MODEL SELECTION IN THE SPACE OF GAUSSIAN MODELS INVARIANT BY SYMMETRY

BY PIOTR GRACZYK[1], HIDEYUKI ISHI[2 *], BARTOSZ KOŁODZIEJEK[3†] AND HÉLÈNE MASSAM[4‡]

[1]Université d'Angers, France, piotr.graczyk@univ-angers.fr

[2]Osaka City University, Japan, hideyuki@sci.osaka-cu.ac.jp

[3]Warsaw University of Technology, Poland, b.kolodziejek@mini.pw.edu.pl

[4]York University, Canada, massamh@mathstat.yorku.ca

arXiv:2004.03503v1 [math.ST] 7 Apr 2020

We consider multivariate centred Gaussian models for the random variable $Z = (Z_1, \ldots, Z_p)$, invariant under the action of a subgroup of the group of permutations on $\{1, \ldots, p\}$. Using the representation theory of the symmetric group on the field of reals, we derive the distribution of the maximum likelihood estimate of the covariance parameter Σ and also the analytic expression of the normalizing constant of the Diaconis-Ylvisaker conjugate prior for the precision parameter $K = \Sigma^{-1}$. We can thus perform Bayesian model selection in the class of complete Gaussian models invariant by the action of a subgroup of the symmetric group, which we could also call complete RCOP models. We illustrate our results with a toy example of dimension 4 and several examples for selection within cyclic groups, including a high dimensional example with $p = 100$.

1. Introduction. Let $V = \{1, \ldots, p\}$ be a finite index set and let $Z = (Z_1, \ldots, Z_p)$ be a multivariate random variable following a centred Gaussian model $\mathrm{N}_p(0, \Sigma)$. Let $\mathfrak{S}_p$ denote the symmetric group on V, that is, the group of all permutations on $\{1, \ldots, p\}$ and let Γ be a subgroup of $\mathfrak{S}_p$. A centred Gaussian model is said to be invariant under the action of Γ if for $g \in \Gamma$, $g \cdot \Sigma \cdot g^\top = \Sigma$ (here we identify a permutation g with its permutation matrix).

Given n data points $Z^{(1)}, \ldots, Z^{(n)}$ from a Gaussian distribution, our aim in this paper is to do Bayesian model selection within the class of models invariant by symmetry, that is, invariant under the action of some subgroup Γ of $\mathfrak{S}_p$ on V. Given the data, our aim is therefore to identify the subgroup $\Gamma \subset \mathfrak{S}_p$ such that the model invariant under Γ has the highest posterior probability.

Gaussian models invariant by symmetry have been considered by Andersson (1975) and Andersson, Brøns and Jensen (1983). Gaussian models invariant under the action of a subgroup $\Gamma \subset \mathfrak{S}_p$ have also been considered in Andersson and Madsen (1998), Madsen (2000) and Højsgaard and Lauritzen (2008), but in this paper, we will consider complete Gaussian models without predefined conditional independencies. In Andersson and Madsen (1998), the reader can also find references to earlier works dealing with particular symmetry models such as, for example, the circular symmetry model of Olkin and Press (1969) that we will consider further (Section 5). These works were concentrating on the derivation of the

*Supported by JSPS KAKENHI Grant Number 16K05174 and JST PRESTO.

†Supported by grant 2016/21/B/ST1/00005 of the National Science Center, Poland.

‡Supported by an NSERC Discovery Grant.

MSC 2010 subject classifications: Primary 62H99, 62F15 ; secondary 20C35

Keywords and phrases: colored graph, conjugate prior, covariance selection, invariance, permutation symmetry

maximum likelihood estimate of Σ and on testing the hypothesis that models were of a particular type. Our work is a first step towards Bayesian model selection in the class of models invariant by symmetry.

Just like the classical papers mentioned above, the fundamental algebraic tool we use in this work is the irreducible decomposition theorem for the matrix representation of the group Γ, which in turn means that, through an adequate change of basis, any matrix X in $\mathcal{P}_\Gamma$, the cone of positive definite matrices invariant under the subgroup Γ of $\mathfrak{S}_p$, can be written as

$$X = U_\Gamma \cdot \begin{pmatrix} M_{\mathbb{K}_1}(x_1) \otimes I_{k_1/d_1} & & & \\ & M_{\mathbb{K}_2}(x_2) \otimes I_{k_2/d_2} & & \\ & & \ddots & \\ & & & M_{\mathbb{K}_L}(x_L) \otimes I_{k_L/d_L} \end{pmatrix} \cdot U_\Gamma^\top, \tag{1}$$

where U_Γ is an orthogonal matrix, k_i, d_i, r_i, $i = 1, \ldots, L$, are integer constants called structure constants we will define later, such that k_i/d_i are also integers and $M_{\mathbb{K}_i}(x_i),\ i = 1, \ldots, L$ is a real matrix representation of an $r_i \times r_i$ Hermitian matrix x_i with entries in $\mathbb{K}_i = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ with $d_i = \dim_{\mathbb{R}} \mathbb{K}_i = 1, 2, 4$. Moreover the empty entries in the middle matrix of (1) are all equal to 0 so that this middle matrix is block diagonal. Finally, $A \otimes B$ denotes the Kronecker product of matrices A and B.

Even though some of the results in the theoretical part of the present paper were already developed in Andersson (1975), we decided to present full proofs for two reasons. First, we think that our arguments are more concrete and should be easier to understand for the reader who is not familiar with representation theory. Second, our results are more explicit and, in particular, we are able to completely solve the case when Γ is a cyclic group (*i.e.* Γ has one generator), that is, we show how U_Γ and all structure constants $(k_i, d_i, r_i)_{i=1}^L$ can be computed explicitly.

Let us consider the following example.

EXAMPLE 1. For $p = 3$ and $\Gamma = \mathfrak{S}_3$, the cone of positive definite matrices X invariant under Γ, that is, such that $X_{ij} = X_{\sigma(i)\sigma(j)}$ for all $\sigma \in \Gamma$, is

$$\mathcal{P}_\Gamma = \left\{ \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix} ;\ a > 0 \text{ and } b \in (-a/2, a) \right\}.$$

The decomposition (1) yields $U_\Gamma := \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \in \mathrm{O}(3)$ with

$$v_1 := \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \quad v_2 := \begin{pmatrix} \sqrt{2/3} \\ -1/\sqrt{6} \\ -1/\sqrt{6} \end{pmatrix}, \quad v_3 := \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1\sqrt{2} \end{pmatrix},$$

and

$$\begin{pmatrix} a & b & b \\ b & a & b \\ a & a & b \end{pmatrix} = U_\Gamma \cdot \begin{pmatrix} a + 2b & & \\ & a - b & \\ & & a - b \end{pmatrix} \cdot U_\Gamma^\top,$$

Here $L = 2$, $k_1/d_1 = 1$, $k_2/d_2 = 2$, $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$, $d_1 = d_2 = 1$.

We see immediately in the example above that, following the decomposition (1), the trace $\mathrm{Tr}\,[X] = a + 2b + 2(a - b)$ and the determinant $\mathrm{Det}\,(X) = (a + 2b)(a - b)^2$ can be readily obtained. Similarly, using (1) allows us to easily obtain $\mathrm{Det}\,(X)$ and $\mathrm{Tr}\,[X]$ in general.

In Section 3, we will see that having the explicit formulas for $\mathrm{Det}\,(X)$ and $\mathrm{Tr}\,[X]$, in turn, allows us to derive the analytic expression of the Gamma function on $\mathcal{P}_\Gamma$, defined as

$$\Gamma_{\mathcal{P}_\Gamma}(\lambda) := \int_{\mathcal{P}_\Gamma} \mathrm{Det}\,(X)^\lambda\, e^{-\mathrm{Tr}[X]} \varphi_\Gamma(X)\,\mathrm{d}X,$$

where $\varphi_\Gamma(X)\,\mathrm{d}X$ is the invariant measure on $\mathcal{P}_\Gamma$ (see Definition 13 and Proposition 9) and $\mathrm{d}X$ denotes the Euclidean measure on the space $\mathcal{Z}_\Gamma$ with the trace inner product.

With the Gamma integral on $\mathcal{P}_\Gamma$, we can derive the analytic expression of the normalizing constant $I_\Gamma(\delta, D)$ of the Diaconis-Ylvisaker conjugate prior on $K = \Sigma^{-1}$ with density, with respect to the Euclidean measure on $\mathcal{P}_\Gamma$, equal to

$$f(K;\delta,D) = \frac{1}{I_\Gamma(\delta,D)} \mathrm{Det}\,(K)^{(\delta-2)/2}\, e^{-\frac{1}{2}\mathrm{Tr}[K\cdot D]} \mathbf{1}_{\mathcal{P}_\Gamma}(K)$$

for appropriate values of the scalar hyper-parameter δ and the matrix hyper-parameter $D \in \mathcal{P}_\Gamma$. By analogy with the G-Wishart distribution, defined in the context of the graphical Gaussian models, Markov with respect to an undirected graph G on the cone P_G of positive definite matrices with zero entry (i,j) whenever there is no edge between the vertices i and j in G, (see Maathuis et al. (2018)), we can call the distribution with density $f(K;\delta,D)$, the RCOP-Wishart (RCOP is the name coined in Højsgaard and Lauritzen (2008) for graphical Gaussian models with restrictions generated by permutation symmetry). It is important to note here that if Σ is in $\mathcal{P}_\Gamma$, so is $K = \Sigma^{-1}$ so that K can also be decomposed according to (1). Equipped with all these results, we compute the Bayes factors comparing models pairwise and perform model selection. We will indicate in Section 4 how to travel through the space of cyclic subgroups of the symmetric group.

In Section 3, we also derive the distribution of the maximum likelihood estimate (henceforth abbreviated MLE) of Σ and show that for $n \geq \max_{i=1,\ldots,L}\left\{\frac{r_i d_i}{k_i}\right\}$ it has a density equal to

$$\frac{\mathrm{Det}\,(X)^{n/2}\, e^{-\frac{1}{2}\mathrm{Tr}[X\cdot\Sigma^{-1}]}}{\mathrm{Det}\,(2\Sigma)^{n/2}\,\Gamma_{\mathcal{P}_\Gamma}(\frac{n}{2})} \varphi_\Gamma(X) \mathbf{1}_{\mathcal{P}_\Gamma}(X).$$

Clearly, the key to computing the Gamma integral on $\mathcal{P}_\Gamma$, the normalizing constant $I_\Gamma(\delta, D)$ or the density of the MLE of Σ is, for each $\Gamma \subset \mathfrak{S}_p$, to obtain the block diagonal matrix with diagonal block entries $M_{\mathbb{K}_i}(x_i) \otimes I_{k_i/d_i}$, $i = 1,\ldots,L$, in the decomposition (1). In principle, we have to derive the invariant measure φ_Γ and find the structure constants k_i, d_i, r_i, $i = 1,\ldots,L$. This goal can be achieved by constructing an orthogonal matrix U_Γ and using (1). However, doing so for every Γ visited during the model selection process is computationally heavy.

We will show that for small to moderate dimensions, we can obtain the constants k_i, d_i, r_i, $i = 1,\ldots,L$ as well as the expression of $\mathrm{Det}\,(X)$ and $\varphi_\Gamma(X)$ without having to compute U_Γ. Indeed, as indicated in Remark 7, for any $X \in \mathcal{P}_\Gamma$, $\mathrm{Det}\,(X)$ admits a unique irreducible factorization of the form

$$\mathrm{Det}\,(X) = \prod_{i=1}^{L} \mathrm{Det}\,(M_{\mathbb{K}_i}(x_i))^{k_i/d_i} = \prod_{j=1}^{L} f_j(X)^{a_j} \quad (X \in \mathcal{Z}_\Gamma), \tag{2}$$

where each a_j is a positive integer, each $f_j(X)$ is an irreducible polynomial of $X \in \mathcal{Z}_\Gamma$, and $f_i \neq f_j$ if $i \neq j$. The constants k_i, d_i, r_i are obtained by identification of the two expressions of $\mathrm{Det}\,(X)$ in (2). Factorization of a homogeneous polynomial $\mathrm{Det}\,(X)$ can be performed using standard software such as either MATHEMATICA or PYTHON.

Due to computational complexity, for bigger dimensions, it is difficult to obtain the irreducible factorization of $\mathrm{Det}\,(X)$. For special cases such as the case where the subgroup Γ is a cyclic group, we give (Section 2) a simple construction of the matrix U_Γ and thus, for any dimension p, we can do model selection in the space of models invariant under the action of a cyclic group. We argue that restriction to cyclic groups is not as limiting as it may look. The formula for the number of different colorings $c_p = \#\{\mathcal{P}_\Gamma\,;\,\Gamma\subset\mathfrak{S}_p\}$ for given p is unknown. Obviously, it is bounded from above by the number of all subgroups of $\mathfrak{S}_p$, because different subgroups may produce the same coloring (e.g. in Example 1 we have $\mathcal{P}_{\mathfrak{S}_3}=\mathcal{P}_{\langle(1,2,3)\rangle}$). On the other hand, it is known (see Lemma 17) that c_p is bounded from below by the number of distinct cyclic subgroups, which grows rapidly with p (see OEIS[1] sequence A051625). In particular, for $p=18$[2], we have $c_p\in(7.1\cdot 10^{14}, 7.6\cdot 10^{18})$, see also Table 1. The lower bound for c_p indicates that the colorings obtained from cyclic subgroups form a rich subfamily of all possible colorings. Let us consider the more general situation of Gaussian graphical models with conditional independence structure encoded by a non complete graph G. Then one can introduce symmetry restrictions (RCOP) by requiring that the precision matrix K is invariant under some subgroup Γ of $\mathfrak{S}_p$. However, when G is not complete, not all subgroups are suited to the problem. In such cases, one has to require that Γ belongs to the automorphism group $\mathrm{Aut}(G)$ of G. If a graph G is sparse, then $\mathrm{Aut}(G)$ may be very small and it is natural to expect that the vast majority of subgroups of $\mathrm{Aut}(G)$ are actually cyclic. Moreover, finding the structure constants for a general group is much more expensive and in some situations it may not be worth to consider the problem in its full generality. We consider our work as a first step towards the rigorous analytical treatment of Bayesian model selection in the space of graphical Gaussian models invariant under the action of $\Gamma\subset\mathfrak{S}_p$ when conditional independencies are allowed. Finally, we expect the statistical interpretability of cyclic models to be easier than that of general groups.

The procedure to do model selection will be described in Section 4 and we will illustrate this procedure with Frets' data (see Frets (1921)) and several examples for selection within cyclic groups, including a high dimensional example with $p=100$ (Section 5).

Most technical parts of the paper are postponed to the Appendix.

2. Preliminaries.

Representation theory has long been known to be very useful in statistics, cf. Diaconis (1988). However, the representation theory over $\mathbb{R}$ that we need in this paper, is less known to the statisticians than the standard one over $\mathbb{C}$. We will thus now recall the basic notions that we need.

2.1. *Notation.* Let $\mathrm{Mat}(n,m;\mathbb{R})$, $\mathrm{Sym}(n;\mathbb{R})$ denote the linear spaces of real $n\times m$ matrices and symmetric real $n\times n$ matrices, respectively. Let $\mathrm{Sym}^+(n;\mathbb{R})$ be the cone of symmetric positive definite real $n\times n$ matrices. $A^\top$ denotes the transpose of a matrix A. Det and Tr denote the usual determinant and trace in $\mathrm{Mat}(n,n;\mathbb{R})$.

For $A\in\mathrm{Mat}(m,n;\mathbb{R})$ and $B\in\mathrm{Mat}(m',n';\mathbb{R})$, we denote by $A\oplus B$ the matrix $\begin{pmatrix}A & O\\ O & B\end{pmatrix}\in\mathrm{Mat}(m+m',n+n';\mathbb{R})$, and by $A\otimes B$ the Kronecker product of A and B. For a positive integer r, we write $B^{\oplus r}$ for $I_r\otimes B\in\mathrm{Mat}(rm',rn';\mathbb{R})$

Let p denote the fixed number of vertices of a graph and let $\mathfrak{S}_p$ denote the symmetric group. We write permutations in cycle notation, meaning that $(i_1,i_2,\ldots,i_n)$ maps i_j to i_{j+1} for $j=1,\ldots,r-1$ and i_n to i_1. By $\langle\sigma_1,\ldots,\sigma_k\rangle$ we denote the group generated by permutations $\sigma_1,\ldots,\sigma_k$. The composition (product) of permutations $\sigma,\sigma'\in\mathfrak{S}_p$ will be denoted by $\sigma\circ\sigma'$.

[1]The On-Line Encyclopedia of Integer Sequences, https://oeis.org/.

[2]The number of subgroups of $\mathfrak{S}_p$ is unknown for $p>18$, see Holt (2010) and OEIS sequence A005432.

DEFINITION 2. For a subgroup $\Gamma \subset \mathfrak{S}_p$, we define the space of symmetric matrices invariant under Γ, or the vector space of colored matrices,

$$\mathcal{Z}_\Gamma := \left\{ x \in \mathrm{Sym}(p;\mathbb{R})\,;\, x_{ij} = x_{\sigma(i)\sigma(j)} \text{ for all } \sigma \in \Gamma \right\},$$

and the cone of positive definite matrices valued in $\mathcal{Z}_\Gamma$,

$$\mathcal{P}_\Gamma := \mathcal{Z}_\Gamma \cap \mathrm{Sym}^+(p;\mathbb{R}).$$

We note that the same colored space and cone can be generated by two different subgroups: in Example 1, the subgroup $\Gamma' = \langle(1,2,3)\rangle$ generated by the permutation $\sigma = (1,2,3)$ is such that $\Gamma' \neq \Gamma$ but $\mathcal{Z}_{\Gamma'} = \mathcal{Z}_\Gamma$. Let us define

$$\Gamma^* = \left\{ \sigma^* \in \mathfrak{S}_p\,;\, x_{ij} = x_{\sigma^*(i)\sigma^*(j)} \text{ for all } x \in \mathcal{Z}_\Gamma \right\}.$$

Clearly, Γ is a subgroup of Γ^* and Γ^* is the unique largest subgroup of $\mathfrak{S}_p$ such that $\mathcal{Z}_{\Gamma^*} = \mathcal{Z}_\Gamma$ or, equivalently, such that the Γ^*- and Γ- orbits in $\{\{v_1, v_2\}\,;\, v_i \in V,\, i = 1,2\}$ are the same. The group Γ^* is called the 2^*-closure of Γ. The group Γ is said to be 2^*-closed if $\Gamma = \Gamma^*$. Subgroups which are 2^*-closed are in bijection with the set of colored spaces. These concepts have been investigated in Wielandt (1969); Siemons (1982) along with a generalization to regular colorings in Siemons (1983). The combinatorics of 2^*-closed subgroups is very complicated and little is known in general, (Graham, Grötschel and Lovász, 1995, p. 1502). In particular, the number of such subgroups is not known, but brute-force search for small p indicates that this number is much less than the number of all subgroups of $\mathfrak{S}_p$ (see Table 1). Even though cyclic subgroups of $\mathfrak{S}_p$ are in general not 2^*-closed, each cyclic group corresponds to a different coloring (see Lemma 17).

For a permutation $\sigma \in \mathfrak{S}_p$, denote its matrix by

$$R(\sigma) := \sum_{i=1}^{p} E_{\sigma(i)i}, \tag{3}$$

where E_{ab} is the $p \times p$ matrix with 1 in the (a,b)-entry and 0 in other entries. The condition $x_{\sigma(i)\sigma(j)} = x_{ij}$ is then equivalent to $R(\sigma)\cdot x \cdot R(\sigma)^\top = x$. Consequently,

$$\mathcal{Z}_\Gamma = \left\{ x \in \mathrm{Sym}(p;\mathbb{R})\,;\, R(\sigma)\cdot x \cdot R(\sigma)^\top = x \text{ for all } \sigma \in \Gamma \right\}. \tag{4}$$

DEFINITION 3. Let $\pi_\Gamma\colon \mathrm{Sym}(p;\mathbb{R}) \to \mathcal{Z}_\Gamma$ be the projection such that for any $x \in \mathrm{Sym}(p;\mathbb{R})$ the element $\pi_\Gamma(x) \in \mathcal{Z}_\Gamma$ is uniquely determined by

$$\mathrm{Tr}\,[x \cdot y] = \mathrm{Tr}\,[\pi_\Gamma(x)\cdot y] \qquad (y \in \mathcal{Z}_\Gamma). \tag{5}$$

In view of (4), it is clear that

$$\pi_\Gamma(x) = \frac{1}{|\Gamma|}\sum_{\sigma\in\Gamma} R(\sigma)\cdot x \cdot R(\sigma)^\top \tag{6}$$

satisfies the above definition. Here $|\Gamma|$ denotes the order of Γ.

2.2. *$\mathcal{Z}_\Gamma$ as a Jordan algebra.* To prove (1), that is, Theorem 4 below, we need to view $\mathcal{P}_\Gamma$ as the cone of squares of a Jordan algebra. We recall here the fundamentals of Jordan algebras, cf. Faraut and Korányi (1994). A Euclidean Jordan algebra is a Euclidean space $\mathcal{A}$ (endowed with the scalar product denoted by $\langle\cdot,\cdot\rangle$) equipped with a bilinear mapping (product)

$$\mathcal{A} \times \mathcal{A} \ni (x,y) \mapsto x \bullet y \in \mathcal{A}$$

such that for all x, y, z in $\mathcal{A}$:

(i) $x \bullet y = y \bullet x$,
(ii) $x \bullet ((x \bullet x) \bullet y) = (x \bullet x) \bullet (x \bullet y)$,
(iii) $\langle x, y \bullet z \rangle = \langle x \bullet y, z \rangle$.

A Euclidean Jordan algebra is said to be simple if it is not a Cartesian product of two Euclidean Jordan algebras of positive dimensions. We have the following result.

PROPOSITION 1. *The Euclidean space $\mathcal{Z}_\Gamma$ with inner product $\langle x, y \rangle = \operatorname{Tr}[x \cdot y]$ and the Jordan product*

$$x \bullet y = \tfrac{1}{2}(x \cdot y + y \cdot x), \tag{7}$$

is a Euclidean Jordan algebra. This algebra is generally non-simple.

PROOF. Since $\mathcal{Z}_\Gamma$ is a subset of the Euclidean Jordan algebra $\mathrm{Sym}(p;\mathbb{R})$, if it is endowed with Jordan product (7), conditions (i)–(iii) are automatically satisfied. Moreover, characterization (4) of $\mathcal{Z}_\Gamma$ implies that the Jordan product is closed in $\mathcal{Z}_\Gamma$, that is, $R(\sigma) \cdot (x \bullet y) = (x \bullet y) \cdot R(\sigma)$ for all $x, y \in \mathcal{Z}_\Gamma$ and $\sigma \in \Gamma$. The result follows. $\square$

Up to linear isomorphism, there are only five kinds of Euclidean simple Jordan algebras. Let $\mathbb{K}$ denote the set of either the real numbers $\mathbb{R}$, the complex ones $\mathbb{C}$ or the quaternions $\mathbb{H}$. Let us write $\mathrm{Herm}(r;\mathbb{K})$ for the space of $r \times r$ Hermitian matrices valued in $\mathbb{K}$. Then $\mathrm{Sym}(r;\mathbb{R})$, $r \geq 1$, $\mathrm{Herm}(r;\mathbb{C})$, $r \geq 2$, $\mathrm{Herm}(r;\mathbb{H})$, $r \geq 2$ are the first three kinds of Euclidean simple Jordan algebras and they are the only ones that will concern us. The determinant and trace in Jordan algebras $\mathrm{Herm}(r;\mathbb{K})$ will be denoted by $\det$ and tr respectively, so that they can be easily distinguished from the determinant and trace in $\mathrm{Mat}(n,n;\mathbb{R})$ which we denote by Det and Tr.

To each Euclidean Jordan algebra $\mathcal{A}$, one can attach the set $\overline{\Omega}$ of Jordan squares, that is, $\overline{\Omega} = \{\, x \bullet x\,;\, x \in \mathcal{A} \}$. The interior Ω of $\overline{\Omega}$ is a symmetric cone, that is, it is self-dual and homogeneous. We say that Ω is irreducible if it is not the Cartesian product of two convex cones. One can prove that an open convex cone is symmetric and irreducible if and only if it is the symmetric cone Ω of some Euclidean simple Jordan algebra. Each simple Jordan algebra corresponds to a symmetric cone. The first three kinds of irreducible symmetric cones are thus, the symmetric positive definite real matrices $\mathrm{Sym}^+(r;\mathbb{R})$ for $r \geq 1$, complex Hermitian positive definite matrices $\mathrm{Herm}^+(r;\mathbb{C})$, and quaternionic Hermitian positive definite matrices $\mathrm{Herm}^+(r;\mathbb{H})$, $r \geq 2$.

It follows from Definition 2 and Proposition 1 that $\mathcal{P}_\Gamma$ is a symmetric cone. In (Faraut and Korányi, 1994, Proposition III.4.5) it is stated that any symmetric cone is a direct sum of irreducible symmetric cones. As it will turn out, only three out of the five kinds of irreducible symmetric cones may appear in this decomposition.

Moreover, we will want to represent the elements of the symmetric cones in their real symmetric matrix representations. So, we recall that both $\mathrm{Herm}^+(r;\mathbb{C})$ and $\mathrm{Herm}^+(r;\mathbb{H})$ can be realized as real symmetric matrices, but of bigger dimension. For $z = a + b\,i \in \mathbb{C}$ define $M_\mathbb{C}(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. The function $M_\mathbb{C}$ is a matrix representation of $\mathbb{C}$. Similarly, any $r \times r$ complex matrix can be realized as a $(2r) \times (2r)$ real matrix by setting the correspondence

$$\mathrm{Mat}(r,r;\mathbb{C}) \ni \big(z_{i,j}\big)_{1 \leq i,j \leq r} \simeq \big(M_\mathbb{C}(z_{i,j})\big)_{1 \leq i,j \leq r} \in \mathrm{Mat}(2r,2r;\mathbb{R}),$$

that is, an (i,j)-entry of a complex matrix is replaced by its 2×2 real matrix representation. Note that $M_\mathbb{C}$ maps the space $\mathrm{Herm}(r;\mathbb{C})$ of Hermitian matrices into the space $\mathrm{Sym}(2r;\mathbb{R})$

of symmetric matrices. For example,

$$M_{\mathbb{C}}\begin{pmatrix} a & c-d\,i \\ c+d\,i & b \end{pmatrix} = \begin{pmatrix} a & 0 & c & d \\ 0 & a & -d & c \\ c & -d & b & 0 \\ d & c & 0 & b \end{pmatrix}.$$

Moreover, by direct calculation one sees that

$$\mathrm{Det}\begin{pmatrix} a & 0 & c & d \\ 0 & a & -d & c \\ c & -d & b & 0 \\ d & c & 0 & b \end{pmatrix} = \det\begin{pmatrix} a & c-d\,i \\ c+d\,i & b \end{pmatrix}^2.$$

It can be shown that, in general,

$$\mathrm{Det}\,(M_{\mathbb{C}}(Z)) = [\det(Z)]^2 \quad \text{and} \quad \mathrm{Tr}\,[M_{\mathbb{C}}(Z)] = 2\,\mathrm{tr}[Z] \qquad (Z \in \mathrm{Herm}(r;\mathbb{C})). \tag{8}$$

Similarly, quaternions can be realized as a 4×4 matrix:

$$a + b\,i + c\,j + d\,k \simeq \begin{pmatrix} a+b\,i & -c+d\,i \\ c+d\,i & a-b\,i \end{pmatrix} \simeq \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}.$$

Then, quaternionic $r \times r$ matrices are realized as $(4r) \times (4r)$ real matrices. Thus, $M_{\mathbb{H}}$ maps $\mathrm{Herm}(r;\mathbb{H})$ into $\mathrm{Sym}(4r;\mathbb{R})$. Moreover, it is true that

$$\mathrm{Det}\,(M_{\mathbb{H}}(Z)) = [\det(Z)]^4 \quad \text{and} \quad \mathrm{Tr}\,[M_{\mathbb{H}}(Z)] = 4\,\mathrm{tr}[Z] \qquad (Z \in \mathrm{Herm}(r;\mathbb{H})). \tag{9}$$

2.3. *Basics of representation theory over real fields.* We will show, in this section, that the correspondence $\sigma \mapsto R(\sigma)$ defined in (3) is a representation of Γ and that, as for all representations of a finite group, through an appropriate change of basis, matrices $R(\sigma)$, $\sigma \in \Gamma$ can be simultaneously written as block diagonal matrices with the number and dimensions of these block matrices being the same for all $\sigma \in \Gamma$. This, in turn, will allow us to write any matrix in $\mathcal{Z}_\Gamma$ under the form (1). To do so, we first need to recall some basic notions and results of the representation theory of groups over the reals. For further details, the reader is referred to Serre (1977).

For a real vector space V, we denote by $\mathrm{GL}(V)$ the group of linear automorphisms on V. Let G be a finite group.

DEFINITION 4. A function $\rho\colon G \to \mathrm{GL}(V)$ is called a representation of G over $\mathbb{R}$ if it is a homomorphism, that is

$$\rho(g\,g') = \rho(g)\,\rho(g') \qquad (g, g' \in G).$$

The vector space V is called the representation space of ρ.

If $\dim V = n$, taking a basis $\{v_1, \ldots, v_n\}$ of V, we can identify $\mathrm{GL}(V)$ with the group $\mathrm{GL}(n;\mathbb{R})$ of all $n \times n$ non-singular real matrices. Then a representation $\rho\colon G \to \mathrm{GL}(V)$ corresponds to a group homomorphism $B\colon G \to \mathrm{GL}(n;\mathbb{R})$ for which

$$\rho(g)v_j = \sum_{i=1}^{n} B_{ij}(g)v_i. \tag{10}$$

We call B the matrix expression of ρ with respect to the basis $\{v_1, \ldots, v_n\}$.

By definition, we have $R\colon \Gamma \to \mathrm{GL}(p;\mathbb{R})$ and $R(\sigma\circ\sigma') = R(\sigma)\cdot R(\sigma')$ for all $\sigma,\sigma'\in\mathfrak{S}_p$. Thus, R is a representation of Γ over $\mathbb{R}$. Regarding $\rho(\sigma) = R(\sigma)\in \mathrm{GL}(\mathbb{R}^p)$ as an operator on $V=\mathbb{R}^p$ via the standard basis $v_i = e_i \in \mathbb{R}^p$, $i=1,\ldots,p$, we see that (10) holds trivially with $B=R$.

Example 6 below gives an illustration of the representation $R(\sigma)$ and also an illustration of all the notions and results we are about to state now, leading to the expression (1) of a matrix in $\mathcal{Z}_\Gamma$.

DEFINITION 5. A linear subspace $W\subset V$ is said to be G-invariant if

$$\rho(g)w\in W \qquad (w\in W,\, g\in G).$$

A representation ρ is said to be irreducible if the only G-invariant subspaces are non-proper, that is, whole V and $\{0\}$. A restriction of ρ to a G-invariant subspace W is a subrepresentation. Two representations, $\rho\colon G\to \mathrm{GL}(V)$ and $\rho'\colon G\to\mathrm{GL}(V')$ are equivalent if there exists an isomorphism of vector spaces $\ell\colon V\mapsto V'$ with

$$\ell(\rho(g)v) = \rho'(g)\ell(v) \qquad (v\in V,\, g\in G).$$

We note that, as we have discussed for the case $B=R$, a group homomorphism $B\colon G\to \mathrm{GL}(n;\mathbb{R})$ defines a representation of G on $\mathbb{R}^n$ naturally. We see that B is a matrix expression of a representation (ρ,V) if and only if B and ρ are equivalent via the map $\ell\colon \mathbb{R}^n\ni (x_i)_{i=1}^n\mapsto \sum_{i=1}^n x_i v_i\in V$, that is, $\ell(B(g)\underline{x}) = \rho(g)\ell(\underline{x})$ for $\underline{x}\in\mathbb{R}^n$. Here $\{v_1,\ldots,v_n\}$ denotes a fixed basis of V. Therefore, two representations (ρ,V) and (ρ',V') are equivalent if and only if they have the same matrix expressions with respect to appropriately chosen bases. We shall write $\rho\sim B$ if ρ has a matrix expression B with respect to some basis.

Let (ρ,V) be a representation of G, and $B\colon G\to\mathrm{GL}(n;\mathbb{R})$ be a matrix expression of ρ with respect to a basis $\{v_1,\ldots,v_n\}$ of V. Then it is known that the function $\chi_\rho\colon G\ni g\mapsto \operatorname{Tr} B(g) = \sum_{i=1}^n B_{ii}(g)\in\mathbb{R}$ is independent of the choice of the basis $\{v_1,\ldots,v_n\}$. The function χ_ρ is called a character of the representation ρ. The function χ_ρ characterizes the representation ρ in the following sense.

LEMMA 2. *Two representations (ρ,V) and (ρ',V') of a group G are equivalent if and only if $\chi_\rho = \chi_{\rho'}$.*

We apply this lemma in practice to know whether two given representations are equivalent or not.

It is known that, for a finite group G, the set $\Lambda(G)$ of equivalence classes of irreducible representations of G is a finite set. We fix the group homomorphisms $B_\alpha\colon G\to\mathrm{GL}(k_\alpha;\mathbb{R})$, $\alpha\in A$, indexed by a finite set A so that $\Lambda(G) = \{\,[B_\alpha]\,;\,\alpha\in A\,\}$, where $[B_\alpha]$ denotes the equivalence class of B_α.

Let (ρ,V) be a representation of G. Then there exists a G-invariant inner product on V. In fact, from any inner product $\langle\cdot,\cdot\rangle_0$ on V, one can define such an invariant inner product $\langle\cdot,\cdot\rangle$ by $\langle v,v'\rangle := \sum_{g\in G}\langle \rho(g)v,\rho(g)v'\rangle_0$ for $v,v'\in V$. In what follows, we fix a G-invariant inner product on V.

If W is a G-invariant subspace, the orthogonal complement $W^\perp$ is also a G-invariant subspace. Thus, any representation ρ can be decomposed into a finite number of irreducible subrepresentations

$$\rho = \rho_1\oplus\ldots\oplus\rho_K \tag{11}$$

along the orthogonal decomposition $V = V_1\oplus\cdots\oplus V_K$, where ρ_i is the restriction of ρ to the G-invariant subspace V_i, $i=1,\ldots,K$. Let r_α be the number of subrepresentations ρ_i such that $\rho_i \sim B_\alpha$. Although the irreducible decomposition (11) of V is not unique in general, r_α is uniquely determined. We have

$$\rho \sim \bigoplus_{r_\alpha > 0} B_\alpha^{\oplus r_\alpha}, \tag{12}$$

where $\sum_{r_\alpha>0} r_\alpha = K$. To see this, let $V(B_\alpha)$ be the direct sum of subspaces V_i for which $\rho_i \sim B_\alpha$. The space $V(B_\alpha)$ is called the B_α-component of V. If $r_\alpha > 0$, gathering an appropriate basis of each V_i, the matrix expression of the subrepresentation of ρ on $V(B_\alpha)$ becomes (recall that $B_\alpha(g) \in \mathrm{GL}(k_\alpha; \mathbb{R})$)

$$B_\alpha(g)^{\oplus r_\alpha} = \begin{pmatrix} B_\alpha(g) & & & \\ & B_\alpha(g) & & \\ & & \ddots & \\ & & & B_\alpha(g) \end{pmatrix} = I_{r_\alpha} \otimes B_\alpha(g) \in \mathrm{GL}(r_\alpha k_\alpha; \mathbb{R}) \qquad (g \in G).$$

Moreover, V is decomposed as $V = \bigoplus_{r_\alpha>0} V(B_\alpha)$. Therefore, taking a basis of V by gathering the bases of $V(B_\alpha)$, we obtain (12).

In particular, for $G = \Gamma \subset \mathfrak{S}_p$ and $(\rho, V) = (R, \mathbb{R}^p)$, if we let $\{\, \alpha \in A\,;\, r_\alpha > 0 \,\} =: \{\alpha_1, \alpha_2, \ldots, \alpha_L\}$ and if we denote by U_Γ an orthogonal matrix whose column vectors form orthonormal bases of $V(B_{\alpha_1}), \ldots, V(B_{\alpha_L})$ successively, then for $\sigma \in \Gamma$, we have

$$U_\Gamma^\top \cdot R(\sigma) \cdot U_\Gamma = \begin{pmatrix} I_{r_1} \otimes B_{\alpha_1}(\sigma) & & & \\ & I_{r_2} \otimes B_{\alpha_2}(\sigma) & & \\ & & \ddots & \\ & & & I_{r_L} \otimes B_{\alpha_L}(\sigma) \end{pmatrix}. \tag{13}$$

Note that, since the left hand side of (13) is an orthogonal matrix, matrices $B_{\alpha_i}(\sigma)$, $i = 1, \ldots, L$, are orthogonal. In the general case, $B_\alpha(g)$ are orthogonal if we work with a G-invariant inner product. In what follows, we will consider only the case $G = \Gamma$ and $\rho = R$. Note that the usual inner product on $V = \mathbb{R}^p$ is clearly Γ-invariant.

The actual formula for $B_{\alpha_i}(\sigma)$ obviously depends on the choice of U_Γ and hence, on the choice of orthonormal basis of $\mathbb{R}^p$. To ensure simplicity of formulation of our next result (Lemma 3), we will work with special orthonormal bases of $V(B_{\alpha_1}), \ldots, V(B_{\alpha_L})$, which together constitute a basis of $\mathbb{R}^p$. Such bases always exist and will be defined in the next section. Usage of these bases is not indispensable for the proof of our main result (Theorem 4), but simplifies it greatly.

EXAMPLE 6. Let $p = 4$ and let $\Gamma = \{\mathrm{id}, (1,2)(3,4)\}$ be the subgroup of $\mathfrak{S}_4$ generated by $\sigma = (1,2)(3,4)$. The matrix representation of σ in the standard basis $(e_i)_i$ of $\mathbb{R}^4$ is

$$R(\sigma) = \begin{pmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix},$$

which has the two eigenvalues 1 and -1 with multiplicity 2 for each. We choose the following orthonormal eigenvectors of $R(\sigma)$:

$$u_1 = \frac{1}{\sqrt{2}}(e_1 + e_2), u_2 = \frac{1}{\sqrt{2}}(e_3 + e_4), u_3 = \frac{1}{\sqrt{2}}(e_1 - e_2), u_4 = \frac{1}{\sqrt{2}}(e_3 - e_4)$$

and let $U_\Gamma = (u_1, u_2, u_3, u_4)$. The corresponding eigenspaces $V_i = \mathbb{R}u_i$ are invariant under $R(\sigma)$ and $R(\mathrm{id}) = I_4$. As V_i, $i = 1, \ldots, 4$, are 1-dimensional, the subrepresentations defined by

$$\rho_i(\gamma) = R(\gamma)|_{V_i} \qquad (\gamma \in \Gamma)$$

are irreducible. We have the decomposition (11) of R:

$$R = \rho_1 \oplus \rho_2 \oplus \rho_3 \oplus \rho_4.$$

The matrix expressions of ρ_1 and ρ_2 are equal to $B_1(\gamma) = (1)$ for all $\gamma \in \Gamma$, since $\rho_i(\gamma)v = v$ for $v \in V_i$, $i = 1, 2$. We have $r_1 = 2$.

The matrix expressions of ρ_3 and ρ_4 are both equal to $B_2(\gamma) = \mathrm{sign}(\gamma)$ for all $\gamma \in \Gamma$, since $\rho_i(\mathrm{id})v = v$ and $\rho_i(\sigma)v = -v$ for $v \in V_i$ for $i = 3, 4$. We have $r_2 = 2$.

The representations ρ_1 and ρ_3 are not equivalent, which can be seen by looking at the characters: $\chi_{\rho_1} = 1$, $\chi_{\rho_3}(\gamma) = \mathrm{sign}(\gamma)$, which are not equal.

In the basis u_1, u_2, u_3, u_4, the matrix of $R(\gamma)$ is (compare with (13))

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \mathrm{sign}(\gamma) & 0 \\ 0 & 0 & 0 & \mathrm{sign}(\gamma) \end{pmatrix} = B_1(\gamma)^{\oplus 2} \oplus B_2(\gamma)^{\oplus 2} = U_\Gamma^\top \cdot R(\gamma) \cdot U_\Gamma.$$

This is the decomposition (12) of R in the basis (u_1, u_2, u_3, u_4).

2.4. *Block diagonal decomposition of $\mathcal{Z}_\Gamma$.* So far, we have shown that through an appropriate change of basis, the representation $(R, \mathbb{R}^p)$ of Γ can be expressed as the direct sum (12) of irreducible subrepresentations. We now want to turn our attention to the elements of $\mathcal{Z}_\Gamma$.

A linear operator $T\colon V \to V$ is said to be an intertwining operator of the representation (ρ, V) if $T \circ \rho(g) = \rho(g) \circ T$ holds for all $g \in G$. In our context, since (4) can be rewritten as

$$\mathcal{Z}_\Gamma = \{\, x \in \mathrm{Sym}(p;\mathbb{R})\,;\, x \cdot R(\sigma) = R(\sigma) \cdot x \text{ for all } \sigma \in \Gamma \,\},$$

$\mathcal{Z}_\Gamma$ is the set of symmetric intertwining operators of the representation $(R, \mathbb{R}^p)$.

Let $\mathrm{End}_\Gamma(\mathbb{R}^p)$ denote the set of all intertwining operators of the representation $(R, \mathbb{R}^p)$ of Γ. Recall that the set A enumerates the elements of $\Lambda(\Gamma)$, the finite set of all equivalence classes of irreducible representations of Γ. From (12) and (13), it is clear that to study $\mathrm{End}_\Gamma(\mathbb{R}^p)$, it is sufficient to study the sets,

$$\mathrm{End}_\Gamma(V_\alpha) = \{\, T \in \mathrm{Mat}(k_\alpha, k_\alpha; \mathbb{R})\,;\, T \cdot B_\alpha(\sigma) = B_\alpha(\sigma) \cdot T \text{ for all } \sigma \in \Gamma \,\},$$

$\alpha \in A$, of all intertwining operators of the irreducible representation B_α, where $V_\alpha := \mathbb{R}^{k_\alpha}$ is the representation space of B_α equipped with a Γ-invariant inner product. Indeed, we have $V(B_\alpha) = I_{r_\alpha} \otimes V_\alpha$.

To reach our main result in this section, Theorem 4, we use the result from (Serre, 1977, Page 108) that, since the representation B_α is irreducible, the space $\mathrm{End}_\Gamma(V_\alpha)$ is isomorphic either to $\mathbb{R}$, $\mathbb{C}$, or the quaternion algebra $\mathbb{H}$. Let

$$f_\alpha \colon \mathbb{K}_\alpha \to \mathrm{End}_\Gamma(V_\alpha),$$

denote this isomorphism, where $\mathbb{K}_\alpha$ is $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. Let

$$d_\alpha := \dim_{\mathbb{R}} \mathrm{End}_\Gamma(V_\alpha) = \dim_{\mathbb{R}} \mathbb{K}_\alpha \in \{1, 2, 4\}.$$

The representation space V_α becomes a vector space over $\mathbb{K}_\alpha$ of dimension k_α/d_α via

$$q \cdot v := f_\alpha(q)v \qquad (q \in \mathbb{K}_\alpha,\ v \in V_\alpha).$$

Clearly the space $\mathbb{R}I_{k_\alpha}$ of scalar matrices is contained in $\mathrm{End}_\Gamma(V_\alpha)$. If $d_\alpha = 1 = \dim_{\mathbb{R}} \mathbb{R}I_{k_\alpha}$, we have $\mathrm{End}_\Gamma(V_\alpha) = \mathbb{R}I_{k_\alpha}$. Further, if $d_\alpha = 2$, take a $\mathbb{C}$-basis $\{v_1,\dots,v_{k_\alpha/2}\}$ of V_α in such a way that $\{v_1,\dots,v_{k_\alpha/2}, i\cdot v_1,\dots,i\cdot v_{k_\alpha/2}\}$ is an orthonormal $\mathbb{R}$-basis of V_α. We identify $\mathbb{R}^{k_\alpha}$ and V_α via this $\mathbb{R}$-basis. Then, the action of $q = a+bi \in \mathbb{C}$ on $w \in \mathbb{R}^{k_\alpha} \simeq V_\alpha$ is expressed as

$$q\cdot w = \begin{pmatrix} aI_{k_\alpha/2} & -bI_{k_\alpha/2} \\ bI_{k_\alpha/2} & aI_{k_\alpha/2} \end{pmatrix} w = \left\{ M_{\mathbb{C}}(a+bi)\otimes I_{k_\alpha/2} \right\} w.$$

Thus, if $d_\alpha = 2$, then

$$\mathrm{End}_\Gamma(V_\alpha) = \left\{ M_{\mathbb{C}}(q)\otimes I_{k_\alpha/2}\,;\, q\in\mathbb{C} \right\} = M_{\mathbb{C}}(\mathbb{C})\otimes I_{k_\alpha/2}.$$

Similarly, when $\mathbb{K}_\alpha = \mathbb{H}$, take an $\mathbb{H}$-basis $\{v_1,\dots,v_{k_\alpha/4}\}$ of V_α so that

$$\{v_1,\dots,v_{k_\alpha/4}, i\cdot v_1,\dots,i\cdot v_{k_\alpha/4}, j\cdot v_1,\dots,j\cdot v_{k_\alpha/4}, k\cdot v_1,\dots,k\cdot v_{k_\alpha/4}\}$$

is an orthonormal $\mathbb{R}$-basis of V_α. The action of $Q\in\mathbb{H}$ on V_α is expressed as $M_{\mathbb{H}}(Q)\otimes I_{k_\alpha/4}$ with respect to this basis.

In this way we have proved the following result.

LEMMA 3. *For each $\alpha\in A$, one has*

$$\mathrm{End}_\Gamma(V_\alpha) = M_{\mathbb{K}_\alpha}(\mathbb{K}_\alpha)\otimes I_{k_\alpha/d_\alpha}. \tag{14}$$

Recall the orthogonal matrix U_Γ in (13). It will be shown in the Appendix that, using Lemma 3, we can represent $\mathcal{Z}_\Gamma$ as follows.

THEOREM 4. *The Jordan algebra $\mathcal{Z}_\Gamma$ is isomorphic to $\bigoplus_{i=1}^L \mathrm{Herm}(r_i;\mathbb{K}_i)$ through the map $\iota\colon \bigoplus_{i=1}^L \mathrm{Herm}(r_i;\mathbb{K}_i) \ni (x_i)_{i=1}^L \mapsto X\in\mathcal{Z}_\Gamma$ given by*

$$X := U_\Gamma \begin{pmatrix} M_{\mathbb{K}_1}(x_1)\otimes I_{k_1/d_1} & & & \\ & M_{\mathbb{K}_2}(x_2)\otimes I_{k_2/d_2} & & \\ & & \ddots & \\ & & & M_{\mathbb{K}_L}(x_L)\otimes I_{k_L/d_L} \end{pmatrix} U_\Gamma^\top. \tag{15}$$

2.5. *Determining the structure constants and invariant measure on $\mathcal{P}_\Gamma$.* As mentioned in the introduction, in order to derive the analytic expression of the Gamma-like functions on $\mathcal{P}_\Gamma$, we need the structure constants r_i, d_i, k_i as well as the invariant measure φ_Γ. However, due to Proposition 9 below, $\varphi_\Gamma(X)$ is expressed in terms of the polynomials $\det(x_i)$, where $x_i\in\mathrm{Herm}(r_i;\mathbb{K}_i)$, $i=1,\dots,L$, coming from decomposition (15). These can be derived from the decomposition of $\mathcal{Z}_\Gamma$ as indicated in Sections 2.1–2.4 above. Let us note that the constants d_i and k_i depend only on the group Γ, while r_i depend on a particular representation of Γ, which is R.

In view of decomposition (15), for $X\in\mathcal{Z}_\Gamma$, define $\phi_i(X) = x_i \in \mathrm{Herm}(r_i;\mathbb{K}_i)$ for $i = 1,\dots,L$.

COROLLARY 5. *For $X\in\mathcal{Z}_\Gamma$, one has*

$$\mathrm{Det}\,(X) = \prod_{i=1}^L \det(\phi_i(X))^{k_i}. \tag{16}$$

PROOF. By (15), we have

$$\begin{aligned}\mathrm{Det}\,(X) &= \prod_{i=1}^{L} \mathrm{Det}\left(M_{\mathbb{K}_i}(x_i) \otimes I_{k_i/d_i}\right) = \prod_{i=1}^{L} \mathrm{Det}\,(M_{\mathbb{K}_i}(x_i))^{k_i/d_i} \\ &= \prod_{i=1}^{L} \left[\det(x_i)^{d_i}\right]^{k_i/d_i} = \prod_{i=1}^{L} \det(x_i)^{k_i},\end{aligned}$$

whence follows the formula. We have used (8) and (9) for the third equality above. □

REMARK 7. Let us note that (16) gives us a simple way to find the structure constants $\{k_i, r_i, d_i\}_{i=1}^{L}$. Indeed, assume that we have an irreducible factorization

$$\mathrm{Det}\,(X) = \prod_{j=1}^{L'} f_j(X)^{a_j} \quad (X \in \mathcal{Z}_\Gamma), \tag{17}$$

where each a_j is a positive integer, each $f_j(X)$ is an irreducible polynomial of $X \in \mathcal{Z}_\Gamma$, and $f_i \neq f_j$ if $i \neq j$. Since the determinant polynomial of a simple Jordan algebra is always irreducible (Upmeier, 1986, Lemma 2.3 (1)), comparing (16) and (17), we obtain that $L = L'$, and that, for each j, there exists i such that $f_j(X)^{a_j} = \det(\phi_i(X))^{k_i}$. Then $k_i = a_j$ and r_i is the degree of $f_j(X) = \det(\phi_i(X))$. From the structure theory of Jordan algebras, we see that, if $E_j \in \mathcal{Z}_\Gamma$ is the gradient of $f_j(X)$ at $X = I_p$, then the linear operator $P_j : \mathcal{Z}_\Gamma \ni x \mapsto E_j \circ x \in \mathcal{Z}_\Gamma$ coincides with the projection $\bigoplus_{i=1}^{L} \mathrm{Herm}(r_i; \mathbb{K}_i) \to \mathrm{Herm}(r_i; \mathbb{K}_i)$. In particular, $\dim_{\mathbb{R}} \mathrm{Herm}(r_i; \mathbb{K}_i) = r_i + d_i r_i (r_i - 1)/2$ is equal to the rank of P_j, from which we can determine d_i when $r_i > 1$.

If $r_i = 1$, the determination of d_i is not needed for writing the block decomposition of $\mathcal{Z}_\Gamma$, since in this case $\mathbb{R} = \mathrm{Herm}(1; \mathbb{R}) = \mathrm{Herm}(1; \mathbb{C}) = \mathrm{Herm}(1; \mathbb{H})$ and, if k_i is divisible by 2 or by 4, we have $M_{\mathbb{K}_i}(x_i) \otimes I_{k_i/d_i} = x_i I_{k_i}$.

The practical significance of the method proposed in this Remark is that neither representation theory nor group theory is used. It is a strong advantage when we consider colorings corresponding to a large number of different groups, for which finding structure constants is very complicated.

REMARK 8. The factorization of multivariate polynomials over an algebraic number field can be done for example in PYTHON (see *sympy.polys.polytools.factor*). Indeed, in our setting, the irreducible factorization over the real number field coincides with the one over the real cyclotomic field

$$\mathbb{Q}\left[\zeta + \frac{1}{\zeta}\right] = \left\{ \sum_{k=0}^{\varphi_E(M)/2-1} q_k \left(\zeta + \frac{1}{\zeta}\right)^k ;\, q_k \in \mathbb{Q},\ k = 0, 1, \ldots, \varphi_E(M)/2 - 1 \right\},$$

where ζ is the primitive M-th root $e^{2\pi i/M}$ of unity with M being the least common multiple of the orders of elements $\sigma \in \Gamma$, and $\varphi_E(M)$ is the number of positive integers up to M that are relatively prime to M (Serre, 1977, Section 12.3).

An example showing the utility of Remark 7 can be found in the Appendix (see Example 18).

2.6. *Construction of the orthogonal matrix U_Γ when Γ is cyclic.* We now show that, when the group Γ is generated by one permutation $\sigma \in \mathfrak{S}_p$, the orthogonal matrix U_Γ can be constructed explicitly, and we obtain the structure constants r_i, k_i and d_i easily.

Let us consider the Γ-orbits in $\{1, 2, \dots, p\}$. Let $\{i_1, \dots, i_C\}$ be a complete system of representatives of the Γ-orbits, and for each $c = 1, \dots, C$, let p_c be the cardinality of the Γ-orbit through i_c. The order N of Γ equals the least common multiple of $p_1, p_2, \dots, p_C$ and one has $\Gamma = \{\mathrm{id}, \sigma, \sigma^2, \dots, \sigma^{N-1}\}$.

THEOREM 6. *Let $\Gamma = \langle\sigma\rangle$ be a cyclic group of order N. For $\alpha = 0, 1, \dots, \lfloor \frac{N}{2} \rfloor$ set*

$$r^*_\alpha = \#\{\, c \in \{1, \dots, C\}\,;\, \alpha\, p_c \text{ is a multiple of } N\,\}$$

$$d^*_\alpha = \begin{cases} 1 & (\alpha = 0 \text{ or } N/2) \\ 2 & (\text{otherwise}). \end{cases}$$

Then we have $L = \#\{\,\alpha\,;\, r^*_\alpha > 0\,\}$, $r = (r^*_\alpha;\, r^*_\alpha > 0)$ and $k = d = (d^*_\alpha;\, r^*_\alpha > 0)$.

For $c = 1, \dots, C$ define $v^{(c)}_1, \dots, v^{(c)}_{p_c} \in \mathbb{R}^p$ by

$$v^{(c)}_1 := \sqrt{\frac{1}{p_c}} \sum_{k=0}^{p_c - 1} e_{\sigma^k(i_c)},$$

$$v^{(c)}_{2\beta} := \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c - 1} \cos\Bigl(\frac{2\pi\beta k}{p_c}\Bigr) e_{\sigma^k(i_c)} \qquad (1 \le \beta < p_c/2),$$

$$v^{(c)}_{2\beta+1} := \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c - 1} \sin\Bigl(\frac{2\pi\beta k}{p_c}\Bigr) e_{\sigma^k(i_c)} \qquad (1 \le \beta < p_c/2),$$

$$v^{(c)}_{p_c} := \sqrt{\frac{1}{p_c}} \sum_{k=0}^{p_c - 1} \cos(\pi k) e_{\sigma^k(i_c)} \qquad (\text{if } p_c \text{ is even}).$$

THEOREM 7. *The orthogonal matrix U_Γ from Theorem 4 can be obtained by arranging column vectors $\{v^{(c)}_k\}$, $1 \le c \le C$, $1 \le k \le p_c$ in the following way: we put $v^{(c)}_k$ earlier than $v^{(c')}_{k'}$ if*
(i) $\frac{[k/2]}{p_c} < \frac{[k'/2]}{p_{c'}}$, *or*
(ii) $\frac{[k/2]}{p_c} = \frac{[k'/2]}{p_{c'}}$ *and $c < c'$, or*
(iii) $\frac{[k/2]}{p_c} = \frac{[k'/2]}{p_{c'}}$ *and $c = c'$ and k is even and k' is odd.*

Proofs of the above results are postponed to the Appendix. We shall see there that $R(\sigma)$ acts on the 2-dimensional space spanned by $v^{(c)}_{2\beta}$ and $v^{(c)}_{2\beta+1}$ as a rotation with the angle $2\pi\beta/p_c$. The condition (i) means that the angle for $v^{(c)}_k$ is smaller than the one for $v^{(c')}_{k'}$.

EXAMPLE 9. Let us consider $\sigma = (1,2,3)(4,5)(6) \in \mathfrak{S}_6$. The three Γ-orbits are $\{1,2,3\}, \{4,5\}$ and $\{6\}$. Set $i_1 = 1$, $i_2 = 4$, $i_3 = 6$. Then $p_1 = 3$, $p_2 = 2$, $p_3 = 1$. We have $N = 6$. We count $r^*_0 = 3$, $r^*_1 = 0$, $r^*_2 = 1$, $r^*_3 = 1$, so that $\mathcal{Z}_\Gamma \simeq \mathrm{Sym}(3;\mathbb{R}) \oplus \mathrm{Herm}(1;\mathbb{C}) \oplus$

$\mathrm{Sym}(1;\mathbb{R})$. According to Theorem 7,

$$U_\Gamma = \left(v_1^{(1)}, v_1^{(2)}, v_1^{(3)}, v_2^{(1)}, v_3^{(1)}, v_2^{(2)}\right) = \begin{pmatrix} 1/\sqrt{3} & 0 & 0 & \sqrt{2/3} & 0 & 0 \\ 1/\sqrt{3} & 0 & 0 & -\sqrt{1/6} & 1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 0 & 0 & -\sqrt{1/6} & -1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 0 & 0 & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 & 0 & 0 & -1/\sqrt{2} \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Then we have (cf. (13))

$$U_\Gamma^\top \cdot R(\sigma^k) \cdot U_\Gamma = \begin{pmatrix} I_3 \otimes B_0(\sigma^k) & & \\ & B_2(\sigma^k) & \\ & & B_3(\sigma^k) \end{pmatrix},$$

where $B_0(\sigma^k) = 1$, $B_2(\sigma^k) = \mathrm{Rot}(\frac{2\pi k}{3}) \in \mathrm{GL}(2;\mathbb{R})$ and $B_3(\sigma^k) = (-1)^k$. Here $\mathrm{Rot}(\theta)$ denotes the rotation matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ for $\theta \in \mathbb{R}$.

The block diagonal decomposition of $\mathcal{Z}_\Gamma$ is

$$U_\Gamma^\top \cdot \mathcal{Z}_\Gamma \cdot U_\Gamma = \left\{ \begin{pmatrix} x_1 & & \\ & x_2 I_2 & \\ & & x_3 \end{pmatrix} ; x_1 \in \mathrm{Sym}(3;\mathbb{R}),\ x_2, x_3 \in \mathbb{R} \right\}.$$

2.7. *The case of a non-cyclic group.* In Section 2.5, we gave the general algorithm for determining the structure constants as well as the invariant measure φ_Γ. In principle, the factorization of a determinant can be done in PYTHON, however there are some limitations regarding the dimension of a matrix. If the $p \times p$ matrix is not sparse, then the number of terms in the usual Laplace expansion of a determinant produces a polynomial with $p!$ terms. The RAM memory requirements for calculating such a polynomial would be in excess of $p!$, which cannot be handled on a standard PC even for moderate p. Depending on the subgroup and the method of calculating the determinant, we were able to obtain the determinant for models of dimensions up to 10-20. In order to factorize the determinant for moderate to high dimensions, we want to find an orthogonal matrix U such that $U^\top \cdot X \cdot U$ is sparse enough for a computer to calculate its determinant $\mathrm{Det}\left(U^\top \cdot X \cdot U\right) = \mathrm{Det}\,(X)$. The matrix U_Γ from (1) is in general very hard to obtain, but we propose an easy surrogate.

Let Γ be a subgroup of $\mathfrak{S}_p$, which is not necessarily a cyclic group. Let $\{i_1, i_2, \ldots, i_C\}$ be a complete family of representatives of the Γ-orbits in $V = \{1,\ldots,p\}$. Take $\sigma_0 \in \mathfrak{S}_p$ for which the cyclic group $\Gamma_0 := \langle\sigma_0\rangle$ generated by σ_0 has the same orbits in V as Γ does. In general, there are no inclusion relations between Γ and Γ_0. However, we observe that a vector $v \in \mathbb{R}^p$ is Γ-invariant, (i.e. $R(\sigma)v = v$ for all $\sigma \in \Gamma$) if and only if $R(\sigma_0)v = v$ if and only if v is constant on Γ-orbits (i.e. $v_i = v_j$ if i and j belong to the same orbit of Γ); see Example 10 below.

Let U_{Γ_0} be the orthogonal matrix constructed as in Theorem 7 from the cyclic group Γ_0. Note that the first C column vectors of U_{Γ_0} are $v_1^{(1)}, v_1^{(2)}, \ldots, v_1^{(C)}$, which are Γ-invariant. The space $V_1 := \mathrm{span}\left\{ v_1^{(c)} ; c = 1,\ldots,C \right\} \subset \mathbb{R}^p$ is the trivial-representation-component of Γ as explained after (12). Therefore, if α_1 is the trivial representation of Γ, then $r_1 = C$ and $d_1 = k_1 = 1$.

The orthogonal complement $V_1^\perp$ of V_1 is spanned by the rest of $v_\beta^{(c)}$, $1 \le c \le C$, $1 < \beta \le 2[p_c/2]$. For $X \in \mathcal{Z}_\Gamma$, we see that $X \cdot v \in V_1$ for $v \in V_1$ and that $X \cdot w \in V_1^\perp$ for $w \in V_1^\perp$. In other words,

$$U_{\Gamma_0}^{\top} \cdot X \cdot U_{\Gamma_0} = \begin{pmatrix} x_1 & 0 \\ 0 & y \end{pmatrix} \tag{18}$$

with $x_1 \in \mathrm{Sym}(C;\mathbb{R})$ and $y \in \mathrm{Sym}(p-C;\mathbb{R})$. Actually, the correspondence $\phi_1 \colon \mathcal{Z}_\Gamma \ni X \mapsto x_1 \in \mathrm{Sym}(C;\mathbb{R})$ is exactly the Jordan algebra homomorphism defined before Corollary 5. Therefore

$$\mathrm{Det}\,(X) = \mathrm{Det}\,(x_1)\,\mathrm{Det}\,(y)\,, \tag{19}$$

while the factor $\mathrm{Det}\,(x_1) = \det(\phi_1(X))$ is an irreducible polynomial of degree $r_1 = C$. In this way, for any subgroup Γ, we are able to factor out the polynomial of degree equal to the number of Γ-orbits in V easily. On the other hand, the factorization of $\mathrm{Det}\,(y)$ requires study of the subrepresentation R of Γ on $V_1^{\perp}$, where the group Γ_0 is useless in general.

EXAMPLE 10. Let $\Gamma = \langle (1,2,3),\,(4,5,6)\rangle \subset \mathfrak{S}_6$, which is not a cyclic group. The space $\mathcal{Z}_\Gamma$ consists of symmetric matrices of the form

$$X = \begin{pmatrix} a & b & b & e & e & e \\ b & a & b & e & e & e \\ b & b & a & e & e & e \\ e & e & e & c & d & d \\ e & e & e & d & c & d \\ e & e & e & d & d & c \end{pmatrix}$$

and moreover, $\mathcal{Z}_\Gamma$ does not coincide with $\mathcal{Z}_{\langle\sigma\rangle}$ for any $\sigma \in \mathfrak{S}_6$. Noting that the group Γ has two orbits: $\{1,2,3\}$ and $\{4,5,6\}$, we define $\sigma_0 := (1,2,3)(4,5,6)$. Taking $i_1 = 1$ and $i_2 = 4$, we have

$$U_{\Gamma_0} = \begin{pmatrix} 1/\sqrt{3} & 0 & \sqrt{2/3} & 0 & 0 & 0 \\ 1/\sqrt{3} & 0 & -1/\sqrt{6} & 1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{3} & 0 & -1/\sqrt{6} & -1/\sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{3} & 0 & 0 & \sqrt{2/3} & 0 \\ 0 & 1/\sqrt{3} & 0 & 0 & -1/\sqrt{6} & 1/\sqrt{2} \\ 0 & 1/\sqrt{3} & 0 & 0 & -1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix}.$$

Note that the first two column vectors $\left(1/\sqrt{3}\; 1/\sqrt{3}\; 1/\sqrt{3}\; 0\; 0\; 0\right)^{\top}$ and $\left(0\; 0\; 0\; 1/\sqrt{3}\; 1/\sqrt{3}\; 1/\sqrt{3}\right)^{\top}$ of U_{Γ_0} are Γ-invariant. By direct calculation we verify that $U_{\Gamma_0}^{\top} \cdot X \cdot U_{\Gamma_0}$ is of the form

$$\begin{pmatrix} A & B & 0 & 0 & 0 & 0 \\ B & C & 0 & 0 & 0 & 0 \\ 0 & 0 & D & 0 & 0 & 0 \\ 0 & 0 & 0 & D & 0 & 0 \\ 0 & 0 & 0 & 0 & E & 0 \\ 0 & 0 & 0 & 0 & 0 & E \end{pmatrix},$$

where $A, B, \cdots, E$ are linear functions of $a, b, \cdots, e$. The matrices x_1 and y are $\begin{pmatrix} A & B \\ B & C \end{pmatrix}$ and $\begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & E & 0 \\ 0 & 0 & 0 & E \end{pmatrix}$ respectively. The matrix y is of such simple form, because Γ_0 is a subgroup of Γ in this case.

We cannot expect that the matrix y in (18) is always of a nice form, as in the example above. However, we note that in many examples we considered, the matrix y was sparse, which also makes the problem of calculating $\mathrm{Det}\,(X)$ much more feasible on a standard PC.

In general Γ_0 defined above is not a subgroup of Γ. As we argue below, valuable insight about the factorization of $\mathcal{Z}_\Gamma$ can be obtained by studying cyclic subgroups of Γ. In general, if Γ_1 is a subgroup of Γ, then $\mathcal{Z}_\Gamma$ is a subspace of $\mathcal{Z}_{\Gamma_1}$.

Let Γ_1 be a cyclic subgroup of Γ and let U_{Γ_1} be the orthogonal matrix constructed in Theorem 7. By Theorem 4, for any $X \in \mathcal{Z}_\Gamma \subset \mathcal{Z}_{\Gamma_1}$ the matrix $U_{\Gamma_1}^\top \cdot X \cdot U_{\Gamma_1}$ belongs to the space

$$\left\{ \begin{pmatrix} M_{\mathbb{K}_1}(x_1') \otimes I_{\frac{k_1}{d_1}} & & & \\ & M_{\mathbb{K}_2}(x_2') \otimes I_{\frac{k_2}{d_2}} & & \\ & & \ddots & \\ & & & M_{\mathbb{K}_L}(x_L') \otimes I_{\frac{k_L}{d_L}} \end{pmatrix} ;\ \begin{matrix} x_i' \in \mathrm{Herm}(r_i;\mathbb{K}_i) \\ i=1,\ldots,L \end{matrix} \right\},$$

where the structure constants $(k_i,d_i,r_i)_{i=1}^L$ can be easily calculated using Theorem 6. In particular, we have $k_1 = d_1 = 1$ and r_1 is the number of Γ_1-orbits in $\{1,\ldots,p\}$. Thus, we have $M_{\mathbb{K}_1}(x_1') \otimes I_{k_1/d_1} = x_1' \in \mathrm{Sym}(r_1;\mathbb{R})$. In contrast to (18), x_1' in general can be further factorized and we know that $\mathrm{Det}\,(x_1)$ from (19) is an irreducible factor of $\mathrm{Det}\,(x_1')$. In conclusion, each cyclic subgroup of the general group Γ brings various information about the factorization.

EXAMPLE 11. We continue Example 10. Let $\Gamma_1 = \langle (1,2,3) \rangle$, which is a subgroup of Γ. There are four Γ_1-orbits in V, that is, $\{1,2,3\}$, $\{4\}$, $\{5\}$, and $\{6\}$. We have

$$U_{\Gamma_1} = \begin{pmatrix} 1/\sqrt{3} & 0 & 0 & 0 & \sqrt{2/3} & 0 \\ 1/\sqrt{3} & 0 & 0 & 0 & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 & 0 & 0 & -1/\sqrt{6} & -1/\sqrt{2} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

For $X \in Z_\Gamma$, we see that $U_{\Gamma_1}^\top \cdot X \cdot U_{\Gamma_1}$ is of the form

$$\begin{pmatrix} A_{11} & A_{21} & A_{31} & A_{41} & 0 & 0 \\ A_{21} & A_{22} & A_{32} & A_{42} & 0 & 0 \\ A_{31} & A_{32} & A_{33} & A_{43} & 0 & 0 \\ A_{41} & A_{42} & A_{43} & A_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & D & 0 \\ 0 & 0 & 0 & 0 & 0 & D \end{pmatrix},$$

where A_{ij} are linear functions of $a,b,\ldots,e$, but they are not linearly independent. Indeed, we have

$$\mathrm{Det}\begin{pmatrix} A_{11} & A_{21} & A_{31} & A_{41} \\ A_{21} & A_{22} & A_{32} & A_{42} \\ A_{31} & A_{32} & A_{33} & A_{43} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} = E^2 \det\begin{pmatrix} A & B \\ B & C \end{pmatrix},$$

which exemplifies the fact that $\mathrm{Det}\,(x_1)$ is an irreducible factor of $\mathrm{Det}\,(x_1')$.

3. Gamma integrals.

3.1. *Gamma integrals on irreducible symmetric cones.* Let Ω be one of the first three kinds of irreducible symmetric cones, that is, $\Omega = \mathrm{Herm}^+(r;\mathbb{K})$, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$. Determinant and trace on corresponding Euclidean Jordan algebras will be denoted by $\det$ and tr. Then, we have the relation

$$\dim \Omega = r + \frac{r(r-1)}{2} d,$$

where $d = 1$ if $\mathbb{K} = \mathbb{R}$, $d = 2$ if $\mathbb{K} = \mathbb{C}$ and $d = 4$ if $\mathbb{K} = \mathbb{H}$.

Let $m(\mathrm{d}x)$ denote the Euclidean measure associated with the Euclidean structure defined on $\mathcal{A} = \mathrm{Herm}(r;\mathbb{K})$ by $\langle x, y\rangle = \mathrm{tr}[x \bullet y] = \mathrm{tr}[x \cdot y]$. The Gamma integral

$$\Gamma_\Omega(\lambda) := \int_\Omega \det(x)^\lambda e^{-\mathrm{tr}[x]} \det(x)^{-\dim\Omega/r} m(\mathrm{d}x)$$

is finite if and only if $\lambda > \frac{1}{2}(r-1)d = \dim\Omega/r - 1$ and in such case

$$\Gamma_\Omega(\lambda) = (2\pi)^{(\dim\Omega - r)/2}\Gamma(\lambda)\Gamma(\lambda - d/2)\ldots\Gamma(\lambda-(r-1)d/2). \tag{20}$$

Moreover, one has

$$\int_\Omega \det(x)^\lambda e^{-\mathrm{tr}[x\bullet y]} \det(x)^{-\dim\Omega/r} m(\mathrm{d}x) = \Gamma_\Omega(\lambda)\det(y)^{-\lambda} \tag{21}$$

for any $y \in \Omega$.

Let us explain the role of the measure $\mu_\Omega(\mathrm{d}x) = \det(x)^{-\dim\Omega/r} m(\mathrm{d}x)$. Let $\mathrm{G}(\Omega)$ be the linear automorphism group of Ω, that is, the set $\{\, g \in \mathrm{GL}(\mathcal{A})\,;\, g\Omega = \Omega\,\}$, where $\mathcal{A}$ is the associated Euclidean Jordan algebra. Then, the measure μ_Ω is a $\mathrm{G}(\Omega)$-invariant measure in the sense that for any Borel measurable set B one has

$$\mu_\Omega(g^{-1}B) = \mu_\Omega(B) \qquad (g \in \mathrm{G}(\Omega)).$$

3.2. *Gamma integrals on the cone $\mathcal{P}_\Gamma$.* We endow the space $\mathcal{Z}_\Gamma$ with the scalar product

$$\langle x, y\rangle = \mathrm{Tr}\,[x\cdot y] \qquad (x, y \in \mathcal{Z}_\Gamma).$$

Let $\mathrm{d}X$ denote the Euclidean measure on the Euclidean space $(\mathcal{Z}_\Gamma, \langle\cdot,\cdot\rangle)$. Let us note that this normalization is not important in the model selection procedure as there we always consider quotients of integrals.

EXAMPLE 12. Consider $p = 3$ and $\Gamma = \mathfrak{S}_3$. The space $\mathcal{Z}_\Gamma$ is 2-dimensional and it consists of matrices of the form (see Example 1)

$$X = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}$$

for $a, b \in \mathbb{R}$. Since $\|X\|^2 = \mathrm{Tr}\left[X^2\right] = 3a^2 + 6b^2 = v^\top v$ with $v^\top = (\sqrt{3}a, \sqrt{6}b)$, we have $\mathrm{d}X = \sqrt{3}\sqrt{6}\,\mathrm{d}a\,\mathrm{d}b = 3\sqrt{2}\,\mathrm{d}a\,\mathrm{d}b$.

Generally, if m_i denotes the Euclidean measure on $\mathcal{A}_i := \mathrm{Herm}(r_i;\mathbb{K}_i)$ with the inner product defined from the Jordan algebra trace (recall (8) and (9)), then (1) implies that for $X \in \mathcal{Z}_\Gamma$ we have

$$\|X\|^2 = \langle X, X\rangle = \sum_{i=1}^L \frac{k_i}{d_i}\mathrm{Tr}\left[M_{\mathbb{K}_i}(x_i)^2\right] = \sum_{i=1}^L k_i \mathrm{tr}[(x_i)^2],$$

which implies that

$$\mathrm{d}X = \prod_{i=1}^{L} \left(\sqrt{k_i}\right)^{\dim\Omega_i} m_i(\mathrm{d}x_i) = \prod_{i=1}^{L} k_i^{\dim\Omega_i/2} m_i(\mathrm{d}x_i). \tag{22}$$

The key ingredient to compute the Gamma integral on $\mathcal{P}_\Gamma$ is Theorem 4. Let Ω_i denote the symmetric cone of the simple Jordan algebra $\mathcal{A}_i = \mathrm{Herm}(r_i;\mathbb{K}_i)$ and $d_i = \dim_{\mathbb{R}} \mathbb{K}_i$, $i = 1,\ldots,L$. Recall that $\dim\Omega_i = r_i + r_i(r_i-1)d_i/2$. We start with the following more general integral. Recall that, for $X \in \mathcal{Z}_\Gamma$ represented as in (1), we write $\phi_i(X) = x_i \in \mathcal{A}_i$ for $i = 1,\ldots,L$.

LEMMA 8. *For any $Y \in \mathcal{P}_\Gamma$ and $\lambda_i > -1$, $i = 1,\ldots,L$, we have*

$$\int_{\mathcal{P}_\Gamma} \prod_{i=1}^{L} \det(\phi_i(X))^{\lambda_i} e^{-\mathrm{Tr}[Y\cdot X]}\,\mathrm{d}X = e^{-B_\Gamma}\left(\prod_{i=1}^{L} k_i^{-r_i\lambda_i}\right)\prod_{i=1}^{L}\frac{\Gamma_{\Omega_i}(\lambda_i+\dim\Omega_i/r_i)}{\det(\phi_i(Y))^{\lambda_i+\dim\Omega_i/r_i}}, \tag{23}$$

where

$$B_\Gamma := \frac{1}{2}\sum_{i=1}^{L}(\dim\Omega_i)(\log k_i). \tag{24}$$

The proof is postponed to the Appendix.

DEFINITION 13. Let $\mathrm{G}(\mathcal{P}_\Gamma) = \{\, g \in \mathrm{GL}(p;\mathbb{R})\,;\, g\mathcal{P}_\Gamma = \mathcal{P}_\Gamma \,\}$ be the linear automorphism group of $\mathcal{P}_\Gamma$. We define the $\mathrm{G}(\mathcal{P}_\Gamma)$-invariant measure $\varphi_\Gamma(X)\,\mathrm{d}X$ by

$$\varphi_\Gamma(X) = e^{B_\Gamma}\int_{\mathcal{P}_\Gamma^*} e^{-\mathrm{Tr}[X\cdot Z]}\,\mathrm{d}Z.$$

PROPOSITION 9. *We have*

$$\varphi_\Gamma(X) = \prod_{i=1}^{L}\det(\phi_i(X))^{-\dim\Omega_i/r_i}. \tag{25}$$

PROOF. Recall that $\mathcal{P}_\Gamma$ is a symmetric cone, so that it coincides with its dual cone, $\mathcal{P}_\Gamma^*$. Thus,

$$\varphi_\Gamma(Y) = e^{B_\Gamma}\int_{\mathcal{P}_\Gamma} e^{-\mathrm{Tr}[Y\cdot X]}\,\mathrm{d}X \qquad (Y \in \mathcal{P}_\Gamma).$$

Setting $\lambda_1 = \ldots = \lambda_L = 0$ in (23) we obtain the expression of $\varphi_\Gamma(Y)$ above. □

DEFINITION 14. The Gamma function of $\mathcal{P}_\Gamma$ is defined by the following integral

$$\Gamma_{\mathcal{P}_\Gamma}(\lambda) := \int_{\mathcal{P}_\Gamma} \mathrm{Det}\,(X)^{\lambda}\, e^{-\mathrm{Tr}[X]}\varphi_\Gamma(X)\,\mathrm{d}X, \tag{26}$$

whenever it converges.

THEOREM 10. *The integral (26) converges if and only if*

$$\lambda > \max_{i=1,\ldots,L}\left\{\frac{(r_i-1)d_i}{2k_i}\right\} \tag{27}$$

and, for these values of λ*, we have*

$$\Gamma_{\mathcal{P}_\Gamma}(\lambda) = e^{-A_\Gamma \lambda + B_\Gamma} \prod_{i=1}^{L} \Gamma_{\Omega_i}(k_i \lambda) \tag{28}$$

where Γ_{Ω_i} *is given in* (20)*,* B_Γ *in* (24) *and*

$$A_\Gamma := \sum_{i=1}^{L} r_i\, k_i \log k_i. \tag{29}$$

Moreover, if $Y \in \mathcal{P}_\Gamma$ *and* (27) *holds true, then*

$$\int_{\mathcal{P}_\Gamma} \mathrm{Det}\,(X)^{\lambda}\, e^{-\mathrm{Tr}[Y\cdot X]} \varphi_\Gamma(X)\, \mathrm{d}X = \Gamma_{\mathcal{P}_\Gamma}(\lambda) \mathrm{Det}\,(Y)^{-\lambda}. \tag{30}$$

We also have the following result

THEOREM 11. *If* $Y \in \mathcal{P}_\Gamma$ *and*

$$\lambda > \max_{i=1,\ldots,L} \left\{ -\frac{1}{k_i} \right\}, \tag{31}$$

then

(32)

$$\int_{\mathcal{P}_\Gamma} \mathrm{Det}\,(X)^{\lambda}\, e^{-\mathrm{Tr}[Y\cdot X]}\, \mathrm{d}X = e^{-A_\Gamma \lambda - B_\Gamma} \prod_{i=1}^{L} \Gamma_{\Omega_i}\left(k_i\, \lambda + \frac{\dim \Omega_i}{r_i}\right) \mathrm{Det}\,(Y)^{-\lambda}\, \varphi_\Gamma(Y).$$

PROOF OF THEOREM 10 AND THEOREM 11. Recall that for $X \in \mathcal{P}_\Gamma$ we have $\mathrm{Det}\,(X) = \prod_{i=1}^{L} \det(\phi_i(X))^{k_i}$, where the map $\phi_i : \mathcal{Z}_\Gamma \to \mathrm{Herm}(r_i; \mathbb{K}_i)$ is a Jordan algebra homomorphism, $i = 1, \ldots, L$.

If $\lambda_i = k_i \lambda - \dim \Omega_i / r_i$ and (27) holds, then (23) implies

$$\prod_{i=1}^{L} k_i^{-r_i \lambda_i} = e^{-A_\Gamma \lambda + 2B_\Gamma} \quad \text{and} \quad \prod_{i=1}^{L} \det(\phi_i(Y))^{-\lambda_i - \dim \Omega_i / r_i} = \left(\prod_{i=1}^{L} \det(\phi_i(Y))^{k_i} \right)^{-\lambda}$$

If $\lambda_i = k_i \lambda$ with (31), then by (23) we obtain

$$\prod_{i=1}^{L} k_i^{-r_i \lambda_i} = e^{-A_\Gamma \lambda} \quad \text{and} \quad \prod_{i=1}^{L} \det(\phi_i(Y))^{-\lambda_i - \dim \Omega_i / r_i} = \left(\prod_{i=1}^{L} \det(\phi_i(Y))^{k_i} \right)^{-\lambda} \varphi_\Gamma(Y).$$

□

3.3. *RCOP-Wishart laws on* $\mathcal{P}_\Gamma$. Let $\Sigma \in \mathcal{P}_\Gamma \subset \mathrm{Sym}^+(p; \mathbb{R})$ and consider i.i.d. random vectors $Z^{(1)}, \ldots, Z^{(n)}$ following the $\mathrm{N}_p(0, \Sigma)$ distribution. Define $U_i = Z^{(i)} \cdot {Z^{(i)}}^\top$, $i = 1, \ldots, n$, and $U = \sum_{i=1}^{n} U_i$.

Our aim is to analyse the probability distribution of the random matrix

$$W_n = \pi_\Gamma(U) = \pi_\Gamma(U_1 + \cdots + U_n) = \pi_\Gamma(U_1) + \ldots + \pi_\Gamma(U_n).$$

In the rest of this section, we find n_0 such that for $n \geq n_0$ the random matrix W_n follows an absolutely continuous law, and we compute its density. Further, we extend the shape parameter to a continuous range and define the RCOP-Wishart law on $\mathcal{P}_\Gamma$.

We start with the following easy result.

LEMMA 12. *For any $\theta \in \mathrm{Sym}^+(p;\mathbb{R})$ we have*

$$\mathbb{E}e^{-\mathrm{Tr}[\theta\cdot\pi_\Gamma(U_1)]} = \mathrm{Det}\,(I_p + 2\,\Sigma\cdot\pi_\Gamma(\theta))^{-1/2}.$$

PROOF. Using (5) repeatedly we have

$$\mathrm{Tr}\,[\theta\cdot\pi_\Gamma\,(U_1)] = \mathrm{Tr}\,[\pi_\Gamma(\theta)\cdot\pi_\Gamma\,(U_1)] = \mathrm{Tr}\,[\pi_\Gamma(\theta)\cdot U_1]\,.$$

The assertion follows from the usual multivariate Gauss integral. □

PROPOSITION 13. *The law of W_n is absolutely continuous on $\mathcal{P}_\Gamma$ if and only if*

$$n \ge n_0 := \max_{i=1,\dots,L}\left\{\frac{r_i d_i}{k_i}\right\}. \tag{33}$$

If $n \ge n_0$, then its density function with respect to $\mathrm{d}X$ is given by

$$\frac{\mathrm{Det}\,(X)^{n/2}\,e^{-\frac{1}{2}\mathrm{Tr}[X\cdot\Sigma^{-1}]}}{\mathrm{Det}\,(2\Sigma)^{n/2}\,\Gamma_{\mathcal{P}_\Gamma}(\frac{n}{2})}\varphi_\Gamma(X)\mathbf{1}_{\mathcal{P}_\Gamma}(X). \tag{34}$$

PROOF. With $\lambda = n/2$, condition (27) becomes

$$n > \max_{i=1,\dots,L}\left\{\frac{(r_i-1)d_i}{k_i}\right\}.$$

Since the quotient k_i/d_i is an integer, the last condition is equivalent to (33).

In view of Lemma 12, it is enough to show that W_n has density (34) if and only if for any $\theta \in \mathcal{P}_\Gamma$,

$$\mathbb{E}e^{-\mathrm{Tr}[\theta\cdot W_n]} = \mathrm{Det}\,(I_p + 2\,\Sigma\cdot\theta)^{-n/2}. \tag{35}$$

This follows directly from (30). □

It is known that the MLE exists and is unique if and only if the sufficient statistic lies in the interior of its convex support, see Barndorff-Nielsen (2014). It is clear that if (33) is not satisfied, then the support of W_n is contained in the boundary of $\mathcal{P}_\Gamma$.

COROLLARY 14. *If the number of samples n satisfies (33), then the MLE of Σ exists and is given by*

$$\hat{\Sigma} = \frac{1}{n}\pi_\Gamma(U_1 + \cdots + U_n).$$

The above result has been already proven in (Andersson, 1975, Theorem 5.9) (see also (Andersson and Madsen, 1998, Sec. A.3, A.4)). Moreover, in (Andersson and Madsen, 1998, Sec. A.5) a formula for $\mathbb{E}\left[\mathrm{Det}\left(\hat{\Sigma}\right)^\alpha\right]$ has been given. The right hand side of their formula (A.4) coincides with $\frac{2^{\alpha p}\Gamma_{\mathcal{P}_\Gamma}(\alpha+\frac{n}{2})}{n^{\alpha p}\Gamma_{\mathcal{P}_\Gamma}(\frac{n}{2})}$ in our notation (substitute $(d_i, nk_i/d_i, r_i)_i$ for their $(d_\mu, n_\mu, p_\mu)_\mu$). However, their formula does not imply an explicit expression for $I_\Gamma(\delta, D)$.

Let us recall that the MLE of Σ in the standard normal model exists if and only if $n \ge p$. We recover this case for $\Gamma = \{\mathrm{id}\}$, since then we have $L = 1$, $r_1 = p$ and $k_1 = d_1 = 1$.

When $n < n_0$, the law of W_n is singular, and it can be described as a direct product of the singular Wishart laws on the irreducible symmetric cones Ω_i, see e.g. Hassairi and Lajmi (2001).

DEFINITION 15. Let $\eta > \max\left\{ (r_i - 1)\frac{d_i}{k_i}\,;\, i = 1, \ldots, L \right\}$ and $\Sigma \in \mathcal{P}_\Gamma$. The RCOP–Wishart law $W^\Gamma_{\eta,\Sigma}$ is defined by its density

$$W^\Gamma_{\eta,\Sigma}(\mathrm{d}X) = \frac{\mathrm{Det}\,(X)^{\eta/2}\, e^{-\frac{1}{2}\mathrm{Tr}[X\cdot\Sigma^{-1}]}}{\mathrm{Det}\,(2\Sigma)^{\eta/2}\,\Gamma_{\mathcal{P}_\Gamma}(\frac{\eta}{2})} \varphi_\Gamma(X)\mathbf{1}_{\mathcal{P}_\Gamma}(X)\,\mathrm{d}X. \tag{36}$$

With this new notation, we see that if (33) is satisfied, then $W_n \sim W^\Gamma_{n,\Sigma}$.

LEMMA 15. *The Jacobian of the transformation*

$$\mathcal{P}_\Gamma \ni X \mapsto X^{-1} \in \mathcal{P}_\Gamma \tag{37}$$

equals $\varphi_\Gamma(X^{-1})^2$.

Proof of the lemma is postponed to the Appendix. The lemma gives the following result.

PROPOSITION 16. *Let* $W \sim W^\Gamma_{\eta,\Sigma}$ *with* $\eta > \max\{(r_i - 1)d_i/k_i\,;\, i = 1, \ldots, L\}$ *and* $\Sigma \in \mathcal{P}_\Gamma$. *Then its inverse* $Y = W^{-1}$ *has density*

$$\frac{\mathrm{Det}\,(y)^{-\eta/2}\, e^{-\frac{1}{2}\mathrm{Tr}[y^{-1}\cdot\Sigma^{-1}]}}{\mathrm{Det}\,(2\Sigma)^{\eta/2}\,\Gamma_{\mathcal{P}_\Gamma}(\frac{\eta}{2})} \varphi_\Gamma(y)\mathbf{1}_{\mathcal{P}_\Gamma}(y).$$

3.4. *The Diaconis-Ylvisaker conjugate prior for* K. The Diaconis-Ylvisaker conjugate prior (Diaconis and Ylvisaker (1979)) for the canonical parameter $K = \Sigma^{-1}$ is given by

$$f(K;\delta,D) = \frac{1}{I_\Gamma(\delta,D)} \mathrm{Det}\,(K)^{(\delta-2)/2}\, e^{-\frac{1}{2}\mathrm{Tr}[K\cdot D]}\mathbf{1}_{\mathcal{P}_\Gamma}(K)\,\mathrm{d}K,$$

for hyper-parameters $\delta > 2\max\{1 - 1/k_i\,;\, i = 1, \ldots, L\}$ and $D \in \mathcal{P}_\Gamma$. By (32), the normalizing constant is equal to

$$I_\Gamma(\delta,D) = e^{-A_\Gamma(\delta-2)/2 - B_\Gamma} \prod_{i=1}^{L} \Gamma_{\Omega_i}\left(k_i \frac{\delta-2}{2} + \frac{\dim \Omega_i}{r_i}\right) \mathrm{Det}\,(D)^{-(\delta-2)/2} \varphi_\Gamma(D), \tag{38}$$

where A_Γ, B_Γ and φ_Γ are given in (29), (24) and (25).

4. Model selection. Bayesian model selection on all colored spaces seems at the moment intractable. This is due in great part to a poor combinatorial description of the colored spaces $\mathcal{Z}_\Gamma$. In particular, the number of such spaces, that is, $\#\{\mathcal{Z}_\Gamma\,;\, \Gamma \in \mathfrak{S}_p\}$ is generally unknown for large p. It was shown in Gehrmann (2011) that these colorings constitute a lattice with respect to the usual inclusion of subspaces. However the structure of this lattice is rather complicated and is unobtainable for big p. This, in turn, does not allow to define a Markov chain with known transition probabilities on such colorings. Finally, the fundamental problem which prevents us from doing Bayesian model selection on all colored spaces for arbitrary p is the following. In order to compute Bayes factors, one has to be able to find the structure constants $(k_i, d_i, r_i)_{i=1}^L$ for arbitrary subgroups of $\mathfrak{S}_p$. This is equivalent to finding irreducible representations over reals for an arbitrary finite group, which is very hard in general, although general algorithms have been developed for this issue (see Plesken and Souvignier (1996)).

In this section, we are making a step forward in the problem of model selection for colored models in two ways. In Section 4.1, we use the results of Section 2.6, to obtain the

TABLE 1
Number of all subgroups of a symmetric group, number of their conjugacy classes, number of different colorings and a number of cyclic groups

p	#subgroups of $\mathfrak{S}_p$	#conjugacy classes of $\mathfrak{S}_p$	#different $\mathcal{Z}_\Gamma$	#cyclic groups
1	1	1	1	1
2	2	2	2	2
3	6	4	5	5
4	30	11	22	17
5	156	19	93	67
6	1 455	56	739	362
7	11 300	96	4 508	2039
8	151 221	296	?	14 170
9	1 694 723	554	?	109 694
10	29 594 446	1 593	?	976 412
18	$\approx 7.6\cdot 10^{18}$	$7.3\cdot 10^{6}$	?	$\approx 7.1\cdot 10^{14}$

structure constants when we restrict our search to the space of colored models generated by a cyclic group, that is, when $\Gamma=\langle\sigma\rangle$ for $\sigma\in\mathfrak{S}_p$ and we propose a model selection procedure restricted to the cyclic colorings. In Section 4.2, we use Remarks 7 and 8 to obtain the irreducible representations of $\mathcal{Z}_\Gamma$ and the structure constants by factorization of the determinant. We apply this technique to do model selection for the four-dimensional example given by Frets' data since, in that case, there are only 22 models and we can compute all the Bayes factors.

4.1. *Model selection within cyclic groups.* The smaller space of cyclic colorings has a much better combinatorial description. In particular, the following result can be proved.

LEMMA 17. *If $\mathcal{Z}_{\langle\sigma\rangle}=\mathcal{Z}_{\langle\sigma'\rangle}$ for some $\sigma,\sigma'\in\mathfrak{S}_p$, then $\langle\sigma\rangle=\langle\sigma'\rangle$.*

This result allows us to calculate the number of different colorings corresponding to cyclic groups, that is, the number of labeled cyclic subgroups of the symmetric group $\mathfrak{S}_p$, which can be found in OEIS, sequence A051625 (see the last column of Table 1).

We will present two applications of the Metropolis-Hastings algorithm. In the first one, the Markov chain will move on the space of cyclic groups. The drawback of this first approach is that we need to compute the proposal distribution g, whose computational complexity grows faster than quadratically as p increases (see (40)). In the second algorithm, we consider a larger state space $\mathfrak{S}_p$, which allows us to consider an easy proposal distribution. However, this comes at the cost of slower convergence of the posterior probabilities (see Theorem 18).

4.1.1. *First approach.* Each cyclic subgroup Γ can be uniquely represented by a permutation, which is minimal in the lexicographic order within permutations generating Γ. Let $\nu(\Gamma)\in\mathfrak{S}_p$ be such a permutation, that is,

$$\nu(\Gamma)=\min\{\,\sigma\in\mathfrak{S}_p\,;\,\langle\sigma\rangle=\Gamma\,\}\,.$$

Define

$$c_t:=\langle\nu(c_{t-1})\circ x_t\rangle\,, \tag{39}$$

where c_0 is a fixed cyclic subgroup and $(x_t)_{t\in\mathbb{N}}$ is a sequence of i.i.d. random transpositions distributed uniformly, that is, $\mathbb{P}(x_t=\alpha)=1/\binom{p}{2}$ for any $\alpha\in\mathcal{T}:=\{(i,\,j)\in\mathfrak{S}_p\}$. Clearly, the sequence $(c_t)_t$ is a Markov chain. Its state space is the set of all cyclic subgroups of $\mathfrak{S}_p$. Moreover, the trivial subgroup $\{\mathrm{id}\}$ can be reached from any subgroup c_t (and vice versa) in a finite number of steps with positive probability. Thus the chain $(c_t)_t$ is irreducible. The proposal distribution in the Metropolis-Hastings algorithm is the conditional distribution of $c_t|c_{t-1}$. It is proportional to the number of possible transitions from c to c', that is,

$$g\left(c'|c\right) := \frac{\#\{\,(i,\,j)\in\mathfrak{S}_p\,;\,c'=\langle\nu(c)\circ(i,\,j)\rangle\,\}}{\binom{p}{2}}, \tag{40}$$

where c and c' are cyclic subgroups.

We follow the principles of Bayesian model selection for graphical models, presented, for example, in (Maathuis et al., 2018, Chapter 10, p.247). Let Γ be uniformly distributed on the set $\mathcal{C} := \{\,\langle\sigma\rangle\,;\,\sigma\in\mathfrak{S}_p\,\}$ of cyclic subgroups of $\mathfrak{S}_p$. We assume that $K|\{\Gamma=c\}$, $c\in\mathcal{C}$, follows the Diaconis-Ylvisaker conjugate prior distribution on $\mathcal{P}_c$ with hyper-parameters δ and D, that is,

$$f_{K|\Gamma=c}(k)=\frac{1}{I_c(\delta,D)}\mathrm{Det}\,(k)^{(\delta-2)/2}\,e^{-\frac{1}{2}\mathrm{Tr}[D\cdot k]}\mathbf{1}_{\mathcal{P}_c}(k),$$

where the normalizing constant is given in (38). Suppose that $Z_1,\dots,Z_n$ given $\{K=k,\Gamma=c\}$ are i.i.d. $\mathrm{N}_p(0,k^{-1})$ random vectors with $k\in\mathcal{P}_c$. Then, it is easily seen that we have

$$\mathbb{P}\left(\Gamma=c|Z_1,\dots,Z_n\right)\propto\frac{I_c(\delta+n,D+U)}{I_c(\delta,D)}\qquad(c\in\mathcal{C}) \tag{41}$$

with $U=\sum_{i=1}^n Z_i\cdot Z_i^\top$. These derivations allow us to run the Metropolis-Hastings algorithm restricted to cyclic groups, as follows.

ALGORITHM 16. Starting from a cyclic group $C_0\in\mathcal{C}$, repeat the following two steps for $t=1,2,\dots$:

1. Sample x_t uniformly from the set $\mathcal{T}$ of all transpositions and set $c'=\langle\nu(C_{t-1})\circ x_t\rangle$;
2. Accept the move $C_t=c'$ with probability
$$\min\left\{1,\frac{I_{c'}(\delta+n,D+U)\;I_{C_{t-1}}(\delta,D)}{I_{c'}(\delta,D)\;I_{C_{t-1}}(\delta+n,D+U)}\;\frac{g\left(C_{t-1}|c'\right)}{g\left(c'|C_{t-1}\right)}\right\}$$
If the move is rejected, set $C_t=C_{t-1}$.

4.1.2. *Second approach.* It is known that $\langle\sigma\rangle=\langle\sigma'\rangle$ if and only if $\sigma'=\sigma^k$ for some $k\in\beta(|\sigma|)$, where

$$\beta(n)=\{\,k\in\{1,\dots,n\}\,;\,k\text{ and }n\text{ are relatively prime}\,\} \tag{42}$$

and $|\sigma|$ denotes the order of σ. Let $\mathcal{C}=\{\,\langle\sigma\rangle\,;\,\sigma\in\mathfrak{S}_p\,\}$ denote the set of cyclic subgroups of $\mathfrak{S}_p$. For $c\in\mathcal{C}$ we define $\Phi(c):=\#\beta(|c|)$ and $\mathcal{C}_c:=\{\,\sigma\in\mathfrak{S}_p\,;\,\langle\sigma\rangle=c\,\}$, the set of permutations, which generate the cyclic subgroup c. We have

$$\Phi(c)=\#\mathcal{C}_c\qquad(c\in\mathcal{C}).$$

For $c\in\mathcal{C}$, we denote

$$\pi_c=\mathbb{P}(\Gamma=c|Z_1,\dots,Z_n),$$

which we want to approximate. In our model we have (see (41))

$$\pi_c\propto\frac{I_c(\delta+n,D+U)}{I_c(\delta,D)}\qquad(c\in\mathcal{C}). \tag{43}$$

In order to find $\pi = (\pi_c;\, c \in \mathcal{C})$ let us consider $\tilde{\pi} = (\tilde{\pi}_\sigma;\, \sigma \in \mathfrak{S}_p)$, a probability distribution on $\mathfrak{S}_p$ such that

$$\tilde{\pi}_\sigma \propto \frac{I_{\langle\sigma\rangle}(\delta+n, D+U)}{I_{\langle\sigma\rangle}(\delta, D)} \qquad (\sigma \in \mathfrak{S}_p). \tag{44}$$

Since (43) and (44) imply that $\tilde{\pi}_\sigma \propto \pi_{\langle\sigma\rangle}$, we have

$$\tilde{\pi}_\sigma = \frac{\pi_{\langle\sigma\rangle}}{\sum_{c\in\mathcal{C}} \Phi(c)\pi_c} \qquad (\sigma \in \mathfrak{S}). \tag{45}$$

As before, let $(x_t)_{t\in\mathbb{N}}$ be a sequence of i.i.d random transpositions distributed uniformly on $\mathcal{T} = \{(i,\, j) \in \mathfrak{S}_p\}$. We define a random walk on $\mathfrak{S}_p$ by

$$s_{t+1} = s_t \circ x_{t+1}, \qquad (t = 0, 1, \ldots).$$

Then, $(s_t)_t$ is an irreducible Markov chain with symmetric transition probability

$$g(\sigma'|\sigma) = \begin{cases} \frac{1}{\binom{p}{2}}, & \text{if } \sigma^{-1}\circ\sigma' \in \mathcal{T}, \\ 0, & \text{if } \sigma^{-1}\circ\sigma' \notin \mathcal{T}. \end{cases}$$

We note that $(\langle s_t\rangle)_t$ is not a Markov chain on the space of cyclic subgroups. Indeed, it can be shown that the necessary conditions for $(f(s_t))_t$ to be a Markov chain (see (Burke and Rosenblatt, 1958, Eq. (3))) are not satisfied for $f(\sigma) := \langle\sigma\rangle$ if $p > 4$. A remedy for this fact was introduced in (39). Indeed, the sequence $(\langle s_t\rangle)_t$ is very similar to the sequence $(c_t)_t$ defined previously. Both move along cyclic subgroups and their definitions are very similar. However, $(\langle s_t\rangle)_t$ is not a Markov chain, whereas $(c_t)_t$ is a Markov chain. We took care of this problem by using the minimal generator $\nu(\cdot)$ as in definition (39) of c_t.

We use the Metropolis-Hastings algorithm with the above proposal distribution to approximate $\tilde{\pi}$.

ALGORITHM 17. Starting from a permutation $\sigma_0 \in \mathfrak{S}_p$, repeat the following two steps for $t = 1, 2, \ldots$:

1. Sample x_t uniformly from the set $\mathcal{T}$ of all transpositions and set $\sigma' = \sigma_{t-1} \circ x_t$;
2. Accept the move $\sigma_t = \sigma'$ with probability

$$\min\left\{1, \frac{I_{\langle\sigma'\rangle}(\delta+n, D+U)\; I_{\langle\sigma_{t-1}\rangle}(\delta, D)}{I_{\langle\sigma'\rangle}(\delta, D)\; I_{\langle\sigma_{t-1}\rangle}(\delta+n, D+U)}\right\}.$$

If the move is rejected, set $\sigma_t = \sigma_{t-1}$.

By the ergodicity of the Markov chain $(\sigma_t)_t$ constructed above, as the number of steps $T \to \infty$, we have

$$\frac{\sum_{t=1}^T \mathbf{1}_{\sigma=\sigma_t}}{T} \xrightarrow{a.s.} \tilde{\pi}_\sigma \qquad (\sigma \in \mathfrak{S}_p). \tag{46}$$

This fact allows us to develop a scheme for approximating the posterior probability π.

THEOREM 18. *We have as $T \to \infty$,*

$$\frac{\frac{1}{\Phi(c)}\sum_{t=1}^T \mathbf{1}_{c=\langle\sigma_t\rangle}}{\sum_{t=1}^T \frac{1}{\Phi(\langle\sigma_t\rangle)}} \xrightarrow{a.s.} \pi_c \qquad (c \in \mathcal{C}). \tag{47}$$

PROOF. Let us denote $n_\sigma^{(T)} = \sum_{t=1}^T \mathbf{1}_{\sigma=\sigma_t}$, $\sigma \in \mathfrak{S}_p$. We have $T = \sum_{\sigma\in\mathfrak{S}_p} n_\sigma^{(T)}$ and $n_\sigma^{(T)}/T \xrightarrow{a.s.} \tilde{\pi}_\sigma$. Moreover,

$$\frac{\frac{1}{\Phi(c)}\sum_{t=1}^T \mathbf{1}_{c=\langle\sigma_t\rangle}}{\sum_{t=1}^T \frac{1}{\Phi(\langle\sigma_t\rangle)}} = \frac{\frac{1}{\Phi(c)}\sum_{\sigma\in\mathcal{C}_c} n_\sigma^{(T)}}{\sum_{t=1}^T\sum_{\gamma\in\mathcal{C}} \frac{1}{\Phi(\gamma)}\mathbf{1}_{\gamma=\langle\sigma_t\rangle}}$$
$$= \frac{\frac{1}{\Phi(c)}\sum_{\sigma\in\mathcal{C}_c} \frac{n_\sigma^{(T)}}{T}}{\sum_{\gamma\in\mathcal{C}} \frac{1}{\Phi(\gamma)}\sum_{\sigma\in\mathcal{C}_\gamma}\frac{n_\gamma^{(T)}}{T}}$$
$$\xrightarrow{a.s.} \frac{\frac{1}{\Phi(c)}\sum_{\sigma\in\mathcal{C}_c}\tilde{\pi}_\sigma}{\sum_{\gamma\in\mathcal{C}} \frac{1}{\Phi(\gamma)}\sum_{\sigma\in\mathcal{C}_\gamma}\tilde{\pi}_\gamma}.$$

Finally, by (45) we have

$$\frac{1}{\Phi(c)}\sum_{\sigma\in\mathcal{C}_c}\tilde{\pi}_\sigma = \frac{\pi_c}{\sum_{\gamma\in\mathcal{C}}\Phi(\gamma)\pi_\gamma} \propto \pi_c,$$

which completes the proof. □

In order to approximate the posterior probability π, we allowed the Markov chain to travel on the larger space $\mathfrak{S}_p$. In particular, each state $c \in \mathcal{C}$ was multiplied $\Phi(c) \geq 1$ times, where $\Phi(c)$ is the number of permutations generating c. This procedure should result in slower convergence to the stationary distribution in (47). By comparing with (46), we see that (47) can be interpreted as follows: let us assign to each cyclic subgroup c a weight $1/\Phi(c) \leq 1$. Then, the denominator $N_T := \sum_{t=1}^T 1/\Phi(\langle\sigma_t\rangle)$ can be thought of as an "effective" number of steps and the numerator is the number of "effective" steps spent in state c. In general, for large T we expect $N_T \ll T$ (see an example in Section 5.2).

4.2. *Model selection for* $p = 4$. Our numbering of colored models on four vertices is in accordance with (Gehrmann, 2011, Figures 15 and 16, p 674–675). However, we identify models by the largest group with the same coloring Γ^* rather than the smallest as in Gehrmann (2011). There are 30 different subgroups of $\mathfrak{S}_4$, which generate 22 different colored spaces. Up to conjugacy (renumbering of vertices), there are 8 different conjugacy classes. Within a conjugacy class, the sets of constants $\{k_i, r_i, d_i\}_{i=1}^L$ remain the same. Groups Γ_k^* for $k = 1, \ldots, 17$ correspond to cyclic colorings.

We apply our results and methods in order to do Bayesian model selection for the celebrated example of Frets' heads, Frets (1921); Whittaker (1990). The head dimensions (length L_i and breadth B_i, $i = 1, 2$) of 25 pairs of first and second sons were measured. Thus we have $n = 25$ and $p = 4$. The following sample covariance matrix is obtained (we have $Z = (L_1, B_1, L_2, B_2)^\top$),

$$U = \sum_{i=1}^n Z^{(i)} \cdot Z^{(i)\top} = \begin{pmatrix} 2287.04 & 1268.84 & 1671.88 & 1106.68 \\ 1268.84 & 1304.64 & 1231.48 & 841.28 \\ 1671.88 & 1231.48 & 2419.36 & 1356.96 \\ 1106.68 & 841.28 & 1356.96 & 1080.56 \end{pmatrix}.$$

We perform Bayesian model selection within all RCOP models, not just the ones corresponding to cyclic subgroups. In Table 2 we list all RCOP models on full graph with four vertices, along with corresponding structure constants. Structure constants remain the same within a conjugacy class, however the invariant measure φ_Γ is always different. Since there are only

TABLE 2
Structure constants for all colorings with four vertices

Group	(k_i)	(r_i)	(d_i)
$\Gamma_1^* = \{\mathrm{id}\}$	(1)	(4)	(1)
$\Gamma_2^* = \langle(1,2)\rangle$	(1,1)	(3,1)	(1,1)
$\Gamma_3^* = \langle(1,3)\rangle$			
$\Gamma_4^* = \langle(1,4)\rangle$			
$\Gamma_5^* = \langle(2,3)\rangle$			
$\Gamma_6^* = \langle(2,4)\rangle$			
$\Gamma_7^* = \langle(3,4)\rangle$			
$\Gamma_8^* = \langle(1,2,3),\ (1,2)\rangle$	(1,2)	(2,1)	(1,1)
$\Gamma_9^* = \langle(1,2,4),\ (1,2)\rangle$			
$\Gamma_{10}^* = \langle(1,3,4),\ (1,3)\rangle$			
$\Gamma_{11}^* = \langle(2,3,4),\ (2,3)\rangle$			
$\Gamma_{12}^* = \langle(1,2)(3,4)\rangle$	(1,1)	(2,2)	(1,1)
$\Gamma_{13}^* = \langle(1,3)(2,4)\rangle$			
$\Gamma_{14}^* = \langle(1,4)(2,3)\rangle$			
$\Gamma_{15}^* = \langle(1,2,3,4),\ (1,3)\rangle$	(1,1,2)	(1,1,1)	(1,1,1)
$\Gamma_{16}^* = \langle(1,2,4,3),\ (1,4)\rangle$			
$\Gamma_{17}^* = \langle(1,3,2,4),\ (1,2)\rangle$			
$\Gamma_{18}^* = \langle(1,2),\ (3,4)\rangle$	(1,1,1)	(2,1,1)	(1,1,1)
$\Gamma_{19}^* = \langle(1,3),\ (2,4)\rangle$			
$\Gamma_{20}^* = \langle(1,4),\ (2,3)\rangle$			
$\Gamma_{21}^* = \langle(1,2)(3,4),\ (1,4)(2,3)\rangle$	(1,1,1,1)	(1,1,1,1)	(1,1,1,1)
$\Gamma_{22}^* = \mathfrak{S}_4$	(1,3)	(1,1)	(1,1)

TABLE 3
Posterior probabilities in Frets' heads for three best models, $\delta = 3$ and given D.

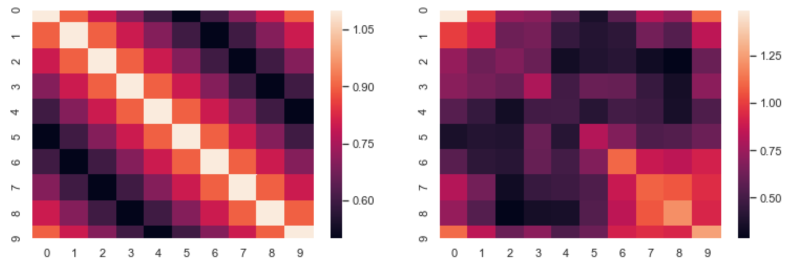
D	Best model		2nd best		3rd best	
I_4	Γ_{22}^*	(95.2%)	Γ_{16}^*	(2.5%)	Γ_{17}^*	(1.3%)
$50I_4$	Γ_{19}^*	(33.8%)	Γ_{13}^*	(29.6%)	Γ_8^*	(13.3%)
$100I_4$	Γ_{13}^*	(39.6%)	Γ_{19}^*	(29.8%)	Γ_8^*	(7.2%)
$1000I_4$	Γ_1^*	(38.9%)	Γ_{13}^*	(10.5%)	Γ_3^*	(10.3%)

22 such models, we calculate all exact posterior probabilities. The Table 2 and the invariant measures φ_Γ were obtained by using Remark 7.

In Table 3 we summarize the results when $\delta = 3$, giving the three best coloring models with the highest posterior probability, for each given D. Results are very similar for $\delta = 10$ and the given values of D.

For different values of $D = dI_4$, the only models that have highest posterior probability are the 4 models: $\Gamma_{22}^* = \mathfrak{S}_4$, $\Gamma_{19}^* = \langle(1,3),\ (2,4)\rangle$, $\Gamma_{13}^* = \langle(1,3)(2,4)\rangle$, $\Gamma_1^* = \{\mathrm{id}\}$. These four subgroups form a path in the Hasse diagram of subgroups of $\mathfrak{S}_4^*$, i.e. $\Gamma_{22}^* \supset \Gamma_{19}^* \supset \Gamma_{13}^* \supset \Gamma_1^*$. Thus the four selected colorings, corresponding to the permutation groups are in some way consistent. Moreover, each of them has a good statistical interpretation. Let us interpret models Γ_{13}^* and Γ_{19}^*. Recall the enumeration of vertices $(1,2,3,4) = (L_1, B_1, L_2, B_2)$. The invariance with respect to the transposition $(1,3)$ means that L_1 is exchangeable with L_2 and, similarly, the invariance with respect to the transposition $(2,4)$ implies exchangeability of B_1 and B_2. Both together correspond to the fact that sons should be exchangeable in some way.

We observe that only the Γ_{22}^* model appeared in former attempts of model selection for Frets' heads data. It was selected in Højsgaard and Lauritzen (2008), by a likelihood ratio test. Note that the only complete RCOP model selected in Gehrmann (2011) (who used the Edwards-Havranek model selection procedure) among the 9 minimally accepted models on p. 676 of her article is Γ_{10}^*, which is not selected by our exact Bayesian procedure for any choice



(a) (b)

FIG 1. *Heat map of matrix Σ (a) and matrix U/n (b).*

TABLE 4
Five most visited cyclic subgroups

generator of a cyclic group	number of visits
(1, 2, 3, 4, 5, 6, 7, 8, 9,10)	457 725
(1, 6, 2, 7)(3, 5, 9)(4, 8, 10)	110 677
(1, 6)(2, 7)(3, 5, 9)(4, 8, 10)	51 618
(1, 7)(2, 6)(3, 5, 9)(4, 8, 10)	40 895
(1, 2, 6, 7)(3, 5, 9)(4, 8, 10)	34 883

of $D = d\,I_4$. Coloured graphical model selection using the colored G-Wishart distribution as the prior on $K = \Sigma^{-1}$ was also done in the supplementary file of Massam, Li and Gao (2018).

5. Simulations. Let the covariance matrix Σ be the symmetric circulant matrix

$$\Sigma = \begin{pmatrix} c_0 & c_1 & \dots & c_2 & c_1 \\ c_1 & c_0 & c_1 & \dots & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_2 & & \ddots & \ddots & c_1 \\ c_1 & c_2 & \dots & c_1 & c_0 \end{pmatrix},$$

with $c_0 = 1 + 1/p$ and $c_k = 1 - k/p$ for $k = 1, \dots, \lfloor p/2 \rfloor$. It is easily seen that this matrix belongs to $\mathcal{P}_{\langle \sigma^* \rangle}$ with $\sigma^* = (1, 2, \dots, p-1, p)$.

5.1. *First approach.* For $p = 10$ and $n = 20$, we sampled $Z^{(1)}, \dots, Z^{(n)}$ from the $\mathrm{N}_p(0, \Sigma)$ distribution and obtained $U = \sum_{i=1}^{n} Z^{(i)} \cdot Z^{(i)\top}$ depicted in Fig. 1 (b).

We run the Metropolis-Hastings algorithm starting from the group $\langle \sigma_0 \rangle = \{\mathrm{id}\}$ with hyper-parameters $\delta = 3$ and $D = I_{10}$. After 1 000 000 steps, the five most visited states are given in the Tab. 4.

The Metropolis-Hastings (M-H) algorithm recovered the true pattern of the covariance matrix. The acceptance rate was 2.5% and the Markov chain visited 746 different cyclic groups. The acceptance rate can be increased by a suitable choice of the hyper-parameters (e.g. for $D = 10 I_{10}$ the acceptance rate is around 10%).

In order to grasp how randomness may influence results, we performed 100 simulations, where each time we sample $Z^{(1)}, \dots, Z^{(n)}$ from $\mathrm{N}_p(0, \Sigma)$ and we run M-H for $100\,000$ steps with the same parameters as before. In Table 5 we present how many times a given cyclic subgroup was most visited during these 100 simulations (second column). There were 53 distinct cyclic subgroups, which were most visited at least in one of the 100 simulations; below we present 10 such subgroups. The average acceptance rate is 1.4% (see the histogram

TABLE 5
Cyclic subgroups which were chosen by M-H algorithm most often

generator of a cyclic group	#most visited	ARI
(1, 2, 3, 4, 5, 6, 7, 8, 9,10)	25	1.00
(1, 3, 5, 7, 9)(2, 4, 6, 8,10)	13	0.60
(1, 2, 4, 3, 5, 6, 7, 9, 8, 10)	3	0.43
(1, 2, 4, 3, 5, 6, 7, 8, 9, 10)	2	0.46
(1, 3, 2, 4, 5, 6, 8, 7, 9, 10)	2	0.43
(1, 3, 5, 9, 2, 6, 8, 10, 4, 7)	2	0.43
(1, 4, 3, 5, 2, 6, 9, 8, 10, 7)	2	0.35
(1, 4, 5, 7, 8)(2, 3, 6, 9, 10)	2	0.24
(1, 8, 10, 9)(2, 7)(3, 5, 4, 6)	2	0.19
(1, 2, 10, 3)(4, 9)(5, 8, 6, 7)	2	0.19

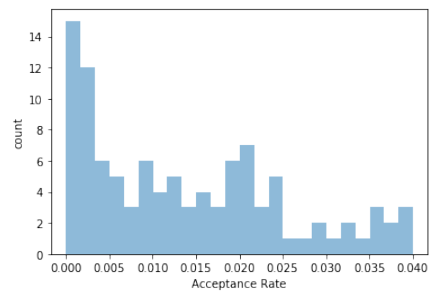


FIG 2. *Histogram of acceptance rates in* 100 *simulations of Metropolis-Hastings algorithm.*

in Fig. 2). When we regard colorings as partitions of the set $V \cup E$ according to group orbit decomposition, the two colorings may be compared using the so-called adjusted Rand index (ARI, see Hubert and P. Arabie (1985)), a similarity measure comparing partitions which takes values between -1 and 1, where 1 stands for perfect match and independent random labelings have score close to 0. In the third column of Table 5, we give the adjusted Rand index between the colorings generated by given cyclic subgroup and the true coloring.

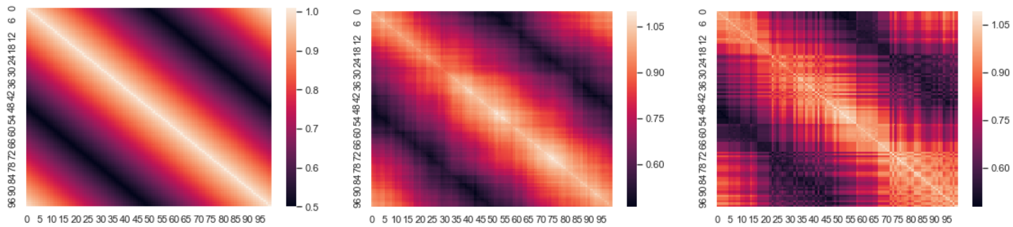
We see that groups which were most visited by the Markov chain have positive ARI and the true pattern was recovered in a quarter of cases. We stress that even though the colorings generated by $\langle(1,2,3,4,5,6,7,8,9,10)\rangle$ and $\langle(1,3,5,7,9)(2,4,6,8,10)\rangle$ are very similar, the distance between these subgroups is 9, that is, the Markov chain $(C_t)_t$ needs at least 9 steps to get from one subgroup to the other.

This indicates that the Markov chain may encounter many local maxima and one should always tune the hyper parameters in order to have higher acceptance rate or to allow the Markov chain $(C_t)_t$ to make bigger steps.

5.2. *Second approach.* We performed $T = 100\,000$ steps of Algorithm 17 with $\sigma_0 = \mathrm{id}$, $p = 100$, $n = 200$, $\delta = 3$ and $D = I_{100}$. Let us note that for $p = 100$, there are about $4 \cdot 10^{155}$ cyclic subgroups and this is the number of models we consider in our model search.

We have used Theorem 18 to approximate the posterior probability distribution $(\pi_c;\, c \in \mathcal{C})$ (see (41)). The highest estimated posterior probability was obtained for $c^* := \langle\sigma^*\rangle$, where

$$\begin{aligned}\sigma^* =&(1,2,3,4)(6,8,15)(7,10,9)(11,16,12)(13,17,14)(18,19,20,22,21)(23,26)\\&(24,42,28,44)(25,31,30,32)(27,34)(29,37)(33,45)(35,39,36,40)\\&(38,47,41,48)(43,51,46,49)(50,52,53,54)(56,58,57)(59,66,67)\\&(60,65,63)(61,62,64)(68,71,72,70,69)(73,93)(74,77)(75,98,81,100)\end{aligned}$$



(a) (b) (c)

FIG 3. *Heat map of matrix Σ (a) and matrix U/n (b) and projection of U/n onto $\mathcal{Z}_{c^*}$.*

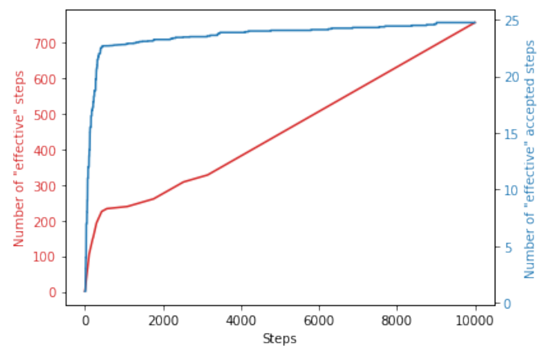


FIG 4. *Number of "effective" steps (red) and number of "effective" accepted steps (blue).*

$$(76,84,78,83)(79,85)(80,94,82,91)(86,92,87,90)(88,96,89,97)(95,99).$$

The order of c^* is $|c^*| = 60$ and $\Phi(c^*) = 16$. The estimate of the posterior probability π_{c^*} is equal to (recall (47))

$$\frac{\frac{1}{\Phi(c^*)}\sum_{t=1}^{T}\mathbf{1}_{c^*=\langle\sigma_t\rangle}}{\sum_{t=1}^{T}\frac{1}{\Phi(\langle\sigma_t\rangle)}} \approx \frac{2\,361.5}{6\,381.5} \approx 37\%.$$

The true covariance matrix Σ, the data matrix U/n and the projection $\Pi_{c^*}(U/n)$ are illustrated in Fig. 3.

We visualize the performance of the algorithm on Fig 4. In red color, a sequence $\left(\sum_{t=1}^{k}\frac{1}{\Phi(\langle\sigma_t\rangle)}\right)_k$ is depicted, which can be thought of as an "effective" number of steps of the algorithm (for an explanation, see the paragraph at the end of Subsection 4.1.2). In blue, we present a sequence $\left(\sum_{t=1}^{k}\frac{1}{\Phi(\langle\sigma_t\rangle)}\mathbf{1}_{\langle\sigma_t\rangle\neq\langle\sigma_{t-1}\rangle}\right)_k$, which represents the number of weighted accepted steps, where the weight of the kth step equals $\frac{1}{\Phi(\langle\sigma_k\rangle)}$. We restricted the plot to steps $k = 1,\dots,10\,000$, because after $10\,000$ steps, the Markov chain $(\sigma_t)_{10\,000\leq t\leq 100\,000}$ changed its state only 9 times. For $k = 100\,000$, the value of the blue curve is 25.75, while the value of red one is $6\,381.5$.

The model suffers from poor acceptance rate, which could be improved by an appropriate choice of the hyper-parameter D or by allowing the Markov chain to do bigger steps.

Acknowledgements. The authors would like to thank Steffen Lauritzen for his interest and encouragements.

SUPPLEMENTARY MATERIAL

Appendix: Proofs of Theorems 4, 6 and 7. Example for Remark 7.

For the proof of Theorem 4, we will need the following Lemma.

LEMMA 19. *Let $i, j = 1, 2, \dots, L$, and assume that $Y \in \mathrm{Mat}(r_i k_i,\, r_j k_j; \mathbb{R})$ satisfies the condition*

$$[I_{r_i} \otimes B_i(\sigma)] \cdot Y = Y \cdot [I_{r_j} \otimes B_j(\sigma)] \qquad (\sigma \in \Gamma). \tag{48}$$

If $i = j$, then there exists $C \in \mathrm{Mat}(r_i,\, r_i; \mathbb{K}_i)$ such that $Y = M_{\mathbb{K}_i}(C) \otimes I_{k_i/d_i}$. On the other hand, if $i \neq j$, then $Y = 0$.

PROOF. Let us consider a block decomposition of Y as

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1,r_j} \\ Y_{21} & Y_{22} & \dots & Y_{2,r_j} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{r_i,1} & Y_{r_i,2} & \dots & Y_{r_i,r_j} \end{pmatrix},$$

where each Y_{ab} is a $k_i \times k_j$ matrix. Then (48) implies that

$$B_i(\sigma) \cdot Y_{ab} = Y_{ab} \cdot B_j(\sigma) \qquad (\sigma \in \Gamma) \tag{49}$$

for all a, b. If $i = j$, then $Y_{ab} \in \mathrm{End}_\Gamma(\mathbb{R}^{k_i})$, so that there exists $C_{ab} \in \mathbb{K}_i$ for which $Y_{ab} = M_{\mathbb{K}_i}(C_{ab}) \otimes I_{k_i/d_i}$ thanks to Lemma 3. Let us consider the case $i \neq j$. Eq. (49) tells us that $\mathrm{Ker}\, Y_{ab} \subset \mathbb{R}^{k_j}$ is a Γ-invariant subspace, which then equals $\{0\}$ or $\mathbb{R}^{k_j}$ because of the irreducibility of B_j. Similarly, since $\mathrm{Image}\, Y_{ab} \subset \mathbb{R}^{k_i}$ is a Γ-invariant subspace by (49), $\mathrm{Image}\, Y_{ab}$ equals $\{0\}$ or $\mathbb{R}^{k_i}$. Now suppose that $Y_{ab} \neq 0$. Then $\mathrm{Ker}\, Y_{ab} = \{0\}$ and $\mathrm{Image}\, Y_{ab} = \mathbb{R}^{k_i}$ by the argument above, and it means that Y_{ab} induces an isomorphism from $(B_j, \mathbb{R}^{k_j})$ onto $(B_i, \mathbb{R}^{k_i})$. But this contradicts the fact that the representations B_i and B_j are not equivalent for $i \neq j$. Hence we get $Y_{ab} = 0$. □

PROOF OF THEOREM 4. Take $y \in U_\Gamma^\top \cdot \mathcal{Z}_\Gamma \cdot U_\Gamma$ and consider the block decomposition of y as

$$y = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1L} \\ Y_{21} & Y_{22} & \dots & Y_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{L1} & Y_{L2} & \cdots & Y_{LL} \end{pmatrix}$$

with $Y_{ij} \in \mathrm{Mat}(r_i k_i,\, r_j k_j; \mathbb{R})$. Then $x := U_\Gamma \cdot y \cdot U_\Gamma^\top$ belongs to $\mathcal{Z}_\Gamma$, so that (4) implies

$$R(\sigma) \cdot U_\Gamma \cdot y \cdot U_\Gamma^\top \cdot R(\sigma)^\top = U_\Gamma \cdot y \cdot U_\Gamma^\top$$

for $\sigma \in \Gamma$, and this equality is rewritten as

$$[U_\Gamma^\top \cdot R(\sigma) \cdot U_\Gamma] \cdot y = y \cdot [U_\Gamma^\top \cdot R(\sigma) \cdot U_\Gamma].$$

By (13), we have

$$[I_{r_i} \otimes B_i(\sigma)] \cdot Y_{ij} = Y_{ij} \cdot [I_{r_j} \otimes B_j(\sigma)].$$

Lemma 19 tells us that $Y_{ij} = 0$ if $i \neq j$, and that $Y_{ii} = M_{\mathbb{K}_i}(x_i) \otimes I_{k_i/d_i}$ with some $x_i \in \mathrm{Mat}(r_i,\, r_i; \mathbb{K}_i)$. Since y is a symmetric matrix, the block Y_{ii} is also symmetric, which implies that $x_i \in \mathrm{Herm}(r_i; \mathbb{K}_i)$. Hence, the map $\iota\colon \bigoplus_{i=1}^L \mathrm{Herm}(r_i; \mathbb{K}_i) \ni (x_i)_{i=1}^L \mapsto X \in \mathcal{Z}_\Gamma$ given by (15) gives a Jordan algebra isomorphism. □

EXAMPLE 18. In this example we present a colored space $\mathcal{Z}_\Gamma \subset \mathrm{Sym}(16;\mathbb{R})$, which has a component $\mathrm{Herm}(2;\mathbb{H})$. Let $\Gamma = \langle \sigma_1, \sigma_2 \rangle$ be the subgroup of $\mathfrak{S}_{16}$ generated by the two permutations

$$\sigma_1 = (1,2,5,6)(3,4,7,8)(9,10,13,14)(11,12,15,16),$$
$$\sigma_2 = (1,3,5,7)(2,8,6,4)(9,11,13,15)(10,16,14,12).$$

The space $\mathcal{Z}_\Gamma$ consists of matrices of the form

$$X = \begin{pmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_5 & \gamma_6 & \gamma_7 & \gamma_8 \\
\alpha_2 & \alpha_1 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_5 & \alpha_4 & \alpha_3 & \gamma_6 & \gamma_1 & \gamma_8 & \gamma_3 & \gamma_2 & \gamma_5 & \gamma_4 & \gamma_7 \\
\alpha_3 & \alpha_4 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_2 & \gamma_7 & \gamma_4 & \gamma_1 & \gamma_6 & \gamma_3 & \gamma_8 & \gamma_5 & \gamma_2 \\
\alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_5 & \gamma_8 & \gamma_7 & \gamma_2 & \gamma_1 & \gamma_4 & \gamma_3 & \gamma_6 & \gamma_5 \\
\alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \gamma_5 & \gamma_6 & \gamma_7 & \gamma_8 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\
\alpha_2 & \alpha_5 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_4 & \alpha_3 & \gamma_2 & \gamma_5 & \gamma_4 & \gamma_7 & \gamma_6 & \gamma_1 & \gamma_8 & \gamma_3 \\
\alpha_3 & \alpha_4 & \alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_1 & \alpha_2 & \gamma_3 & \gamma_8 & \gamma_5 & \gamma_2 & \gamma_7 & \gamma_4 & \gamma_1 & \gamma_6 \\
\alpha_4 & \alpha_3 & \alpha_2 & \alpha_5 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \gamma_4 & \gamma_3 & \gamma_6 & \gamma_5 & \gamma_8 & \gamma_7 & \gamma_2 & \gamma_1 \\
\gamma_1 & \gamma_6 & \gamma_7 & \gamma_8 & \gamma_5 & \gamma_2 & \gamma_3 & \gamma_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_2 & \beta_3 & \beta_4 \\
\gamma_2 & \gamma_1 & \gamma_4 & \gamma_7 & \gamma_6 & \gamma_5 & \gamma_8 & \gamma_3 & \beta_2 & \beta_1 & \beta_4 & \beta_3 & \beta_2 & \beta_5 & \beta_4 & \beta_3 \\
\gamma_3 & \gamma_8 & \gamma_1 & \gamma_2 & \gamma_7 & \gamma_4 & \gamma_5 & \gamma_6 & \beta_3 & \beta_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_2 \\
\gamma_4 & \gamma_3 & \gamma_6 & \gamma_1 & \gamma_8 & \gamma_7 & \gamma_2 & \gamma_5 & \beta_4 & \beta_3 & \beta_2 & \beta_1 & \beta_4 & \beta_3 & \beta_2 & \beta_5 \\
\gamma_5 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_1 & \gamma_6 & \gamma_7 & \gamma_8 & \beta_5 & \beta_2 & \beta_3 & \beta_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\
\gamma_6 & \gamma_5 & \gamma_8 & \gamma_3 & \gamma_2 & \gamma_1 & \gamma_4 & \gamma_7 & \beta_2 & \beta_5 & \beta_4 & \beta_3 & \beta_2 & \beta_1 & \beta_4 & \beta_3 \\
\gamma_7 & \gamma_4 & \gamma_5 & \gamma_6 & \gamma_3 & \gamma_8 & \gamma_1 & \gamma_2 & \beta_3 & \beta_4 & \beta_5 & \beta_2 & \beta_3 & \beta_4 & \beta_1 & \beta_2 \\
\gamma_8 & \gamma_7 & \gamma_2 & \gamma_5 & \gamma_4 & \gamma_3 & \gamma_6 & \gamma_1 & \beta_4 & \beta_3 & \beta_2 & \beta_5 & \beta_4 & \beta_3 & \beta_2 & \beta_1
\end{pmatrix}.$$

The irreducible factorization of the determinant is given by

$$\begin{aligned}
\mathrm{Det}\,(X) = & \left((\gamma_1-\gamma_5)^2 + (\gamma_2-\gamma_6)^2 + (\gamma_3-\gamma_7)^2 + (\gamma_4-\gamma_8)^2 - (\alpha_1-\alpha_5)(\beta_1-\beta_5)\right)^4 \\
& \cdot \left((\alpha_1 - 2(\alpha_2+\alpha_3-\alpha_4)+\alpha_5)(\beta_1 - 2(\beta_2+\beta_3-\beta_4)+\beta_5) - (\gamma_1-\gamma_2-\gamma_3+\gamma_4+\gamma_5-\gamma_6-\gamma_7+\gamma_8)^2\right) \\
& \cdot \left((\alpha_1 - 2(\alpha_2-\alpha_3+\alpha_4)+\alpha_5)(\beta_1 - 2(\beta_2-\beta_3+\beta_4)+\beta_5) - (\gamma_1-\gamma_2+\gamma_3-\gamma_4+\gamma_5-\gamma_6+\gamma_7-\gamma_8)^2\right) \\
& \cdot \left((\alpha_1 + 2(\alpha_2-\alpha_3-\alpha_4)+\alpha_5)(\beta_1 + 2(\beta_2-\beta_3-\beta_4)+\beta_5) - (\gamma_1+\gamma_2-\gamma_3-\gamma_4+\gamma_5+\gamma_6-\gamma_7-\gamma_8)^2\right) \\
& \cdot \left((\alpha_1 + 2(\alpha_2+\alpha_3+\alpha_4)+\alpha_5)(\beta_1 + 2(\beta_2+\beta_3+\beta_4)+\beta_5) - (\gamma_1+\gamma_2+\gamma_3+\gamma_4+\gamma_5+\gamma_6+\gamma_7+\gamma_8)^2\right).
\end{aligned}$$

Thus, Remark 7 gives us that $L = 5$ and

$$r = (2,2,2,2,2), \quad k = (4,1,1,1,1), \quad d = (4,1,1,1,1).$$

This in turn implies

$$\mathcal{Z}_\Gamma \simeq \mathrm{Herm}(2;\mathbb{H}) \oplus \mathrm{Sym}(2;\mathbb{R})^{\oplus 4}.$$

As a matter of fact, the group Γ has four 1-dimensional representations and one 4-dimensional real irreducible representation, and each representation appears twice in $\mathbb{R}^{16}$.

PROOF OF THEOREM 6 AND THEOREM 7. Let $M := \left\lfloor \frac{N}{2} \right\rfloor$ and denote the irreducible representations of Γ by

$$B_0 \colon \Gamma \ni \sigma^k \mapsto 1 \in \mathrm{GL}(1;\mathbb{R}),$$
$$B_\alpha \colon \Gamma \ni \sigma^k \mapsto \mathrm{Rot}\left(\frac{2\pi\alpha k}{N}\right) \in \mathrm{GL}(2;\mathbb{R}) \quad (1 \le \alpha < N/2),$$
$$B_{N/2} \colon \Gamma \ni \sigma^k \mapsto (-1)^k \in \mathrm{GL}(1;\mathbb{R}) \quad (\text{if } N \text{ is even}),$$

where $\mathrm{Rot}(\theta)$ denotes the rotation matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ for $\theta \in \mathbb{R}$. Then all the equivalence classes of the irreducible representations of Γ are $[B_0], [B_1], \ldots, [B_M]$ whether $N = 2M$ or $N = 2M+1$. We have $k_\alpha = d_\alpha = \begin{cases} 1 & (\alpha = 0 \text{ or } N/2) \\ 2 & (\text{otherwise}). \end{cases}$

Recall that $\{i_1, \ldots, i_C\}$ is a complete system of representatives of the Γ-orbits, and, for each $c = 1, \ldots, C$, p_c is the cardinality of the Γ-orbit through i_c. Let $\zeta_c := \exp(2\pi\sqrt{-1}/p_c)$. When $1 \le \beta < p_c/2$, we have

$$v_{2\beta}^{(c)} + \sqrt{-1} v_{2\beta+1}^{(c)} = \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c-1} \zeta_c^{\beta k} e_{\sigma^k(i_c)}.$$

Thus

$$R(\sigma)(v_{2\beta}^{(c)} + \sqrt{-1} v_{2\beta+1}^{(c)}) = \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c-1} \zeta_c^{\beta k} e_{\sigma^{k+1}(i_c)} = \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c-1} \zeta_c^{\beta(k-1)} e_{\sigma^k(i_c)}$$

$$= \zeta_c^{-\beta}(v_{2\beta}^{(c)} + \sqrt{-1} v_{2\beta+1}^{(c)})$$

$$= \left\{ \cos\left(\frac{2\pi\beta}{p_c}\right) v_{2\beta}^{(c)} + \sin\left(\frac{2\pi\beta}{p_c}\right) v_{2\beta+1}^{(c)} \right\} + \sqrt{-1} \left\{ -\sin\left(\frac{2\pi\beta}{p_c}\right) v_{2\beta}^{(c)} + \cos\left(\frac{2\pi\beta}{p_c}\right) v_{2\beta+1}^{(c)} \right\},$$

where we have used $\sigma^{p_c}(i_c) = i_c$ and $\zeta_c^{p_c} = 1$ at the second equality. It follows that

$$R(\sigma) \begin{pmatrix} v_{2\beta}^{(c)} & v_{2\beta+1}^{(c)} \end{pmatrix} = \begin{pmatrix} v_{2\beta}^{(c)} & v_{2\beta+1}^{(c)} \end{pmatrix} \mathrm{Rot}\left(\frac{2\pi\beta}{p_c}\right) = \begin{pmatrix} v_{2\beta}^{(c)} & v_{2\beta+1}^{(c)} \end{pmatrix} B_\alpha(\sigma) \tag{50}$$

with $\frac{\beta}{p_c} = \frac{\alpha}{N}$. Similarly, we have

$$R(\sigma) v_1^{(c)} = v_1^{(c)} = B_0(\sigma) v_1^{(c)},$$
$$R(\sigma) v_{p_c}^{(c)} = -v_{p_c}^{(c)} = B_{N/2}(\sigma) v_{p_c}^{(c)} \quad (\text{if } p_c \text{ and } N \text{ are even}).$$

Therefore, for $\alpha = 0, \ldots, [N/2]$, the multiplicity r_α of the representation B_α of Γ in $(R, \mathbb{R}^p)$ is equal to the number of c such that $\frac{\beta}{p_c} = \frac{\alpha}{N}$ with some $\beta \in \mathbb{N}$. In other words,

$$r_\alpha = \#\{\, c;\, \alpha p_c \text{ is a multiple of } N \,\} \quad (0 \le \alpha \le [N/2]). \tag{51}$$

Then we have

$$\mathcal{Z}_\Gamma \simeq \bigoplus_{r_\alpha > 0} \mathrm{Herm}(r_\alpha; \mathbb{K}_\alpha).$$

□

PROOF OF LEMMA 8. Denote the left hand side of (23) by I. Let us change variables $x_i = \phi_i(X)$ for $i = 1, \ldots, L$. By (16) and (22) we obtain

$$I = e^{B_\Gamma} \prod_{i=1}^{L} \int_{\Omega_i} \det(x_i)^{\lambda_i} e^{-k_i \mathrm{tr}[\phi_i(Y) \bullet x_i]} m_i(\mathrm{d}x_i).$$

Each integral can be calculated using (21) for $\lambda_i > -1$ and $\phi_i(Y) \in \Omega_i$, $i = 1, \ldots, L$. Hence,

$$I = e^{B_\Gamma} \prod_{i=1}^{L} \Gamma_{\Omega_i}(\lambda_i + \dim\Omega_i/r_i) \det(k_i\, \phi_i(Y))^{-\lambda_i - \dim\Omega_i/r_i}$$

and so we obtain (23). □

PROOF OF LEMMA 15. First observe that

$$(X+h)^{-1} - X^{-1} = (X+h)^{-1} \cdot [X - (X+h)] \cdot X^{-1} = -X^{-1} \cdot h \cdot X^{-1} + o(h),$$

so that, the Jacobian of (37) equals $\mathrm{Det}_{\mathrm{End}}(\mathbb{P}_{X^{-1}})$, where $\mathrm{Det}_{\mathrm{End}}$ is the determinant in the space of endomorphisms of $\mathcal{Z}_\Gamma$ and for any $X \in \mathcal{Z}_\Gamma$ by $\mathbb{P}_X$ we denote the linear map on $\mathcal{Z}_\Gamma$ to itself defined by $\mathbb{P}_X Y = X \cdot Y \cdot X$. It is easy to see that for any $X \in \mathcal{P}_\Gamma$ we have $\mathbb{P}_X \in \mathrm{G}(\mathcal{P}_\Gamma)$. Indeed, since $\mathbb{P}_X Y$ is positive definite for $Y \in \mathcal{P}_\Gamma$, it is enough to verify that

$$R(\sigma) \cdot [\mathbb{P}_X Y] = [\mathbb{P}_X Y] \cdot R(\sigma) \qquad (\sigma \in \Gamma).$$

This follows quickly by the fact that $X, Y \in \mathcal{P}_\Gamma$. Further, by the $\mathrm{G}(\mathcal{P}_\Gamma)$ invariance of φ_Γ, we have

$$\varphi_\Gamma(gX) = |\mathrm{Det}_{\mathrm{End}}(g)|^{-1} \varphi_\Gamma(X) \qquad (g \in \mathrm{G}(\mathcal{P}_\Gamma)).$$

Taking $g = \mathbb{P}_{X^{-1}}$, we eventually obtain

$$\mathrm{Det}_{\mathrm{End}}(\mathbb{P}_{X^{-1}}) = \frac{\varphi_\Gamma(X)}{\varphi_\Gamma(X^{-1})} = [\varphi_\Gamma(X)]^2,$$

where the latter inequality can be easily verified by (25). □

().

REFERENCES

ANDERSSON, S. A. (1975). Invariant normal models. *Ann. Statist.* **3** 132–154.

ANDERSSON, S. A., BRØNS, H. K. and JENSEN, S. T. (1983). Distribution of eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11** 392–415.

ANDERSSON, S. A. and MADSEN, J. (1998). Symmetry and lattice conditional independence in a multivariate normal distribution. *Ann. Statist.* **26** 525–572.

BARNDORFF-NIELSEN, O. (2014). *Information and exponential families in statistical theory*. *Wiley Series in Probability and Statistics*. John Wiley & Sons, Ltd., Chichester.

BURKE, C. J. and ROSENBLATT, M. (1958). A Markovian function of a Markov chain. *Ann. Math. Statist.* **29** 1112–1122.

DIACONIS, P. (1988). *Group representations in probability and statistics*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **11**. Institute of Mathematical Statistics, Hayward, CA.

DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281.

FARAUT, J. and KORÁNYI, A. (1994). *Analysis on symmetric cones*. *Oxford Mathematical Monographs*. The Clarendon Press, Oxford University Press, New York Oxford Science Publications.

FRETS, G. P. (1921). Heredity of head form in man. *Genetica* **41** 193–400.

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

GRAHAM, R. L., GRÖTSCHEL, M. and LOVÁSZ, L., eds. (1995). *Handbook of combinatorics. Vol. 1, 2*. Elsevier Science B.V., Amsterdam; MIT Press, Cambridge, MA.

HASSAIRI, A. and LAJMI, S. (2001). Riesz exponential families on symmetric cones. *J. Theoret. Probab.* **14** 927–948.

HØJSGAARD, S. and LAURITZEN, S. L. (2008). Graphical Gaussian models with edge and vertex symmetries. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **70** 1005–1027.

HOLT, D. F. (2010). Enumerating subgroups of the symmetric group. In *Computational group theory and the theory of groups, II*. *Contemp. Math.* **511** 33–37. Amer. Math. Soc., Providence, RI.

HUBERT, L. and P. ARABIE, P. (1985). Comparing partitions. *J. Classif.* **2** 193âĂŞ-218.

MAATHUIS, M., DRTON, M., LAURITZEN, S. and WAINWRIGHT, M. (2018). *Handbook of Graphical Models*. Chapman and Hall - CRC Chapman and Hall - CRC Handbooks of Modern Statistical Methods.

MADSEN, J. (2000). Invariant normal models with recursive graphical Markov structure. *Ann. Statist.* **28** 1150–1178.

MASSAM, H., LI, Q. and GAO, X. (2018). Bayesian precision and covariance matrix estimation for graphical Gaussian models with edge and vertex symmetries. *Biometrika* **105** 371–388.

OLKIN, I. and PRESS, S. J. (1969). Testing and estimation for a circular and stationary model. *Ann. Math. Statist.* **40** 1358–1373.

PLESKEN, W. and SOUVIGNIER, B. (1996). Constructing rational representations of finite groups. *Experiment. Math.* **5** 39–47.

SERRE, J. P. (1977). *Linear representations of finite groups*. Springer-Verlag, New York-Heidelberg Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.

SIEMONS, J. (1982). On partitions and permutation groups on unordered sets. *Arch. Math. (Basel)* **38** 391–403.

SIEMONS, J. (1983). Automorphism groups of graphs. *Arch. Math. (Basel)* **41** 379–384.

UPMEIER, H. (1986). Jordan algebras and harmonic analysis on symmetric spaces. *Amer. J. Math.* **108** 1–25.

WHITTAKER, J. (1990). *Graphical models in applied multivariate statistics*. Wiley.

WIELANDT, H. (1969). *Permutation groups through invariant relations and invariant functions*. *Lect. Notes Dept. Math.* Ohio St. Univ. Columbus.