation of all suitable Latin square designs leads to non-trivial combinatorial considerations. Those were omitted for the sake of brevity.

Possible extensions of our work on estimation of model means consists in the application of our results to settings different from comparative experiments.

A further interesting direction for further work is the quantification of the effect on the estimation of the mean if it 'almost' satisfies the constraints which ensure equality of estimators (where 'almost' will need to be given appropriate meaning).

7.2. Model Selection

In Chapter 5 we showed four model classes within the sets of RCON and RCOR models, each satisfying desirable statistical properties and being more readily interpretable than RCON and RCOR models in general, to form complete lattices and identified their meet and join operations. Exploiting the fact that each model class corresponds to a colouring regularity condition and Brendan McKay's program `nauty`, our computer program `GraphCheck` can determine whether a coloured graph represents a model lying in any of the model classes.

As all four model classes form complete non-distributive lattices, each of them qualifies for the Edwards-Havránek model selection procedure. In Chapter 6 we argued that two of them, those represented by edge regular and permutation generated colourings respectively, are most suitable for such a model search. For the former model class we developed an Edwards-Havránek model selection algorithm, and demonstrated an encouraging performance for a data set on five variables.

We further developed a systematic search procedure in agreement with the Edwards-Havránek principles for the lattice of models with constraints on Σ^{-1} induced by permutation symmetry, which we illustrated with the example of Frets' heads data on four variables. Here as well, the algorithm performed in a satisfactory fashion. In order to fully develop an Edwards-Havránek search algorithm for this class of models, further investigations into the relationship between permutation groups acting on V and their orbits in V and $V \times V$ need to be undertaken.

While the algorithm's performance in the above examples was encouraging, it has to be taken into account that the number of variables is rather small in both cases. Further, it is at this stage unknown how much this behaviour relied on strong/weak conditional independence and symmetry relations in the data sets considered. It may be that the number of models to be tested can still grow in an unmanageable fashion. Therefore other, and larger, examples should be studied.

A general concern with the Edwards-Havránek model selection procedure is that the procedure does not control the overall error rate. It may therefore be worthwhile to explore alternative model selection approaches. While neighbourhood selection with the lasso (Meinshausen and Bühlmann, 2006), stability selection (Meinshausen and Bühlmann, 2010), and the SINful approach (Drton and Perlman, 2008) all do control for the overall error rate, we showed them not to be directly applicable to the lattices of models studied here due to their non-distributivity. Modified variants may still be feasible, which could be investigated.

One further viable alternative may be a variant of the graphical lasso

(Friedman et al., 2008; Ravikumar et al., 2008) for models with equality constraints. In its original form the graphical lasso seeks to maximise the penalised log-likelihood

$$\log\det\Sigma^{-1} - \mathrm{tr}(S\Sigma^{-1}) - \rho\|\Sigma^{-1}\|_1, \tag{7.1}$$

over non-negative definite matrices Σ^{-1}, with S denoting the empirical covariance matrix of the observations, $\|\cdot\|_1$ being the l_1 norm giving the sum of the absolute values of the elements in the argument matrix and ρ denoting the penalisation parameter. Testing for equality constraints on the entries of $\Sigma^{-1} = (k_{\alpha\beta})_{\alpha,\beta\in V}$ can be enforced by replacing (7.1) by the following function

$$\log\det\Sigma^{-1} - \mathrm{tr}(S\Sigma^{-1}) - \rho_1\|\Sigma^{-1}\|_1 - \rho_2\sum_{\alpha,\beta\in V}|k_{\alpha\alpha} - k_{\beta\beta}| - \rho_3\sum_{\substack{\alpha,\beta,\gamma,\delta\in V,\\ \alpha\neq\beta,\gamma\neq\delta}}|k_{\alpha\beta} - k_{\gamma\delta}|.$$

This lies in direct analogy to the development of the fused lasso (Tibshirani et al., 2005) from the standard lasso for linear regression. However, it has to be mentioned that how to maximise the function over a given symmetry class, for example over models with edge regular colourings to ensure scale invariance, seems non-trivial.

7.3. Likelihood Ratio Tests

A direction for further work of relevance to both main topics addressed in this thesis is the study of likelihood ratio tests for RCON, RCOR and RCOP models, with and without constraints on the means. Most interesting is the characterisation of models which allow likelihood ratio statistics in closed form. Hylleberg et al. (1993) showed this to hold for mean-zero RCOP models generated by compound symmetry (Votaw, 1948), see (2.13), which are represented by decomposable graphs, i.e. graphs which do not have cycles of length greater than 3 as induced subgraphs.

It would be interesting to know whether there are other RCON, RCOR and RCOP models with this property, and further which assumptions on the means are permissible. This will facilitate exact model comparison for the models studied in this thesis, including the comparison of competing restrictions on the means for a given covariance structure, and may give rise to new model selection procedures.

Conceptually, knowledge about explicit likelihood ratio test will provide an appealing bridge to results for Gaussian models without conditional independence relations, for which hypothesis testing was among the first studied topics, see Section 2.3.1.

Appendix A.

Tables of Stable Subspaces in Section 4.3.4

	(α,β,γ)	Ω_{v_1}	Relation between μ_1,μ_2,μ_3	X^{v_1}
1	$(0,0,0)$	0	$\mu_1=\mu_2=\mu_3=0$	
2	$(0,0,1)$	$\langle s\rangle$	$\mu_i=cs_i,\ c\in\mathbb{R}$	$\begin{pmatrix}s_1\\s_2\\s_3\end{pmatrix}$
3	$(0,1,0)$	$\langle r\rangle$	$\mu_i=cr_i,\ c\in\mathbb{R}$	$\begin{pmatrix}r_1\\r_2\\r_3\end{pmatrix}$
4	$(0,1,1)$	$\left(\langle\begin{pmatrix}1\\1\\1\end{pmatrix}\rangle\right)^{\perp}$	$\mu_1+\mu_2+\mu_3=0$	$\begin{pmatrix}r_1&s_1\\r_2&s_2\\r_3&s_3\end{pmatrix}$
5	$(1,0,0)$	$\langle\begin{pmatrix}1\\1\\1\end{pmatrix}\rangle$	$\mu_1=\mu_2=\mu_3$	$\begin{pmatrix}1\\1\\1\end{pmatrix}$
6	$(1,0,1)$	$(\langle r\rangle)^{\perp}$	$r_1\mu_1+r_2\mu_2+r_3\mu_3=0$	$\begin{pmatrix}1&s_1\\1&s_2\\1&s_3\end{pmatrix}$
7	$(1,1,0)$	$(\langle s\rangle)^{\perp}$	$s_1\mu_1+s_2\mu_2+s_3\mu_3=0$	$\begin{pmatrix}1&r_1\\1&r_2\\1&r_3\end{pmatrix}$
8	$(1,1,1)$	$\mathbb{R}^{v_1}$	$\mu_1,\mu_2,\mu_3\in\mathbb{R}$	I_3

Table A.1.: All $\mathcal{G}^{[u_1\{v_1v_1\}]}$-stable spaces Ω_{v_1}.

	$(\alpha,\beta,\gamma,\delta)$	Ω_{v_1}	Ω_{v_2}
1	$(0,0,0,0)$	0	0
2	$(0,1,0,0)$	0	C_0
3	$(1,0,0,0)$	B_0	0
4	$(1,1,0,0)$	B_0	C_0
5	$(0,0,1,0)$	B_1	C_1
6	$(0,1,1,0)$	B_1	$C_0 \oplus C_1 = C_2^\perp$
7	$(1,0,1,0)$	$B_0 \oplus B_1 = B_2^\perp$	C_1
8	$(1,1,1,0)$	$B_0 \oplus B_1 = B_2^\perp$	$C_0 \oplus C_1 = C_2^\perp$
9	$(0,0,0,1)$	B_2	C_2
10	$(0,1,0,1)$	B_2	$C_0 \oplus C_2 = C_1^\perp$
11	$(1,0,0,1)$	$B_0 \oplus B_2 = B_1^\perp$	C_2
12	$(1,1,0,1)$	$B_0 \oplus B_2 = B_1^\perp$	$C_0 \oplus C_2 = C_1^\perp$
13	$(0,0,1,1)$	$B_1 \oplus B_2 = B_0^\perp$	$C_1 \oplus C_2 = C_0^\perp$
14	$(0,1,1,1)$	$B_1 \oplus B_2 = B_0^\perp$	$C_0 \oplus C_1 \oplus C_2 = \mathbb{R}^{v_2}$
15	$(1,0,1,1)$	$B_0 \oplus B_1 \oplus B_2 = \mathbb{R}^{v_1}$	$C_1 \oplus C_2 = C_0^\perp$
16	$(1,1,1,1)$	$B_0 \oplus B_1 \oplus B_2 = \mathbb{R}^{v_1}$	$C_0 \oplus C_1 \oplus C_2 = \mathbb{R}^{v_2}$

Table A.2.: Description of $\Phi(\mathcal{G}^{[u_1\{v_1v_2\}]})$.

	Relations between $\mu_1,\dots,\mu_6$	X^{v_1}	X^{v_2}
1	$\mu_1=\mu_2=\mu_3=0,\ \mu_4=\mu_5=\mu_6=0$		
2	$\mu_1=\mu_2=\mu_3=0,\ \mu_4=-\mu_5,\ \mu_6=0$		$\begin{pmatrix}1\\-1\\0\end{pmatrix}$
3	$\mu_1=0,\ \mu_2=-\mu_3,\ \mu_4=\mu_5=\mu_6=0$	$\begin{pmatrix}0\\1\\-1\end{pmatrix}$	
4	$\mu_1=0,\ \mu_2=-\mu_3,\ \mu_4=-\mu_5,\ \mu_6=0$	$\begin{pmatrix}0\\1\\-1\end{pmatrix}$	$\begin{pmatrix}1\\-1\\0\end{pmatrix}$
5	$-\mu_1=\mu_2=\mu_3,\ \mu_4=\mu_5=-\mu_6$	$\begin{pmatrix}-1\\1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\-1\end{pmatrix}$
6	$-\mu_1=\mu_2=\mu_3,\ \mu_4+\mu_5+2\mu_6=0$	$\begin{pmatrix}-1\\1\\1\end{pmatrix}$	$\begin{pmatrix}-1&1\\1&1\\0&-1\end{pmatrix}$
7	$2\mu_1+\mu_2+\mu_3=0,\ \mu_4=\mu_5=-\mu_6$	$\begin{pmatrix}0&-1\\-1&1\\1&1\end{pmatrix}$	$\begin{pmatrix}1\\1\\-1\end{pmatrix}$
8	$2\mu_1+\mu_2+\mu_3=0,\ \mu_4+\mu_5+2\mu_6=0$	$\begin{pmatrix}0&-1\\-1&1\\1&1\end{pmatrix}$	$\begin{pmatrix}-1&1\\1&1\\0&-1\end{pmatrix}$
9	$\mu_1=2\mu_2=2\mu_3,\ 2\mu_4=2\mu_5=\mu_6$	$\begin{pmatrix}2\\1\\1\end{pmatrix}$	$\begin{pmatrix}1\\1\\2\end{pmatrix}$
10	$\mu_1=2\mu_2=2\mu_3,\ \mu_4+\mu_5-\mu_6=0$	$\begin{pmatrix}2\\1\\1\end{pmatrix}$	$\begin{pmatrix}-1&1\\1&1\\0&2\end{pmatrix}$
11	$\mu_1-\mu_2-\mu_3=0,\ 2\mu_4=2\mu_5=\mu_6$	$\begin{pmatrix}0&2\\-1&1\\1&1\end{pmatrix}$	$\begin{pmatrix}1\\1\\2\end{pmatrix}$
12	$\mu_1-\mu_2-\mu_3=0,\ \mu_4+\mu_5-\mu_6=0$	$\begin{pmatrix}0&2\\-1&1\\1&1\end{pmatrix}$	$\begin{pmatrix}-1&1\\1&1\\0&2\end{pmatrix}$
13	$\mu_1\in\mathbb{R},\ \mu_2=\mu_3,\ \mu_4=\mu_5,\ \mu_6\in\mathbb{R}$	$\begin{pmatrix}-1&2\\1&1\\1&1\end{pmatrix}$	$\begin{pmatrix}1&1\\1&1\\-1&2\end{pmatrix}$
14	$\mu_1\in\mathbb{R},\ \mu_2=\mu_3,\ \mu_4,\mu_5,\mu_6\in\mathbb{R}$	$\begin{pmatrix}-1&2\\1&1\\1&1\end{pmatrix}$	I_3
15	$\mu_1,\mu_2,\mu_3\in\mathbb{R},\ \mu_4=\mu_5,\ \mu_6\in\mathbb{R}$	I_3	$\begin{pmatrix}1&1\\1&1\\-1&2\end{pmatrix}$
16	$\mu_1,\mu_2,\mu_3\in\mathbb{R},\ \mu_4,\mu_5,\mu_6\in\mathbb{R}$	I_3	I_3

Table A.3.: Relations between $\mu_1,\dots,\mu_6$ and design matrices for $\Phi(\mathcal{G}^{[u_1\{v_1v_2\}]})$.

	(α,β,γ)	Ω_{v_2}	Ω_{v_3}
1	$(0,0,0)$	0	0
2	$(0,0,1)$	$D_{\sqrt{2-\sqrt{2}}}$	$F_{\sqrt{2-\sqrt{2}}}$
3	$(0,1,0)$	$D_{\sqrt{2+\sqrt{2}}}$	$F_{\sqrt{2+\sqrt{2}}}$
4	$(0,1,1)$	$D_{\sqrt{2-\sqrt{2}}} \oplus D_{\sqrt{2+\sqrt{2}}} = D_0^{\perp}$	$F_{\sqrt{2-\sqrt{2}}} \oplus F_{\sqrt{2+\sqrt{2}}} = \mathbb{R}^{v_3}$
5	$(1,0,0)$	D_0	0
6	$(1,0,1)$	$D_0 \oplus D_{\sqrt{2-\sqrt{2}}} = D^{\perp}_{\sqrt{2+\sqrt{2}}}$	$F_{\sqrt{2-\sqrt{2}}}$
7	$(1,1,0)$	$D_0 \oplus D_{\sqrt{2+\sqrt{2}}} = D^{\perp}_{\sqrt{2-\sqrt{2}}}$	$F_{\sqrt{2+\sqrt{2}}}$
8	$(1,1,1)$	$D_0 \oplus D_{\sqrt{2-\sqrt{2}}} \oplus D_{\sqrt{2+\sqrt{2}}} = \mathbb{R}^{v_2}$	$F_{\sqrt{2-\sqrt{2}}} \oplus F_{\sqrt{2+\sqrt{2}}} = \mathbb{R}^{v_3}$

Table A.4.: Description of $\Phi(\mathcal{G}^{[u_2\{v_2v_3\}]})$.

	Relations between $\mu_4,\dots,\mu_8$	X^{v_2}	X^{v_3}
1	$\mu_4=\mu_5=\mu_6=0,\ \mu_7=\mu_8=0$		
2	$\sqrt{2}\mu_4=\sqrt{2}\mu_5=-\mu_6,\ (1+\sqrt{2})\mu_7=-\mu_8$	$\begin{pmatrix}1\\1\\-\sqrt{2}\end{pmatrix}$	$\begin{pmatrix}1\\-1-\sqrt{2}\end{pmatrix}$
3	$\sqrt{2}\mu_4=\sqrt{2}\mu_5=\mu_6,\ (1-\sqrt{2})\mu_7=-\mu_8$	$\begin{pmatrix}1\\1\\\sqrt{2}\end{pmatrix}$	$\begin{pmatrix}1\\-1+\sqrt{2}\end{pmatrix}$
4	$\mu_4=\mu_5,\ \ \mu_6\in\mathbb{R},\ \mu_7,\mu_8\in\mathbb{R}$	$\begin{pmatrix}1&1\\1&1\\\sqrt{2}&-\sqrt{2}\end{pmatrix}$	I_2
5	$\mu_4=-\mu_5,\ \mu_6=0,\ \ \mu_7=\mu_8=0$	$\begin{pmatrix}-1\\1\\0\end{pmatrix}$	
6	$\mu_4+\mu_5+\sqrt{2}\mu_6=0,\ (1+\sqrt{2})\mu_7=-\mu_8$	$\begin{pmatrix}-1&1\\1&1\\0&-\sqrt{2}\end{pmatrix}$	$\begin{pmatrix}1\\-1-\sqrt{2}\end{pmatrix}$
7	$\mu_4+\mu_5-\sqrt{2}\mu_6=0,\ (1-\sqrt{2})\mu_7=-\mu_8$	$\begin{pmatrix}-1&1\\1&1\\0&\sqrt{2}\end{pmatrix}$	$\begin{pmatrix}1\\-1+\sqrt{2}\end{pmatrix}$
8	$\mu_4,\mu_5,\mu_6\in\mathbb{R},\ \mu_7,\mu_8\in\mathbb{R}$	I_3	I_2

Table A.5.: Relations between $\mu_4,\dots,\mu_8$ and design matrices for $\Phi(\mathcal{G}^{[u_2\{v_2v_3\}]})$.

Appendix B.

Sizes of Colouring Classes

We observe first that each partition of $V = [4]$ can be obtained by applying a permutation in $S([4])$ to one of

(a) $\{\{1\},\{2\},\{3\},\{4\}\}$ (b) $\{\{1,2\},\{3\},\{4\}\}$ (c) $\{\{1,2\},\{3,4\}\}$

(d) $\{\{1,2,3\},\{4\}\}$ (e) $\{\{1,2,3,4\}\}$.

B.1. Edge Regular Colourings $\mathcal{B}_V$

By definition of edge regularity, no edge colour may appear in two different components. Given the vertex colouring of a graph, the total number of edge colourings which conform with the criterion is given by the product of the number of possible edge colourings of each component.

In case (a), there are four singleton vertex colour classes and therefore $\binom{4}{2} = 6$ $u\{vw\}$-components and no $u\{vv\}$-components. Each $u\{vw\}$-component may either contain an edge or not which, by (5.9), gives $\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right] = B_2 = 2$ possibilities.

The total number of graph colourings in $\mathcal{B}_{[4]}$ in case (a) is therefore

$$\text{(a)} \quad \left[\sum_{k=0}^{1} \binom{1}{k} B_k\right]^6 = B_2^6 = 64.$$

Analogously, for the other cases we get

$$\begin{array}{ll}
\text{(b)} & \left[\sum_{k=0}^{1} \binom{1}{k} B_k\right]^2 \left[\sum_{k=0}^{2} \binom{2}{k} B_k\right]^2 = B_2^2 B_3^2 = 100 \\
\text{(c)} & \left[\sum_{k=0}^{1} \binom{1}{k} B_k\right]^2 \left[\sum_{k=0}^{4} \binom{4}{k} B_k\right] = B_2^2 B_5 = 208 \\
\text{(d)} & \left[\sum_{k=0}^{3} \binom{3}{k} B_k\right]^2 = B_4^2 = 225 \\
\text{(e)} & \sum_{k=0}^{6} \binom{6}{k} B_k = B_7 = 877.
\end{array}$$

Accounting for the number of different vertex colourings of types (a) through (e), the above give a total of

$$|\mathcal{B}_{[4]}| = \binom{4}{0}64 + \binom{4}{2}100 + \frac{1}{2}\binom{4}{2}208 + \binom{4}{3}225 + \binom{4}{4}877 = 3,065.$$

We further obtain $|\mathcal{B}_{[5]}|$. Each partition of $V = [5]$ can be obtained by applying a permutation in $S([5])$ to one of

(a) $\{\{1\},\{2\},\{3\},\{4\},\{5\}\}$ (b) $\{\{1,2\},\{3\},\{4\},\{5\}\}$ (c) $\{\{1,2\},\{3,4\},\{5\}\}$

(d) $\{\{1,2\},\{3,4,5\}\}$ (e) $\{\{1,2,3\},\{4\},\{5\}\}$ (f) $\{\{1,2,3,4\},\{5\}\}$

(g) $\{\{1,2,3,4,5\}\}$.

Thus, repeating the above steps for $V = [5]$, the numbers of edge colourings for each of the partitions of V are given by

(a) $\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right]^{\binom{5}{2}} = B_2^{10} = 1,024$

(b) $\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right]\left[\sum_{k=0}^{2}\binom{2}{k}B_k\right]^3\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right]^3 = B_2^4B_3^3 = 2,000$

(c) $\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right]^2\left[\sum_{k=0}^{2}\binom{2}{k}B_k\right]^2\left[\sum_{k=0}^{4}\binom{4}{k}B_k\right] = B_2^2B_3^2B_5 = 5,200$

(d) $\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right]\left[\sum_{k=0}^{3}\binom{3}{k}B_k\right]\left[\sum_{k=0}^{6}\binom{6}{k}B_k\right] = B_2B_4B_7 = 26,310$

(e) $\left[\sum_{k=0}^{3}\binom{3}{k}B_k\right]\left[\sum_{k=0}^{3}\binom{3}{k}B_k\right]\left[\sum_{k=0}^{3}\binom{3}{k}B_k\right]\left[\sum_{k=0}^{1}\binom{1}{k}B_k\right] = B_4^3B_2 = 6,750$

(f) $\left[\sum_{k=0}^{6}\binom{6}{k}B_k\right]\left[\sum_{k=0}^{4}\binom{4}{k}B_k\right] = B_7B_5 = 45,604$

(g) $\sum_{k=0}^{10}\binom{10}{k}B_k = B_{11} = 678,570.$

which, accounting for the number of different vertex colourings of types (a) through (g), gives

$$\begin{aligned}|\mathcal{B}_{[5]}| &= \binom{5}{0}1,024 + \binom{5}{2}2,000 + \frac{1}{2}\binom{5}{2}\binom{3}{2}5,200 + \binom{5}{2}26,310 + \\ &\qquad + \binom{5}{3}6,750 + \binom{5}{4}45,604 + \binom{5}{5}678,570 \\ &= 1,336,214.\end{aligned}$$

B.2. Regular Colourings $\mathcal{R}_V$

We repeat the counting process described for $\mathcal{B}_V$ on $\mathcal{R}_V$. All graphs in case (a) have a regular colouring and therfore contribute $2^6 = 64$ regular graph colourings to $\mathcal{R}_{[4]}$.

The regular colourings for (b) can all be obtained by dropping edge colour classes from the graph displayed in Figure B.1, in total giving $2^4 = 16$ graphs as there are four edge colour classes.

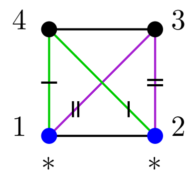


Figure B.1.: Regular colouring for $V = [4]$ and $\mathcal{V}$ of type (b).

All regular colourings in case (c) can be obtained from the graph in Figure B.2. This time we may drop edge colour classes, giving $2^4 = 16$ graphs, and/or merge the green and purple edge colour class, which gives another 4 graphs.

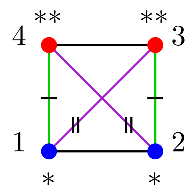


Figure B.2.: Regular colouring for $V = [4]$ and $\mathcal{V}$ of type (c).

All regular colourings in case (d) can be obtained by dropping edge colour classes from the graph in Figure B.3, thus giving $2^2 = 4$ graphs.

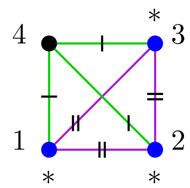


Figure B.3.: Regular colouring for $V = [4]$ and $\mathcal{V}$ of type (d).

All regular colourings in case (e) may be obtained from the graph in Figure B.4 by dropping and merging of edge colour classes, giving a total number of graphs of

$$\sum_{k=0}^{3} \binom{3}{k} B_k = B_4 = 15.$$

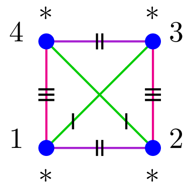


Figure B.4.: Regular colouring for $V = [4]$ and $\mathcal{V}$ of type (e).

Accounting for the number of different vertex colourings of types (a) through (e), the above give a total of

$$|\mathcal{R}_{[4]}| = \binom{4}{0}64 + \binom{4}{2}16 + \frac{1}{2}\binom{4}{2}20 + \binom{4}{3}4 + \binom{4}{4}15 = 251.$$

B.3. Vertex Regular Colourings $\mathcal{P}_V$

$|\mathcal{P}_V|$ can be found by counting the number of graphs which can be obtained from graphs in $\mathcal{R}_V$ by merging edge colour classes.

The number of edge colour classes in each of the cases (a)-(e) in $\mathcal{R}_V$ is: (a) 6, (b) 4, (c) 4, (d) 2 and (e) 3. Thus the total number of permissible edge colourings in each case is given by

$$\text{(a)} \sum_{k=0}^{6} \binom{6}{k} B_k = B_7 = 877 \qquad \text{(b)} \sum_{k=0}^{4} \binom{4}{k} B_k = B_5 = 52$$

$$\text{(c)} \sum_{k=0}^{4} \binom{4}{k} B_k = B_5 = 52 \quad \text{(d)} \sum_{k=0}^{2} \binom{2}{k} B_k = B_3 = 5 \quad \text{(e)} \sum_{k=0}^{3} \binom{3}{k} B_k = B_4 = 15.$$

We thus obtain

$$|\mathcal{P}_{[4]}| = \binom{4}{0}877 + \binom{4}{2}52 + \frac{1}{2}\binom{4}{2}52 + \binom{4}{3}5 + \binom{4}{4}15 = 1,380.$$

B.4. Permutation Generated Colourings Π_V

By Lemma 5.8, all colourings in $\Pi_{[4]}$ can be found by (i) determining all colourings $(\mathcal{V}_{\langle\sigma\rangle}, \mathcal{E}_{\langle\sigma\rangle})$ for $\sigma \in S([4])$, (ii) determining all colourings which can be obtained from (i) by the meet operation $\wedge$, and (iii) finding all colourings which can be obtained from (i) and (ii) by dropping edge colour classes.

For (i), $S([4])$ contains $4! = 24$ permutations, which generate 17 distinct colourings of the complete graph, given in Figure B.5. (The difference of 7 is due to the fact that $\langle\sigma\rangle = \langle\sigma^{-1}\rangle$ for all σ. For 17 permutations $\sigma = \sigma^{-1}$, while for $\sigma = (123), (124), (134), (234), (1234), (1243), (1324)$ the two are different.)

For (ii), we obtain 5 additional colourings of the complete graph, given in Figure B.6 together with one of their generating groups.

We obtain:

$$\begin{aligned}
|\Pi_{[4]}| &= N_1 + 6N_2 + 4N_8 + 3N_{12} + 3(N_{15} - 1) + 3(N_{18} - 4) + (N_{21} - 4) + N_{22} \\
&= 2^6 + 6 \cdot 2^4 + 4 \cdot 2^2 + 3 \cdot 2^4 + 3(2^2 - 1) + 3(2^3 - 4) + (2^3 - 4) + 2 \\
&= 251
\end{aligned}$$

where N_i denotes the number of graphs one can obtain from graph $\mathcal{G}_i$ by dropping edge colour classes. The subtracted correction terms prevent some graphs to be counted more than once.

Note that we have established that $\Pi_{[4]} = \mathcal{R}_{[4]}$, as $\Pi_{[4]} \subseteq \mathcal{R}_{[4]}$ and $|\Pi_{[4]}| = |\mathcal{R}_{[4]}|$.

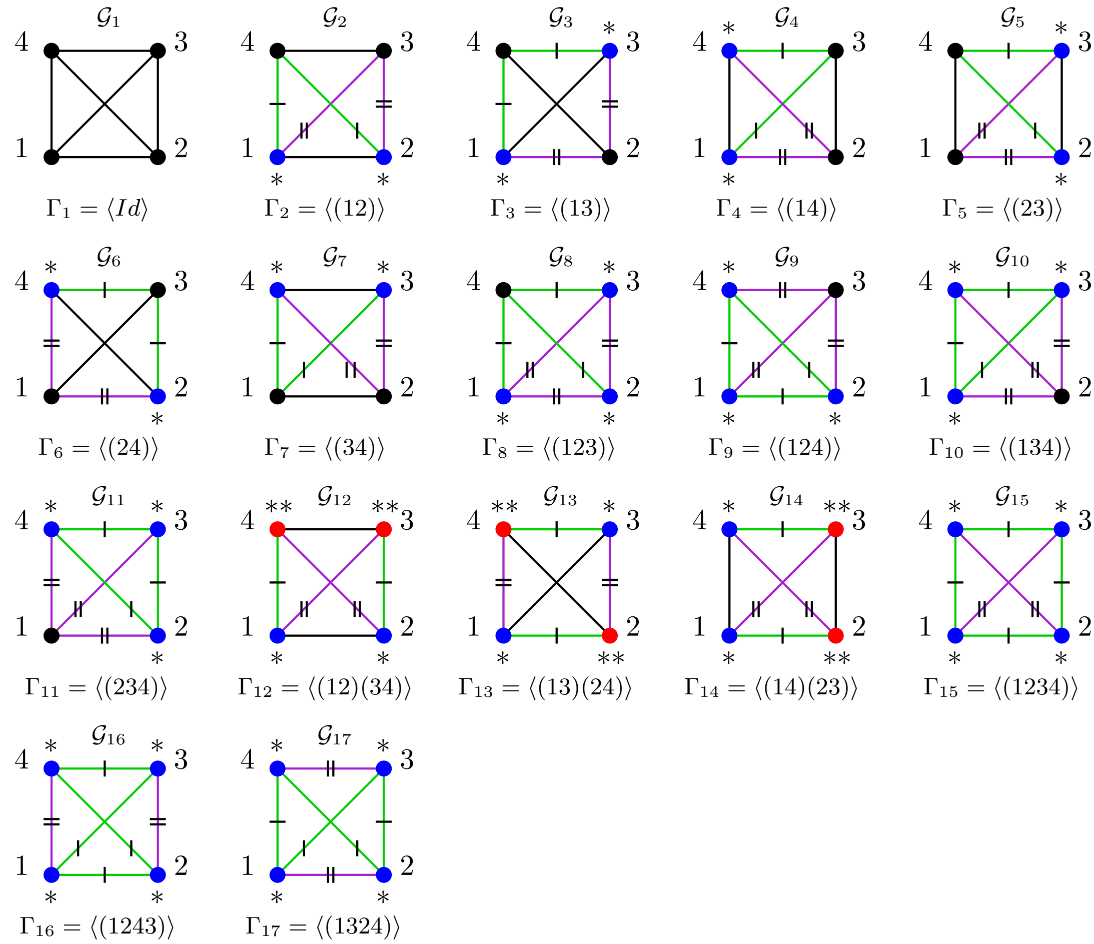


Figure B.5.: $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$ with $\Gamma = \langle \sigma \rangle$ for some $\sigma \in S([4])$.

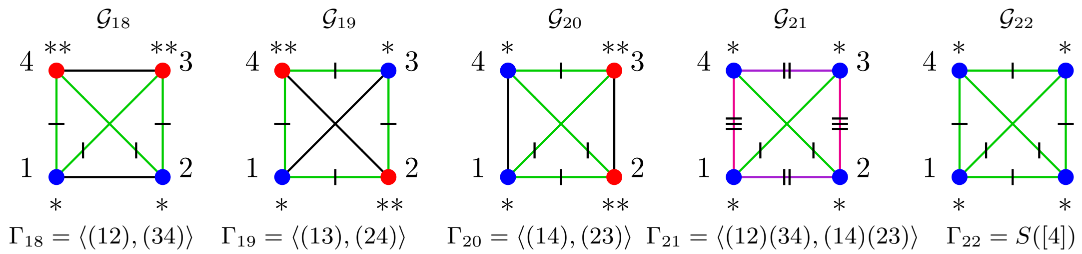


Figure B.6.: $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$ with $\Gamma \leq S([4])$ generated by more than one permutation.

Appendix C.

`GraphCheck` Algorithms

`GraphCheck` is a program we have written in Java which tests vertex and edge colourings of graphs for the four colouring regularities defined in this thesis. If a graph colouring does not fall into the desired colouring class, `GraphCheck` finds the coarsest refinement of the colouring which does.

Core computations are performed by Brendan McKay's computer program `nauty 2.4` (McKay, 2010). The `nauty` functions `GraphCheck` calls are (i) `nauty()` and (ii) `refine()`, which both take a vertex coloured graph as their argument and return (i) its automorphism group and (ii) the coarsest refinement of the vertex colouring which is equitable with respect to the entered graph.

Below we give the implemented algorithms for the four colouring classes, together with the output of each algorithm for the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{\{0, 2\}, \{1, 3\}\}$ and $\mathcal{E} = \{\{01, 13, 23\}, \{03\}, \{12\}\}$ in Figure C.1. As `nauty` numbers vertices of graphs starting from zero, so does `GraphCheck`.

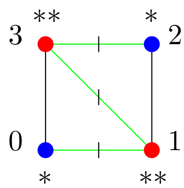


Figure C.1.: Vertex and edge coloured graph $\mathcal{G}$ used for illustration.

$\mathcal{G}$ is passed to `GraphCheck` with the commands given below. `n` stands for the number of vertices, `g` indicates that edge set of the graph is to be given next, `f` precedes the vertex colouring, which is given in brackets `[]` with colour classes separated by `|`, and `F` precedes the edge colouring, also given in brackets `[]` with colour classes separated by `|` and end vertices of an edge separated by `-`.

'`../data/graph3.txt`', appearing in all outputs, is the path to the corresponding input file. All four computed graph colourings are displayed in Figure C.2.

```
n = 4
g
0:1,3
1:2,3
2:3
f
[0,2 | 1,3]
F
[0-1,2-3,1-3]
```

C.1. Edge Regular Colourings $\mathcal{B}_V$

For a given vertex and edge coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$:

1. Test whether edges inside the same edge colour class connect the same vertex colour classes.
2. If the answer to 1. is 'Yes', return 'Colouring is edge regular!' and stop.
3. If the answer to 1. is 'No', refine $\mathcal{E}$ by splitting an edge colour class whenever it lies in more than one component, return obtained graph and stop.

```
******************************************
Output for graph ../data/graph3.txt
******************************************
Received graph colouring is not edge regular.
Found coarsest colouring refinement which is edge regular.
Vertex colour classes:
[ 0, 2 | 1, 3 ]
Edge colour classes:
[ 0-1, 2-3 | 0-3 | 1-2 | 1-3 ]
```

C.2. Regular Colourings $\mathcal{R}_V$

For a given vertex and edge coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$:

1. Transform $\mathcal{G}$ to its corresponding coloured factor graph $\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F)$.
2. Call the `nauty` function `refine()` with argument $\mathcal{G}_F$ and let it return the coarsest refinement $(\mathcal{V}' \cup \mathcal{N}')$ of $(\mathcal{V} \cup \mathcal{N})$ which is equitable with respect to G_F.
3. Compare $(\mathcal{V}' \cup \mathcal{N}')$ to $(\mathcal{V} \cup \mathcal{N})$.
4. If the colourings in 3. agree, return 'Colouring is regular!' and stop.

5. If the colourings in 3. do not agree, compute the vertex and edge coloured graph $\mathcal{G}'$ corresponding to $\mathcal{G}'_F = (\mathcal{V}' \cup \mathcal{N}', E_F)$, return $\mathcal{G}'$ and stop.

```
*****************************************
Output for graph ../data/graph3.txt
*****************************************
Received graph colouring is not regular.
Found coarsest colouring refinement which is regular.
Vertex colour classes:
[ 0 | 2 | 3 | 1 ]
Edge colour classes:
[ 1-3 | 0-1 | 2-3 | 0-3 | 1-2 ]
```

C.3. Vertex Regular Colourings $\mathcal{P}_V$

For a given vertex and edge coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$:

1. Transform $\mathcal{G}$ to its coloured factor graph $\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F)$.
2. Call the `nauty` function `refine()` with argument $\mathcal{G}_F$ and let it return the coarsest refinement $(\mathcal{V}' \cup \mathcal{N}')$ of $(\mathcal{V} \cup \mathcal{N})$ which is equitable with respect to G_F.
3. Compare $\mathcal{V}'$ to $\mathcal{V}$.
4. If the colourings in 3. agree, return 'Colouring is vertex regular!' and stop.
5. If the colourings in 3. do not agree, compute the vertex and edge coloured graph $\mathcal{G}'$ corresponding to $\mathcal{G}'_F = (\mathcal{V}' \cup \mathcal{N}, E_F)$, return $\mathcal{G}'$ and stop.

```
****************************************
Output for graph ../data/graph3.txt
****************************************
Received graph colouring is not vertex regular.
Found coarsest colouring refinement which is vertex regular.
Vertex colour classes:
[ 0 | 2 | 3 | 1 ]
Edge colour classes:
[ 0-1, 1-3, 2-3 | 0-3 | 1-2 ]
```

C.4. Permutation Generated Colourings Π_V

For a given vertex and edge coloured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$:

1. Transform $\mathcal{G}$ to its corresponding coloured factor graph $\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F)$.
2. Call the `nauty` function `nauty()` with argument $\mathcal{G}_F$ and let it return the orbits $(\mathcal{V}' \cup \mathcal{N}')$ of $\Gamma = \mathrm{Aut}(\mathcal{V} \cup \mathcal{N}, E_F)$ in $(\mathcal{V} \cup \mathcal{N})$.
3. Compare $(\mathcal{V}' \cup \mathcal{N}')$ to $(\mathcal{V} \cup \mathcal{N})$.
4. If the colourings in 3. agree, return 'Colouring is permutation generated!' and stop.
5. If the colourings in 3. do not agree, find the vertex and edge coloured graph $\mathcal{G}'$ corresponding to $\mathcal{G}'_F = (\mathcal{V}' \cup \mathcal{N}', E_F)$, return $\mathcal{G}'$ and stop.

```
****************************************
Output for graph ../data/graph3.txt
****************************************
Received graph colouring is not permutation generated.
Found coarsest colouring refinement which is permutation generated.
Vertex colour classes:
[ 0 | 1 | 2 | 3 ]
Edge colour classes:
[ 0-1 | 0-3 | 1-2 | 1-3 | 2-3 ]
```

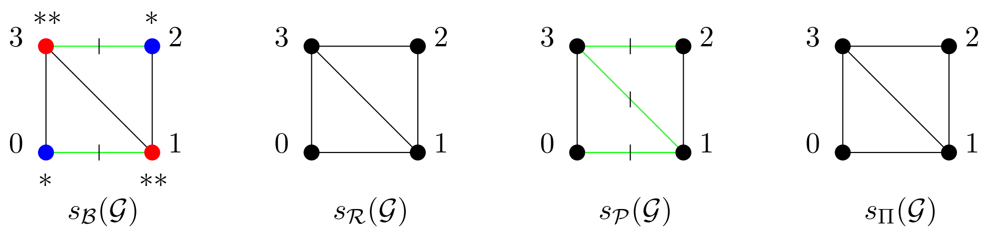


Figure C.2.: Colourings of $\mathcal{G}$ obtained by `GraphCheck`.

Note that the computed $s_{\mathcal{R}}(\mathcal{G})$ and $s_{\Pi}(\mathcal{G})$ are identical. This is precisely right as $\Pi_{[4]} = \mathcal{R}_{[4]}$, as shown in Appendix B.

Appendix D.

Models Fitted in Section 6.2.3

We denote the variables Algebra, Analysis, Mechanics, Statistics and Vectors by `al`, `an`, `me`, `st` and `ve` respectively. A model is defined through its differences from the saturated unrestricted model, using the following notation. If a model is represented by the complete graph in which each vertex and each edge forms a singleton colour class except for vertices `al` and `an` which are assigned the same colour, we denote the model by '`al+an`'. If additionally the edge between `me` and `st` is missing, the model is denoted by '`al+an, -me:st`'. If a further constraint is that edges '`an:st`' and '`al:st`' are of equal colour, we write '`al+an, -me:st, an:st+al:st`'.

We group the models by stage of the algorithm and indicate by an 'A' or 'R' whether the model was accepted or rejected. The minimally accepted models are underlined.

Stage 0

Fit of saturated model without constraints.

Stage 1

20 models fitted: 6 accepted, 14 rejected.

- 10 with one vertex pair of equal colour.
- 10 with one missing edge.

1.	`me+ve`	R
2.	`me+al`	R
3.	`me+an`	R
4.	`me+st`	A
5.	`ve+al`	R
6.	`ve+an`	A
7.	`ve+st`	R
8.	`al+an`	R
9.	`al+st`	R
10.	`an+st`	R
11.	`-me:ve`	R
12.	`-me:al`	R
13.	`-me:an`	A
14.	`-me:st`	A
15.	`-ve:al`	R
16.	`-ve:an`	A
17.	`-ve:st`	A
18.	`-al:an`	R
19.	`-al:st`	R
20.	`-an:st`	R

Stage 2

21 models fitted: 19 accepted, 2 rejected.

- 1 with two vertex pairs of equal colour.

- 6 with one vertex pair of equal colour and one edge pair of equal colour.
- 8 with one vertex pair of equal colour and one missing edge.
- 6 with two missing edges.

1.	`me+st, ve+an`	A
2.	`me+st, me:ve+st:ve`	R
3.	`me+st, me:al+st:al`	A
4.	`me+st, me:an+st:an`	A
5.	`ve+an, ve:me+an:me`	A
6.	`ve+an, ve:st+an:st`	A
7.	`ve+an, ve:al+an:al`	R
8.	`me+st, -me:an`	A
9.	`me+st, -me:st`	A
10.	`me+st, -ve:an`	A
11.	`me+st, -ve:st`	A
12.	`ve+an, -me:an`	A
13.	`ve+an, -me:st`	A
14.	`ve+an, -ve:an`	A
15.	`ve+an, -ve:st`	A
16.	`-me:an, -me:st`	A
17.	`-me:an, -ve:an`	A
18.	`-me:an, -ve:st`	A
19.	`-me:st, -ve:an`	A
20.	`-me:st, -ve:st`	A
21.	`-ve:an, -ve:st`	A

Stage 3

41 models fitted: 40 accepted, 1 rejected.

- 6 with two vertex pairs of equal colour and one edge pair of equal colour.
- 4 with two vertex pairs of equal colour and one missing edge.

- 2 with one vertex pair of equal colour and two edge pairs of equal colour.
- 13 with one vertex pair of equal colour, one edge pair of equal colour and one missing edge.
- 12 with one vertex pair of equal colour and two missing edges.
- 4 with three missing edges.

1.	`me+st, ve+an, me:al+st:al`	A
2.	`me+st, ve+an, me:an+st:an`	A
3.	`me+st, ve+an, ve:me+an:me`	A
4.	`me+st, ve+an, ve:st+an:st`	A
5.	`me+st, ve+an, me:ve+st:an`	A
6.	`me+st, ve+an, an:me+st:ve`	A
7.	`me+st, ve+an, -me:an`	A
8.	`me+st, ve+an, -me:st`	A
9.	`me+st, ve+an, -ve:an`	A
10.	`me+st, ve+an, -ve:st`	A
11.	`me+st, me:al+st:al, me:an+st:an`	R
12.	`ve+an, ve:st+an:st, ve:me+an:me`	A
13.	`me+st, me:al+st:al, -me:an`	A
14.	`me+st, me:al+st:al, -me:st`	A
15.	`me+st, me:al+st:al, -ve:an`	A
16.	`me+st, me:al+st:al, -ve:st`	A
17.	`me+st, me:an+st:an, -me:st`	A
18.	`me+st, me:an+st:an, -ve:an`	A
19.	`me+st, me:an+st:an, -ve:st`	A
20.	`ve+an, ve:me+an:me, -me:st`	A
21.	`ve+an, ve:me+an:me, -ve:an`	A
22.	`ve+an, ve:me+an:me, -ve:st`	A
23.	`ve+an, ve:st+an:st, -me:an`	A
24.	`ve+an, ve:st+an:st, -me:st`	A
25.	`ve+an, ve:st+an:st, -ve:an`	A
26.	`me+st, -me:an, -me:st`	A
27.	`me+st, -me:an, -ve:an`	A

28.	`me+st, -me:an, -ve:st`	A
29.	`me+st, -me:st, -ve:an`	A
30.	`me+st, -me:st, -ve:st`	A
31.	`me+st, -ve:an, -ve:st`	A
32.	`ve+an, -me:an, -me:st`	A
33.	`ve+an, -me:an, -ve:an`	A
34.	`ve+an, -me:an, -ve:st`	A
35.	`ve+an, -me:st, -ve:an`	A
36.	`ve+an, -me:st, -ve:st`	A
37.	`ve+an, -ve:an, -ve:st`	A
38.	`-me:an, -me:st, -ve:an`	A
39.	`-me:an, -me:st, -ve:st`	A
40.	`-me:an, -ve:an, -ve:st`	A
41.	`-me:st, -ve:an, -ve:st`	A

Stage 4

56 models fitted: 56 accepted, 0 rejected.

- 6 with two vertex pairs of equal colour and two edge pairs of equal colour.
- 20 with two vertex pairs of equal colour, one edge pair of equal colour and one missing edge.
- 6 with two vertex pairs of equal colour and two missing edges.
- 15 with one vertex pair of equal colour, one edge pair of equal colour and two missing edges.
- 8 with one vertex pair of equal colour and three missing edges.
- 1 with four missing edges.

1.	`me+st, ve+an, ve:me+an:me, me:al+st:al`	A
2.	`me+st, ve+an, ve:me+an:me+st:an`	A

3.	`me+st, ve+an, me:an+ve:st+st:an`	A
4.	`me+st, ve+an, ve:st+an:st, me:al+st:al`	A
5.	`me+st, ve+an, me:ve+st:an, me:al+st:al`	A
6.	`me+st, ve+an, an:me+st:ve, me:al+st:al`	A
7.	`me+st, ve+an, ve:me+an:me, -me:st`	A
8.	`me+st, ve+an, ve:me+an:me, -ve:an`	A
9.	`me+st, ve+an, ve:me+an:me, -ve:st`	A
10.	`me+st, ve+an, ve:st+an:st, -me:an`	A
11.	`me+st, ve+an, ve:st+an:st, -me:st`	A
12.	`me+st, ve+an, ve:st+an:st, -ve:an`	A
13.	`me+st, ve+an, me:al+st:al, -me:an`	A
14.	`me+st, ve+an, me:al+st:al, -me:st`	A
15.	`me+st, ve+an, me:al+st:al, -ve:an`	A
16.	`me+st, ve+an, me:al+st:al, -ve:st`	A
17.	`me+st, ve+an, me:an+st:an, -me:st`	A
18.	`me+st, ve+an, me:an+st:an, -ve:an`	A
19.	`me+st, ve+an, me:an+st:an, -ve:st`	A
20.	`me+st, ve+an, me:ve+st:an, -me:an`	A
21.	`me+st, ve+an, me:ve+st:an, -me:st`	A
22.	`me+st, ve+an, me:ve+st:an, -ve:an`	A
23.	`me+st, ve+an, me:ve+st:an, -ve:st`	A
24.	`me+st, ve+an, -an:me, -st:ve`	A
25.	`me+st, ve+an, an:me+st:ve, -me:st`	A
26.	`me+st, ve+an, an:me+st:ve, -ve:an`	A
27.	`me+st, ve+an, -me:an, -me:st`	A
28.	`me+st, ve+an, -me:an, -ve:an`	A
29.	`me+st, ve+an, -me:an, -ve:st`	A
30.	`me+st, ve+an, -me:st, -ve:an`	A
31.	`me+st, ve+an, -me:st, -ve:st`	A
32.	`me+st, ve+an, -ve:an, -ve:st`	A
33.	`me+st, me:al+st:al, -me:an, -me:st`	A
34.	`me+st, me:al+st:al, -me:an, -ve:an`	A
35.	`me+st, me:al+st:al, -me:an, -ve:st`	A
36.	`me+st, me:al+st:al, -me:st, -ve:an`	A
37.	`me+st, me:al+st:al, -me:st, -ve:st`	A
38.	`me+st, me:al+st:al, -ve:an, -ve:st`	A

39.	`me+st, me:an+st:an, -me:st, -ve:an`	A
40.	`me+st, me:an+st:an, -me:st, -ve:st`	A
41.	`me+st, me:an+st:an, -ve:an, -ve:st`	A
42.	`ve+an, ve:me+an:me, -me:st, -ve:an`	A
43.	`ve+an, ve:me+an:me, -me:st, -ve:st`	A
44.	`ve+an, ve:me+an:me, -ve:an, -ve:st`	A
45.	`ve+an, ve:st+an:st, -me:an, -me:st`	A
46.	`ve+an, ve:st+an:st, -me:an, -ve:an`	A
47.	`ve+an, ve:st+an:st, -me:st, -ve:an`	A
48.	`me+st, -me:an, -me:st, -ve:an`	A
49.	`me+st, -me:an, -me:st, -ve:st`	A
50.	`me+st, -me:an, -ve:an, -ve:st`	A
51.	`me+st, -me:st, -ve:an, -ve:st`	A
52.	`ve+an, -me:an, -me:st, -ve:an`	A
53.	`ve+an, -me:an, -me:st, -ve:st`	A
54.	`ve+an, -me:an, -ve:an, -ve:st`	A
55.	`ve+an, -me:st, -ve:an, -ve:st`	A
56.	`-me:an, -me:st, -ve:an, -ve:st`	A

Stage 5

55 models fitted: 55 accepted, 0 rejected.

- 18 with two vertex pairs of equal colour, two edge pairs of equal colour and one missing edge.
- 24 with two vertex pairs of equal colour, one edge pair of equal colour and two missing edges.
- 4 with two vertex pairs of equal colour and three missing edges.
- 7 with one vertex pair of equal colour, one edge pair of equal colour and three missing edges.
- 2 with one vertex pair of equal colour and four missing edges.

1.	`me+st, ve+an, ve:me+an:me, me:al+st:al, -me:st`	A
2.	`me+st, ve+an, ve:me+an:me, me:al+st:al, -ve:an`	A
3.	`me+st, ve+an, ve:me+an:me, me:al+st:al, -ve:st`	A
4.	`me+st, ve+an, me:an+ve:st+st:an, -me:st`	A
5.	`me+st, ve+an, me:an+ve:st+st:an, -ve:an`	A
6.	`me+st, ve+an, ve:me+me:an+st:an, -me:st`	A
7.	`me+st, ve+an, ve:me+an:me+st:an, -ve:an`	A
8.	`me+st, ve+an, ve:me+me:an+st:an, -ve:st`	A
9.	`me+st, ve+an, ve:st+an:st, me:al+st:al, -me:an`	A
10.	`me+st, ve+an, ve:st+an:st, me:al+st:al, -me:st`	A
11.	`me+st, ve+an, ve:st+an:st, me:al+st:al, -ve:an`	A
12.	`me+st, ve+an, me:ve+st:an, me:al+st:al, -me:an`	A
13.	`me+st, ve+an, me:ve+st:an, me:al+st:al, -me:st`	A
14.	`me+st, ve+an, me:ve+st:an, me:al+st:al, -ve:an`	A
15.	`me+st, ve+an, me:ve+st:an, me:al+st:al, -ve:st`	A
16.	`me+st, ve+an, -an:me, -st:ve, me:al+st:al`	A
17.	`me+st, ve+an, an:me+st:ve, me:al+st:al, -me:st`	A
18.	`me+st, ve+an, an:me+st:ve, me:al+st:al, -ve:an`	A
19.	`me+st, ve+an, ve:me+an:me, -me:st, -ve:an`	A
20.	`me+st, ve+an, ve:me+an:me, -me:st, -ve:st`	A
21.	`me+st, ve+an, ve:me+an:me, -ve:an, -ve:st`	A
22.	`me+st, ve+an, ve:st+an:st, -me:an, -me:st`	A
23.	`me+st, ve+an, ve:st+an:st, -me:an, -ve:an`	A
24.	`me+st, ve+an, ve:st+an:st, -me:st, -ve:an`	A
25.	`ve+an, me+st, me:al+st:al, -me:an, -me:st`	A
26.	`ve+an, me+st, me:al+st:al, -me:an, -ve:an`	A
27.	`ve+an, me+st, me:al+st:al, -me:an, -ve:st`	A
28.	`ve+an, me+st, me:al+st:al, -me:st, -ve:an`	A
29.	`ve+an, me+st, me:al+st:al, -me:st, -ve:st`	A
30.	`ve+an, me+st, me:al+st:al, -ve:an, -ve:st`	A
31.	`ve+an, me+st, me:an+st:an, -me:st, -ve:an`	A
32.	`ve+an, me+st, me:an+st:an, -me:st, -ve:st`	A
33.	`ve+an, me+st, me:an+st:an, -ve:an, -ve:st`	A
34.	`me+st, ve+an, me:ve+st:an, -me:an, -me:st`	A
35.	`me+st, ve+an, me:ve+st:an, -me:an, -ve:an`	A
36.	`me+st, ve+an, me:ve+st:an, -me:an, -ve:st`	A

37.	`me+st, ve+an, me:ve+st:an, -me:st, -ve:an`	A
38.	`me+st, ve+an, me:ve+st:an, -me:st, -ve:st`	A
39.	`me+st, ve+an, me:ve+st:an, -ve:an, -ve:st`	A
40.	`me+st, ve+an, -an:me, -st:ve, -me:st`	A
41.	`me+st, ve+an, -st:ve, -me:an, -ve:an`	A
42.	`me+st, ve+an, an:me+st:ve, -me:st, -ve:an`	A
43.	`me+st, ve+an, -me:an, -me:st, -ve:an`	A
44.	`me+st, ve+an, -me:an, -me:st, -ve:st`	A
45.	`me+st, ve+an, -me:an, -ve:an, -ve:st`	A
46.	`me+st, ve+an, -me:st, -ve:an, -ve:st`	A
47.	`me+st, me:al+st:al, -me:an, -me:st, -ve:an`	A
48.	`me+st, me:al+st:al, -me:an, -me:st, -ve:st`	A
49.	`me+st, me:al+st:al, -me:an, -ve:an, -ve:st`	A
50.	`me+st, me:al+st:al, -me:st, -ve:an, -ve:st`	A
51.	`me+st, me:an+st:an, -me:st, -ve:an, -ve:st`	A
52.	`ve+an, ve:me+an:me, -me:st, -ve:an, -ve:st`	A
53.	`ve+an, ve:st+an:st, -me:an, -me:st, -ve:an`	A
54.	`me+st, -me:an, -me:st, -ve:an, -ve:st`	A
55.	`ve+an, -me:an, -me:st, -ve:an, -ve:st`	A

Stage 6

29 models fitted: 29 accepted, 0 rejected.

- 17 with two vertex pairs of equal colour, two edge pairs of equal colour and two missing edges.
- 10 with two vertex pairs of equal colour, one edge pair of equal colour and three missing edges.
- 1 with two vertex pairs of equal colour and four missing edges.
- 1 with one vertex pair of equal colour, one edge pair of equal colour and four missing edges.

1. `me+st, ve+an, ve:me+an:me, me:al+st:al, -me:st, -ve:an` A
2. `me+st, ve+an, ve:me+an:me, me:al+st:al, -me:st, -ve:st` A
3. `me+st, ve+an, ve:me+an:me, me:al+st:al, -ve:an, -ve:st` A
4. `me+st, ve+an, ve:me+an:me+st:an, -me:st, -ve:an` A
5. `me+st, ve+an, ve:me+an:me+st:an, -me:st, -ve:st` A
6. `me+st, ve+an, ve:me+an:me+st:an, -ve:an, -ve:st` A
7. `me+st, ve+an, me:an+ve:st+st:an, -me:st, -ve:an` A
8. `me+st, ve+an, ve:st+an:st, me:al+st:al, -me:an, -me:st` A
9. `me+st, ve+an, ve:st+an:st, me:al+st:al, -me:an, -ve:an` A
10. `me+st, ve+an, ve:st+an:st, me:al+st:al, -me:st, -ve:an` A
11. `me+st, ve+an, me:ve+st:an, me:al+st:al, -me:an, -ve:st` A
12. `me+st, ve+an, me:ve+st:an, me:al+st:al, -me:st, -ve:an` A
13. `me+st, ve+an, me:ve+st:an, me:al+st:al, -me:st, -ve:st` A
14. `me+st, ve+an, me:ve+st:an, me:al+st:al, -ve:an, -ve:st` A
15. `ve+an, me+st, me:al+st:al, -an:me, -st:ve, -me:st` A
16. `ve+an, me+st, me:al+st:al, -st:ve, -me:an, -ve:an` A
17. `ve+an, me+st, me:al+st:al, an:me+st:ve, -me:st, -ve:an` A
18. `me+st, ve+an, ve:me+an:me, -me:st, -ve:an, -ve:st` A
19. `me+st, ve+an, ve:st+an:st, -me:an, -me:st, -ve:an` A
20. `ve+an, me+st, me:al+st:al, -me:an, -ve:an, -ve:st` A
21. `ve+an, me+st, me:al+st:al, -me:st, -ve:an, -ve:st` A
22. `ve+an, me+st, me:an+st:an, -me:st, -ve:an, -ve:st` A
23. `me+st, ve+an, me:ve+st:an, -me:an, -me:st, -ve:an` A
24. `me+st, ve+an, me:ve+st:an, -me:an, -me:st, -ve:st` A
25. `me+st, ve+an, me:ve+st:an, -me:an, -ve:an, -ve:st` A
26. `me+st, ve+an, me:ve+st:an, -me:st, -ve:an, -ve:st` A
27. `me+st, ve+an, -an:me, -st:ve, -me:st, -ve:an` A
28. `me+st, ve+an, -me:an, -me:st, -ve:an, -ve:st` A
29. `me+st, me:al+st:al, -me:an, -me:st, -ve:an, -ve:st` A

Stage 7

9 models fitted: 9 accepted, 0 rejected.

- 8 with two vertex pairs of equal colour, two edge pairs of equal colour and three missing edges.

- 1 with two vertex pairs of equal colour, one edge pair of equal colour and four missing edges.

1.	`me+st, ve+an, ve:me+an:me, me:al+st:al, -me:st, -ve:an, -ve:st`	A
2.	`me+st, ve+an, ve:me+an:me+st:an, -me:st, -ve:an, -ve:st`	A
3.	`me+st, ve+an, ve:st+an:st, me:al+st:al, -me:an, -me:st, -ve:an`	A
4.	`me+st, ve+an, me:al+st:al, me:ve+st:an, -me:an, -me:st, -ve:an`	A
5.	`me+st, ve+an, me:al+st:al, me:ve+st:an, -me:an, -me:st, -ve:st`	A
6.	`me+st, ve+an, me:al+st:al, me:ve+st:an, -me:an, -ve:an, -ve:st`	A
7.	`me+st, ve+an, me:al+st:al, me:ve+st:an, -me:st, -ve:an, -ve:st`	A
8.	`me+st, ve+an, me:al+st:al, -st:ve, -me:an, -me:st, -ve:an`	A
9.	`me+st, ve+an, me:ve+st:an, -me:an, -me:st, -ve:an, -ve:st`	A

Stage 8

1 model fitted: 1 accepted, 0 rejected.

- 1 with two vertex pairs of equal colour, two edge pairs of equal colour and four missing edges.

1.	`me+st, ve+an, me:al+st:al, me:ve+st:an, -me:an, -me:st, -ve:an, -ve:st`	A

232 models in total: 215 accepted, 17 maximally rejected, 4 minimally accepted.

Appendix E.

Models Fitted in Section 6.3.2

We label the graphs representing the fitted models by $\mathcal{H}_i$ for $0 \leq i \leq 48$. The labelling by $\mathcal{G}_j$ for $1 \leq j \leq 22$ is the labelling used in Section B.1 in Appendix B. We circle elements in the given Hasse diagrams to indicate that the corresponding model was minimally accepted in the current stage of the algorithm.

Stage 0 consisted of fitting the saturated model with singleton colour classes, whose graph $\mathcal{H}_0$ is given in Figure E.1.

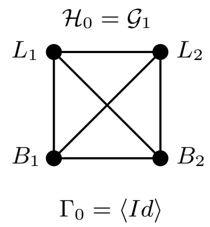


Figure E.1.: Stage 0: the saturated model $\mathcal{H}_0$.

The Hasse diagram of $\mathcal{K}_{[4]}$ is given in Figure E.2. $\mathcal{G}_1$ is circled as after stage 0 it is the only minimally accepted model in $\mathcal{K}_{[4]}$.

In stage 1 a total of 15 models was fitted: the models represented by the 6 graphs linked to $\mathcal{H}_0 = \mathcal{G}_1$ in $D(\mathcal{E}_{\langle Id \rangle})$ and those represented by the 9 graphs linked to

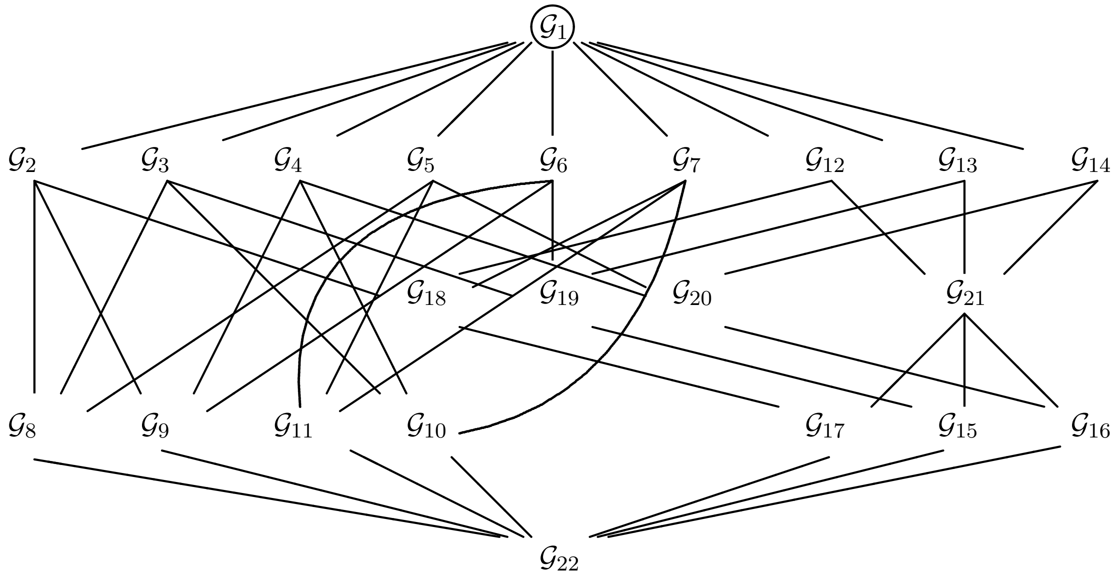


Figure E.2.: Hasse diagram of $\mathcal{K}_{[4]}$.

$\mathcal{H}_0 = \mathcal{G}_1$ in $\mathcal{K}_{[4]}$, of which 9 were accepted in total. We display the corresponding 15 graphs in Figure E.3. The graphs of the accepted models are given in colour, while the graphs representing the rejected models appear in gray. Their colouring may be inferred from the asterisks placed next to the vertices and the dashes on the edges.

The Hasse diagram of $\mathcal{K}_{[4]}$ in Figure E.4 illustrates the results of stage 1 for $\mathcal{K}_{[4]}$. The graphs of the rejected models, and their links, are coloured in gray and all minimally accepted models are circled.

Stage 2 continued to explore $\mathcal{K}_{[4]}$ by fitting models represented by graphs linked to the previously accepted ones, and moved down $D(\mathcal{E}_\Gamma)$ for all $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \in \mathcal{K}_{[4]}$ which have previously been accepted. 16 models were fitted in total, of which 4 lie in $D(\mathcal{E}_{\langle Id \rangle}) \setminus \mathcal{K}_{[4]}$, 2 lie in $\mathcal{K}_{[4]}$ and 10 lie in $D(\mathcal{E}_\Gamma) \setminus \mathcal{K}_{[4]}$ for $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \in \mathcal{K}_{[4]}$ accepted in stage 1. All models were accepted. The graphs are given in Figure E.5. Graph $\mathcal{H}_{22}$ is circled as it represents a minimally accepted model, i.e. all models smaller than

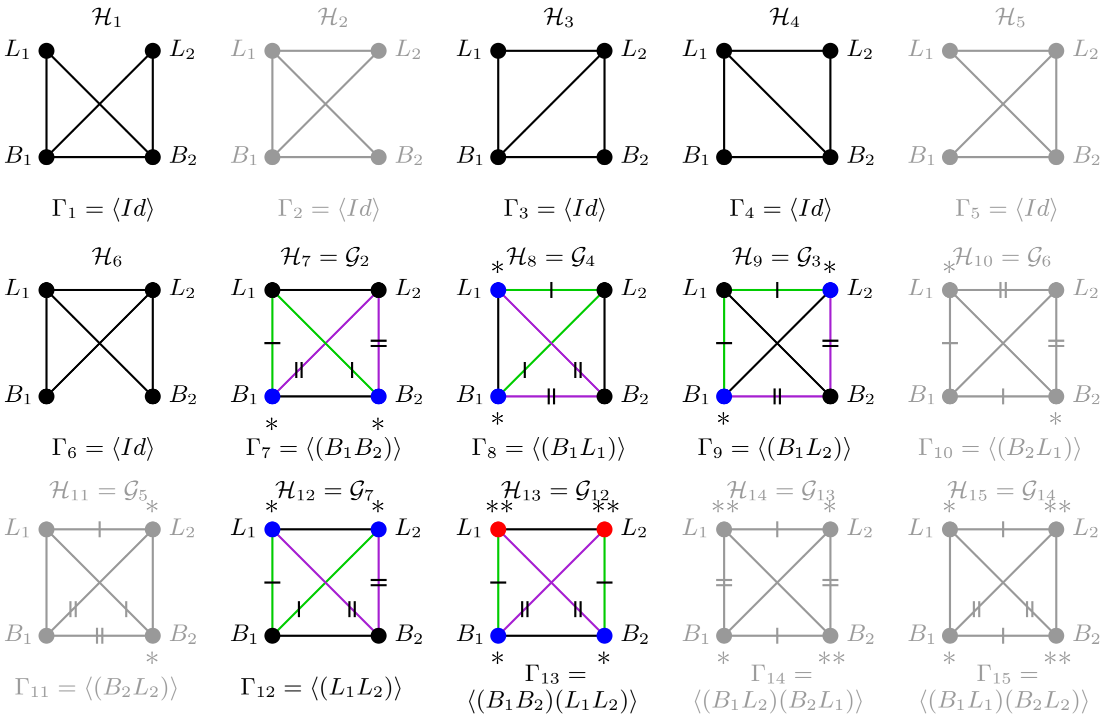


Figure E.3.: Stage 1: 15 models fitted of which 9 accepted and 6 rejected.

the one represented by $\mathcal{H}_{22}$ were rejected at some stage by the algorithm.

We give the updated Hasse diagram of $\mathcal{K}_{[4]}$ after stage 2 of the algorithm in Figure E.6. Now all of $\mathcal{K}_{[4]}$ has been covered by the search as each model is either rejected or accepted. All next steps of the algorithm lie in $D(\mathcal{E}_\Gamma)$ for $\Gamma \leq S(V)$ such that $(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma) \in \mathcal{K}_{[4]}$ has been previously accepted.

In total 13 models were fitted in stage 3, all of which were accepted and whose graphs are given in Figure E.7. 5 models are minimally accepted and are therefore circled. In stage 4, a total of 4 models was fitted, all displayed in Figure E.8, only 1 of which was rejected and the remaining 3 minimally accepted. The algorithm terminated with 9 minimally accepted models, whose graphs are displayed in Figure 6.6 in Section 6.3.2, after fitting 48 in total.

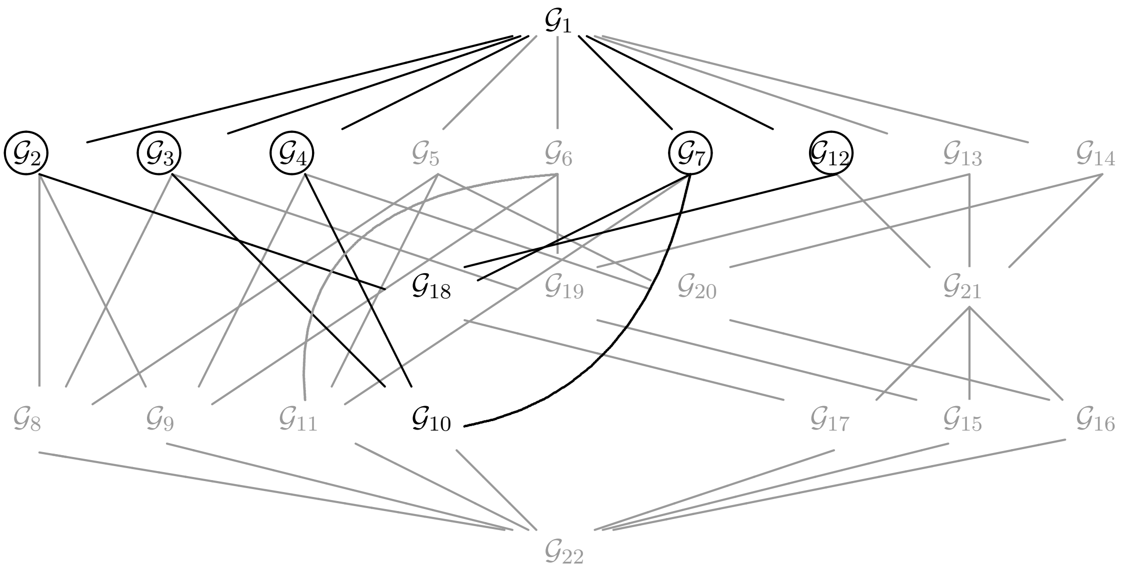


Figure E.4.: Minimally accepted colourings in $\mathcal{K}_{[4]}$ after stage 1.

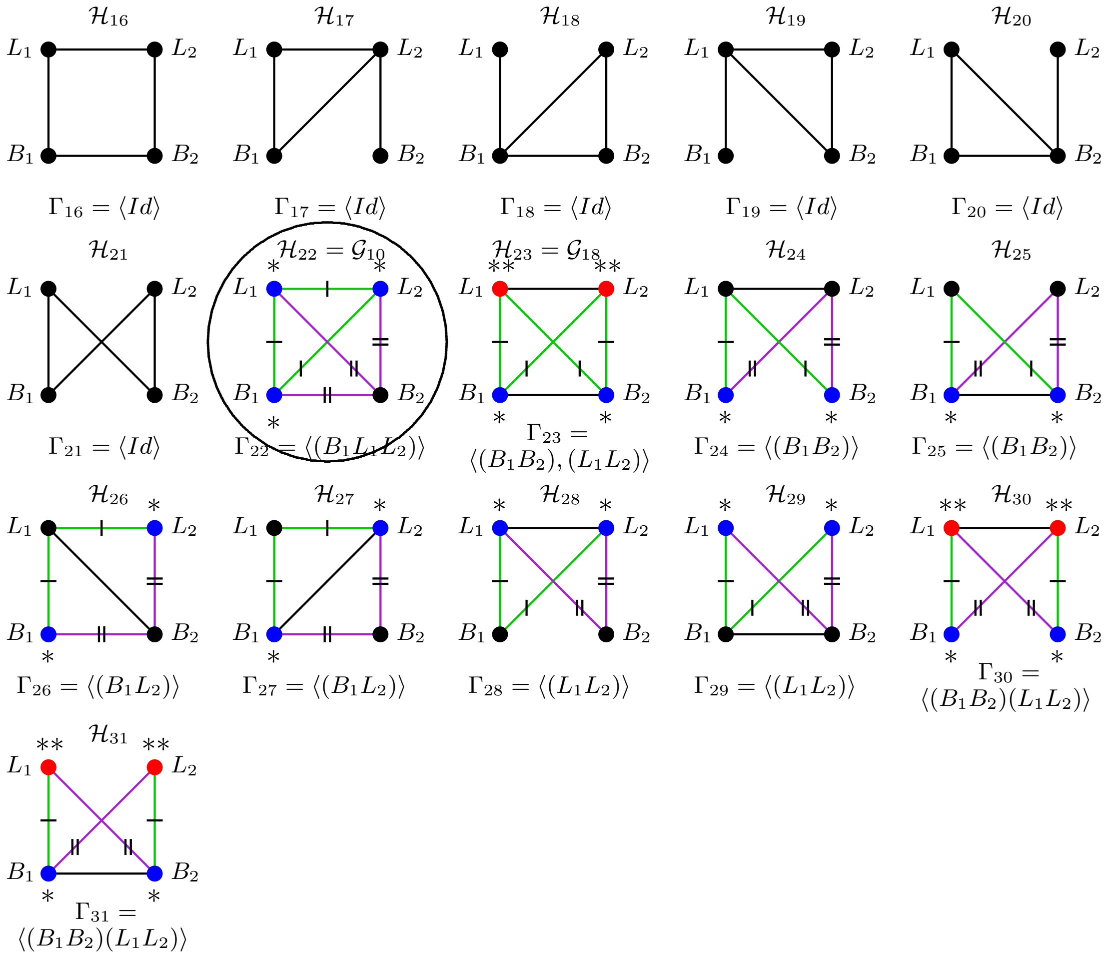


Figure E.5.: Stage 2: 16 models fitted of which 16 accepted and 0 rejected.

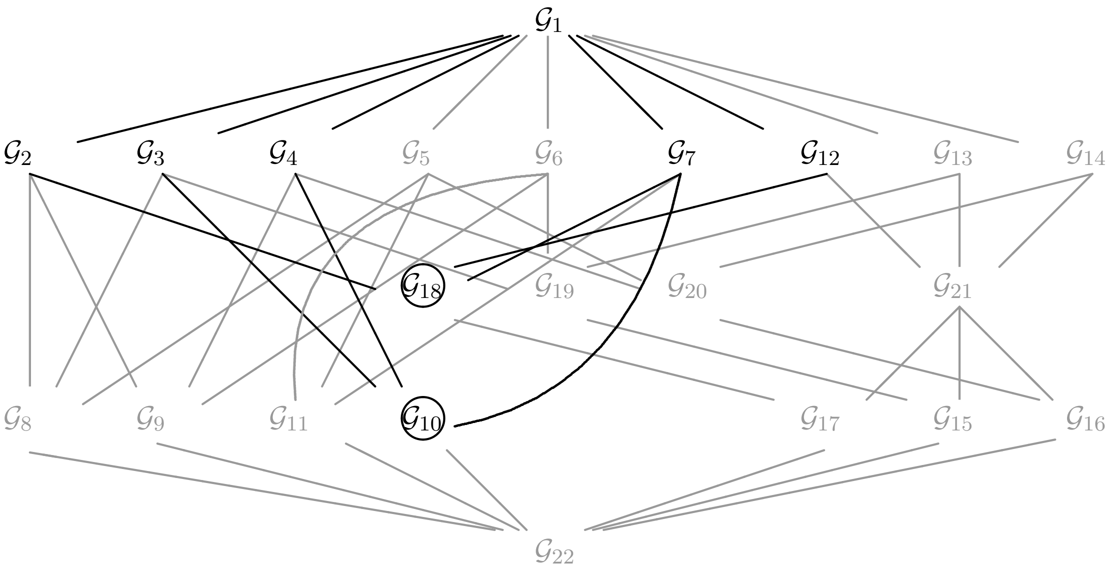


Figure E.6.: Minimally accepted colourings in $\mathcal{K}_{[4]}$ after stage 2.

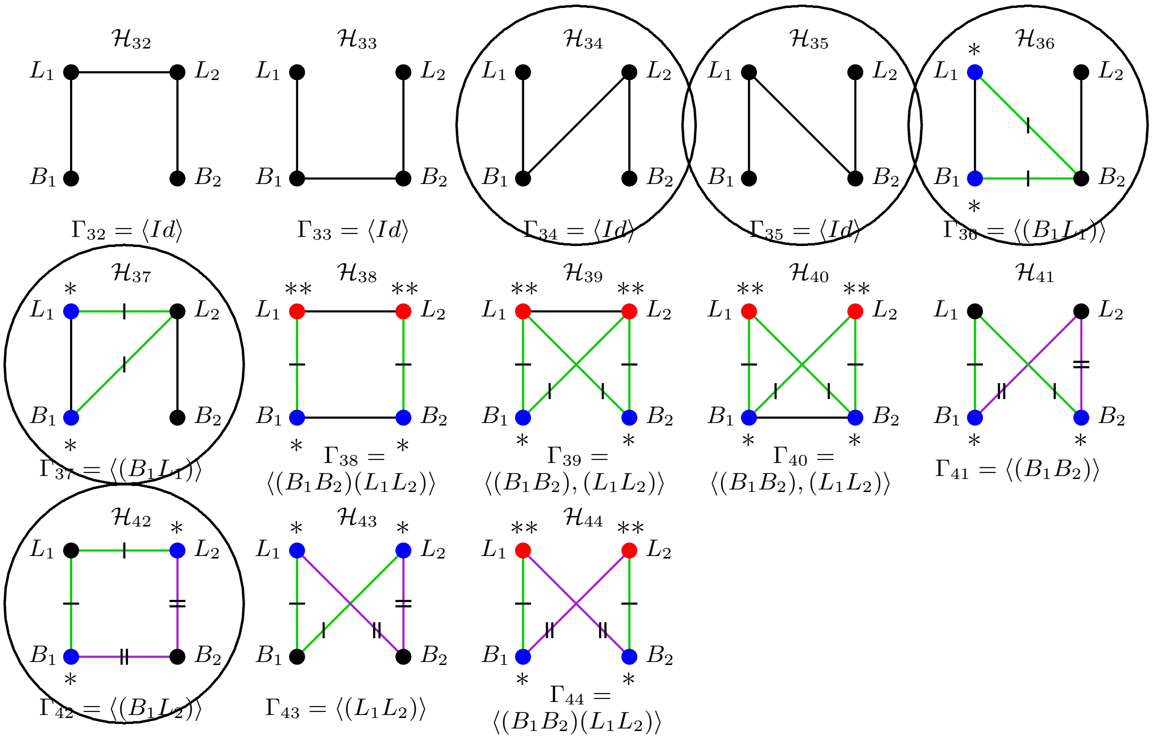


Figure E.7.: Stage 3: 13 models fitted of which 13 accepted and 0 rejected.

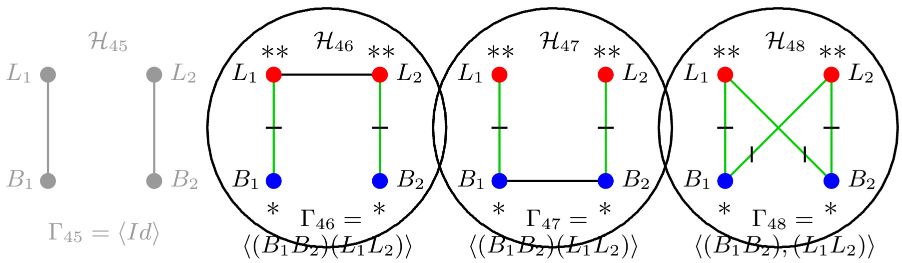


Figure E.8.: Stage 4: 4 models fitted of which 3 accepted and 1 rejected.

References

Andersen, H., M. Højbjerre, D. Sørensen, and P. Eriksen (1995). *Linear and Graphical Models for the Multivariate Complex Normal Distribution.* New York, NY, USA: Springer.

Anderson, T. (1984). *An Introduction to Multivariate Statistical Analysis.* New York, NY, USA: Wiley.

Andersson, S., H. Brøns, and T. Jensen (1983). Distribution of eigenvalues in multivariate statistical analysis. *Ann. Statist. 11*, 392–415.

Andersson, S. and J. Madsen (1998). Symmetry and lattice conditional independence in a multivariate normal distribution. *Ann. Statist. 26*, 525–572.

Andersson, S. A. (1975). Invariant normal models. *Ann. Statist. 3*, 132–154.

Arnold, S. (1973). Application of the theory of products of problems to certain patterned covariance matrices. *Ann. Statist. 1*, 682–699.

Bailey, R. (2008). *Design of Comparative Experiments.* Cambridge, UK: Cambridge University Press.

Bartlett, M. (1955). *An Introduction to Stochastic Processes.* Cambridge, UK: Cambridge University Press.

Behrens, W. (1929). Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen. *Landwirtschaftliche Jahrbücher 68*, 807–837.

Bell, E. (1934). Exponential numbers. *American Mathematical Monthly 1*, 411–419.

Bennett, B. (1951). Note on a solution of the generalized Behrens–Fisher problem. *Annals of the Institute of Statistical Mathematics 2*, 87–90.

Berend, D. and T. Tassa (2010). Improved bounds on Bell numbers and on moments of sums of random variables. *Probability and Mathematical Statistics 30*, 185–205.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. R. Statist. Soc. B 36*, 192–236.

Besag, J. and D. Higdon (1999). Bayesian analysis of agricultural field experiments. *J. R. Statist. Soc. B 61*, 691–746.

Besag, J. and P. Moran (1975). On the estimation and testing of spatial interaction in Gaussian lattice processes. *Biometrika 62*, 555–562.

Brown, L. D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*, Volume 9 of *Lecture Notes–Monograph Series*. Hayward, CA, USA: Institute of Mathematical Statistics.

Buot, M.-L., S. Hoşten, and D. Richards (2007). Counting and locating the solutions of polynomial systems of maximum likelihood equations ii: The Behrens-Fisher problem. *Statistica Sinica 17*, 1343–1354.

Castelo, R. and A. Roverato (2006). A robust procedure for Gaussian graphical model search from microarray data with p larger than n. *Journal of Machine Learning Research 7*, 2621–2650.

Chan, A. and C. Godsil (1997). Symmetry and eigenvectors. In *Graph Symmetry. Algebraic methods and applications*, Volume 497 of *NATO ASI Series C: Mathematical and Physical Sciences*, pp. 75–106. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cox, D. and N. Wermuth (1993). Linear dependencies represented by chain graphs (with discussion). *Statistical Science 8*, 204–218, 247–277.

Dawid, A. (1979). Conditional independence in statistical theory. *J. R. Statist. Soc. B 41*, 1–31.

Dawid, A. (1988). Symmetry models and hypotheses for structured data layouts. *J. R. Statist. Soc. B 50, Part 1*, 1–34.

Dempster, A. P. (1972). Covariance selection. *Biometrics 28*, 157–175.

Drton, M. (2008). Multiple solutions to the likelihood equations in the Behrens–Fisher problem. *Statistics & Probability Letters 78*, 3288–3293.

Drton, M. and D. Perlman (2008). A SINful approach to gaussian graphical model selection. *Journal of Statistical Planning and Inference 138*, 1179–1200.

Eaton, M. (1970). Gauss-Markov estimation for multivariate linear models: A coordinate free approach. *Ann. Math. Statist. 41*, 528–538.

Eaton, M. (1983). *Multivariate Statistics: A Vector Space Approach.* New York, NY, USA: Wiley.

Eaton, M. (1989). *Group Invariance Applications in Statistics*, Volume 1 of *Regional Conference Series in Probability and Statistics.* Hayward, CA, USA: Institute of Mathematical Statistics.

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

Edwards, D. and T. Havránek (1987). A fast model selection procedure for large families of models. *J. Am. Statist. Ass. 82, No. 397*, 205–213.

Fisher, R. (1939). The comparison of samples with possibly unequal variances. *Annals of Human Genetics 9*, 174–180.

Frets, G. (1921). Heredity of head form in man. *Genetica 3*, 193–400.

Frey, B., F. Kschischang, H. Loelinger, and N. Wiberg (1998). Factor graphs and algorithms. In *Proceedings of the* 35^{th} *Allerton Conference on Communication, Control and Computing 1997.*

Friedberg, S., A. Insel, and L. Spence (1979). *Linear Algebra.* Englewood Cliffs, NJ, USA: Prentice Hall.

Friedman, F., T. Hastie, and R. Tibshirani (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics 9*, 432–441.

Gabriel, K. (1969). Simultaneous test procedures - some theory of multiple comparisons. *Ann. Math. Statist. 40*, 224–250.

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

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

Grätzer, G. (1998). *General Lattice Theory.* Basel, Switzerland: Birkhäuser Verlag.

Haberman, S. (1975). How much do Gauss–Markov and least-squares estimates differ? A coordinate-free approach. *Ann. Statist. 3*, 982 – 990.

Halmos, P. (1958). *Finite-Dimensional Vector Spaces.* Princeton, NJ, USA: D. Van Nostrand Company.

Hammersley, J. and P. Clifford (1971). Markov fields in finite graphs and lattices. Unpublished manuscript.

Højsgaard, S. and S. Lauritzen (2007). Inference in graphical Gaussian models with edge and vertex symmetries with the gRc package for R. *Journal of Statistical Software 23*.

Højsgaard, S. and S. Lauritzen (2008). Graphical Gaussian models with edge and vertex symmetries. *J. R. Statist. Soc. B 70*, 1005–1027.

Hylleberg, B., M. Jensen, and E. Ørnbøl (1993). Graphical symmetry models. Master's thesis, Aalborg University, Aalborg, Denmark.

Jensen, S. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist. 16*, 302–322.

Kruskal, W. (1968). When are Gauss–Markov and least squares estimators identical? A coordinate-free approach. *Ann. Math. Statist. 39*, 70–75.

Lauritzen, S. (1989). *Lectures on Multivariate Analysis*. University of Copenhagen, Denmark: Institute of Mathematical Statistics.

Lauritzen, S. (1996). *Graphical Models*. Oxford, UK: Clarendon Press.

Madsen, J. (2000). Invariant normal models with recursive graphical Markov structure. *Ann. Statist. 28*, 1150–1178.

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

McKay, B. (1976). Backtrack programming and the graph isomorphism problem. Master's thesis, University of Melbourne, Melbourne, Australia.

McKay, B. (2010). `http://cs.anu.edu.au/~bdm/nauty/`.

Meinshausen, N. and P. Bühlmann (2006). High dimensional graphs and variable selection with the lasso. *Ann. Statist. 34*, 1436–1462.

Meinshausen, N. and P. Bühlmann (2010). Stability selection. *J. R. Statist. Soc. B 72*, 417–473.

Olkin, I. (1972). Testing and estimation for structures which are circularly symmetric in blocks. Technical report, Educational Testing Service, Princeton, NJ, USA.

Olkin, I. (2000). Symmetrically dependent models arising in visual assessment data. *Biometrics 56*, 1188–1191.

Olkin, I. and S. J. Press (1969). Testing and estimation for a circular stationary model. *Ann. Math. Statist. 40*, 1358—1373.

Pearl, J. and A. Paz (1987). Graphoids: a graph based logic for reasoning about relevancy relations. In *Advances in Artificial Intelligence – II*, Amsterdam, The Netherlands, pp. 357–363. Elsevier Science Ltd.

Pitman, J. (1997). Probabilistic aspects of set partitions. *The American Mathematical Monthly 104*, 201–209.

Ravikumar, P., M. Wainwright, G. Raskutti, and B. Yu (2008, December). High-dimensional covariance estimation by minimizing l_1-penalized log-determinant divergence. In *NIPS Conference*, Vancouver, Canada.

Ripley, B. (1981). *Spatial Statistics.* New York, NY, USA: Wiley.

Ripley, B. (1988). *Statistical Inference for Spatial Processes.* Cambridge, UK: Cambridge University Press.

Rotmann, J. (1995). *An Introduction to the Theory of Groups.* New York, NY, USA: Springer.

Sachs, H. (1966). Über Teiler, Faktoren und charakteristische Polynome von Graphen. *Teil I. Wissenschaftliche Zeitschrift Technische Hochschune Ilmenau 12*, 7–12.

Schäfer, J. and K. Strimmer (2005). Learning large-scale graphical Gaussian models for genomic data. In *Science of Complex Networks: From Biology to the WWW*, Volume 776, pp. 263–276.

Scheffé, H. (1943). On solutions for the Behrens–Fisher problem, based on the t-distribution. *Ann. Math. Statist. 14*, 35–44.

Scheffé, H. (1944). A note on the Behrens–Fisher problem. *Ann. Math. Statist. 15*, 430–432.

Schmidt, R. (1994). *Subgroup Lattices of Groups.* Berlin, Germany: Walter de Gruyter.

Siemons, J. (1983). Automorphism groups of graphs. *Arch. Math. 41*, 379–384.

Tibshirani, R., M. Saunders, S. Rosset, J. Zhu, and K. Knight (2005). Sparsity and smoothness via the fused lasso. *J. R. Statist. Soc. B 67*, 91–108.

Uhler, C. (2010). Geometry of maximum likelihood estimation in Gaussian graphical models. `http://arxiv.org/abs/1012.2643`.

Viana, M. (2008). *Symmetry Studies - An Introduction to the Analysis of Structured Data in Applications.* Cambridge, UK: Cambridge University Press.

Votaw, D., A. Kimball, and J. Rafferty (1950). Compound symmetry tests in the multivariate analysis of medical experiments. *Biometrics 6*, 259–281.

Votaw, D. F. (1948). Testing compound symmetry in a normal multivariate distribution. *Ann. Math. Statist. 19*, 447–473.

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

Whittle, P. (1954). On stationary processes in the plane. *Biometrika 41*, 434–449.

Whittle, P. (1963). Stochastic processes in several dimensions. *Bulletin Of The International Statistical Institute 40*, 974–994.

Wielandt, H. (1969). Permutation groups through invariant relations. Lecture Notes, Department of Mathematics, The Ohio State University, Columbus, OH, USA.

Wilks, S. S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Ann. Math. Statist. 17*, 257–281.

Wu, X., Y. Ye, and K. Subramanian (2003). Interactive analysis of gene interaction using graphical Gaussian models. In *3rd ACM SIGKDD Workshop on Data Mining in Bioinformatics*, Washington, DC, USA.

Young, D. (1975). Inference concerning the mean vector when the matrix is totally reducible. Technical report, Department of Statistics, Stanford University, Stanford, CA, USA.

List of Figures

List of Tables

Glossary

The following is a list of most symbols used in this thesis.

Undirected Graphs

V	Discrete finite set.
$V^{(2)}$	$\{\{\alpha,\beta\} \mid \alpha,\beta \in V\}$.
$V^{(2)}_{\neq}$	$\{\{\alpha,\beta\} \mid \alpha,\beta \in V, \alpha \neq \beta\}$.
$\alpha\beta$	For $\alpha,\beta \in V$ with $\alpha \neq \beta$, short form for $\{\alpha,\beta\}$.
E	Subset of $V^{(2)}_{\neq}$.
$G = (V,E)$	Undirected graph with vertex set V and edge set E.
U_V	Set of undirected graphs with vertex set V.
$\sim$	Adjacent.
$\mathrm{ne}(\alpha)$	$\{\beta \in V \mid \alpha\beta \in E\}$.
$\mathrm{ne}_u(\alpha)$	$\{\beta \in V \mid \alpha\beta \in u\}$ where $u \subset E$.

Permutation Groups

$S(V)$	Symmetric group of V.
σ	Element of $S(V)$.
Γ	Subgroup of $S(V)$.
J_Γ	Generating set of Γ.
$\Gamma_1 \times \Gamma_2$	For subgroups Γ_1 and Γ_2 of $S(V)$, direct product of Γ_1 and Γ_2.
$\mathrm{Aut}(G)$	$\{\sigma \in S(V) \mid \alpha\beta \in E \Leftrightarrow \sigma(\alpha)\sigma(\beta) \in E\}$.

Sets, Partitions and Colourings

$[d]$	$\{1, 2, \ldots, d\}$, for $d \in \mathbb{N}$.
$P(V)$	Power set of V.
$\mathcal{P}(V)$	Set of partitions of V.
$\mathcal{P}_1 \Delta \mathcal{P}_2$	$\{C \mid C \in \mathcal{P}_1, C \subseteq D$ for some $D \in \mathcal{P}_2\}$, where $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions of not necessarily the same set.
$\mathcal{M}$, $\mathcal{V}$	Elements of $\mathcal{P}(V)$.
$\mathcal{E}$	Element of $\mathcal{P}(E)$.
$\alpha \equiv \beta\ (\mathcal{M})$	α and β lie in the same set in $\mathcal{M}$.
$(\mathcal{V}, \mathcal{E})$	Graph colouring.
$\mathcal{G} = (\mathcal{V}, \mathcal{E})$	Vertex and edge coloured graph.
$(\mathcal{V}_\Gamma, \mathcal{E}_\Gamma)$	Graph colouring of $(V, V^{(2)}_{\neq})$ given by coarsest partitions which are invariant under the action of Γ on V and $V^{(2)}_{\neq}$.

Special Matrices

I_A	For $A \subseteq V$, identity matrix in $\mathbb{R}^{A \times A}$. Sometimes written as $I_{\lvert A \rvert}$.
T^u	For $u \in \mathcal{V}$: matrix in $\mathbb{R}^{V \times V}$ given by $T^u_{\alpha\alpha} = 1$ for $\alpha \in u$ and $T^u_{\alpha\beta} = 0$ otherwise. For $u \in \mathcal{E}$: matrix in $\mathbb{R}^{V \times V}$ given by $T^u_{\alpha\beta} = 1$ for $\alpha\beta \in u$ and $T^u_{\alpha\beta} = 0$ otherwise.
$\mathcal{T}(\mathcal{V})$	$\{T^u \mid u \in \mathcal{V}\}$.
$\mathcal{T}(\mathcal{E})$	$\{T^u \mid u \in \mathcal{E}\}$.
$\mathcal{T}(\mathcal{G})$	$\mathcal{T}(\mathcal{V}) \cup \mathcal{T}(\mathcal{E})$.
$T^{u\{vw\}}$	$T^v T^u T^w + T^w T^u T^v$ for $u \in \mathcal{E}$, $v, w \in \mathcal{V}$, $v \neq w$; $T^v T^u T^v$ for $u \in \mathcal{E}$, $v, w \in \mathcal{V}$, $v = w$.
$T^{[u\{vw\}]}$	Submatrix of $T^{u\{vw\}}$ with rows and columns indexed by $v \cup w$.
$G(\sigma)$	Matrix in $\mathbb{R}^{V \times V}$ given by $G(\sigma)_{\alpha\beta} = 1$ if $\sigma(\beta) = \alpha$ and $G(\sigma)_{\alpha\beta} = 0$ otherwise.

Parameter Sets

$\mathbb{R}^V$	Set of vectors indexed by the elements of V with entries in $\mathbb{R}$.
$\mathbb{R}^{V\times V}$	Set of matrices whose rows and columns are indexed by the elements of V with entries in $\mathbb{R}$.
$\mathcal{S}$	Set of symmetric matrices.
$\mathcal{S}^+$	Set of symmetric positive definite matrices.
$\mathcal{S}^+(G)$	Set of symmetric positive definite matrices in $\mathbb{R}^{V\times V}$ which have zeros in entries which correspond to edges missing from E.
$\mathcal{S}^+(\mathcal{V},\mathcal{E})$	$\mathcal{S}^+ \cap \mathrm{span}(\mathcal{T}(\mathcal{G}))$.
$\mathcal{R}^+(\mathcal{V},\mathcal{E})$	$\mathcal{S}^+ \cap \{ACA \mid A \in \mathrm{span}(\mathcal{T}(\mathcal{V})),\ C = I_V + B,\ B \in \mathrm{span}(\mathcal{T}(\mathcal{E}))\}$.
$\mathcal{S}(\Gamma)$	$\mathcal{S} \cap \{S \mid G(\sigma)^T S G(\sigma)^T = S \ \forall\ \sigma \in \Gamma\}$.
Ω	Subspace of $\mathbb{R}^V$.
$\Phi(\mathcal{G})$	$\{\Omega \mid T\Omega \subseteq \Omega \ \forall\ T \in \mathcal{T}(\mathcal{G})\}$.
$\Omega(\mathcal{M})$	$\{\mu \in \mathbb{R}^V \mid \mu_\alpha = \mu_\beta$ whenever $\alpha \equiv \beta\ (\mathcal{M})\}$.

Statistics

$X \perp\!\!\!\perp Y$	X independent of Y.
$X \perp\!\!\!\perp Y \mid Z$	X conditionally independent of Y given Z.
Y^i	i^{th} observation.
$\mathcal{N}_V(\mu,\Sigma)$	Multivariate Gaussian distribution on $\mathbb{R}^V$ with $\mu \in \mathbb{R}^V$ and $\Sigma \in \mathcal{S}^+ \cap \mathbb{R}^{V\times V}$.
Σ	Covariance matrix.
Σ^{-1}	Concentration matrix.
μ	Mean vector.
$\hat{\mu}$	Maximum likelihood estimator of μ.
$\tilde{\mu}$	Gauss-Markov estimator of μ.
μ^*	Least squares estimator of μ.
$\mathcal{S}_V^G$	Set of graphical Gaussian models with variables indexed by V.
$\mathcal{S}_V^{\mathcal{G}}$	Set of RCON models with variables indexed by V.
$\mathcal{R}_V^{\mathcal{G}}$	Set of RCOR models with variables indexed by V.

Further Graphs

$G_F = (V \cup N, E_F)$	Factor graph.
F_V	Set of factor graphs with vertex set V.
$\mathcal{G}_F = (\mathcal{V} \cup \mathcal{N}, E_F)$	Coloured factor graph.
$\mathcal{F}_V$	Set of coloured factor graphs with vertex set V.
$\mathcal{C}_V$	Set of coloured graphs with vertex set V.
$\mathcal{B}_V$	Set of graphs in $\mathcal{C}_V$ with edge regular colouring.
$\mathcal{P}_V$	Set of graphs in $\mathcal{C}_V$ with vertex regular colouring.
$\mathcal{R}_V$	Set of graphs in $\mathcal{C}_V$ with regular colouring.
Π_V	Set of graphs in $\mathcal{C}_V$ with permutation generated colouring.
$\mathcal{G}^{u\{vw\}}$	Graph whose adjacency matrix is given by $T^{u\{vw\}}$.
$E^{u\{vw\}}$	Edge set of $\mathcal{G}^{u\{vw\}}$.
$\mathcal{G}^{[u\{vw\}]}$	Graph whose adjacency matrix is given by $T^{[u\{vw\}]}$.

Partially Ordered Sets, Lattices and Orderings

$\langle A; \leq \rangle$	Partially ordered set A with partial ordering $\leq$.
$\langle A; \wedge, \vee \rangle$	Lattice A with meet operation $\wedge$ and join operation $\vee$.
$\subseteq$	Subset of.
$\leq$	Partial ordering symbol, with precise meaning depending on the type of objects it is applied to. For $a \leq b$ with $a, b \in S$, (i) $S = \mathbb{R}$: 'less than or equal to', (ii) $S =$ the set of linear subspaces of $\mathbb{R}^V$: 'subspace of', (iii) $S = S(V)$: 'subgroup of', (iv) $S = \mathcal{P}(V)$: 'finer than', (v) $S = U_V$: 'subgraph of'.
$\preceq$	Symbol of partial ordering on $\mathcal{C}_V$.

Spectral Analysis

$\Lambda^{[u\{vw\}]}$	Spectrum of $T^{[u\{vw\}]}$.
$\Lambda^{[u\{vw\}]}_{\geq 0}$	Non-negative elements of $\Lambda^{[u\{vw\}]}$.
$E^{[u\{vw\}]}_\lambda$	Eigenspace of $T^{[u\{vw\}]}$ corresponding to $\lambda \in \Lambda^{[u\{vw\}]}$.
$A^{[u\{vw\}]}_\lambda$	Subspace of $E^{[u\{vw\}]}_\lambda$.

Projections and Embeddings

Π_b^A	For $b \subseteq A \subseteq V$, orthogonal projection from $\mathbb{R}^A$ to $\mathbb{R}^A$, given by $\left(\Pi_b^A x\right)_\alpha = x_\alpha$ for $\alpha \in b$ and $\left(\Pi_b^A x\right)_\alpha = 0$ otherwise.
$\tilde{\Omega}_b$	Image of Ω under Π_b^V.
$\mathcal{I}_b^A$	For $b \subseteq A \subseteq V$, embedding of $\mathbb{R}^A$ in $\mathbb{R}^b$, given by $\left(\Pi_b^A x\right)_\alpha = x_\alpha$ for $\alpha \in b$.
Ω_b	Image of Ω under $\mathcal{I}_b^A$.