ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 8, 2004

The Wishart distributions on homogeneous cones

STEEN A. ANDERSSON AND G. GERARD WOJNAR

ABSTRACT. The classical family of Wishart distributions on a cone of positive definite matrices and its fundamental features are extended to a family of generalized Wishart distributions on a *homogeneous cone* using the theory of exponential families. The generalized Wishart distributions include all known families of Wishart distributions as special cases. The relations to graphical models and Bayesian statistics are indicated.

1. Introduction

The classical Wishart distribution arises as the distribution of the maximum likelihood (ML) estimator of the covariance matrix in a multivariate sample as follows: Let I be a finite set. Whereas the context always precludes misunderstanding, I will also denote the cardinality of the set I. Let the observable $x \in \mathbb{R}^I$ follow the multivariate normal distribution $\mathrm{N}_I(0,\Sigma) \equiv \mathrm{N}_I(\Sigma)$ on $\mathbb{R}^I$ with expectation $0 \in \mathbb{R}^I$ and unknown covariance matrix $\Sigma \in \mathcal{P}(I,\mathbb{R})$, the open cone* of *positive definite* $I \times I$ matrices in the vector space $\mathcal{S}(I,\mathbb{R})$ of *symmetric* $I \times I$ matrices with real-valued entries. Thus one considers the statistical model†

$$(\mathrm{N}_I(\Sigma) \in \mathbf{P}(\mathbb{R}^I) | \Sigma \in \mathcal{P}(I,\mathbb{R})), \tag{1.1}$$

where $\mathbf{P}(\Omega)$ denotes the set of probability measures on a sample space Ω.

Received September 9, 2003.

2000 *Mathematics Subject Classification.* Primary 62-09, 62H05, 62H99; Secondary 06F99, 15A30.

Key words and phrases. Wishart distributions, homogeneous cones, Vinberg algebras, exponential familes, posets, matrices with zeroes, normal graphical models.

Research supported by U.S. National Science Foundation Grants DMS 97-04516 and DMS 00-71920.

*A cone is closed under addition, closed under multiplication with positive scalars, and contains no lines.

†A *statistical model* $(P_\theta \in \mathbf{P}(\Omega)|\theta \in \Theta)$ is a family (*not* a set) of probability measures P_θ, $\theta \in \Theta$, on the same *sample space* Ω, parametrized by the *parameter set* Θ.

Let N be a finite set and let $x_\nu \in \mathbb{R}^I$, $\nu \in N$, be N independent identically distributed (i.i.d.) observables from the model (1.1) indexed by N, where again N denotes both an index set and its cardinality. Thus the statistical model under consideration consists of N independent repetitions of the model (1.1), i.e., the normal model

$$(\mathrm{N}_I(\Sigma)^{\otimes N} \in \mathbf{P}((\mathbb{R}^I)^N)|\Sigma \in \mathcal{P}(I,\mathbb{R})), \tag{1.2}$$

where $\mathrm{N}_I(\Sigma)^{\otimes N}$ denotes the distribution[‡] of the observable $y \equiv (x_\nu|\nu \in N) \in (\mathbb{R}^I)^N \equiv \mathbb{R}^{I\times N} \equiv \mathcal{M}(I \times N, \mathbb{R})$, the vector space of $I \times N$ matrices with real-valued entries.

The ML estimator $\hat{\Sigma}$ of $\Sigma \in \mathcal{P}(I,\mathbb{R})$ exists with probability one if and only if $N \geq I$.[§] In the case $N \geq I$ it is uniquely given by $\hat{\Sigma}(y) = \frac{1}{N}yy'$, where y' denotes the transposed matrix of y. The distribution of the ML estimator $\hat{\Sigma}$ is the *classical Wishart distribution* with *multivariate scale* $\frac{1}{N}\Sigma$ and $f \equiv N$ *degrees of freedom*. This distribution of $\hat{\Sigma}$ was first derived by J. Wishart (1928). The Wishart distributions are concentrated on the open cone $\mathcal{P}(I,\mathbb{R})$. The family of Wishart distributions is usually parametrized by the multivariate scale $\Upsilon \in \mathcal{P}(I,\mathbb{R})$ and the degrees of freedom $f \in \{I, I+1, I+2, \cdots\}$. The expectation of the Wishart distribution is $f\Upsilon$ and the ML estimator $\hat{\Sigma}$ is thus unbiased.

In the present work it is more convenient to replace the multivariate scale Υ by the *expectation* $\Sigma := f\Upsilon \in \mathcal{P}(I,\mathbb{R})$ and replace the degrees of freedom f by the *shape parameter* $\lambda := \frac{1}{2}f \in \{\frac{I}{2}, \frac{I+1}{2}, \frac{I+2}{2}, \cdots\}$. The Wishart distribution with expectation Σ and shape parameter λ is thus denoted by $W_{\Sigma,\lambda}$ and is given by

$$dW_{\Sigma,\lambda}(S) = \frac{\lambda^{I\lambda}\det(S)^{\lambda-\frac{I+1}{2}}\exp\{-\lambda\mathrm{tr}(\Sigma^{-1}S)\}}{\pi^{\frac{I(I-1)}{4}}\prod(\Gamma(\lambda-\frac{i-1}{2})|i=1,\cdots,I)\det(\Sigma)^\lambda}dS, \tag{1.3}$$

where dS denotes the standard Lebesgue measure[¶] on $\mathcal{P}(I,\mathbb{R})$, $\det(\cdot)$ denotes the determinant of a square matrix, and $\mathrm{tr}(\cdot)$ denotes the trace of a square matrix. In fact, the right-hand side of (1.3) defines a probability measure for any $\lambda > \frac{I-1}{2}$ and any $\Sigma \in \mathcal{P}(I,\mathbb{R})$. This extension of the possible values of the shape parameter reduces in the case $I = 1$ to the usual inclusion of

[‡]Here $P^{\otimes N}$ is used for the product measure $\otimes(P|\nu \in N) \equiv P \otimes \cdots \otimes P$ (N times) on $\Omega^N \equiv \times(\Omega|\nu \in N) \equiv \Omega \times \cdots \times \Omega$ (N times) when P is a probability measure on a sample space Ω.

[§]In the opposite case $N < I$ the ML estimator does not exist for any observation $y \in \mathbb{R}^{I\times N}$.

[¶]More precisely: the restriction of the standard Lebesgue measure on the vector space $\mathcal{S}(I,\mathbb{R})$ to the open subset $\mathcal{P}(I,\mathbb{R})$.

the family of χ^2 distributions with integer degrees of freedom and positive scale within the family of gamma distributions.

Remark 1.1. It is, of course, also possible to define the Wishart distributions in terms of their Laplace transforms and/or their characteristic functions. In these cases it could be argued that a further extension of the family is natural. This extension corresponds to I extra values $\frac{I-1}{2}, \frac{I-2}{2}, \cdots, \frac{1}{2}, 0$ of the shape parameter λ. In these cases there is no density with respect to (wrt) to the Lebesgue measure on $\mathcal{P}(I, \mathbb{R})$ and for $\lambda = \frac{I-i}{2}$ the Wishart distribution is concentrated on the positive semidefinite $I \times I$ matrices of rank $I - i$, $i = 1, \cdots, I$. These I cases are therefore called the *singular* cases (as opposed to the *regular* cases). For details of the Laplace transform approach see Casalis and Letac (1996), Letac and Massam (1998), and their references. Another possible approach to this classical extension to the singular cases uses the existence of relatively invariant measures to define the extension. The latter method of extension to the singular cases has a generalization to generalized Wishart distributions in the present paper, cf. Remark 2.7 for a brief description of the method. Nevertheless, the theory presented in this paper concerns the regular cases only. □

The family of Wishart distributions on the sample space $\mathcal{P}(I, \mathbb{R})$ with a fixed shape parameter constitutes a statistical model in its own right:

$$(W_{\Sigma,\lambda} \in \mathbf{P}(\mathcal{P}(I, \mathbb{R})) | \Sigma \in \mathcal{P}(I, \mathbb{R})), \tag{1.4}$$

called the *classical Wishart model.* We emphasize that the shape parameter λ is considered as known and that the sample space and parameter set are identical. Such families of Wishart distributions provide the basis for construction of a very flexible class of Bayesian prior distributions[||] on the parameter set $\mathcal{P}(I, \mathbb{R})$ in the models (1.1), (1.2), and (1.4).

Any subset $\mathcal{P} \subseteq \mathcal{P}(I, \mathbb{R})$ defines by restriction of the parameter set to $\mathcal{P}$ in the models (1.1), (1.2), and (1.4), (a) the general normal model

$$(\mathrm{N}_I(\Sigma) \in \mathbf{P}(\mathbb{R}^I) | \Sigma \in \mathcal{P}), \tag{1.5}$$

(b) the N independent repetitions of the general normal model, i.e., the model

$$(\mathrm{N}_I(\Sigma)^{\otimes N} \in \mathbf{P}((\mathbb{R}^I)^N) | \Sigma \in \mathcal{P}), \tag{1.6}$$

and (c) the *restricted classical Wishart model* $(W_{\Sigma,\lambda} \in \mathbf{P}(\mathcal{P}(I, \mathbb{R})) | \Sigma \in \mathcal{P})$, respectively. The estimation of the parameter $\Sigma \in \mathcal{P}$ in such models is an important part of what is usually called *inference for the covariance matrix.*[**]

[||] Through the inverse Wishart distribution.

[**] Test procedures among such models are also included in the general terminology "inference for the covariance".

We now present twelve examples of such models, i.e., examples of subsets $\mathcal{P}$. These examples also illustrate many of the classical covariance hypotheses (models) in the literature.

In most of these examples $\mathcal{P}$ is an open subcone[††] of the open cone $\mathcal{P}(I,\mathbb{R})$. However, in *all* the examples the subset $\mathcal{P}$ can be *parametrized* in a one-to-one fashion by an open cone C, i.e.,

$$\begin{array}{ccc} C & \leftrightarrow & \mathcal{P} \\ c & \mapsto & \Sigma(c) \\ c(\Sigma) & \leftarrow\!\shortmid & \Sigma. \end{array} \tag{1.7}$$

In all the cases where $\mathcal{P}$ is a subcone of $\mathcal{P}(I,\mathbb{R})$, the parametrization of $\mathcal{P}$ by the open cone C is semilinear.[‡‡] When the parameter set C replaces $\mathcal{P}$ in the model (1.5) we obtain the model

$$(\mathrm{N}_I(\Sigma(c)) \in \mathbf{P}(\mathbb{R}^I)|c \in C). \tag{1.8}$$

Interest in the distribution of the ML estimator of $c \in C$ in the model (1.8), generalization of the classical Wishart model (1.4) to *the general Wishart model* with the cone C as sample and parameter space, and a flexible class of Bayesian prior distributions on the parameter set C constitute the basic motivations for our generalization of the Wishart distributions on $\mathcal{P}(I,\mathbb{R})$ to the generalized Wishart distributions on more general cones.

(i) **Real normal models.** When $\mathcal{P} = \mathcal{P}(I,\mathbb{R})$ one obtains the normal model (1.1) and the classical Wishart model in (1.4). The classical Wishart distribution is also called the *real Wishart distribution.* A cone isomorphic[*] to $\mathcal{P}(I,\mathbb{R})$ for some finite set I is said to be of *type* $\mathbb{R}$ or of *real type.*

(ii) **Complex normal models.** Let $\mathcal{P} = \mathcal{P}_{\mathbb{C}}(I,\mathbb{R})$ be the subset consisting of all $\Sigma \in \mathcal{P}(I,\mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Gamma & -\Delta \\ \Delta & \Gamma \end{pmatrix}. \tag{1.9}$$

That is, $I = J\dot{\cup}J$, where J is a finite set and $\dot{\cup}$ denotes disjoint union of sets, $\Gamma \in \mathcal{P}(J,\mathbb{R})$, and $\Delta \in \mathcal{K}(J,\mathbb{R})$, the vector space of $J \times J$ *antisymmetric* matrices with real-valued entries. The subset $\mathcal{P}$ is a subcone of $\mathcal{P}(I,\mathbb{R})$, and can be parametrized by the open cone $C := \mathcal{P}(J,\mathbb{C})$ of *positive definite (Hermitian)* $J \times J$ matrices in the vector space $\mathcal{H}(J,\mathbb{C})$ over $\mathbb{R}$ of *Hermitian* $J \times J$ matrices with entries from $\mathbb{C}$, the complex numbers[†]. The inverse parametrization in (1.7) is given by $c(\Sigma) \equiv \Phi := \Gamma + \mathbf{i}\Delta$. An observable

[††]A subcone will be defined precisely in Section 2.

[‡‡]A semilinear mapping is a linear mapping of the enveloping vector spaces restricted to the cones, cf. Section 2.

[*]Defined in Section 2.

[†]Note that the set of Hermitian matrices is *not* a vector space over $\mathbb{C}$.

from the normal statistical model (1.5) given by (1.9) is then equivalent to an observable from a complex normal distribution on $\mathbb{C}^J$ with expectation 0 and unknown complex covariance matrix $\Phi \in \mathcal{P}(J,\mathbb{C})$. Distribution of the ML estimator $\hat{\Phi}$ for $\Phi \in \mathcal{P}(J,\mathbb{C})$ for $N \geq \frac{I}{2} \equiv J$ independent repetitions of this model becomes the well-known *complex Wishart distribution*. See for example Goodman (1963), Khatri (1965a, 1965b), Andersson (1975a, 1975b, 1976, 1978, 1992), Andersson and Perlman (1984), and Andersson et al. (1975, 1983). Thus the distribution of the ML estimator $\hat{\Sigma}$ of $\Sigma \in \mathcal{P}_{\mathbb{C}}(I,\mathbb{R})$ is easily described through the reparametrization and is also called the complex Wishart distribution, although it is concentrated on the cone $\mathcal{P}_{\mathbb{C}}(I,\mathbb{R})$. The complex Wishart distribution $W_{\Sigma,\lambda}$ on $\mathcal{P}(J,\mathbb{C})$ with expectation $\Sigma \in \mathcal{P}(J,\mathbb{C})$ and shape parameter $\lambda > J-1$ is given by (6.3) with I replaced by J. The complex Wishart model is then given by

$$(W_{\Sigma,\lambda} \in \mathbf{P}(\mathcal{P}(J,\mathbb{C}))|\Sigma \in \mathcal{P}(J,\mathbb{C})). \tag{1.10}$$

A cone isomorphic to $\mathcal{P}(J,\mathbb{C})$ for some finite set J is said to be of *type* $\mathbb{C}$ or of *complex type*.

(iii) **Quaternion normal models.** Let $\mathcal{P} = \mathcal{P}_{\mathbb{H}}(I,\mathbb{R})$ be the subset consisting of all $\Sigma \in \mathcal{P}(I,\mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Gamma & -\Delta_1 & -\Delta_2 & -\Delta_3 \\ \Delta_1 & \Gamma & -\Delta_3 & \Delta_2 \\ \Delta_2 & \Delta_3 & \Gamma & -\Delta_1 \\ \Delta_3 & -\Delta_2 & \Delta_1 & \Gamma \end{pmatrix}. \tag{1.11}$$

That is, $I = J\dot{\cup}J\dot{\cup}J\dot{\cup}J$, where J is a finite set, $\Gamma \in \mathcal{P}(J,\mathbb{R})$, and Δ_1, Δ_2, $\Delta_3 \in \mathcal{K}(J,\mathbb{R})$. The subset $\mathcal{P}$ is a subcone of $\mathcal{P}(I,\mathbb{R})$, and can be parametrized by the open cone $C = \mathcal{P}(J,\mathbb{H})$ of *positive definite (Hermitian)* $J \times J$ matrices in the vector space $\mathcal{H}(J,\mathbb{H})$ over $\mathbb{R}$ of *Hermitian* $J \times J$ matrices with entries from the quaternion numbers $\mathbb{H}$. The inverse parametrization in (1.7) is given by $c(\Sigma) \equiv \Phi := \Gamma + \mathbf{i}\Delta_1 + \mathbf{j}\Delta_2 + \mathbf{k}\Delta_3$. An observable from the normal statistical model (1.5) given by (1.11) is then equivalent to an observable from a quaternion normal distribution on $\mathbb{H}^J$ with expectation 0 and unknown quaternion covariance matrix $\Phi \in \mathcal{P}(J,\mathbb{H})$. The distribution of the ML estimator $\hat{\Phi}$ for $\Phi \in \mathcal{P}(J,\mathbb{H})$ for $N \geq \frac{I}{4} \equiv J$ independent repetitions of this model is the *quaternion Wishart distribution*. See for example Andersson (1975a, 1975b, 1978, 1992) and Andersson et al. (1975). The distribution of the ML estimator $\hat{\Sigma}$ of $\Sigma \in \mathcal{P}_{\mathbb{H}}(I,\mathbb{R})$ is easily described through the reparametrization and is also called the quaternion Wishart distribution. It is concentrated on the cone $\mathcal{P}_{\mathbb{H}}(J,\mathbb{R})$. The quaternion Wishart distributions $W_{\Sigma,\lambda}$ on $\mathcal{P}(J,\mathbb{H})$ with expectation $\Sigma \in \mathcal{P}(J,\mathbb{H})$ and shape parameter $\lambda > 2J-2$ is given by (6.5) with I replaced by J. The quaternion Wishart model is then given by

$$(W_{\Sigma,\lambda} \in \mathbf{P}(\mathcal{P}(J,\mathbb{H}))|\Sigma \in \mathcal{P}(J,\mathbb{H})). \tag{1.12}$$

A cone isomorphic to $\mathcal{P}(J,\mathbb{H})$ for some finite set J is said to be of *type* $\mathbb{H}$ or of *quaternion type.*

(iv) **Group symmetry normal models.** Let H be a closed subgroup[‡] of $\mathcal{O}(I,\mathbb{R})$, the group of *orthogonal* $I \times I$ matrices, and let

$$\mathcal{P} = \mathcal{P}_H(I,\mathbb{R}) := \{\Sigma \in \mathcal{P}(I,\mathbb{R})|h\Sigma h' = \Sigma \text{ for all } h \in H\}. \tag{1.13}$$

The subset $\mathcal{P} \equiv \mathcal{P}_H(I,\mathbb{R})$ is a subcone of $\mathcal{P}(I,\mathbb{R})$ and the general model (1.5) given by (1.13) is called a *(normal) (group) symmetry model.* The class of symmetry models was first defined by Andersson (1975b). In the same work the structure of the models and one form of the ML estimator were presented. The theory and the complete likelihood analysis of symmetry models are central parts of an unpublished general algebraic theory of normal models developed by Andersson, Brøns, and Tolver Jensen in the years 1972-1985. Several parts of this algebraic theory are reported in several sets of unpublished lecture notes in Danish, Andersson (1975a, 1976), Tolver Jensen (1973, 1974, 1977, 1983), Brøns (1969), Andersson *et al.* (1975). These notes contain a complete solution for the likelihood inference for symmetry models, including ML estimation, distribution of ML estimators, likelihood ratio (LR) statistics and their central distributions. A small but central part of this work can be found in Andersson (1978, 1992) and Andersson and Madsen (1998), Appendix A. See also Perlman (1987).

The interpretation of the symmetry model (1.5) given by (1.13) is that the distribution of the observable x is *invariant* under each symmetry transformation $h \in H$, i.e., x and hx have the same distribution. In many important examples the symmetry group H is induced by a subgroup, denoted by $\mathcal{S}$, of the full permutation group $\mathcal{S}(I)$ of the index set I. For $\sigma \in \mathcal{S}$, a symmetry $h \equiv h(\sigma) \in H$ is given by $h^{-1}x := (x_{\sigma(i)}|i \in I)$, where $x \equiv (x_i|i \in I) \in \mathbb{R}^I$ and h^{-1} denotes the inverse of $h \in H$.

Note that $\mathcal{P}$ in (i), (ii), and (iii) are special cases of $\mathcal{P}_H(I,\mathbb{R})$. In fact the corresponding three types of models are the "building stones" of all symmetry models, as described in Andersson (1975b). The main result therein is that the symmetry model (1.5) given by (1.13), by an orthogonal change of basis in $\mathbb{R}^I$, becomes an independent product of normal models[§] *each* consisting of independent repetitions of a real normal model, a complex normal model, or a quaternion normal model. This decomposition of a symmetry

[‡]More precisely, consider an orthogonal continuous representation on $\mathbb{R}^I$ of a compact group H.

[§]A product of a finite family of statistical models $(P_{\theta_\mu} \in \mathbf{P}(\Omega_\mu)|\theta_\mu \in \Theta_\mu)$, $\mu \in M$, is the statistical model $(P_\theta \in \mathbf{P}(\Omega)|\theta \in \Theta)$, where $\Omega := \times(\Omega_\mu|\mu \in M)$, $\Theta := \times(\Theta_\mu|\mu \in M)$, and $P_\theta := \otimes(P_{\theta_\mu}|\mu \in M)$, $\theta \equiv (\theta_\mu|\mu \in M)$.

model is unique up to an uninteresting isomorphism[¶]. In particular, it follows from Andersson (1975b) that any group-symmetry cone $\mathcal{P}_H(I,\mathbb{R})$ is isomorphic to a cone $C = \times(C_\kappa|\kappa \in K)$, where for each $\kappa \in K$, $C_\kappa = \mathcal{P}(I_\kappa,\mathbb{R})$, $C_\kappa = \mathcal{P}(I_\kappa,\mathbb{C})$, *or* $C_\kappa = \mathcal{P}(I_\kappa,\mathbb{H})$. Thus when the parametrization by C is used, the distribution of the ML estimator $\hat{c}$ of $c \in C$ can be described as a product of Wishart distributions on C_κ, $\kappa \in K$, each being a classical, a complex, or a quaternion Wishart distribution. Andersson *et al.* (1975) give a simple definition and description of the family of Wishart distributions on the cone $\mathcal{P}_H(I,\mathbb{R})$, see also Andersson (1978, 1992). This definition is the inspiration for the definition of the generalized Wishart distribution in Section 3. The distribution of the ML estimator $\hat{\Sigma}$ of $\Sigma \in \mathcal{P}_H(I,\mathbb{R})$ is then obtained as a generalized Wishart distribution.

It should be mentioned that in the examples[‖] of symmetry models in the literature, the distribution of $\hat{\Sigma}$ is not described. We now list the classical examples of symmetry models from the literature together with some comments. For each example the generalized Wishart distributions will provide a corresponding new family of Wishart distributions defined on the cone defining the model.

(v) **Multivariate complete symmetry models (multivariate interclass correlation).** Let $\mathcal{P}$ be the subset consisting of all $\Sigma \in \mathcal{P}(I,\mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Gamma & \Delta & \cdots & \Delta \\ \Delta & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \Delta \\ \Delta & \cdots & \Delta & \Gamma \end{pmatrix}. \tag{1.14}$$

That is, $I = J\dot{\cup}J\dot{\cup}J\cdots\dot{\cup}J$ (say k times), $\Gamma \in \mathcal{P}(J,\mathbb{R})$, and $\Delta \in \mathcal{S}(J,\mathbb{R})$. The normal model (1.5) given by (1.14) is a group symmetry model with the the group H induced by the group $\mathcal{S} = \mathcal{S}(\{1,\cdots,k\})$ of all permutations of the k equal-sized J-blocks of I. For $J = 1$ this model was first defined and analyzed under the name of *complete symmetry* by Wilks (1946). His student Votaw (1948) defined and analysed this general model under the name of *compound symmetry of type* II[**]. Votaw's covariance matrix Σ appears to be different from the above, but this is only because his ordering of the coordinates is different from the one presented above. It was probably because of this superficial difference that the model reappeared again in Arnold (1973) in the present form without any reference to Votaw. This history shows the

[¶]Two statistical models $(P_\theta \in \mathbf{P}(E)|\theta \in \Theta)$ and $(Q_\xi \in \mathbf{P}(F)|\xi \in \Xi)$ are called *isomorphic* if there exists a bimeasurable bijective mapping $t : E \to F$ and a bijective mapping $\psi : \Theta \to \Xi$ such that $t(P_\theta) = Q_{\psi(\theta)}$, $\theta \in \Theta$.

[‖]Except for the complex normal model in (ii) above.

[**]Votaw's compound symmetry of type I is also an interesting symmetry model.

importance of invariant formulation in statistics, in particular in multivariate statistical analysis. It is well known that the normal model (1.5) determined by (1.14) decomposes into an independent product of two models of the type (1.2).

(vi) **Multivariate circular symmetry models.** Let $\mathcal{P}$ be the subset consisting of all $\Sigma \in \mathcal{P}(I, \mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Gamma & \Delta_1 & \Delta_2 & \cdots & \Delta_{k-2} & \Delta_{k-1} \\ \Delta_{k-1} & \Gamma & \Delta_1 & \ddots & \ddots & \Delta_{k-2} \\ \Delta_{k-2} & \Delta_{k-1} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \Delta_1 & \Delta_2 \\ \Delta_2 & \ddots & \ddots & \Delta_{k-1} & \Gamma & \Delta_1 \\ \Delta_1 & \Delta_2 & \cdots & \Delta_{k-2} & \Delta_{k-1} & \Gamma \end{pmatrix}. \tag{1.15}$$

That is, $I = J \dot{\cup} J \dot{\cup} J \cdots \dot{\cup} J$ (k times), $\Gamma \in \mathcal{P}(J, \mathbb{R})$, and $\Delta_i = \Delta'_{k-i} \in \mathcal{M}(J, \mathbb{R}) := \mathcal{M}(J \times J, \mathbb{R})$, $i = 1, \cdots, k-1$. Thus, when k is odd there are $\frac{k-1}{2}$ different restricting equations among $\Delta_1, \cdots, \Delta_{k-1}$, but when k is even there are $\frac{k}{2}$ different restricting equations, the last one being $\Delta_{\frac{k}{2}} = \Delta'_{\frac{k}{2}}$, i.e., $\Delta_{\frac{k}{2}} \in \mathcal{S}(I, \mathbb{R})$.

The normal model (1.5) determined by (1.15) is a symmetry model with the group H induced by the group $S = \mathcal{C}(k)$ of all the *cyclic* permutations of the k equal-sized J-blocks of I. For $J = 1$ this model was first defined and analyzed under the name of *circular symmetry* by Olkin and Press (1968). For $J > 1$ the multivariate circular symmetry model has yet to appear in the literature. Nevertheless, Olkin (1973) did define and analyze a normal model under the name *circular symmetry in blocks*: the covariance matrix Σ was defined as in (1.15) *but* with the more restrictive condition that $\Delta_i = \Delta_{k-i} \in \mathcal{S}(J, \mathbb{R})$, $i = 1, \cdots, k-1$. For $J > 1$ this model is a proper submodel of the multivariate circular symmetry model. In fact a closer examination shows that Olkin's model is the dihedral†† symmetry model. Thus Olkin's model should properly not be called circular symmetry in blocks but rather the *multivariate dihedral symmetry model*.

It can be established that when k is odd the multivariate circular symmetry model (1.5) given by (1.15) decomposes into an independent product of one model of the type (1.1) and $\frac{k-1}{2}$ complex normal models, cf. (ii). In the case when k is even there will be two models of type (1.1) and $\frac{k}{2} - 1$ complex normal models in the decomposition. For the dihedral symmetry model there will be, when k is odd, one model of type (1.1) and $\frac{k-1}{2}$ models

††The dihedral group is generated by the cyclic group and the permutation that reverses the order of $\{1, \cdots, k\}$.

of type (1.2) with $N = 2$. When k is even there will be two models of type (1.1) and $\frac{k}{2} - 1$ models of type (1.2) with $N = 2$.

(vii) **Multivariate block independence.** Let $\mathcal{P}$ be the subset consisting of all $\Sigma \in \mathcal{P}(I, \mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Gamma_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Gamma_k \end{pmatrix} =: \operatorname{diag}(\Gamma_\kappa | \kappa = 1, \cdots, k). \tag{1.16}$$

That is, $I = J_1 \dot{\cup} \cdots \dot{\cup} J_k$, where J_κ is a finite set, and $\Gamma_\kappa \in \mathcal{P}(J_\kappa, \mathbb{R})$, $\kappa = 1, \cdots, k$. The normal model (1.5) given by (1.16) is a symmetry model with the group H induced by variation-independent sign changes of the k components of the observable $(x_1, \cdots, x_k) \equiv x \in \mathbb{R}^I$. The model is well-known, cf. Anderson (1984), Chapter 9. Trivially, the model (1.5) given by (1.16) decomposes into a product of k models of the type (1.1).

(viii) **Multivariate spherical symmetry (the i.i.d. normal model).** Let $\mathcal{P}$ be the subset consisting of all $\Sigma \in \mathcal{P}(I, \mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Gamma & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Gamma \end{pmatrix} \equiv \operatorname{diag}(\Gamma | \kappa = 1, \cdots, k). \tag{1.17}$$

That is, $I = J \dot{\cup} \cdots \dot{\cup} J$ (k times), where J is a finite set, and $\Gamma \in \mathcal{P}(J, \mathbb{R})$. The normal model (1.5) given by (1.17) is a symmetry model with the group H generated by the groups in (v) and (vii). The model is also the model (1.2) with $N = k$. This shows that the model (1.2) is in fact also itself a symmetry model. By further combining with a general group H from (iv) it is seen that N independent repetitions of a symmetry model is itself again a symmetry model.

(ix) **The Jordan normal models.** The model consisting of N independent repetitions of the model (1.5) can be considered as a curved[‡‡] exponential family with $\theta := \Sigma^{-1}$, $\Sigma \in \mathcal{P} \subseteq \mathcal{P}(I, \mathbb{R})$, as the canonical parameter. Anderson (1969, 1970, and 1973) has studied the likelihood equation when $\mathcal{P}$ is a cone, obtaining its solution via iterative procedures, and the asymptotic distribution of the ML estimator. To obtain further and more explicit results using the powerful theory of full regular exponential families, it is natural to also require linearity in the canonical parameter, i.e., that $\mathcal{P}^{-1} := \{\Sigma^{-1} | \Sigma \in \mathcal{P}\}$ is a cone. Under this requirement the likelihood

[‡‡]This requires at least that $\mathcal{P}$ be a differentiable manifold, for example a cone.

equation becomes a linear equation and the solution (if any) becomes a linear mapping of yy', $y \in \mathbb{R}^{I \times N}$. By assuming (without loss of generality) that $\mathcal{P}$ contains the identity matrix $1_I := \operatorname{diag}(1|i \in I)$, one obtains the requirement that $\mathcal{P}$ is a cone and that $\mathcal{P} = \mathcal{P}^{-1}$. These assumptions seem to be rather general, but Tolver Jensen (1988) established that this class of normal models is almost the same as the class of symmetry models in (v) above*. He noted that the linear span $\mathcal{J} \subseteq \mathcal{S}(I, \mathbb{R})$ of $\mathcal{P}$ is a formally real Jordan algebra† (see also Seely (1971,1972)) and that $\mathcal{P}$ is the set $\mathcal{J}^+$ of positive elements in $\mathcal{J}$, i.e., the interior of the set $\{A^2 | A \in \mathcal{J}\}$ of non-negative elements in $\mathcal{J}$. Using the well-known structure theorems for formally real Jordan algebras and their linear representation, Tolver Jensen obtained the structure of the so-called *Jordan normal model* as now described.

There exists a basis for $\mathbb{R}^I$ such that the Jordan model becomes a product of normal models, each factor in the product being one of the following four types: independent repetitions of a real normal model, independent repetitions of a complex normal model, independent repetitions of a quaternion normal model, or almost‡ independent repetitions of the so-called *Lorentz* or *Clifford normal model.* Each of the models in the decomposition corresponds to a simple Jordan algebra together with a representation of the simple Jordan algebra, cf. Tolver Jensen (1988), Section 4 and 5.

The sample space $\mathbb{R}^J$ for a Lorentz normal model must have $J = 2^n$ for some integer n. Within this new class of Lorentz models Tolver Jensen obtained explicitly the ML estimator with its distribution and the likelihood ratio (LR) statistics with their central distributions.

The adjectives "Lorentz" or "Clifford" are justified as follows: the corresponding simple Jordan algebra $\mathcal{J}$ is closely related to the *Clifford algebra*, cf. Braun and Koecher (1966), and the set of positive elements in this simple Jordan algebra is often called the *Lorentz cone*, cf. for example Faraut and Korányi (1994), page 7.

Although the Lorentz cone, cf. (x) below, can be represented§ as a cone $\mathcal{P}$ of positive definite matrices and the Wishart distribution can be introduced as the distribution of the ML estimator as in Tolver Jensen (1988), in the present paper we shall define the extended class of *Lorentz* Wishart distributions directly on the Lorentz cone itself, cf. Example 6.6, and in turn obtain the *Lorentz Wishart model.* A cone is said to be of *Lorentz type* if it is isomorphic to a Lorentz cone, i.e., a cone

$$C := \{(\alpha, x) \in \mathbb{R} \times W | \alpha > 0,\ \alpha^2 - \|x\|^2 > 0\}, \tag{1.18}$$

*Note that the symmetry models satisfy these conditions.

†Sometimes called *quadratic subspaces* in statistical literature.

‡See Tolver Jensen (1988), Theorem 6, page 318.

§The representation entails the restriction of a representation of the simple Jordan algebra to its positive elements.

where W is a Euclidean space with inner product denoted by $x \cdot y$, and $\|x\|^2 := x \cdot x$, $x, y \in W$. In particular, it follows from Tolver Jensen (1988) that any subcone $\mathcal{P} \subseteq \mathcal{P}(I, \mathbb{R})$ with the properties $1_I \in \mathcal{P}$ and $\mathcal{P} = \mathcal{P}^{-1}$ is isomorphic to a product of indecomposable* cones, each of which is one of the following four types: real type, complex type, quaternion type, or Lorentz type.

(x) **Symmetric cones.** A self-dual homogeneous cone† is called a *symmetric cone.* Such cones are studied by Faraut and Korányi (1994). Any symmetric cone is isomorphic to a unique product of indecomposable symmetric cones, each of which is one of the following five types: real type, complex type, quaternion type, Lorentz type, or *the* so-called *exceptional* cone $\mathcal{P}(3, \mathbb{O})$ of 3×3 positive definite‡ Hermitian matrices with entries from the *octonions* $\mathbb{O}$, cf. Example 6.8. Note that the fifth type consists of *only one* cone.

The theory of symmetric cones is closely related to, indeed almost equivalent to, the theory of formally real Jordan algebras. In a Jordan algebra $\mathcal{J}$ one may consider the positive part $\mathcal{J}^+$, the interior of the non-negative part $\{a^2 | a \in \mathcal{J}\}$. This forms a symmetric cone, and all symmetric cones can be realized in this way. But this realization of a symmetric cone requires distinct status to an arbitrary point in the cone, thus the Jordan algebra formulation is less invariant than the cone formulation. Casalis and Letac (1996), Letac and Massam (1998, 2000), and Massam and Neher (1997) have discussed many aspects of Wishart distributions on $\mathcal{J}^+$, in particular, on the five types of indecomposable symmetric cones above, cf. also Faraut and Korányi (1994), Chapter XVI, Section 1.

The name "symmetric" cone might be confused with a cone given by a symmetry group H i.e., isomorphic to the cone $\mathcal{P}_H(I, \mathbb{R})$, cf. (v) above. In fact, from a structural point of view the difference between these two classes is that the decomposition of the more general symmetric cones into the product of indecomposable cones may contain indecomposable components isomorphic to a cone of Lorentz type or *the* cone $\mathcal{P}(3, \mathbb{O})$, cf. (iv) above.

(xi) **Marginal independence model.** Let $\mathcal{P}$ be the subset consisting of all $\Sigma \in \mathcal{P}(I, \mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & 0 & \Sigma_{a1} \\ 0 & \Sigma_{bb} & \Sigma_{b1} \\ \Sigma_{1a} & \Sigma_{b1} & \Sigma_{11} \end{pmatrix}. \tag{1.19}$$

In particular $I = I_a \dot{\cup} I_b \dot{\cup} I_1$, where I_a, I_b, and I_1 are finite sets, $\Sigma_{aa} \in \mathcal{P}(I_a, \mathbb{R})$, $\Sigma_{bb} \in \mathcal{P}(I_b, \mathbb{R})$, $\Sigma_{11} \in \mathcal{P}(I_1, \mathbb{R})$, $\Sigma_{a1} = \Sigma_{1a}^t \in \mathcal{M}(I_a \times I_1, \mathbb{R})$, and

*Explained in Section 2.

†These properties will be defined in Remark 2.4.

‡This concept will be defined in Example 6.8

$\Sigma_{b1} = \Sigma_{1b}^t \in \mathcal{M}(I_b \times I_1, \mathbb{R})$. The subset $\mathcal{P}$ is a subcone of $\mathcal{P}(I, \mathbb{R})$. The corresponding decomposition of an observable $x \in \mathbb{R}^I$ is $x \equiv (x_a, x_b, x_1) \in \mathbb{R}^I \equiv \mathbb{R}^{I_a} \times \mathbb{R}^{I_b} \times \mathbb{R}^{I_1}$. The interpretation of the normal model (1.5) given by (1.19) is that x_a and x_b are independent, i.e., the marginal distribution on $\mathbb{R}^{I_a} \times \mathbb{R}^{I_b}$ is the product distribution $N(\Sigma_{aa}) \otimes N(\Sigma_{bb})$. This model is a special case of the *lattice* models introduced by Andersson and Perlman (1993). In Example 6.14 it will be indicated that some of the general lattice normal model is given by a homogeneous cone.

(xii) **Conditional independence model.** Let $\mathcal{P}$ be the subset consisting of all $\Sigma \in \mathcal{P}(I, \mathbb{R})$ of the form

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{a0}\Sigma_{00}^{-1}\Sigma_{0b} & \Sigma_{a0} \\ \Sigma_{b0}\Sigma_{00}^{-1}\Sigma_{0a} & \Sigma_{bb} & \Sigma_{b0} \\ \Sigma_{0a} & \Sigma_{0b} & \Sigma_{00} \end{pmatrix}. \tag{1.20}$$

In particular $I = I_a \dot\cup I_b \dot\cup I_0$, where I_a, I_b, and I_0 are finite sets, $\Sigma_{aa} \in \mathcal{P}(I_a, \mathbb{R})$, $\Sigma_{bb} \in \mathcal{P}(I_b, \mathbb{R})$, $\Sigma_{00} \in \mathcal{P}(I_0, \mathbb{R})$, $\Sigma_{0a} = \Sigma_{a0}^t \in \mathcal{M}(I_0 \times I_a, \mathbb{R})$, and $\Sigma_{0b} = \Sigma_{b0}^t \in \mathcal{M}(I_0 \times I_b, \mathbb{R})$. However the subset $\mathcal{P}$ is *not* a subcone of $\mathcal{P}(I, \mathbb{R})$. The corresponding decomposition of an observable $x \in \mathbb{R}^I$ is $x \equiv (x_a, x_b, x_0) \in \mathbb{R}^I \equiv \mathbb{R}^{I_a} \times \mathbb{R}^{I_b} \times \mathbb{R}^{I_0}$. The interpretation of the normal model (1.5) given by (1.20) is that x_a and x_b are independent given x_0, i.e., the conditional distribution on $\mathbb{R}^{I_a} \times \mathbb{R}^{I_b}$ is the product distribution $N(\Sigma_{a0}\Sigma_{00}^{-1}x_0, \Sigma_{a\bullet}) \otimes N(\Sigma_{b0}\Sigma_{00}^{-1}x_0, \Sigma_{b\bullet})$, where $\Sigma_{a\bullet} := \Sigma_{aa} - \Sigma_{a0}\Sigma_{00}^{-1}\Sigma_{0a}$ and $\Sigma_{b\bullet} := \Sigma_{bb} - \Sigma_{b0}\Sigma_{00}^{-1}\Sigma_{0b}$.

The parameter set $\mathcal{P}$ can be parametrized by the cone C of all arrays of the form

$$c = \begin{pmatrix} \Sigma_{aa} & & \Sigma_{a0} \\ & \Sigma_{bb} & \Sigma_{b0} \\ \Sigma_{0a} & \Sigma_{0b} & \Sigma_{00} \end{pmatrix}$$

where

$$\begin{pmatrix} \Sigma_{aa} & \Sigma_{a0} \\ \Sigma_{0a} & \Sigma_{00} \end{pmatrix} \in \mathcal{P}(I_a \dot\cup I_0, \mathbb{R}) \quad \text{and} \quad \begin{pmatrix} \Sigma_{bb} & \Sigma_{b0} \\ \Sigma_{0b} & \Sigma_{00} \end{pmatrix} \in \mathcal{P}(I_b \dot\cup I_0, \mathbb{R}),$$

i.e., two overlapping positive definite matrices. The parametrization $c \mapsto \Sigma(c)$ is given in (1.20). The cone C is again an example of a homogeneous cone[§] and the normal model is again a special case of the lattice models of Andersson and Perlman (1993).

In Section 2 we define cones and mappings between cones, in particular the dual cone. Homogeneous cones are defined and several simple facts

[§]In fact, a homogeneous cone that is not self-dual, i.e., not a symmetric cone.

are stated. The exponential family based on a relatively invariant measure on a homogeneous cone are presented in the same section. The generalized Wishart distribution is defined in Section 3 by changing the canonical parametrization of the exponential family from Section 2 into the parametrization by the expectation parameter. Several properties enjoyed by this new class of generalized Wishart distributions are then easily obtained.

Homogeneous cones can be described in several different ways, all introduced by Vinberg (1960, 1963). Vinberg's original aim was to describe all homogeneous convex domains (= a cut in a homogeneous cone) and their automorphism groups, cf. also Vinberg (1962, 1965). We have chosen the so-called T-algebras introduced by Vinberg as the instrument for the description of the homogeneous cones. These algebras have since been called *Vinberg algebras* and we shall adopt this name. From a pure mathematical point of view, Vinberg algebras do not provide the most efficient description[¶] of homogeneous cones, *but* lie close to one of the standard definitions of the cone of positive definite matrices. Thus the classical Wishart distribution on the cone of positive definite matrices becomes a straightforward special case. The elements of a Vinberg algebra are arrays $A \equiv (a_{ij} | (i,j) \in I \times I)$ with entries a_{ij} indexed by an index set of the form $I \times I$, where I is a finite set. The algebra $\mathcal{M}(I, \mathbb{R})$ of $I \times I$ matrices is with standard matrix multiplication a special case of a Vinberg algebra. Several of the standard manipulations with square matrices used in classical multivariate statistical analysis can then be generalized to a Vinberg algebra.

We have thus chosen clarity and familiarity of presentation over mathematical efficiency of the formulation. Even with this choice, viz. Vinberg algebras, a fair amount of abstract mathematics from different areas is required. In Section 4 we introduce Vinberg algebras in a manner slightly different from Vinberg's original formulation. The main result in this section is that all homogeneous cones are cones of "positive definite matrices" in a certain generalized matrix algebra. With this description of all homogeneous cones, the generalized Wishart distribution is described in complete detail in Section 5. In Section 6 many examples are presented in detail from the point of view of a Vinberg algebra and several new Wishart distributions with their homogeneous cones are presented.

It then becomes obvious that the generalized Wishart distributions apply to many statistical models involving unknown covariance matrices with structures defined by symmetry and/or graphical relations. In particular many special cases are of interest in their own right and require a special treatment beyond the present paper. This situation is analogous to the theory of classical linear models, where special cases such as two-way analysis

[¶] Vinberg's nilpotent N-algebras are most efficient.

of variance, Latin square, etc., require special attention beyond the general theory.

The paper also opens several new avenues to be investigated, for which some results are already obtained, but beyond the scope of the present paper, cf. Section 7.

2. Homogeneous cones and exponential families

Let $\mathbb{R}_+$, $\mathbb{R}_{+0}$, $\mathbb{Z}$, $\mathbb{N}$, and $\mathbb{N}_0$ denote the positive real numbers, the non-negative real numbers, the integers, the positive integers, and the non-negative integers, respectively.

Definition 2.1. A *finite-dimensional closed proper convex cone* is a pair (C_0, V) consisting of a finite-dimensional vector space V over $\mathbb{R}$ and a closed subset $C_0 \subseteq V$ with the properties

(i) $\alpha_1 v_1 + \alpha_2 v_2 \in C_0$, for all $\alpha_1, \alpha_2 \in \mathbb{R}_{+0}$ and all $v_1, v_2 \in C_0$.

(ii) $C_0 + (-C_0) = V$.

(iii) $C_0 \cap (-C_0) = \{0\}$.

Remark 2.1. In the literature it is customary to call a (closed) subset C_0 of a vector space a cone if it closed under multiplication by the non-negative real numbers. A cone C_0 is then called convex if $\alpha_1 v_1 + \alpha_2 v_2 \in C_0$ for all $\alpha_1, \alpha_2 \in \mathbb{R}_{+0}$ with $\alpha_1 + \alpha_2 = 1$ and all $v_1, v_2 \in C_0$. Condition (i) in Definition 2.1 ensures that the subset C_0 is a convex cone in V. Condition (ii) together with (i) ensures that the interior of the closed subset C_0, denoted by C, is non-empty. A closed convex cone C_0 is usually called proper if it contains no line through 0. Condition (iii) ensures that the subset C_0 is such a proper convex cone in V. We shall abbreviate "finite-dimensional closed proper convex cone" in Definition 2.1 to just *closed cone* and usually we shall subsume the vector space V and refer to C_0 itself as a closed cone. The vector space is called the *enveloping* vector space. Finite-dimensional *open* proper convex cones are also often considered in the literature. They are defined by replacing "closed subset", "$\mathbb{R}_{+0}$", and (iii) in Definition 2.1 by "open non-empty subset", "$\mathbb{R}_+$", and (iii)$'$, respectively, where (iii)$'$ states that the open set, now denoted by C, contains no affine line. We shall again abbreviate "finite dimensional open proper convex cone" to just *open cone* or even just *cone*, and again we shall often subsume the vector space V and simply refer to the (open) cone C. Note that the interior of a closed cone is an open cone and the closure of an open cone is a closed cone. The *zero cone* $(\{0\}, \{0\})$ is an open and closed cone and the only cone where $C = V$. For all other closed cones, $C_0 \subset V$. For all open cones C except the zero cone, $0 \notin C$. □

Definition 2.2. A *homomorphism* from a closed cone (C_{01}, V_1) to a closed cone (C_{02}, V_2) is a linear mapping $f : V_1 \to V_2$ with the property $f(C_{01}) \subseteq C_{02}$.

Remark 2.2. Any mapping $f : C_{01} \to C_{02}$ with the property

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2), \tag{2.1}$$

for all $\alpha_1, \alpha_2 \in \mathbb{R}_{0+}$ and all $v_1, v_2 \in C_{01}$ can uniquely be extended to a homomorphism, also denoted by f, from (C_{01}, V_1) to (C_{02}, V_2). Furthermore, the restriction of a homomorphism f from (C_{01}, V_1) to (C_{02}, V_2) uniquely defines a mapping, also denoted by $f : C_{01} \to C_{02}$, with the property (2.1). Thus there is a one-to-one correspondence between all homomorphisms from (C_{01}, V_1) to (C_{02}, V_2) and all mappings from C_{01} to C_{02} satisfying (2.1). We shall therefore identify the homomorphisms and the mappings with the property (2.1) and refer to the homomorphism as a linear mapping from C_{01} to C_{02}.

If in the above, C_{01} and C_{02} are replaced by the corresponding open cones C_1 and C_2, respectively, and $\mathbb{R}_{+0}$ is replaced by $\mathbb{R}_+$ in (2.1) then a similar one-to-one correspondence still holds. □

The set of all linear mappings from C_{01} to C_{02} is denoted by $\mathrm{Hom}_0(C_{01}, C_{02})$. The subset $\mathrm{Hom}_0(C_{01}, C_{02}) \subseteq \mathrm{Hom}(V_1, V_2)$, where $\mathrm{Hom}(V_1, V_2)$ is the vector space of linear mappings from the vector space V_1 to the vector space V_2, is itself a closed cone with $\mathrm{Hom}(V_1, V_2)$ as the enveloping vector space. The interior of the closed cone $\mathrm{Hom}_0(C_{01}, C_{02})$ is $\{f \in \mathrm{Hom}_0(C_{01}, C_{02}) | f(C_1) \subseteq C_2\} =: \mathrm{Hom}(C_1, C_2)$, the open cone of homomorphisms from C_1 to C_2. A composition of linear mappings of closed cones (open cones) is again a linear mapping of closed (open) cones and the identity mapping Id_{C_0} (Id_C) of the *set* C_0 (C) is the identity mapping of the closed cone C_0 (open cone C). A linear mapping f from C_{01} to C_{02} (C_1 to C_2) is called *surjective* if $f(C_{01}) = C_{02}$ ($f(C_1) = C_2$) and is called *injective* if f is one-to-one. A linear mapping f from C_{01} to C_{02} is surjective if and only if f is a surjective linear mapping from C_1 to C_2. A linear mapping is an isomorphism between the closed cones C_{01} and C_{02} (open cones C_1 and C_2) if it is surjective and injective. A linear mapping from C_{01} to C_{02} is an isomorphism if and only if it is an isomorphism between C_1 and C_2. Closed cones (open cones) are *isomorphic* if there exists an isomorphism between them. Note that two closed cones are isomorphic if and only if the corresponding open cones are isomorphic.

A pair (C_{00}, u) consisting of a closed cone C_{00} with enveloping vector space V_0 and a injective linear mapping $u \in \mathrm{Hom}_0(C_{00}, C_0)$ is called a *closed subcone* of the closed cone C_0. When $C_{00} \subseteq C_0$, $V_0 \subseteq V$ is a subspace, and u is the embedding mapping, we subsume u and call $C_{00} \subseteq C_0$ a *standard* *closed subcone*. Note that if $u : V_0 \to V$ is bijective, $u^{-1} : V \to V_0$ is not in general an element in $\mathrm{Hom}_0(C_0, C_{00})$.

The corresponding concepts of *open subcone* and *standard open subcone*, or simply *subcone*, are similarly formulated.

Let $(C_{0i}|i \in I)$ be a finite family of closed cones and let V_i denote the enveloping vector space for the cone C_{0i}, $i \in I$. The *closed product cone* of this family is then the cone $C_0 := \times(C_{0i}|i \in I)$, with the enveloping vector space of C_0 being $V := \times(V_i|i \in I)$. Note that $C = \times(C_i|i \in I)$.

A cone (open or closed) is called *indecomposable* if it is not isomorphic to a nontrivial product of cones. Any closed (open) cone splits *uniquely* into a product of indecomposable closed (open) cones.

The *dual closed cone* to a closed cone C_0 is given by $C_0^* := \mathrm{Hom}_0(C_0, \mathbb{R}_{+0})$ with enveloping vector space V^* (the dual vector space). Note that $C_0^* = \{x^* \in V^* |\, x^*(p) \geq 0, \text{ for all } p \in C_0\}$ and $C^* := \mathrm{Int}(C_0^*) = \{x^* \in V^* |\, x^*(p) > 0, \text{ for all } p \in C\backslash\{0\}\}$. Any linear mapping $f \in \mathrm{Hom}_0(C_{01}, C_{02})$ determines the dual linear mapping ${}^t f \in \mathrm{Hom}_0(C_{02}^*, C_{01}^*)$ given by ${}^t f(p_2^*) := p_2^* \circ f$, $p_2^* \in C_{02}^*$. If $f \in \mathrm{Hom}(C_1, C_2)$ then ${}^t f \in \mathrm{Hom}(C_2^*, C_1^*)$. Note that if f is surjective then ${}^t f$ is injective. *But* if f is injective we cannot in general conclude that f is surjective.

Unlike the category of vector spaces and linear mappings, C_0 and C_0^* are not in general isomorphic cones. Nevertheless, the natural identification $(x \leftrightarrow (x^* \mapsto x^*(x)))$ of V and V^{**} identifies the closed cones C_0 and C_0^{**} (and the open cones C and C^{**}) by the natural cone isomorphism. The natural identification

$$
\begin{array}{rcl}
\times(V_i^*|i \in I) & \leftrightarrow & (\times(V_i|i \in I))^* \\
(x_i^*|i \in I) & \mapsto & ((x_i|i \in I) \mapsto \sum(x_i^*(x_i)|i \in I))
\end{array}
\tag{2.2}
$$

also identifies the closed cones $\times(C_{0i}^*|i \in I)$ and $(\times(C_{0i}|i \in I))^*$ (open cones $\times(C_i^*|i \in I)$ and $(\times(C_i|i \in I))^*$) by the natural cone isomorphism.

The group of isomorphisms of a closed (open) cone C_0 (C) is denoted by $\mathrm{Aut}(C_0)$ $(\mathrm{Aut}(C))$. The group $\mathrm{Aut}(C_0) = \mathrm{Aut}(C)$ is a Lie subgroup of the Lie group $GL(V)$, the general linear group of the enveloping vector space V. The mapping $f \mapsto {}^t f^{-1}$ of $\mathrm{Aut}(C)$ into $\mathrm{Aut}(C^*)$ is a group isomorphism. Thus $\mathrm{Aut}(C)$ and $\mathrm{Aut}(C^*)$ are isomorphic Lie groups. Note that $\mathrm{Aut}(C^*) = {}^t\mathrm{Aut}(C)$.

Let $G(C) \equiv G \subseteq \mathrm{Aut}(C)$ denote the connected component in $\mathrm{Aut}(C)$ of the identity mapping Id_C of C. When $C = \times(C_i|i \in I)$ we have $\mathrm{Aut}(C) \supseteq \times(\mathrm{Aut}(C_i)|i \in I) \supseteq \times(G_i|i \in I) = G$, where G_i denotes the connected component of the identity in $\mathrm{Aut}(C_i)$. The group isomorphism between $\mathrm{Aut}(C)$ and $\mathrm{Aut}(C^*)$ defines by restriction the group isomorphism

$$\begin{array}{rcl} G(C) & \leftrightarrow & G(C^*) \\ g & \mapsto & {}^t g^{-1}. \end{array} \tag{2.3}$$

Note that $G(C^*) = {}^tG(C)$. Consider the action

$$\begin{array}{rcl} G \times C & \to & C \\ (g,p) & \mapsto & g(p) =: gp \end{array} \tag{2.4}$$

of G on C. The mapping (2.4) is differentiable, and induces an action of G on the set of all measures μ on C, denoted by $g\mu$, $g \in G$.

Proposition 2.1. *The isotropy subgroups (or stabilizers) $G_p := \{g \in G | gp = p\} \subseteq G$, $p \in C$, for the action (2.4), are compact.*

Proof. From the first half of Proposition 1.1.8 in Faraut and Korányi (1994) it follows that the extended action

$$\begin{array}{rcl} \mathrm{Aut}(C) \times C & \to & C \\ (f,p) & \mapsto & f(p) =: fp \end{array} \tag{2.5}$$

has compact isotropy subgroups. Since $G \subseteq \mathrm{Aut}(C)$ is a closed subgroup it follows that the isotropy subgroups of G are also compact. See also Vinberg (1963), Proposition 12. □

Definition 2.3. A cone C is called *homogeneous* if the action (2.4) is transitive.

Remark 2.3. The action (2.4) is transitive if and only if the extended action (2.5) is transitive, cf. Faraut and Korányi (1994), page 5. □

Remark 2.4. A cone is said to be *self-dual* if there exists a *Euclidean* isomorphism between C and C^*, i.e., a bijective linear mapping $f : V \to V^*$ with the properties $f(C) = C^*$ and $f(x)(x) > 0$ for all $x \in V$ with $x \neq 0$. Note that C and C^* can be isomorphic cones without C (and C^*) being self-dual. A homogeneous self-dual cone is called a *symmetric* cone. The study of the special class of symmetric cones is the subject of Faraut and Korányi (1994). There are only five different types of indecomposable symmetric cones, cf. (x) in our Introduction. □

Proposition 2.2. *Let C be a homogeneous cone. Then the action (2.4) is a proper action, i.e., the mapping*

$$\begin{array}{rcl} G \times C & \to & C \times C \\ (g,p) & \mapsto & (gp,p) \end{array} \tag{2.6}$$

is proper[||].

Proof. See Bourbaki (1963), Appendice I, Lemme 2[**], and Bourbaki (1971), Ch. III, §4, Corollaire to Proposition 5. □

Remark 2.5. By the same argument, the action (2.5) is proper when C is a homogeneous cone. Also, a cone C is homogeneous if and only if C^* is homogeneous.

□

Let C be a homogeneous cone with enveloping vector space V and let G be the connected component of the identity in $\mathrm{Aut}(C)$. Let $\chi : G \to \mathbb{R}_+$ be a *multiplier* on the group G, i.e., χ is continuous, $\chi(\mathrm{Id}_C) = 1$, and $\chi(g_1 g_2) = \chi(g_1)\chi(g_2)$ for all $g_1, g_2 \in G$. Since G is a Lie group, χ is also differentiable. Since the action (2.4) is transitive and proper (cf. Proposition 2.2) there exists for each multiplier χ one and (up to multiplication by a positive constant) only one relatively invariant measure ν^χ on the open subset $C \subseteq V$ with multiplier χ under the action (2.4), cf. Bourbaki (1963), Ch. 7, §2, Theorem 3. That is,

$$g^{-1}\nu^\chi = \chi(g)\nu^\chi \tag{2.7}$$

for all $g \in G$. The integral

$$\phi^\chi(p^*) := \int_C \exp\{-p^*(p)\} d\nu^\chi(p) \tag{2.8}$$

is well-defined for all $p^* \in C^*$ (in fact for all $x^* \in V^*$) and, by (2.7),

$$\phi^\chi({}^t g^{-1} p^*) = \chi(g)\phi^\chi(p^*), \tag{2.9}$$

$g \in G$, $p^* \in C^*$. Since C^* is homogeneous it follows from (2.3) that convergence of the integral (2.8) does not depend on $p^* \in C^*$ but only on χ. Thus either $\phi^\chi(p^*) < \infty$ for all $p^* \in C^*$ or $\phi^\chi(p^*) = \infty$ for all $p^* \in C^*$. If $\chi = |\det(\cdot)|$, i.e. $\chi(g) = |\det(g)|$, $g \in G$, where $\det(g)$ denotes the determinant of the linear mapping $g \in GL(V)$, then $\nu^{|\det(\cdot)|}$ is the restriction of a Lebesgue measure to the open set C and (2.8) converges for all $p^* \in C^*$, cf. for example Vinberg (1963), §2.

From Bourbaki (1963), Ch. 7, §2, Proposition 7 a), it follows that for any multiplier $\chi : G \to \mathbb{R}_+$ there exists a positive continuous function $n^\chi : C \to \mathbb{R}_+$ with the property

$$n^\chi(gp) = \chi(g)n^\chi(p), \tag{2.10}$$

[||]A mapping is proper if it is continous and the inverse image of any compact set is compact

[**]This result shows that C is a homogeneous space under the action (2.4) (and the action (2.5)) and thus the cone C deserves the adjective "homogeneous".

$g \in G, p \in C$. Since the action (2.4) is transitive, such a function is unique up to multiplication by a positive constant. The measure $d\nu(p) = \frac{1}{n^{\chi}(p)} d\nu^{\chi}(p)$ is thus an invariant measure on C. The integral in (2.8) can then be re-expressed in terms of the invariant measure as

$$\phi^{\chi}(p^*) = \int_C \exp\{-p^*(p)\} n^{\chi}(p) d\nu(p). \tag{2.11}$$

From the standard theory of multivariate Laplace transforms, cf. for example Barndorff-Nielsen (1978), we can obtain the following collection of facts:

Proposition 2.3. *Let $\chi : G \to \mathbb{R}_+$ be a multiplier on the Lie group G and assume that* (2.8) *converges. Then*

$$C^* = \{x^* \in V^* |\ \phi^{\chi}(x^*) < \infty\}$$

and $\phi^{\chi}(p^*) \to \infty$ when $p^* \in C^*$ converges to a boundary point of C^*. The positive function $\phi^{\chi} : C^* \to \mathbb{R}_+$ is an analytic and strictly logarithmic-convex function on the open cone C^*, and its derivatives can be obtained by differentiation under the integral sign in (2.8). *In particular*

$$-D\log(\phi^{\chi}(p^*)) = \int_C p \frac{1}{\phi^{\chi}(p^*)} \exp\{-p^*(p)\} d\nu^{\chi}(p) \in C, \tag{2.12}$$

$p^* \in C^*$, under the natural identification of V and V^{**}, and

$$D^2\log(\phi^{\chi}(p^*))(x^*, x^*) = \int_C x^*(p + D\log\phi^{\chi}(p^*))^2 \frac{1}{\phi^{\chi}(p^*)} \exp\{-p^*(p)\} d\nu^{\chi}(p), \tag{2.13}$$

$p^* \in C^*$, $x^* \in V^*$, defines a positive definite form on V^*.

Remark 2.6. Note that $dE_{p^*}(p) := \phi^{\chi}(p^*)^{-1}\exp\{-p^*(p)\} d\nu^{\chi}(p)$ is a probability measure on V concentrated on the cone C, $-D\log(\phi^{\chi}(p^*))$ is its expectation, and $D^2\log(\phi^{\chi}(p^*))$ is its variance[††], $p^* \in C^*$. These depend of the choice of the relatively invariant measure ν^{χ} on C only through χ. The statistical model $(E_{p^*} \in \mathbf{P}(V) | p^* \in C^*)$ is thus a regular canonical full exponential family. Except for the parametrization by $p^* \in C^*$, this is the family of generalized Wishart distributions on C to be defined in Section 3. □

Remark 2.7. The extended action of G on the closed cone $C_0 (\neq \{0\})$, is no longer transitive and $C \subset C_0$ becomes one of the orbits. Let $O \subset C_0$ be another orbit and consider the restriction of the action to this orbit, i.e. the transitive action,

[††]The *variance* of a probability measure on V is a positive semidefinite form on V^*.

$$\begin{array}{rcl} G \times O & \to & O \\ (g, c) & \mapsto & gc := g(c) \end{array}$$

The above considerations could then be repeated with C replaced by O, but the action is *not* proper when $O \neq C$. In particular, a relatively invariant measure ν^χ on O only exists for certain multipliers χ, cf. Bourbaki (1963), Ch. 7, §2, Theorem 3. This leads to a definition of generalized *singular* Wishart distributions. In the classical cases where $C = \mathcal{P}(I, D)$, $D = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, the above definition of singular Wishart distribution yields the usual singular Wishart distributions corresponding to the discrete part of the Gyndikin set, cf. for example Letac and Massam (1998). Further consideration of the singular case is beyond the scope of the present basic introduction to generalized Wishart distributions. □

For $p^* \in C^*$, define

$$(p^*)^\chi := -D\log\phi^\chi(p^*) \tag{2.14}$$

and call $(p^*)^\chi$ the *χ-inverse* of p^*. Under the natural identification of V^{**} with V, $(p^*)^\chi \in C$. The mapping

$$\begin{array}{rcl} C^* & \to & C \\ p^* & \mapsto & (p^*)^\chi \end{array} \tag{2.15}$$

is a bijection and has the property

$$({}^t g p^*)^\chi = g^{-1}((p^*)^\chi), \tag{2.16}$$

$g \in G$, $p^* \in C^*$. The inverse mapping to (2.15) is denoted by

$$\begin{array}{rcl} C & \to & C^* \\ p & \mapsto & p^{-\chi} \end{array} \tag{2.17}$$

and has the property

$$(g^{-1}p)^{-\chi} = {}^t g p^{-\chi}, \tag{2.18}$$

$g \in G$ and $p \in C$. It follows from (2.18) that the positive number $\alpha := p^{-\chi}(p)$ is independent of $p \in C$. The value of α is given below in (3.14).

Proposition 2.4. *If* (2.8) *converges, then it also converges when χ is replaced by any multiplier of the form χ^λ, $\lambda \geq 1$.*

Proof. This follows from Remark 5.5 below. □

Remark 2.8. Conversely, let $\phi^\chi : C \to \mathbb{R}_+$ satisfy (2.9) and suppose that (2.15) (defined by (2.14)) is a bijection. Then it also a bijection when χ is replaced with χ^λ, $\lambda > 0$. □

Proposition 2.5. *If* (2.8) *converges for two multipliers* χ_1 *and* χ_2, *then it also converges for* $\chi = \chi_1^{\lambda_1}\chi_2^{\lambda_2}$, *where* $\lambda_i > 0$, $i = 1, 2$, *and* $\lambda_1 + \lambda_2 = 1$.

Proof. Consider the alterate form of (2.8) given in (2.11). Then

$$\infty > \phi^{\chi_1}(\lambda_1^{-1}p^*)^{\lambda_1}\phi^{\chi_2}(\lambda_2^{-1}p^*)^{\lambda_2}$$

$$= \left(\int_C \exp\{-\frac{1}{\lambda_1}p^*(p)\}n^{\chi_1}(p)d\nu(p)\right)^{\lambda_1}\left(\int_C \exp\{-\frac{1}{\lambda_2}p^*(p)\}n^{\chi_2}(p)d\nu(p)\right)^{\lambda_2}$$

$$\geq \int_C (\exp\{-\frac{1}{\lambda_1}p^*(p)\}n^{\chi_1}(p))^{\lambda_1}(\exp\{-\frac{1}{\lambda_2}p^*(p)\}n^{\chi_2}(p))^{\lambda_2}d\nu(p)$$

$$= \int_C \exp\{-p^*(p)\}n^\chi(p)d\nu(p) = \phi^\chi(p^*),$$

where the inequality follows from Hölder's inequality. □

Corollary 2.1. *If* (2.8) *converges for two multipliers* χ_1 *and* χ_2, *then it also converges for* $\chi = \chi_1\chi_2$.

Proof. The result follows from Proposition 2.4 and Proposition 2.5 and the relation $\chi = (\chi_1^{\frac{1}{2}}\chi_2^{\frac{1}{2}})^2$ □

Remark 2.9. Because of the isomorphism $g \leftrightarrow {}^tg^{-1}$ between the groups $G(C)$ and $G(C^*)$ and the bijection between C and C^* given by (2.15) and (2.17), several natural questions arise. For example, what is the connection, if any, via the isomorphism $G(C)$ and $G(C^*)$, between the set of multipliers on $G \equiv G(C)$ that make (2.8) convergent and the set of multipliers on $G(C^*)$ that make the "dual" integral to (2.8) convergent? This question and other related questions are resolved and important for the further development of this theory, but are beyond the scope of the present introductory paper. □

3. Wishart distributions on a homogeneous cone

Let C be a homogeneous cone with enveloping vector space V, G the connected component of the identity mapping in Aut(C), $\chi : G \to \mathbb{R}_+$ a multiplier on the Lie group G, and ν^χ the relatively invariant measure on C with multiplier χ wrt the action (2.4). (The measure ν^χ is unique up to multiplication with a positive constant.) Assume that the integral (2.8) converges and set

$$n^{\chi}(\sigma) := \int_C \exp\{-\sigma^{-\chi}(s)\} d\nu^{\chi}(s), \tag{3.1}$$

$\sigma \in C$. Then the positive function $n^{\chi} : C \to \mathbb{R}_+$ has the property

$$n^{\chi}(g\sigma) = \chi(g) n^{\chi}(\sigma), \tag{3.2}$$

$g \in G$, $\sigma \in C$, cf. (2.8) and (2.18). The property (3.2) determines the positive function n^{χ} on C uniquely up to multiplication by a positive constant. Thus the notation in (3.2) is consistent with the notation in (2.10).

The probability measure $W_{\sigma,\chi}$ on C defined by

$$dW_{\sigma,\chi}(s) = \frac{1}{n^{\chi}(\sigma)} \exp\{-\sigma^{-\chi}(s)\} d\nu^{\chi}(s) \tag{3.3}$$

is called the *(generalized) Wishart distribution* on C (or C_0) with *parameter* $\sigma \in C$ and *multiplier* χ. Note that the definition of the Wishart distribution $W_{\sigma,\chi}$ only depends on ν^{χ} through χ[‡‡]. From the standard properties of the Laplace transform it also follows that the parametrization of the Wishart distributions on C by the pair (σ, χ) is one-to-one, i.e., $W_{\sigma_1,\chi_1} = W_{\sigma_2,\chi_2}$ implies that $(\sigma_1, \chi_1) = (\sigma_2, \chi_2)$.

The measure $d\nu(s) := n^{\chi}(s)^{-1} d\nu^{\chi}(s)$ is invariant under the action (2.4) and independent of the choice of the relatively invariant measure ν^{χ}. The Wishart distribution in (3.3) can thus also be represented as

$$dW_{\sigma,\chi}(s) = \frac{n^{\chi}(s)}{n^{\chi}(\sigma)} \exp\{-\sigma^{-\chi}(s)\} d\nu(s). \tag{3.4}$$

Since $n^{\chi}(s)/n^{\chi}(\sigma)$ only depends on ν^{χ} through χ and $W_{\sigma,\chi}$ is a probability measure, the invariant measure ν in (3.4) is unique.

When $\chi(g) = |\det(g)|$, $g \in G$, the positive function $n^{|\det(\cdot)|}$ defined by (2.10) is denoted simply by n (and is unique up to multiplication by a positive constant). The measure $d\lambda(s) := n(s) d\nu(s)$ is then relatively invariant with multiplier $g \mapsto |\det(g)|$ under the action (2.4) and is thus the restriction of a Lebesgue measure on V to the open subset $C \subseteq V$. The Wishart distribution in (3.3) can thus also be be represented as

$$dW_{\sigma,\chi}(s) = \frac{n^{\chi}(s) n(s)^{-1}}{n^{\chi}(\sigma)} \exp\{-\sigma^{-\chi}(s)\} d\lambda(s). \tag{3.5}$$

Finally, if $V \supseteq C$ is a Euclidean space, i.e., is equipped with an inner product, then the Lebesgue measure λ can be chosen such that the unit ball in V has the standard measure $\pi^{\frac{n}{2}}/\Gamma(\frac{\dim(V)}{2} + 1)$, where $\dim(V)$ is the

[‡‡]I.e., any other relatively invariant measure with multiplier χ will yield the same Wishart distribution.

dimension of V and Γ denotes the gamma function. In this case $d\lambda(s)$ is simply denoted by ds.

Remark 3.1. It is well known that the interior of the closed cone $C_0 = \mathcal{P}(I, \mathbb{R})_0$ of positive semidefinite $I \times I$ matrices is the homogeneous cone $C = \mathcal{P}(I, \mathbb{R})$ of positive definite $I \times I$ matrices. The classical Wishart distributions on $\mathcal{P}(I, \mathbb{R})$ have the form (1.3). Thus in this case $n^{\chi}(\Sigma) = \det(\Sigma)^{\lambda}$ and $n(\Sigma) = \det(\Sigma)^{(I+1)/2}$, $\Sigma \in \mathcal{P}(I, \mathbb{R})$ (up to multiplication by positive constants). A detailed description of this classical case can be found in Example 6.2 below. It is standard in multivariate statistical analysis to call the positive function $\det(\Sigma)$, $\Sigma \in \mathcal{P}(I, \mathbb{R})$ the *generalized variance.* Nevertheless, in the present work on the generalized Wishart distribution we will call $n^{\chi}(\sigma)$ the *(theoretical) generalized variance.* When $s \in C$ follows the Wishart distribution $W_{\sigma,\chi}$ we will call $n^{\chi}(s)$ the *(empirical) generalized variance* wrt the Wishart distribution $W_{\sigma,\chi}$. □

The mapping $x \mapsto \chi(x \cdot \mathrm{Id}_C)$ of $\mathbb{R}_+$ into itself is a multiplier on $\mathbb{R}_+$. Thus it has the form $x \mapsto x^{\alpha}$ for some $\alpha \in \mathbb{R}$, i.e.,

$$\chi(x \cdot \mathrm{Id}_C) = x^{\alpha}, \tag{3.6}$$

$x \in \mathbb{R}_+$.

Proposition 3.1. *Let $W_{\sigma,\chi}$ be a Wishart distribution on C. Then $\sigma^{-\chi}(W_{\sigma,\chi})$ is the gamma distribution on $\mathbb{R}_+$ with shape parameter α given by* (3.6) *and scale 1 , i.e., the distribution P given by*

$$dP(x) := \Gamma(\alpha)^{-1} x^{\alpha-1} \exp\{-x\} dx. \tag{3.7}$$

Proof. The density of the Wishart distribution $W_{\sigma,\chi}$ on C with respect to an invariant measure ν on C is given by

$$\frac{n^{\chi}(s)}{n^{\chi}(\sigma)} \cdot \exp\{-\sigma^{-\chi}(s)\} = \frac{(\sigma^{-\chi}(s))^{\alpha}}{n^{\chi}(\sigma)} \exp\{-\sigma^{-\chi}(s)\} \cdot \frac{n^{\chi}(s)}{(\sigma^{-\chi}(s))^{\alpha}}.$$

The first factor depends on $s \in C$ only through $\sigma^{-\chi}(s) \in \mathbb{R}_+$ and the second factor is invariant under the action of $\mathbb{R}_+$ on C given by ordinary scalar multiplication. The result then readily follows from Andersson *et al.* (1983), Lemma 3, page 397, with X, Y, t, G, ν, ν_0 replaced by C, $\mathbb{R}_+$, $\sigma^{-\chi}$, $\mathbb{R}_+$, ν, and $\frac{1}{x}dx$, respectively *. □

*In the notation of the referenced Lemma the Wishart distribution takes the form $f(t(x))r(\pi(x))d\nu(x)$, where f and r are positive functions on Y and X/G, respectively. The distribution of (t, π) is then $f(y)r(z)d\nu_0(y)d\kappa(z)$, $(y, z) \in Y \times X/G$. Thus the distribution of t is proportional to $f(y)d\nu_0(y)$.

Let $(C_i|i \in I)$ be a finite family of homogeneous cones indexed by the finite set I and let G_i be the connected component of the identity in $\mathrm{Aut}(C_i)$. Then $C := \times(C_i|i \in I)$ is again a homogeneous cone, and $G := \times(G_i|i \in I)$ is the connected component of the identity of $\mathrm{Aut}(C)$. Every multiplier χ on G has the form

$$\chi(g) = \prod(\chi_i(g_i)|i \in I), \tag{3.8}$$

$g \equiv (g_i|i \in I) \in G$, where χ_i, $i \in I$, are multipliers on G_i uniquely determined by χ through (3.8). Conversely, if χ_i is a multiplier on G_i, $i \in I$, then χ defined by (3.8) is a multiplier on G, denoted by $\chi \equiv \otimes(\chi_i|i \in I)$. Any relatively invariant measure ν^χ on C with multiplier χ thus has the form $\nu^\chi = \otimes(\nu^{\chi_i}|i \in I)$, where ν^{χ_i} is a relatively invariant measure on C_i, with multiplier χ_i, $i \in I$ and $\chi = \otimes(\chi_i|i \in I)$. Conversely if ν^{χ_i} is a relatively invariant measure on C_i with multiplier χ_i then the product measure on C is relatively invariant with multiplier $\chi = \otimes(\chi_i|i \in I)$. The integral (2.8) over C converges for $\chi = \otimes(\chi_i|i \in I)$ if and only if for all $i \in I$ the integral over C_i converges for χ_i. In this case we have $(p^*)^\chi = ((p_i^*)^{\chi_i}|i \in I)$, $p^* \equiv (p_i^*|i \in I) \in C^* \equiv \times(C_i^*|i \in I)$, cf. (2.2) and the definition (2.14). Thus we have that

$$\sigma^{-\chi} = (\sigma_i^{-\chi_i}|i \in I), \tag{3.9}$$

$\sigma \in C$. Note that $\sigma^{-\chi}(s) = \sum(\sigma_i^{-\chi_i}(s_i)|i \in I)$, $s \equiv (s_i|i \in I) \in C$, cf. (2.2). The following two propositions now follow directly from the above considerations.

Proposition 3.2 (Product of Wishart distributions). *Let W_{σ_i,χ_i} be Wishart distribution on the cones C_i, $i \in I$, respectively. Then*

$$\otimes(W_{\sigma_i,\chi_i}|i \in I) = W_{\sigma,\chi}, \tag{3.10}$$

where $W_{\sigma,\chi}$ is the Wishart distribution on $C = \times(C_i|i \in I)$ with multiplier $\chi := \otimes(\chi_i|i \in I)$ and parameter $\sigma := (\sigma_i|i \in I) \in C$.

Conversely:

Proposition 3.3 (Decomposition of a Wishart distribution). *Let $W_{\sigma,\chi}$ be a Wishart distribution on $C = \times(C_i|i \in I)$. Then*

$$W_{\sigma,\chi} = \otimes(W_{\sigma_i,\chi_i}|i \in I), \tag{3.11}$$

where $\sigma_i \in C_i$ and χ_i are determined by $\sigma =: (\sigma_i|i \in I)$ and $\chi = \otimes(\chi_i|i \in I)$.

Let C be a homogeneneous cone, $(P_i|i \in I)$ a family of probability measures on C, and $P := \otimes(P_i|i \in I)$ the product probability measure on C^I. The transformation of P by the *addition mapping* $(s_i|i \in I) \mapsto \sum(s_i|i \in I) =: s_\bullet$ of C^I onto C is usually called the *convolution* of the family $(P_i|i \in I)$, denoted by $*(P_i|i \in I)$. If $(\chi_i|i \in I)$ is a family of multipliers on G, the function $\chi : G \to \mathbb{R}_+$ defined by $\chi(g) = \prod(\chi_i(g)|i \in I)$, $g \in G$, is also a multiplier on G, usually called the *product multiplier*, and is denoted by $\chi = \prod(\chi_i|i \in I)$. If (3.1) converges for χ_i, $i \in I$, then (3.1) also converges for $\chi = \prod(\chi_i|i \in I)$, cf. Corollary 2.1.

Proposition 3.4 (Convolution of Wishart distributions). *Suppose that* (3.1) *converges for* χ_i, $i \in I$. *Let* $\sigma \in C$ *and define* $\sigma_i := (\sigma^{-\chi})^{\chi_i} \in C$, $i \in I$, *where* $\chi = \prod(\chi_i|i \in I)$. *Then the convolution*

$$*(W_{\sigma_i,\chi_i}|i \in I) = W_{\sigma,\chi}. \tag{3.12}$$

Proof. The density of the Wishart distribution $\otimes(W_{\sigma,\chi_i}|i \in I)$ on C^I with respect to $\nu^{\otimes I}$, where ν is an invariant measure on C, is given by

$$\begin{aligned}&\prod\left(\frac{n^{\chi_i}(s_i)}{n^{\chi_i}(\sigma)}\Big|i \in I\right) \cdot \exp\{-\sigma^{-\chi}\left(\sum(s_i|i \in I)\right)\}\\ &\quad = \frac{n^{\chi}(s_\bullet)}{n^{\chi}(\sigma)}\exp\{-\sigma^{-\chi}(s_\bullet)\} \cdot \frac{\prod(n^{\chi_i}(s_i)|i \in I)}{n^{\chi}(s_\bullet)}.\end{aligned}$$

The first factor depends on $(s_i|i \in I) \in C^I$ only through $s_\bullet \in C$ and the second factor is invariant under the action of G on C^I given by $g(s_i|i \in I) := (gs_i|i \in I)$, $g \in G$, $(s_i|i \in I) \in C^I$. As in the proof of Proposition 3.1 the result readily follows from Andersson *et al.* (1983), Lemma 3, page 397, with X, Y, t, G, ν, ν_0 replaced by C^I, C, the addition mapping, G, $\nu^{\otimes I}$ and ν, respectively. □

From (2.12) and the definition of $\sigma^{-\chi}$, cf. (2.17), it follows that the expectation $\mathbb{E}(W_{\sigma,\chi})$ is given by

$$\mathbb{E}(W_{\sigma,\chi}) := \int_C s\, dW_{\sigma,\chi}(s) = \sigma, \tag{3.13}$$

$\sigma \in C$. From (3.13) and Proposition 3.1, $\sigma^{-\chi}(\sigma)$ is the expectation of the gamma distribution (3.7), i.e.,

$$\sigma^{-\chi}(\sigma) = \alpha, \tag{3.14}$$

where α is given by (3.6). It follows by Vinberg (1963), page 347, (or from (5.6) together with (5.2)) that the integral (2.8) converges for $\chi = |\det(\cdot)|$. For this multiplier we thus obtain $\alpha = \dim(V)$.

A *generalized Wishart model* is the statistical model, with observation space C and parameter space C, given by

$$(W_{\sigma,\chi} \in \mathbf{P}(C)|\sigma \in C). \tag{3.15}$$

It follows from (2.18), (3.2), and the representation (3.4) that the Wishart model is invariant under the action (2.4) on the observation space and parameter space, i.e.,

$$gW_{\sigma,\chi} = W_{g\sigma,\chi}, \tag{3.16}$$

$g \in G$, $\sigma \in C$.

From the standard theory of exponential families it follows that the ML estimator $\hat{\sigma}$ of $\sigma \in C$ in the model (3.15) exists for *all* observations $s \in C$ and is given by

$$\hat{\sigma}(s) = s. \tag{3.17}$$

The ML estimator $\hat{\sigma}$ for $\sigma \in C$ is thus complete, sufficient, unbiased (cf. (3.13)), and follows the distribution $W_{\sigma,\chi}$. The variance $\mathbb{V}(W_{\sigma,\chi})$ of $W_{\sigma,\chi}$ is calculated explicitly in Andersson *et al.* (2001), and it is in general a rational function of σ.

In Remark 5.7 we shall describe the distribution of the *empirical generalized variance* $n^{\chi}(\hat{\sigma})$, i.e., the ML estimator of the theoretical generalized variance $n^{\chi}(\sigma)$, $\sigma \in C$, in terms of all its r^{th} moments, $r \geq 0$.

The maximum of the likelihood function based on the representation (3.3) is

$$\frac{1}{n^{\chi}(\hat{\sigma})}\exp\{-\frac{\alpha}{2}\}, \tag{3.18}$$

where α is given by (3.6) or (3.14).

Remark 3.2 (Bayesian inference). Consider the model (3.15) and the integral

$$m^{\delta}_{\chi}(\rho) := \int_C \exp\{-\sigma^{-\chi}(\rho)\}d\nu^{\delta}(\sigma),$$

where δ is a multiplier on G, ν^{δ} is a relatively invariant measure on C with multiplier δ under the action (2.4), and $\rho \in C$. As for the integral (2.8), it follows that $m^{\delta}(g\rho) = \delta(g)m^{\delta}(\rho)$, cf. (2.18), and convergence of the integral does not depend on $\rho \in C$, but only on δ. Suppose that the integral converges. The probability measure $IW^{\chi}_{\rho,\delta}$ on C defined by

$$dIW^{\chi}_{\rho,\delta}(\sigma) := \frac{1}{m^{\delta}(\rho)}\exp\{-\sigma^{-\chi}(\rho)\}d\nu^{\delta}(\sigma)$$

is called *the χ-inverse Wishart distribution* with *parameter* $\rho \in C$ and multiplier δ. It is routine to establish that $\sigma \in C$ follows the distribution $IW^{\chi}_{\rho,\delta}$ if and only if $\sigma^{-\chi} \in C^*$ follows the Wishart distribution on C^* with multiplier $\gamma : G(C^*) \to \mathbb{R}_+$ given by $\gamma({}^t g) := \delta(g^{-1})$, $g \in G$, and expectation parameter ρ^{γ}, cf. (2.14), (2.18), and (2.3).

If we now let $IW^{\chi}_{\rho,\delta}$ be a prior distribution for the statistical model (3.15), then the joint distribution of observable and parameter $(s,\sigma) \in C \times C$ is

$$\frac{1}{n^{\chi}(\sigma)m^{\delta}(\rho)}\exp\{-\sigma^{-\chi}(s+\rho)\}d\nu^{\chi}(s)d\nu^{\delta}(\sigma).$$

Since $d\nu^{\delta/\chi} := n^{\chi}(\sigma)^{-1}d\nu^{\delta}(\sigma)$ is relatively invariant with multiplier δ/χ, the marginal distribution of $s \in C$ becomes

$$\frac{m^{\delta/\chi}(s+\rho)}{m^{\delta}(\rho)}d\nu^{\chi}(s),$$

from which it follows that the posterior distribution is $IW^{\chi}_{s+\rho,\delta/\chi}$, the χ-inverse Wishart distribution with parameter $s+\rho$ and multiplier δ/χ. Then the expectation $\mathbb{E}(IW^{\chi}_{s+\rho,\delta/\chi})$ is a Bayes estimator of $\sigma \in C$. The existence of this expectation depends on δ/χ only. □

4. Vinberg algebras

Vinberg algebras, described below with some alternative formulations, were introduced by Vinberg (1960, 1962, 1963, 1965) with the aim of describing all homogeneous cones and homogeneous domains together with their automorphism groups.

Let I be a partially ordered finite set (poset), i.e., a set equipped with a relation denoted $\leq$ which is

(i) reflexive; $\forall i \in I: \ i \leq i$

(ii) antisymmetric; $\forall i,j \in I: \ i \leq j$ and $j \leq i$ implies that $i = j$

(iii) transitive; $\forall i,j,k \in I: \ i \leq j$ and $j \leq k$ implies that $i \leq k$.

We write $i < j$ if $i \leq j$ and $i \neq j$, $i,j \in I$. For all pairs $(i,j) \in I \times I$ with $j < i$, let E_{ij} be a finite-dimensional vector space over $\mathbb{R}$ with $n_{ij} := \dim(E_{ij}) > 0$. Set

$$\mathcal{A}_{ij} := \begin{cases} \mathbb{R} & \text{for } i = j \\ E_{ij} & \text{for } j < i \\ E_{ji} & \text{for } i < j \\ \{0\} & \text{otherwise} \end{cases} \tag{4.1}$$

and $\mathcal{A} := \times(\mathcal{A}_{ij} | (i,j) \in I \times I)$. Thus $\mathcal{A}$ is a vector space of $I \times I$ "matrices" with entries from vector spaces according to (4.1)†

Let $f_{ij} : E_{ij} \to E_{ij}$, $i > j$, be involutional linear mappings, i.e., $f_{ij}^{-1} = f_{ij}$. These induce an *involutional* mapping $(A \mapsto A^*)$ of $\mathcal{A}$ given as follows: $A^* := (a'_{ij} | (i,j) \in I \times I)$, where

†A purer approach is to eliminate the "{0} otherwise" statement in (4.1). Then $\mathcal{A}$ becomes $I \times I$ "matrices with voids".

$$a'_{ij} := \begin{cases} a_{ii} & \text{for } i = j \\ f_{ij}(a_{ji}) =: a^*_{ji} & \text{for } j < i \\ f_{ji}(a_{ji}) =: a^*_{ji} & \text{for } i < j \\ 0 & \text{otherwise,} \end{cases} \tag{4.2}$$

$A \equiv (a_{ij}|(i,j) \in I \times I) \equiv (a_{ij}) \in \mathcal{A}$. Note thus that

(I) $\forall A \in \mathcal{A}: \ (A^*)^* = A$,

and that[‡] $(\mathcal{A}_{ij})^* = \mathcal{A}_{ji}$, $(i,j) \in I \times I$.

We may now define the following subspaces of $\mathcal{A}$:

the *upper triangular matrices* $\mathcal{T}_u := \{A \equiv (a_{ij}) \in \mathcal{A}|\forall i,j \in I : i \not\leq j \Rightarrow a_{ij} = 0\}$;

the *lower triangular matrices* $\mathcal{T}_l := \{A \equiv (a_{ij}) \in \mathcal{A}|\forall i,j \in I : j \not\leq i \Rightarrow a_{ij} = 0\}$;

the *hermitian matrices* $\mathcal{H} := \{A \in \mathcal{A}|A^* = A\}$; and

the *skew-hermitian matrices* $\mathcal{K} := \{A \in \mathcal{A}|A^* = -A\}$.

Note that $\mathcal{A} = \mathcal{H}+\mathcal{K}$ and that the sum is direct, i.e., $\mathcal{H}\cap\mathcal{K} = \{0\}$. The sets of upper and lower triangular matrices with *positive* diagonal elements are denoted by $\mathcal{T}_u^+$ and $\mathcal{T}_l^+$, respectively. The sets of upper and lower triangular matrices with all diagonal elements equal to 1 are denoted by $\mathcal{T}_u^1$ and $\mathcal{T}_l^1$, respectively. The set of diagonal matrices and the set of diagonal matrices with positive entries on the diagonal are denoted by $\mathcal{D}$ and $\mathcal{D}^+$, respectively.

Define $n_{i\cdot} := \sum(n_{i\nu}|\nu < i)$, $n_{\cdot i} := \sum(n_{\nu i}|i < \nu)$, $n_i := 1 + \frac{1}{2}(n_{i\cdot} + n_{\cdot i})$, $i \in I$, and $n_{\cdot} := \sum(n_i|i \in I) = \dim(\mathcal{H})$. Also define $n_{ji} := n_{ij}$ for $i > j$ and $n_{ij} = 0$ when $i \not< j$ and $j \not< i$.

Let $\mathrm{tr}(A) := \sum(a_{ii}|i \in I)$ be the *trace* of $A \equiv (a_{ij}|(i,j) \in I \times I)$ and note that

(TR) $\forall A \in \mathcal{A}: \ \mathrm{tr}(A) = \mathrm{tr}(A^*)$.

We now equip the vector space $\mathcal{A}$ with a "multiplication", denoted by $(A,B) \mapsto AB$, $A, B \in \mathcal{A}$ For this purpose we need to define bilinear mappings $\mathcal{A}_{ij} \times \mathcal{A}_{jk} \to \mathcal{A}_{ik}$, denoted by $(a_{ij}, b_{jk}) \mapsto a_{ij}b_{jk}$, $i,j,k \in I$, and then define $AB = C \equiv (c_{ij}|(i,j) \in I \times I)$ by $c_{ij} := \sum(a_{i\nu}b_{\nu j}|\nu \in I)$ (similar to ordinary matrix multiplication). In all cases other than $i > j > k$, $i > j < k$, $i < j < k$, and $i < j > k$, the desired bilinear mappings are obtained from the structure of $\mathcal{A}$: if $i \not\geq j$, $i \not\leq j$, $j \not\geq k$ *or* $j \not\leq k$ then $a_{ij}b_{jk} \equiv 0$; if $i = j$ and/or $j = k$ then $a_{ij}b_{jk}$ is simply multiplication by $\mathbb{R}$. The multiplication is required to satisfy the following properties:

(TP) $\forall A \in \mathcal{A} : A \neq 0 \Rightarrow \mathrm{tr}(AA^*) > 0$

[‡]As usual $\mathcal{A}_{ij}$ also denotes the subspace of $\mathcal{A}$ given by the natural embedding of $\mathcal{A}_{ij}$ into $\mathcal{A}$.

(IP) $\forall A, B \in \mathcal{A} : (AB)^* = B^* A^*$

(TC) $\forall A, B \in \mathcal{A} : \mathrm{tr}(AB) = \mathrm{tr}(BA)$

(TA) $\forall A, B, C \in \mathcal{A} : \mathrm{tr}(A(BC)) = \mathrm{tr}((AB)C)$

(AT) $\forall S, T, U \in \mathcal{T}_l : (ST)U = S(TU)$

(LT) $\forall T, U \in \mathcal{T}_l : T(UU^*) = (TU)U^*$.

The abbreviations (TP), (IP), (TC), (TA), (AT), and (LT) stand for trace positivity, involution anti-commutativity, trace commutativity, trace associativity, associativity for (lower) triangular matrices, and linearity for (lower) triangular matrices (cf. Remark 4.1 below), respectively. An algebra $\mathcal{A}$ with the above structure and properties is called a *Vinberg algebra*.

The defining properties of a Vinberg algebra determine interplay between and properties of the bilinear mappings $E_{ij} \times E_{jk} \to E_{ik}$, $k < j < i$. It follows for example that (AT) implies the following associativity condition:

(A) $\forall a_{ij} \in E_{ij},\ b_{jk} \in E_{jk},\ c_{kl} \in E_{kl} : (a_{ij}b_{jk})c_{kl} = a_{ij}(b_{jk}c_{kl}), \quad l < k < j < i$.

The conditions (TP) and (TC) define inner products $(\cdot,\cdot)_{ij}$ on E_{ij}, $i > j$ by $\|a_{ij}\|_{ij}^2 := a_{ij}f_{ij}(a_{ij})$, $a_{ij} \in E_{ij}$. Thus instead of specifying the bilinear form $(a_{ij}, b_{ji}) \mapsto a_{ij}b_{ji}$ on E_{ij} one can equivalently specify an inner product $(\cdot,\cdot)_{ij}$ on E_{ij}, $i > j$. It can then be established that the following two conditions also must hold:

(B) $\forall a_{ij} \in E_{ij},\ b_{jk} \in E_{jk} : \|a_{ij}b_{jk}\|_{ik}^2 = \|a_{ij}\|_{ij}^2 \|b_{jk}\|_{jk}^2, \quad k < j < i,$

and

(C) If $a_{ik} \in E_{ik}$, $b_{jk} \in E_{jk}$ with $k < j$, $k < i$ and $(a_{ik}, c_{ij}b_{jk})_{ik} = 0$ for all $c_{ij} \in E_{ij}$, then $(d_{li}a_{ik}, c_{lj}b_{jk})_{lk} = 0$ for all $l \in I$ with $i < l$, $j < l$, and all $d_{li} \in E_{li}$, and all $c_{lj} \in E_{lj}$.

On the other hand, if $(E_{ij} | i, j \in I \text{ with } i > j)$ is a family of Euclidean spaces together with a family of bilinear mappings $E_{ij} \times E_{jk} \to E_{ik}$, $k < j < i$, satisfying (A), (B), and (C), then this algebraic structure, called an *N-algebra* (N for nilpotent) by Vinberg, induces a Vinberg algebra, i.e., all other bilinear mappings needed to define the Vinberg algebra exist and are uniquely determined by the conditions defining an N-algebra. Thus it is enough to define $a_{ij}b_{jk}$ for $k < j < i$, such that the properties (A), (B), and (C) are satisfied. The details can be found in Vinberg (1965), Chapter III, §7.

Definition 4.1. The subset $\mathcal{P} := \{TT^* \in \mathcal{A} | T \in \mathcal{T}_l^+\} \subset \mathcal{H} \subset \mathcal{A}$ will be called the set of *positive-definite matrices* in $\mathcal{A}$.

Proposition 4.1. *a) The positive-definite matrices $\mathcal{P}$ form a convex homogeneous cone with enveloping vector space $\mathcal{H}$.*

b) The mapping

$$\begin{array}{rcl} \mathcal{T}_l^+ & \to & \mathcal{P} \\ T & \mapsto & TT^* \end{array} \tag{4.3}$$

is a diffeomorphism.

c) The mapping $\pi(T) : \mathcal{P} \to \mathcal{P}$ given by $\pi(T)(S) := (TU)(U^*T^*)$, where $S = UU^* \in \mathcal{P}$, $U \in \mathcal{T}_l^+$, is well-defined, an automorphism of the open cone $\mathcal{P}$, and $\pi(T) \in G \equiv G(\mathcal{P})$, $T \in \mathcal{T}_l^+$.

d) The action

$$\begin{array}{rcl} \mathcal{T}_l^+ \times \mathcal{P} & \to & \mathcal{P} \\ (T, S) & \mapsto & \pi(T)(S) \end{array} \tag{4.4}$$

is well-defined, transitive and free.

Proof. These results follow from Vinberg (1965), Chapter III, §2. □

Remark 4.1. Condition (LT) in the definition of a Vinberg algebra is used in c) to show that $\pi(T)$ is linear. The results c) and d) in Proposition 4.1 show that $\pi : \mathcal{T}_l^+ \to G(\mathcal{P})$ is an injective group homomorphism, in particular we have that $\pi(\mathcal{T}_l^+)$ is a subgroup of $G(\mathcal{P})$. □

Remark 4.2. The definition of $\mathcal{P}$ could be changed to the equivalent definition

$$\mathcal{P} := \{TDT^* \in \mathcal{A} | T \in \mathcal{T}_l^1,\ D \in \mathcal{D}^+\}. \tag{4.5}$$

Note that $(TD)T^* = T(DT^*)$ for any triangular matrix T and any $D \in \mathcal{D}$. Thus it is meaningful to write TDT^* in this case. The diffeomorphism (4.3) can then replaced by the diffeomorphism

$$\begin{array}{rcl} \mathcal{T}_l^1 \times \mathcal{D}^+ & \leftrightarrow & \mathcal{P} \\ (T, D) & \mapsto & TDT^* \\ (T(S), D(S)) & \leftarrow\!\shortmid & S \equiv (s_{ij} | (i,j) \in I \times I). \end{array} \tag{4.6}$$

This version is more important for statistical applications. Note that the connection between the two unique decompositions $S = TT^*$, $T \in \mathcal{T}_l^+$, and $S = T_1 D T_1^*$, $T_1 \in \mathcal{T}_l^1$, $D \in \mathcal{D}^+$, is given by $T = T_1\sqrt{D}$, where $\sqrt{D} := \mathrm{diag}(\sqrt{d_i} | i \in I) \in \mathcal{D}^+$ with $D \equiv \mathrm{diag}(d_i | i \in I) \in \mathcal{D}^+$.

The positive functions $S \mapsto s_{[i]\bullet} := D(S)_{ii} \in \mathbb{R}_+$ are *rational* functions in $S \in \mathcal{P}$. We will furthermore define

$$S_{[i>\bullet} \equiv T(S)_{[i>} := (T(S)_{i\nu}|\nu < i) \in \times(E_{i\nu}|\nu < i) =: E_{[i>}, \; S \in \mathcal{P}, \; i \in I.$$

The one to one correspondence (4.6) can then be restated as

$$\begin{array}{rcl} \times(E_{[i>} \times \mathbb{R}_+|i \in I) & \leftrightarrow & \mathcal{P} \\ ((S_{[i>\bullet}, s_{[i]\bullet})|i \in I) & \leftarrow\!\mid & S \equiv (s_{ij}|i,j \in I), \end{array} \tag{4.7}$$

where $S_{[i>\bullet} \equiv (t_{i\nu}|\nu < i)$ and $s_{[i]\bullet} \equiv d_i$, $i \in I$, can be found by step-wise solving the equation system

$$\sum(d_\nu \|t_{j\nu}\|^2|\nu < j) + d_j \;\; = \;\; s_{jj}, \; j \in I \tag{4.8}$$

$$\sum(t_{i\nu}t^*_{j\nu}d_\nu|\nu < j) + t_{ij}d_j \;\; = \;\; s_{ij}, \;\; i,j \in I \text{ with } j < i. \tag{4.9}$$

as follows: Choose a never-decreasing listing $j_1, j_2, \cdots, j_I$ of the elements in I. Solve (4.8) for $j = j_k$, then solve (4.9) for all $i > j \equiv j_k$, $k = 1, \cdots, I$.

This solution can of course be carried out *formally* even for $S \in \mathcal{H} \supset \mathcal{P}$ to define the *real-valued* rational function *expressions*, also denoted by $s_{[i]\bullet} \equiv d_i$, $i \in I$ on $\mathcal{H}$. The cone $\mathcal{P} \subset \mathcal{H}$ can then be characterized as $\mathcal{P} = \{S \in \mathcal{H}|s_{[i]\bullet} > 0, i \in I\}$, where $s_{[i]\bullet} > 0$ means the rational function $s_{[i]\bullet}$ is defined and is positive at the point $S \in \mathcal{H}$. This characterization is directly useful for our purpose. Moreover, one can avoid the problem of undefined points in the rational functions by only considering polynomials: If the above equation system is solved in the new variables $p_j := \prod(d_\nu^{n(\nu,j)}|\nu < j)d_j$ and $c_{ij} := t_{ij}\prod(d_\nu^{n(\nu,j)}|\nu < j)d_j$, where $n(\nu, j)$ is the number of elements in the set $\{\mu \in I|\nu \le \mu < j\}$, then p_j, $j \in I$, are polynomials in $S \in \mathcal{H}$ and the characterization of $S \in \mathcal{P}$ now becomes $p_j > 0$, $j \in I$; cf. Vinberg (1963), §3. □

Let N be a finite set with a partial ordering also denoted by $\le$, and let $\mathcal{B} \equiv \times(\mathcal{B}_{\mu\nu}|(\mu,\nu) \in N \times N)$ be a Vinberg algebra over N, where

$$\mathcal{B}_{\mu\nu} := \begin{cases} \mathbb{R} & \text{for } \; \mu = \nu \\ F_{\mu\nu} & \text{for } \; \nu < \mu \\ F_{\nu\mu} & \text{for } \; \mu < \nu \\ \{0\} & \text{otherwise.} \end{cases} \tag{4.10}$$

and $g_{\mu\nu} : F_{\mu\nu} \to F_{\mu\nu}$ $\mu > \nu$ denote the involutional mappings for $\mathcal{B}$.

Definition 4.2. An *isomorphism* between two Vinberg algebras $\mathcal{A}$ and $\mathcal{B}$ is a family of mappings $(\iota, (\phi_{ij}|(i,j) \in I \times I, \; j < i))$, where $\iota : I \to N$ is an isomorphism between the two posets and $\phi_{ij} : E_{ij} \to F_{\iota(i)\iota(j)}$, $j < i$, are vector space isomorphisms such that

$$\phi_{ik}(a_{ij}b_{jk}) = \phi_{ij}(a_{ij})\phi_{jk}(b_{jk}),\ k < j < i$$

and

$$\phi_{ij} \circ f_{ij} = g_{\iota(i)\iota(j)} \circ \phi_{ij},\ j < i.$$

Two Vinberg algebras are called *isomorphic* if there exists an isomorphism between them.

Remark 4.3. Note that an isomorphism between $\mathcal{A}$ and $\mathcal{B}$ uniquely defines a bijective mapping $\phi : \mathcal{A} \to \mathcal{B}$ with the properties $\phi(A^*) = \phi(A)^*$, $\mathrm{tr}_{\mathcal{B}}(\phi(A)) = \mathrm{tr}_{\mathcal{A}}(A)$, $\phi(AB) = \phi(A)\phi(B)$, $\phi(\mathcal{T}_l(\mathcal{A})) = \mathcal{T}_l(\mathcal{B})$, $\phi(\mathcal{T}_u(\mathcal{A})) = \mathcal{T}_u(\mathcal{B})$, $\phi(\mathcal{H}(\mathcal{A})) = \mathcal{H}(\mathcal{B})$, $\phi(\mathcal{K}(\mathcal{A})) = \mathcal{K}(\mathcal{B})$, $\phi(\mathcal{T}_l^+(\mathcal{A})) = \mathcal{T}_l^+(\mathcal{B})$, $\phi(\mathcal{T}_u^+(\mathcal{A})) = \mathcal{T}_u^+(\mathcal{B})$, $\phi(\mathcal{T}_l^1(\mathcal{A})) = \mathcal{T}_l^1(\mathcal{B})$, $\phi(\mathcal{T}_u^1(\mathcal{A})) = \mathcal{T}_u^1(\mathcal{B})$, and $\phi(\mathcal{P}(\mathcal{A})) = \mathcal{P}(\mathcal{B})$, where $\mathcal{P}(\mathcal{A})$ and $\mathcal{P}(\mathcal{B})$ denote the positive definite matrices in $\mathcal{A}$ and $\mathcal{B}$, respectively. Thus ϕ constitutes an isomorphism between Vinberg algebras in the standard sense. □

Proposition 4.2. *The Vinberg algebras $\mathcal{A}$ and $\mathcal{B}$ are isomorphic if and only if the homogeneous convex cones $\mathcal{P}(\mathcal{A})$ and $\mathcal{P}(\mathcal{B})$ are isomorphic. For every convex homogeneous cone C there exists a unique (up to an isomorphism of Vinberg algebras) Vinberg algebra $\mathcal{A}$ such that C and $\mathcal{P}(\mathcal{A})$ are isomorphic cones.*

Proof. The two results can be found in Vinberg (1965), Chapter III, §2 and §9. □

Let $\mathcal{A} =: \mathcal{A}(\le)$ be the Vinberg algebra as defined in the beginning of this section. Denote $\mathcal{P}(\le) := \mathcal{P}$, $\mathcal{T}_l^+(\le) := \mathcal{T}_l^+$, and $\mathcal{T}_u^+(\le) := \mathcal{T}_u^+$. We shall now construct a Vinberg algebra $\mathcal{A}^{\mathrm{op}}$ corresponding to the dual cone, i.e., the homogeneous cone $\mathcal{P}(\mathcal{A}^{\mathrm{op}})$ of the positive definite matrices in $\mathcal{A}^{\mathrm{op}}$ is a representation of (isomorphic to) the dual cone of $\mathcal{P}(\le)$.

Let $\le^{\mathrm{op}}$ be the opposite ordering on the index set I, i.e., $i \le^{\mathrm{op}} j \Leftrightarrow j \le i$. In the definition of $\mathcal{A}$ replace $\le$ with the opposite ordering $\le^{\mathrm{op}}$ and define the vector space $\mathcal{A}^{\mathrm{op}} := \times(\mathcal{A}_{ij}^{\mathrm{op}} | (i,j) \in I \times I)$, where

$$\mathcal{A}_{ij}^{\mathrm{op}} := \begin{cases} \mathbb{R} & \text{for } i = j \\ E_{ji} & \text{for } j <^{\mathrm{op}} i \\ E_{ij} & \text{for } i <^{\mathrm{op}} j \\ \{0\} & \text{otherwise.} \end{cases} \tag{4.11}$$

It follows that the Vinberg algebra $\mathcal{A}^{\mathrm{op}} =: \mathcal{A}(\le^{\mathrm{op}})$ differs from the Vinberg algebra $\mathcal{A} \equiv \mathcal{A}(\le)$ only in the ordering of the index set I. In particular, $\mathcal{T}_l^+(\le^{\mathrm{op}}) = \mathcal{T}_u^+(\le)$. Thus

$$\mathcal{P}(\leq^{\text{op}}) = \{T^*T \in \mathcal{A} | T \in \mathcal{T}_l^+(\leq)\} \tag{4.12}$$

is the homogeneous cone of positive matrices in $\mathcal{A}(\leq^{\text{op}})$. From Vinberg (1965), Chapter III, §6, it then follows that $\mathcal{P}(\leq^{\text{op}})$ is the dual cone of $\mathcal{P}(\leq)$. The enveloping vector space for both cones is of course $\mathcal{H}$. The inner product $(A, B) \mapsto \text{tr}(AB)$ on $\mathcal{H}$ identifies $\mathcal{H}$ with its dual $\mathcal{H}^*$, i.e.,

$$\begin{array}{rcl} \mathcal{H} & \leftrightarrow & \mathcal{H}^* \\ A & \mapsto & (B \mapsto \text{tr}(AB)), \end{array} \tag{4.13}$$

and this isomorphism identifies $\mathcal{P}(\leq^{\text{op}})$ with the dual cone $\mathcal{P}(\leq)^*$ of $\mathcal{P}(\leq)$.

Let $(I_\nu | \nu \in N)$ be a finite family of finite posets and let $V_\nu \subseteq I_\nu \times I_\nu$ be the partial ordering relation on I_ν, $\nu \in N$. Then constructing the disjoint union $I = \dot{\cup}(I_\nu | \nu \in N)$ also gives a poset in a natural way under the partial ordering relation $V = \dot{\cup}(V_\nu | \nu \in N) \subseteq I \times I$. If $\mathcal{A}_\nu \equiv \times(\mathcal{A}_{i_\nu j_\nu} | (i_\nu, j_\nu) \in I_\nu \times I_\nu)$ is a Vinberg algebra over the poset I_ν then $\mathcal{A} := \times(\mathcal{A}_{ij} | (i,j) \in I)$, where $\mathcal{A}_{ij} = \{0\}$ when $i \in I_\nu$ and $j \in I_{\nu'}$ with $\nu \neq \nu'$, $\nu, \nu' \in N$, is a Vinberg algebra called the *product* of the family $(\mathcal{A}_\nu | \nu \in N)$ of Vinberg algebras indexed by N, and we will write $\mathcal{A} = \times(\mathcal{A}_\nu | \nu \in N)$. It then follows that

$$\mathcal{P} = \times(\mathcal{P}_\nu | \nu \in N), \tag{4.14}$$

where $\mathcal{P}$ is the homogeneous cone of positive definite matrices in $\mathcal{A}$ and $\mathcal{P}_\nu$ is the homogeneous cone of positive definite matrices in $\mathcal{A}_\nu$, $\nu \in N$.

Remark 4.4. The *center* of a Vinberg algebra $\mathcal{A}$ is important for a more specific description of the automorphism group $G(\mathcal{P})$, cf. Remark 5.3, and assists the understanding of a Vinberg algebra. We shall only sketch the construction of the center and leave the details to the reader. Because of our reformulation of Vinberg algebras, the description of the center is different but equivalent to Vinberg (1962), §4.

Consider the relation on I defined by: $i \sim j$ iff i is comparable to j, i.e., $i < j$ or $j < i$, *and* $n_{is} = n_{js}$ for all $s \in I$ with $s \neq i, j$. This is clearly an equivalence relation on I. Now define a refined equivalence relation given by $i \approx j$ iff $i \sim j$ and $i \sim s \sim j$ for all $s \in I$ between i and j, i.e., $i < s < j$ or $j < s < i$. The equivalence relation $\approx$ induces the decomposition $I = \dot{\cup}(I_\kappa | \kappa \in K)$ of I into its equivalence classes. Note that $n_{ii'} = n_{\kappa\kappa'}$ for all $i \in I_\kappa$ and $i' \in I_{\kappa'}$, $\kappa, \kappa' \in K$. This decomposition of I induces the decomposition $\mathcal{A} = \times(\mathcal{A}_{\kappa\kappa'} | \kappa, \kappa' \in K)$, where $\mathcal{A}_{\kappa,\kappa'} := \times(\mathcal{A}_{ii'} | i \in I_\kappa, i' \in I_{\kappa'})$. Set $\mathcal{A}_{\kappa\kappa} =: \mathcal{A}_\kappa$ and note that $\mathcal{A}_\kappa$ itself is a Vinberg algebra based on the *totally* ordered index set I_κ. The Vinberg algebra $\mathcal{A}_c := \times(\mathcal{A}_\kappa | \kappa \in K)$ is called the *center* of $\mathcal{A}$. Note that $\mathcal{A}_c$ corresponds to the block diagonal matrices in $\mathcal{A}$ with respect to the above decomposition of I.

The cone of positive definite matrices $\mathcal{P}_c$ corresponding to $\mathcal{A}_c$ is called the *center cone* of $\mathcal{P}$ and is a subcone of $\mathcal{P}$, since $\mathcal{P}_c = \{S \in \mathcal{P} | S \in \mathcal{A}_c\}$. Note that the center is the same for the opposite ordering $\leq^{\text{op}}$. In particular, it not difficult to establish that $\mathcal{P}_c = \mathcal{P}_c^*$ is a symmetric cone and that $\mathcal{P}_c = \times(P_\kappa | \kappa \in K)$, where each $\mathcal{P}_\kappa$ is the cone of positive definite matrices in the Vinberg algebra $\mathcal{A}_\kappa$.

We shall use the notation $|S_{[\kappa]\bullet}| := \prod(s_{[i]\bullet} | i \in I_\kappa)$, $\kappa \in K$, $S \in \mathcal{P}$, in accordance with standard notation in multivariate statistical analysis, cf. Remark 4.2. □

5. Representations of Wishart distributions

The aim of this section is to use the representation of a homogeneous cone C as the cone $\mathcal{P}$ of positive definite matrices in a Vinberg algebra $\mathcal{A}$ (over the poset I), cf. Proposition 4.2, to obtain an explicit and general expression for the Wishart distribution in (3.3).

Since the homogeneous cone $\mathcal{P}$ is represented in terms of matrices we shall in the present section use uppercase letters for the observations $S \equiv (s_{ij} | (i,j) \in I \times I) \in \mathcal{P}$ and for the parameters $\Sigma \equiv (\sigma_{ij} | (i,j) \in I \times I) \in \mathcal{P}$ in accordance with standard notation in the literature on the classical Wishart distribution.

The first step is to establish a one-to-one correspondence between the set of multipliers $\chi : G \to \mathbb{R}_+$ on the connected component G in $\mathrm{Aut}(\mathcal{P})$ and a subspace of $\mathbb{R}^I$, cf. Remark 5.3 below. In other words a multiplier χ on G can be described by its corresponding point $(\lambda_i | i \in I) \in \mathbb{R}^I$.

The restriction of χ to the (lower) triangular group in G, i.e., $\chi \circ \pi : \mathcal{T}_l^+ \to \mathbb{R}_+$, cf. Remark 4.1, is a multiplier on $\mathcal{T}_l^+$.

Lemma 5.1. *Any multiplier $\tau : \mathcal{T}_l^+ \to \mathbb{R}_+$ has the form*

$$\tau(T) = \prod(t_{ii}^{2\lambda_i} | i \in I), \quad T = (t_{ij} | (i,j) \in I \times I) \in \mathcal{T}_l^+, \tag{5.1}$$

where $(\lambda_i | i \in I) \in \mathbb{R}^I$. Any τ of the above form is a multiplier on $\mathcal{T}_l^+$.

Proof. Since τ is differentiable, its derivative $D\tau$ (at $\mathrm{Id}_{\mathcal{P}}$) is a Lie algebra homomorphism $D\tau : \mathcal{T}_l \to \mathbb{R}$, where $\mathcal{T}_l$ and $\mathbb{R}$ are the Lie algebras corresponding to the Lie groups $\mathcal{T}_l^+$ and $\mathbb{R}_+$, respectively. The problem is thus to find all Lie algebra homomorphisms $f : \mathcal{T}_l \to \mathbb{R}$. The subspace $N \subset \mathcal{T}_l$ of upper triangular matrices with zeros on the diagonal is an ideal with $[\mathcal{T}_l, N] = N$. Thus $f(N) = f([\mathcal{T}_l, N]) = 0$. Any Lie algebra homomorphism $f : \mathcal{T}_l \to \mathbb{R}$ thus has the form $f(T) = \sum(2\lambda_i t_{ii} | i \in I)$, $T \equiv (t_{ij} | (i,j) \in I \times I) \in \mathcal{T}_l$. The corresponding multiplier then has the form (5.1). The converse statement in the lemma is trivial. □

Remark 5.1. From the definition of α in (3.6) one obtains that

$$\alpha = \sum (\lambda_i | i \in I) =: \lambda_\bullet.$$

□

Remark 5.2. For the special case $\chi(g) = |\det(g)|$, $g \in G$, we obtain directly

$$\chi(\pi(T)) = \prod (t_{ii}^{2n_i} | i \in I), \tag{5.2}$$

cf. also Vinberg (1963), Chapter III, §4 (32). This multiplier thus corresponds to $(\lambda_i | i \in I) = (n_i | i \in I)$. □

Remark 5.3. Note that in general *not* every multiplier τ on $\mathcal{T}_l^+$ is a restriction of a multiplier χ on G, i.e., $\tau = \chi \circ \pi$. The necessary and sufficient condition on $(\lambda_i | i \in I)$ to ensure that τ is a restriction of a multiplier χ on G is as follows: Let $I = \dot{\cup}(I_\kappa | \kappa \in K)$ be the decomposition of I established in Remark 4.4. Note that this decomposition only depends on the partial ordering of I and the integer constants n_{ij}, $i, j \in I$. A family $(\lambda_i | i \in I)$ then corresponds to a multiplier on G if and only if there exists a family $(\lambda_\kappa | \kappa \in K) \in \mathbb{R}^K$ such that $\lambda_i = \lambda_\kappa$, $i \in I_\kappa$, $\kappa \in K$, i.e., the λ_i's are constant on the subsets I_κ, $\kappa \in K$.

This result can be established using the almost complete description of the Lie algebra $\mathfrak{g}$ of G established in Vinberg (1965), §4, in particular Theorem 4. The Lie algebra $\mathfrak{g}$ decomposes as $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_c \oplus \mathcal{T}_{lB}^0$, where each component is a Lie subalgebra of $\mathfrak{g}$ corresponding to a subgroup of G, $\mathfrak{g}_0$ is a compact Lie algebra, $\mathfrak{g}_c$ is naturally isomorphic to the Lie algebra of the center cone $\mathcal{P}_c$, and $\mathcal{T}_{lB}^0$ is the Lie algebra of lower block triangular matrices in $\mathcal{A}$ with respect to the decomposition $I = \dot{\cup}(I_\kappa | \kappa \in K)$ and with zero blocks in the block diagonal. Furthermore, $\mathcal{T}_{lB}^0$ is an ideal in $\mathfrak{g}$ with $[\mathfrak{g}_c, \mathcal{T}_{lB}^0] = \mathcal{T}_{lB}^0$, and $\mathfrak{g}_0$ and $\mathfrak{g}_c$ commute. From these facts it follows that the derivative $D\chi$ of any multiplier $\chi : G \to \mathbb{R}_+$ has the property that $D\chi(\mathfrak{g}_0 \oplus \mathcal{T}_{lB}^0) = 0$ and thus is uniquely determined by its restriction, also denoted $D\chi$, to $\mathfrak{g}_c$. Thus it follows that there is a one-to-one correspondence between multipliers $\chi : G \to \mathbb{R}_+$ and Lie algebra homomorphisms $D\chi : \mathfrak{g}_c \to \mathbb{R}$. The Lie algebra $\mathfrak{g}_c$ decomposes into a direct sum $\mathfrak{g}_c = \oplus(\mathfrak{g}_\kappa | \kappa \in K)$ of commuting Lie algebras $\mathfrak{g}_\kappa$, $\kappa \in K$, each corresponding to the connected component G_κ for the indecomposable symmetric cone $\mathcal{P}_\kappa$, cf. Remark 4.4. Thus it is enough to describe the Lie algebra homomorphisms from $\mathfrak{g}_\kappa$ into $\mathbb{R}$, $\kappa \in K$. In Examples 6.2, 6.3, 6.4, 6.6, and 6.8 these homomorphisms will be described and it will be established that each homomorphism naturally corresponds to a real number λ_κ. □

The second step is to find a necessary and sufficient condition, (5.6) below, on the multiplier $\chi : G \to \mathbb{R}_+$ (or equivalently on the corresponding family $(\lambda_i | i \in I)$) to ensure that the integral (2.8) converges.

Let $\chi \circ \pi : \mathcal{T}_l^+ \to \mathbb{R}_+$ be induced from a multiplier χ on G. (In fact the derivation below will work for any multiplier on $\mathcal{T}_l^+$, *but* the distribution in (5.15) will *not* in general be a Wishart distribution on $\mathcal{P}$ in the invariant sense of Section 3.)

The measure ν^χ is also the unique (up to multiplication by a positive constant) relatively invariant measure with multiplier $\chi \circ \pi$ *under the action* (4.4). Thus it is enough to establish conditions for convergence of the integral (2.8) for $p^* = 1_I \in \mathcal{P}^*$, i.e.,

$$\phi^\chi(1_I) = \int_{\mathcal{P}} \exp\{-\mathrm{tr}(S)\} d\nu^\chi(S). \tag{5.3}$$

Since the diffeomorphism (4.3) commutes with the left multiplication on the group $\mathcal{T}_l^+$ and with the action (4.4), the integral (5.3) is equal to

$$\int_{\mathcal{T}_l^+} \exp\{-\mathrm{tr}(TT^*)\} d\nu_{\mathcal{T}}^\chi(T), \tag{5.4}$$

where $\nu_{\mathcal{T}}^\chi$ denotes the corresponding left relatively invariant measure on $\mathcal{T}_l^+$ with multiplier $\chi \circ \pi$.

We now want to express $\nu_{\mathcal{T}}^\chi$ in terms of a density with respect to the restriction of the Lebesgue measure $dT \equiv \prod(dt_{ij} | (i,j) \in I \times I, j \leq i)$ on $\mathcal{T}_l$ to the open subset $\mathcal{T}_l^+ \subset \mathcal{T}_l$. Note that dt_{ij} denotes the standard Lebesgue measure on E_{ij}, $i,j \in I$, $j < i$, and dt_{ii} denotes the standard Lebesgue measure on $\mathbb{R}_+$. The Lebesgue measure is relatively invariant under left multiplication in $\mathcal{T}_l^+$ with multiplier $T \mapsto |\det(L_T)|$, $T \equiv (t_{ij} | (i,j) \in I \times I) \in \mathcal{T}_l^+$, where $L_T : \mathcal{T}_l^+ \to \mathcal{T}_l^+$ is given by $L_T(U) := TU$. Thus an invariant measure $\nu_{\mathcal{T}}$ under left multiplication is $d\nu_{\mathcal{T}}(T) = \det(L_T)^{-1} dT$, and the relatively invariant measure $\nu_{\mathcal{T}}^\chi$ is thus proportional to $\chi(\pi(T))\det(L_T)^{-1}dT$. Now choose $\nu_{\mathcal{T}}^\chi$ (and thus also ν^χ) such that $d\nu_{\mathcal{T}}^\chi(T) = \chi(\pi(T))\det(L_T)^{-1}dT$.

It is easy to obtain that $\det(L_T) = \prod(t_{ii}^{1+n_{i\cdot}} | i \in I)$. Furthermore $\mathrm{tr}(TT^*) = \sum(t_{ii}^2 | i \in I) + \sum(\sum(t_{ij}t_{ij}^* | j \in I, j < i) | i \in I)$. The integral (5.4) can now be rewritten as

$$\int_{\mathcal{T}_l^+} \prod(t_{ii}^{2\lambda_i - n_{i\cdot} - 1} \exp\{-t_{ii}^2\} | i \in I) \prod(\exp\{-t_{ij}t_{ij}^*\} | (i,j) \in I \times I, j < i) dT. \tag{5.5}$$

This integral converges if and only if

$$\lambda_i > \frac{n_{i\cdot}}{2}, \quad i \in I. \tag{5.6}$$

Remark 5.4. The condition (5.6) can equivalently be expressed as

$$\lambda_\kappa > \frac{n_{i\cdot}}{2}, \quad i \in I_\kappa,\ \kappa \in K, \tag{5.7}$$

cf. Remark 5.3.

□

Remark 5.5. The condition (5.6) establishes a proof of Proposition 2.4 since the family $(\alpha\lambda_i|i \in I) \in \mathbb{R}^I$ corresponds to χ^α when $(\lambda_i|i \in I) \in \mathbb{R}^I$ corresponds to χ.

□

Note that when the multiplier is $g \mapsto |\det(g)|$, $g \in G$, we have $\lambda_i = n_i \equiv 1 + \frac{1}{2}(n_{i\cdot} + n_{\cdot i}) > \frac{1}{2}n_{i\cdot}$, $i \in I$, cf. Remark 5.1, so (5.5) converges in this case.

The third and final step is to derive an explicit form, (5.15) below, of the Wishart distributions on $\mathcal{P}$.

When χ satisfies (5.6), the value of (5.5) is

$$2^{-I}(\sqrt{\pi})^{n_{\cdot}-I}\prod(\Gamma(\lambda_i - \frac{n_{i\cdot}}{2})|i \in I). \tag{5.8}$$

The following is therefore, cf. (5.5) and (5.8), the density of a probability measure on $\mathcal{T}_l^+$ with respect to the standard Lebesgue measure:

$$\begin{aligned} 2^I(\sqrt{\pi})^{I-n_{\cdot}}\prod(\Gamma(\lambda_i - \frac{n_{i\cdot}}{2})^{-1}|i \in I)\prod(t_{ii}^{2\lambda_i - n_{i\cdot}-1}\exp\{-t_{ii}^2\}|i \in I) \\ \times\prod(\exp\{-t_{ij}t_{ij}^*\}|(i,j) \in I \times I, j < i). \end{aligned} \tag{5.9}$$

Suppose that the observable $T \equiv (t_{ij}|(i,j) \in I \times I)$ follows the distribution given by the density in (5.9). Since $(1_I^\chi)^{-\chi} = 1_I$ the distribution of $S = TT^*$ is the Wishart distribution $W_{1^\chi,\chi}$ on $\mathcal{P}$.

Remark 5.6. Thus all t_{ij}, $i,j \in I$ with $j \leq i$ are independent, t_{ii}^2 follows a gamma distribution with shape parameter $\lambda_i - \frac{n_{i\cdot}}{2}$, $i \in I$, and t_{ij} follows a n_{ij}-dimensional normal distribution on E_{ij} with expectation $0 \in E_{ij}$ and *precision* $t_{ij} \mapsto 2t_{ij}t_{ij}^*$ on E_{ij}, $i,j \in I$, $j < i$.

□

From Remark 5.6, it follows that the expectation of $S = TT^*$ is given by

$$\mathbb{E}(W_{1^\chi,\chi}) = \operatorname{diag}(\lambda_i\ |i \in I) \in \mathcal{P}. \tag{5.10}$$

Thus we obtain that

$$1^\chi = \operatorname{diag}(\lambda_i\ |i \in I) \quad \text{and} \quad 1_I = \operatorname{diag}(\lambda_i\ |i \in I)^{-\chi}. \tag{5.11}$$

Since $\chi(\pi(\operatorname{diag}(\sqrt{\lambda_i}|i \in I))) = \prod(\lambda_i^{\lambda_i}|i \in I)$ it follows from (5.11) that

$$n^\chi(1_I^\chi) = \prod(\lambda_i^{\lambda_i}|i \in I)n^\chi(1_I). \tag{5.12}$$

For $\Sigma \in \mathcal{P}$ we then obtain, cf. Remark 4.2, that

$$n^\chi(\Sigma) = \prod(\sigma_{[i]\bullet}^{\lambda_i}|i \in I)n^\chi(1_I) = \frac{\prod(\sigma_{[i]\bullet}^{\lambda_i}|i \in I)}{\prod(\lambda_i^{\lambda_i}|i \in I)}n^\chi(1_I^\chi). \tag{5.13}$$

The Jacobian of (4.3) is easily calculated to be $T \mapsto 2^I \prod(t_{ii}^{1+n_{\cdot i}} | i \in I)$, $T \equiv (t_{ij} | (i,j) \in I \times I) \in \mathcal{T}_l^+$. The measure $2^I \prod(t_{ii}^{1+n_{\cdot i}} | i \in I) dT$ is thus by (4.3) transformed into the standard Lebesgue measure on $\mathcal{P}$, denoted by dS. Thus, cf. Remark 4.2, (5.9) is transformed into the probability measure

$$dW_{1_I^\chi, \chi}(S) = (\sqrt{\pi})^{I - n_{\cdot\cdot}} \frac{\prod(s_{[i]\bullet}^{\lambda_i - n_i} | i \in I)}{\prod(\Gamma(\lambda_i - \frac{n_{i\cdot}}{2}) | i \in I)} \exp\{-\mathrm{tr}(S)\} dS \tag{5.14}$$

on $\mathcal{P}$. From (5.13) and (5.12) it then follows that (3.5) takes the form

$$dW_{\Sigma,\chi}(S) = (\sqrt{\pi})^{I - n_{\cdot\cdot}} \frac{\prod(\lambda_i^{\lambda_i} | i \in I) \prod(s_{[i]\bullet}^{\lambda_i - n_i} | i \in I) \exp\{-\mathrm{tr}(\Sigma^{-\chi} S)\}}{\prod(\Gamma(\lambda_i - \frac{n_{i\cdot}}{2}) | i \in I) \prod(\sigma_{[i]\bullet}^{\lambda_i} | i \in I)} dS. \tag{5.15}$$

Since $\prod(s_{[i]\bullet}^{-n_i} | i \in I) dS$ is an invariant measure on $\mathcal{P}$ it follows from (5.15) that (3.4) takes the form

$$dW_{\Sigma,\chi}(S) = \frac{\prod(s_{[i]\bullet}^{\lambda_i} | i \in I)}{\prod(\sigma_{[i]\bullet}^{\lambda_i} | i \in I)} \exp\{-\mathrm{tr}(\Sigma^{-\chi} S)\} d\nu(S), \tag{5.16}$$

where

$$d\nu(S) = (\sqrt{\pi})^{I - n_{\cdot\cdot}} \frac{\prod(\lambda_i^{\lambda_i} | i \in I)}{\prod(\Gamma(\lambda_i - \frac{n_{i\cdot}}{2}) | i \in I)} \prod(s_{[i]\bullet}^{-n_i} | i \in I) dS \tag{5.17}$$

is the unique invariant measure on $\mathcal{P}$.

Remark 5.7. When S follows the Wishart distribution $W_{\Sigma,\chi}$ in (5.15), the distribution of *the normalized empirical generalized variance* $n^\chi(S)/n^\chi(\Sigma)$ does not depend on $\Sigma \in \mathcal{P}$, by (3.2) and (3.16). It is thus enough to find the distribution when $\Sigma = 1_I^\chi$. From (5.13) and Remark 5.6 it follows that $n^\chi(S)/n^\chi(\Sigma)$ has the same distribution as $\prod((\frac{1}{\lambda_i} X_i)^{\lambda_i} | i \in I)$, where X_i, $i \in I$, are independent positive stochastic variables and where X_i follows a gamma distribution with shape parameter $\lambda_i - \frac{n_{i\cdot}}{2}$, $i \in I$. The distribution of the empirical generalized variance $n^\chi(S)$ therefore may be represented as $n^\chi(\Sigma) \prod(\frac{1}{2\lambda_i} \chi^2_{2\lambda_i - n_{i\cdot}} | i \in I)$, where $\beta\chi_f^2$ denotes a χ^2 distribution with scale β and f degrees of freedom, $\beta, f > 0$. The r^{th} moment of $n^\chi(S)$ is given by

$$E(n^\chi(S)^r / n^\chi(\Sigma)^r) = \prod(\lambda_i^{-r\lambda_i} \frac{\Gamma(\lambda_i - \frac{n_{i\cdot}}{2} + r)}{\Gamma(\lambda_i - \frac{n_{i\cdot}}{2})} | i \in I). \tag{5.18}$$

□

From (5.11), (2.16), and (2.18) it follows that

$$\Sigma^{-\chi} = (T^*)^{-1}\mathrm{diag}(\lambda_i|i \in I)T^{-1} = (T^*)^{-1}1_I^{-\chi}T^{-1}, \tag{5.19}$$

where $\Sigma = TT^* \in \mathcal{P}$, and

$$\Xi^{\chi} = T^{-1}\mathrm{diag}(\lambda_i|i \in I)(T^*)^{-1} = T^{-1}1_I^{\chi}(T^*)^{-1}, \tag{5.20}$$

where $\Xi = T^*T \in \mathcal{P}^*$, $T \in \mathcal{T}_l^+$.

We will also use the notations $\Sigma^{-(\lambda_\kappa|\kappa\in K)} := \Sigma^{-(\lambda_i|i\in I)} := \Sigma^{-\chi}$, cf. Remark 5.3. For $\lambda_i = 1$, $i \in I$, cf. Remark 2.8, we write $\Sigma^{-1} := \Sigma^{-(1|i\in I)}$, $\Sigma \in \mathcal{P}$.

6. Examples of Wishart distributions

Example 6.1. Let I be a one-point set $\{1\}$. Then $\mathcal{A} = \mathbb{R}$ and under standard multiplication $\mathbb{R}$ becomes a Vinberg algebra. In this example we shall use lower case letters for elements in $\mathcal{A}$. Note that $n_{1\cdot} = n_{\cdot 1} = 0$, $n_1 = 1$, $n_{\cdot} = 1$, $\mathcal{H} = \mathcal{K} = \mathcal{T}_l = \mathcal{T}_u = \mathbb{R}$, and $\mathcal{T}_l^+ = \mathcal{T}_u^+ = \mathcal{P} = \mathcal{P}^* = \mathbb{R}_+$. The function $s_{[1]\bullet} = s$, $s \in \mathbb{R}_+$, cf. Remark 4.2. Furthermore, the connected component G (in the automorphism group $\mathrm{Aut}(\mathcal{P})$) is isomorphic to the multiplicative group $\mathbb{R}_+$ according to $g \mapsto (s \mapsto gs)$, $g \in \mathbb{R}_+$, $s \in \mathcal{P} \equiv \mathbb{R}_+$. The injective, in fact bijective, group homomorphism $\pi : \mathcal{T}_l^+ \to G$ is thus given by $\pi(t) = t^2$, $t \in \mathcal{T}_l^+ \equiv \mathbb{R}_+$. Every multiplier χ on G has the form $\chi(g) = g^\lambda$ for some $\lambda \in \mathbb{R}$, so we can replace χ by λ whenever χ is used as an index or parameter. The condition (5.6) takes the form $\lambda > 0$. From (5.19) it follows that $\sigma^{-\chi} = \frac{\lambda}{\sigma}$ and the Wishart distribution (5.15) takes the form

$$dW_{\sigma,\lambda}(s) = \frac{\lambda^\lambda s^{\lambda-1}}{\Gamma(\lambda)\sigma^\lambda}\exp\{-\frac{\lambda}{\sigma}s\}ds. \tag{6.1}$$

The Wishart distribution $W_{\sigma,\lambda}$ on $\mathbb{R}_+$ is thus the gamma distribution with scale $\frac{\sigma}{\lambda}$ and shape parameter $\lambda > 0$. The r^{th} moment of the normalized generalized variance is given by, cf. (5.18), $\lambda^{-r\lambda}\Gamma(\lambda + r)/\Gamma(\lambda)$, $r \geq 0$.

Example 6.2. Let I be a finite set with any total ordering. Thus we can assume that $I = \{1, \cdots, I\}$ with the usual ordering[§]. Set $E_{ij} = \mathbb{R}$ and $f_{ij} = \mathrm{Id}_{\mathbb{R}}$, $j < i$. Then $n_{\cdot i} = I - i$, $n_{i\cdot} = i - 1$, $n_i = \frac{I+1}{2}$, and $n_{\cdot} = \frac{I(I+1)}{2}$ The vector space $\mathcal{A}$ is $\mathcal{M}(I, \mathbb{R})$, the vector space of all $I \times I$ matrices with the involution $A^* = A'$, $A \in \mathcal{M}(I, \mathbb{R})$. Note that the center $\mathcal{A}_c = \mathcal{A}$, cf. Remark 4.4. In particular K is a one-point set. The subspaces $\mathcal{T}_l$, $\mathcal{T}_u$, and $\mathcal{H}$, of lower triangular, upper triangular, and hermitian matrices become the usual lower triangular $I \times I$ matrices $\mathcal{T}_l(I, \mathbb{R})$, the usual upper triangular $I \times I$ matrices $\mathcal{T}_u(I, \mathbb{R})$, and the usual symmetric $I \times I$ matrices $\mathcal{S}(I, \mathbb{R})$, respectively. With the standard multiplication and inner product on

[§] Recall that I denotes the set I *and also* its cardinality.

$\mathbb{R}$ the vector space $\mathcal{M}(I,\mathbb{R})$ becomes a Vinberg algebra. The multiplication becomes the standard matrix multiplication of $I \times I$ real matrices. (The special case $I = 1$ is the previous Example 6.1.) The groups $\mathcal{T}_l^+$ and $\mathcal{T}_u^+$ are the usual groups $\mathcal{T}_l^+(I,\mathbb{R})$ and $\mathcal{T}_u^+(I,\mathbb{R})$ of lower and upper triangular $I \times I$ real matrices with positive diagonal entries, respectively. The homogeneous cone $\mathcal{P}$ in this Vinberg algebra is the usual cone of $I \times I$ positive-definite matrices $\mathcal{P}(I,\mathbb{R})$ and $\mathcal{P}^* = \mathcal{P} = \mathcal{P}_c$.

It is well known, cf. for example Vinberg (1965), §3, that the connected component¶ is $G = \{g_A | A \in \mathrm{GL}(I,\mathbb{R})\}$, where $\mathrm{GL}(I,\mathbb{R})$ is the general linear group, i.e., the group of non-singular elements in $\mathcal{M}(I,\mathbb{R})$, and $g_A := (S \mapsto ASA')$. The parametrization of G by $\mathrm{GL}(I,\mathbb{R})$ is not injective and $G = \mathrm{Aut}(\mathcal{P})$. The Lie algebra of G is $\mathfrak{g} = \mathfrak{g}_c = L \oplus \mathbb{R}$, where $L = \mathfrak{sl}(I,\mathbb{R})$ is the simple Lie algebra of $I \times I$ real matrices with trace zero, cf. for example Faraut and Korányi (1994) page 97. Every multiplier $\chi : G \to \mathbb{R}_+$ induces a Lie algebra homomorphism $D\chi : L \oplus \mathbb{R} \to \mathbb{R}$. Since $[L, L] = L$ and $[\mathbb{R},\mathbb{R}] = \{0\}$ it follows that every Lie algebra homomorphism $f : L \oplus \mathbb{R} \to \mathbb{R}$ has the form $f(l, x) = \alpha x$ for some $\alpha \in \mathbb{R}$. Therefore every $\chi : G \to \mathbb{R}_+$ has the form $\chi(g) = |\det(g)|^\alpha$. Thus $\lambda_i = \alpha\frac{I+1}{2} =: \lambda$ does not depend on $i \in I$ and λ can replace χ whenever χ is used as an index or a parameter. The I conditions (5.6) thus reduce to the single condition $\lambda > \frac{I-1}{2}$.

From (5.19) it follows that $\Sigma^{-\chi} = \lambda\Sigma^{-1}$, where Σ^{-1} is the usual inverse of an $I \times I$ real matrix. The Wishart distribution (5.15) takes the form (1.3), since $\det(S) = \prod(s_{[i]\bullet} | i \in I)$, $S \in \mathcal{P}$. This distribution is the usual Wishart distribution on $\mathcal{P}(I,\mathbb{R})$ with shape parameter $\lambda > \frac{I-1}{2}$ and expectation Σ. The r^{th} moment ($r \geq 0$) of the normalized generalized variance is then given by, cf. (5.18),

$$\lambda^{-Ir\lambda}\frac{\prod(\Gamma(\lambda - \frac{i-1}{2} + r) | i = 1, \cdots, I)}{\prod(\Gamma(\lambda - \frac{i-1}{2}) | i = 1, \cdots, I)}. \tag{6.2}$$

Example 6.3. As in Example 6.2 let $I = \{1, \cdots, I\}$ with $I > 1$. Set $E_{ij} = \mathbb{C}$ and $f_{ij}(z) = \bar{z}$, $z \in \mathbb{C}$, where $\bar{z}$ denotes the complex conjugate of z, $j < i$. Then $n_{\cdot i} = 2(I - i)$, $n_{i\cdot} = 2(i-1)$, $n_i = I$, and $n_\cdot = I^2$. The vector space $\mathcal{A}$ is now the real vector space $\mathcal{M}_\mathbb{R}(I,\mathbb{C})$ of $I \times I$ complex matrices with real-valued diagonal entries. The involution in $\mathcal{A}$ is given by $A^* = \overline{A'}$, $A \in \mathcal{A}$. Note that then $\mathcal{A}_c = \mathcal{A}$ and in particular that K is a one-point set. The subspaces $\mathcal{T}_l$, $\mathcal{T}_u$, and $\mathcal{H}$, of lower triangular, upper triangular, and Hermitian matrices become the lower triangular $I \times I$ complex matrices with real-valued diagonal entries, denoted by $\mathcal{T}_{l,\mathbb{R}}(I,\mathbb{C})$, the upper triangular $I \times I$ complex matrices with real-valued diagonal entries, denoted by $\mathcal{T}_{u,\mathbb{R}}(I,\mathbb{C})$, and the Hermitian $I \times I$ complex matrices denoted by $\mathcal{H}(I,\mathbb{C})$, respectively.

¶We subsume "of the identity in the automorphism group $\mathrm{Aut}(\mathcal{P})$".

With the standard multiplication and standard inner product $z \mapsto \|z\|^2$ on $\mathbb{C}$ the vector space $\mathcal{M}_{\mathbb{R}}(I, \mathbb{C})$ becomes a Vinberg algebra. The Vinberg multiplication becomes standard matrix multiplication of $I \times I$ complex matrices followed by taking the real part of diagonal elements. The groups $\mathcal{T}_l^+$ and $\mathcal{T}_u^+$ are the usual groups $\mathcal{T}_l^+(I, \mathbb{C})$ and $\mathcal{T}_u^+(I, \mathbb{C})$ of lower and upper triangular $I \times I$ complex matrices with positive diagonal elements, respectively. The homogeneous cone $\mathcal{P}$ in this Vinberg algebra is the usual cone of $I \times I$ positive definite (Hermitian) complex matrices $\mathcal{P}(I, \mathbb{C})$, and $\mathcal{P}^* = \mathcal{P} = \mathcal{P}_c$.

It is well known, cf. Vinberg (1965) §3, that the connected component is $G = \{g_A | A \in \mathrm{GL}(I, \mathbb{C})\}$, where $\mathrm{GL}(I, \mathbb{C})$ is the complex general linear group, and $g_A := (S \mapsto AS\overline{A'})$. The parametrization of G by $GL(I, \mathbb{C})$ is not injective and $G \subset \mathrm{Aut}(\mathcal{P})$. The Lie algebra of G is $\mathfrak{g} = \mathfrak{g}_c = L \oplus \mathbb{R}$, where $L = \mathfrak{sl}(I, \mathbb{C})$ is the simple Lie algebra of $I \times I$ complex matrices with trace zero, cf. for example Faraut and Korányi (1994), page 97. Similar to Example 6.2 we obtain that every multiplier $\chi : G \to \mathbb{R}_+$ has the form $\chi(g) = |\det(g)|^\alpha$, so $\lambda_i = \alpha I =: \lambda$ does not depend on $i \in I$ and λ can replace χ whenever χ is used as an index or a parameter. The I conditions (5.6) reduce to the single condition $\lambda > I - 1$.

From (5.19) it follows that $\Sigma^{-\chi} = \lambda\Sigma^{-1}$, where Σ^{-1} is the usual inverse of an $I \times I$ complex matrix, and the Wishart distribution (5.15) takes the form

$$dW_{\Sigma,\lambda}(S) = \frac{\lambda^{I\lambda}\det(S)^{\lambda-I}\exp\{-\lambda\mathrm{tr}(\Sigma^{-1}S)\}}{\pi^{\frac{I(I-1)}{2}}\prod(\Gamma(\lambda - i + 1)|i = 1, \cdots, I)\det(\Sigma)^{\lambda}}dS, \tag{6.3}$$

where $\det(S) = \prod(s_{[i]\bullet}|i \in I)$, $S \in \mathcal{P}(I, \mathbb{C})$, is the usual determinant of an $I \times I$ complex matrix. This distribution is the complex Wishart distribution on $\mathcal{P}(I, \mathbb{C})$ with shape parameter $\lambda > I - 1$ and expectation Σ. The r^{th} moment ($r \geq 0$) of the normalized generalized variance is then given by, cf. (5.18),

$$\lambda^{-Ir\lambda}\frac{\prod(\Gamma(\lambda - i + 1 + r)|i = 1, \cdots, I)}{\prod(\Gamma(\lambda - i + 1)|i = 1, \cdots, I)}. \tag{6.4}$$

Example 6.4. As in Example 6.2 let $I = \{1, \cdots, I\}$ with $I > 1$. Set $E_{ij} = \mathbb{H}$, the *quaternion* division algebra over $\mathbb{R}$, and let $f_{ij}(h) = \bar{h}$, $h \in \mathbb{H}$, where $\bar{h}$ denotes the quaternion conjugate of h, $j < i$. Then $n_{\cdot i} = 4(I - i)$, $n_{i\cdot} = 4(i-1)$, $n_i = 2I - 1$, and $n_{\cdot} = 2I^2 - I$. The vector space $\mathcal{A}$ is now the real vector space $\mathcal{M}_{\mathbb{R}}(I, \mathbb{H})$ of all $I \times I$ quaternion matrices with real-valued diagonal entries. The involution in $\mathcal{A}$ is given by $A^* = \overline{A'}$, $A \in \mathcal{A}$. Note that $\mathcal{A}_c = \mathcal{A}$ and in particular that K is a one-point set. The subspaces $\mathcal{T}_l$, $\mathcal{T}_u$, and $\mathcal{H}$ of lower triangular, upper triangular, and Hermitian matrices become the lower triangular $I \times I$ quaternion matrices with real-valued diagonal entries, denoted by $\mathcal{T}_{l,\mathbb{R}}(I, \mathbb{H})$, the upper triangular $I \times I$ quaternion matrices with real-valued diagonal entries, denoted by $\mathcal{T}_{u,\mathbb{R}}(I, \mathbb{H})$, and the usual Hermitian $I \times I$ quaternion matrices denoted by $\mathcal{H}(I, \mathbb{H})$, respectively. With the standard multiplication and and standard inner product $q \mapsto \|q\|^2$ on $\mathbb{H}$ the vector space $\mathcal{M}_{\mathbb{R}}(I, \mathbb{H})$ becomes a Vinberg algebra. The Vinberg multiplication becomes standard matrix multiplication of $I \times I$ quaternion matrices followed by taking the real part of diagonal elements. The groups $\mathcal{T}_l^+$ and $\mathcal{T}_u^+$ are the the groups $\mathcal{T}_l^+(I, \mathbb{H})$ and $\mathcal{T}_u^+(I, \mathbb{H})$ of lower and upper triangular $I \times I$ quaternion matrices with positive diagonal elements, respectively. The homogeneous cone $\mathcal{P}$ in the Vinberg algebra is the cone of $I \times I$ positive definite (Hermitian) quaternion matrices $\mathcal{P}(I, \mathbb{H})$ and $\mathcal{P}^* = \mathcal{P} = \mathcal{P}_c$.

It is well known, cf. again Vinberg (1965), §3, that the connected component is $G = \{g_A | A \in \mathrm{GL}(I, \mathbb{H})\}$, where $\mathrm{GL}(I, \mathbb{H})$ is the quaternion general linear group, and $g_A := (S \mapsto AS\overline{A'})$. The parametrization of G by $\mathrm{GL}(I, \mathbb{H})$ is not injective and $G = \mathrm{Aut}(\mathcal{P})$. The Lie algebra of G is $\mathfrak{g} = \mathfrak{g}_c = L \oplus \mathbb{R}$, where $L = \mathfrak{sl}(I, \mathbb{H})$ is the simple Lie algebra of $I \times I$ quaternion matrices with trace zero, cf. for example Faraut and Korányi (1994), page 97. As in Example 6.2 it is seen that every multiplier $\chi : G \to \mathbb{R}_+$ has the form $\chi(g) = |\det(g)|^\alpha$ and thus $\lambda_i = \alpha(2I - 1) =: \lambda$ does not depend on $i \in I$, so λ can then replace χ whenever χ is used as an index or a parameter. The I conditions (5.6) reduce to the single condition $\lambda > 2I - 2$.

From (5.19) it follows that $\Sigma^{-\chi} = \lambda\Sigma^{-1}$, where Σ^{-1} is the usual inverse of an $I \times I$ quaternion matrix. The Wishart distribution (5.15) takes the form

$$dW_{\Sigma,\lambda}(S) = \frac{\lambda^{I\lambda}\det(S)^{\lambda-2I+1}\exp\{-\lambda\mathrm{tr}(\Sigma^{-1}S)\}}{\pi^{I(I-1)}\prod(\Gamma(\lambda - 2i + 2)|i = 1, \cdots, I)\det(\Sigma)^\lambda}dS, \tag{6.5}$$

where $\det(S) := \prod(s_{[i]\cdot}|i \in I)$ does not depend of the specific choice of a total ordering of I and thus could be called the determinant of $S \in \mathcal{P}(I, \mathbb{H})$. This distribution is the *quaternion Wishart distribution* on $\mathcal{P}(I, \mathbb{H})$ with shape parameter $\lambda > 2I - 2$ and expectation Σ. The r^{th} moment ($r \geq 0$) of the normalized generalized variance is then given by, cf (5.18),

$$\lambda^{-Ir\lambda}\frac{\prod(\Gamma(\lambda - 2i + 2 + r)|i = 1, \cdots, I)}{\prod(\Gamma(\lambda - 2i + 2)|i = 1, \cdots, I)}. \tag{6.6}$$

Example 6.5. Let E be a vector space over $\mathbb{R}$, $\mathrm{GL}(E)$ the group of bijective linear mappings of E onto itself, and $\mathrm{P}(E)$ the open cone of positive definite forms $\delta : E \times E \to \mathbb{R}$ on E, i.e., inner products on E. Set $\delta(x) := \delta(x, x)$, $x \in E$. Since the classical action

$$\begin{array}{rcl} \mathrm{GL}(E) \times \mathrm{P}(E) & \to & \mathrm{P}(E) \\ (f, \delta) & \mapsto & \delta \circ (f \times f) \equiv (x \mapsto \delta(f(x))) \end{array} \tag{6.7}$$

is transitive $\mathrm{P}(E)$ is a homogeneous cone. (By choosing any basis for E we obtain an isomorphism between $\mathrm{P}(E)$ and $\mathcal{P}(I, \mathbb{R})$, where $I = \dim(E)$, cf. Example 6.2.)

Let H be a compact group, and $\rho : H \to \mathrm{GL}(E)$ a continuous group representation. Denote $\rho(h)(x) =: hx$, $h \in H$, $x \in E$,

$$\mathrm{GL}(E, H) := \{f \in \mathrm{GL}(E) | f \circ \rho(h) = \rho(h) \circ f, \ h \in H\}$$

and

$$\mathrm{P}(E, H) := \{\delta \in \mathrm{P}(E) | \delta(hx) = \delta(x), \ x \in E, h \in H\}.$$

Then $\mathrm{P}(E, H)$ is not empty and it is a homogeneous cone since the restriction of the action (6.7) to the subgroup $\mathrm{GL}(E, H) \subseteq \mathrm{GL}(E)$ and to the subcone $\mathrm{P}(E, H) \subseteq \mathrm{P}(E)$ also is transitive, cf. Andersson and Madsen (1998), Lemma A1. (By choosing an orthonormal basis for E wrt an inner product $\delta_0 \in \mathrm{P}(E, H)$ we obtain an isomorphism between $\mathrm{P}(E, H)$ and $\mathcal{P}_H(I, \mathbb{R})$, where H also denotes the group of orthogonal matrices for the mappings in $\rho(H)$, cf. (iv) in Section 1.) From Examples 6.2, 6.3, and 6.4 and from (iv) in Section 1, it follows that the connected component $G = \{g_f | f \in \mathrm{GL}(E, H)\}$, where $g_f(\delta) := \delta \circ (f \times f)$, and that $\mathrm{P}(E, H)$ decomposes into a product of homogeneous cones each being one of the three types in Examples 6.2, 6.3, and 6.4. Thus one has a complete description of Wishart distributions on $\mathrm{P}(E, H)$.

Example 6.6. Let I be a set with $I = 2$ and with any total ordering, e.g., $I = \{1, 2\}$ with the standard order. Let $E_{21} = V$, where V is a vector space with $m - 1 := \dim(V) > 0$, and let $f_{21} = \mathrm{Id}_V$. The vector space $\mathcal{A}$ is thus the vector space $\mathcal{M}_2(V)$ of all 2×2 generalized matrices of the form

$$\begin{pmatrix} \alpha_1 & v_2 \\ v_1 & \alpha_2 \end{pmatrix}, \tag{6.8}$$

where $\alpha_1, \alpha_2 \in \mathbb{R}$, and $v_1, v_2 \in V$. The involution is given by transposition, denoted by $A^* = A'$, $A \in \mathcal{M}_2(V)$, of the array. The vector space of Hermitian matrices $\mathcal{H} \subset \mathcal{M}_2(V)$ is all matrices of the form (6.8) with $v_1 = v_2$. Note that the center $\mathcal{A}_c = \mathcal{A}$ and in particular that K is a one-point set. Then $n_{\cdot 1} = m - 1 = n_{2\cdot}$, $n_{\cdot 2} = 0 = n_{1\cdot}$, $n_1 = n_2 = \frac{m+1}{2}$, and $n_{\cdot} = m + 1$. Since $I = 2$, we only need to specify an inner product, $(v_1, v_2) \mapsto v_1 \cdot v_2$, on V to establish the Vinberg multiplication given by

$$\left(\begin{pmatrix} \alpha_1 & v_2 \\ v_1 & \alpha_2 \end{pmatrix}, \begin{pmatrix} \beta_1 & u_2 \\ u_1 & \beta_2 \end{pmatrix}\right) \mapsto \begin{pmatrix} \alpha_1\beta_1 + v_2 \cdot u_1 & \alpha_1 u_2 + \beta_2 v_2 \\ \beta_1 v_1 + \alpha_2 u_1 & \alpha_2\beta_2 + v_1 \cdot u_2 \end{pmatrix}. \tag{6.9}$$

With this multiplication $\mathcal{M}_2(V)$ becomes a Vinberg algebra.

(For $m = 2$, $m = 3$ and $m = 5$ this example is covered by Examples 6.2, 6.3, and 6.4, respectively. It is thus natural to demand $I > 2$ in these examples, to obtain five non-overlaping different types represented in Examples 6.1-6.4 and the present example.)

The homogeneous cone $\mathcal{P}$ determined by $\mathcal{A}$ is given by

$$\mathcal{P} = \left\{ \begin{pmatrix} t_1 & 0 \\ t & t_2 \end{pmatrix} \begin{pmatrix} t_1 & t \\ 0 & t_2 \end{pmatrix} | t_1 > 0,\ t_2 > 0,\ t \in V \right\} \tag{6.10}$$

or equivalently

$$\mathcal{P}_2(V) := \mathcal{P} = \left\{ \begin{pmatrix} s_1 & s \\ s & s_2 \end{pmatrix} | s_1 > 0,\ s_2 > 0,\ s \in V,\ s_1 s_2 - \|s\|^2 > 0 \right\}. \tag{6.11}$$

Again $\mathcal{P} = \mathcal{P}^* = \mathcal{P}_c$. In Remark 6.1 below it is established that $\mathcal{P}_2(V)$ is a Lorentz cone, cf. (1.18). For

$$S = \begin{pmatrix} s_1 & s \\ s & s_2 \end{pmatrix} \in \mathcal{P} \tag{6.12}$$

we define $\det(S) := s_{[1]\bullet} s_{[2]\bullet} = s_1 s_2 - \|s\|^2$ and call it the determinant. Note that $\det(S)$ does not depend of the choice of the total ordering of I.

The Lie algebra of the connected component G is $\mathfrak{g} = \mathfrak{g}_c = L \oplus \mathbb{R}$, where $L = \mathfrak{o}(1, m)$, one of the classical simple Lie algebras, cf. Faraut and Korányi (1994), page 97. By a similar argument as in Example 6.2, every multiplier has the form $\chi(g) = |\det(g)|^\alpha$, so $\lambda_1 = \lambda_2 = \frac{m+1}{2}\alpha =: \lambda$ does not depend on $i \in \{1, 2\}$ and λ can replace χ whenever χ is used as an index or a parameter. The two conditions (5.6) reduce to the single condition $\lambda > \frac{m-1}{2}$. From (5.19) it follows that $\Sigma^{-\chi} = \lambda\Sigma^{-1}$, where

$$\begin{aligned} \Sigma^{-1} := (T^*)^{-1}T^{-1} &= \det(\Sigma)^{-1} \begin{pmatrix} \sigma_2 & -\gamma \\ -\gamma & \sigma_1 \end{pmatrix}, \\ \Sigma &= \begin{pmatrix} \sigma_1 & \gamma \\ \gamma & \sigma_2 \end{pmatrix} = TT^*,\ T \in \mathcal{T}_l^+. \end{aligned} \tag{6.13}$$

The Wishart distribution (5.15) then takes the form

$$dW_{\Sigma,\lambda}(S) = \frac{\lambda^{2\lambda} \det(S)^{\lambda - \frac{m+1}{2}} \exp\{-\lambda \mathrm{tr}(\Sigma^{-1}S)\}}{\pi^{\frac{m-1}{2}} \Gamma(\lambda)\Gamma(\lambda - \frac{m-1}{2}) \det(\Sigma)^\lambda} dS. \tag{6.14}$$

This distribution will be refered to as the *Lorentz Wishart distribution* on $\mathcal{P}_2(V)$ with shape parameter $\lambda > \frac{m-1}{2}$ and expectation Σ. An equivalent form for this family of Lorenz Wishart distributions was obtained by Casalis and Letac (1996), page 772. The r^{th} moment ($r \geq 0$) of the normalized generalized variance is given by, cf. (5.18),

$$\lambda^{-2r\lambda} \frac{\Gamma(\lambda + r)\Gamma(\lambda + \frac{m-1}{2} + r)}{\Gamma(\lambda)\Gamma(\lambda + \frac{m-1}{2})}. \tag{6.15}$$

In the special case where $4\lambda \in \mathbb{N}$ this distribution was obtained by Tolver Jensen (1988), cf. Remark 6.1 below, as the distribution of the ML estimator in his Clifford normal models.

Remark 6.1. Set $W := \mathbb{R} \times V$ and define the inner product on W by $\|(\alpha, w)\|^2 := \alpha^2 + \|v\|^2$, $(\alpha, w) \in W$. The homogeneous cone $\mathcal{P}_2(V)$ is then isomorphic to the (open) Lorentz cone C given in (1.18) according to the isomorphism

$$\begin{array}{ccc} \mathcal{H} & \to & \mathbb{R} \times W \\ \begin{pmatrix} h_1 & h \\ h & h_2 \end{pmatrix} & \mapsto & (\frac{h_1 + h_2}{2}, (\frac{h_1 - h_2}{2}, h)) \end{array} \tag{6.16}$$

restricted to and corestricted to $\mathcal{P}_2(V)$ and C, respectively. In Tolver Jensen (1988), formula (29), a distribution on C with parameters $(\lambda, w) \equiv (\lambda, (\alpha, v)) \in C$ and $N \in \mathbb{N}$ with $\frac{N}{2} > m - 1$ is presented. This distribution on C is transformed by (6.16) into the Lorentz Wishart distribution on $\mathcal{P}_2(V)$ with shape parameter $\lambda = \frac{N}{4}$ and expectation

$$\Sigma = \begin{pmatrix} \lambda + \alpha & v \\ v & \lambda - \alpha \end{pmatrix}.$$

□

Example 6.7. Let E be a Euclidean space with its inner product $(\ ,\ )$. The vector space $\mathrm{S}(E)$ of all symmetric forms s on E is identified with the vector space of symmetric linear mappings (also denoted by s) through $s(x, x) = (s(x), x)$, $x \in E$. The homogeneous cone of positive definite forms $\mathrm{P}(E)$ then corresponds to the positive mappings. Following Tolver Jensen (1988), let $L \subseteq \mathrm{S}(E)$ be a subspace of symmetric linear mappings with $\mathrm{Id}_E \in L$. Set $\mathrm{P}_L := L \cap \mathrm{P}(E)$. Thus P_L is an open cone (in L) with L as enveloping vector space. Then $(\mathrm{P}_L)^{-1} = \mathrm{P}_M$ for some subspace $M \subseteq \mathrm{S}(E)$ if and only if L is closed under Jordan multiplication ($\frac{1}{2}(AB + BA) \in L$, for all $A, B \in L$) and in this case $L = M$, cf. Tolver Jensen (1988), Lemma 1.

Consider, therefore, cones $\mathrm{P} \equiv \mathrm{P}_L$ with the essential properties $\mathrm{Id}_E \in \mathrm{P}$ and $\mathrm{P}^{-1} = \mathrm{P}$. As described in (ix) in Section 1 the cone P decomposes into a product of homogeneous cones each being of one of the five types each described in the Examples 6.1, 6.2 ($I > 2$), 6.3 ($I > 2$), 6.4 ($I > 2$), and 6.6. In particular, P is a homogeneous cone whose connected group G is a product of the connected Lie groups each one being from one of these five examples. Thus one has a complete description of Wishart distributions on all cones P of positive mappings with the properties $\mathrm{Id}_E \in P$ and $\mathrm{P}^{-1} = \mathrm{P}$.

Example 6.8. Let I be a set with $I = 3$ and any total ordering, e.g., $I = \{1, 2, 3\}$ with the standard ordering. Let $E_{ij} = \mathbb{O}$, $j < i$, where $\mathbb{O}$ denotes the division algebra over $\mathbb{R}$ of *octonions*, and $f_{ij}(q) = \overline{q}$, where $\overline{q}$ denotes the octonion conjugate of $q \in \mathbb{O}$. Let $\mathbf{Re}(q) := \frac{1}{2}(q + \bar{q})$ denote the real part of $q \in \mathbb{O}$. The vector space $\mathcal{A}$ over $\mathbb{R}$ is $\mathcal{M}_{\mathbb{R}}(3, \mathbb{O})$ of all 3×3 octonion matrices with real-valued diagonal entries. The involution on $\mathcal{A}$ is given by $A^* = \overline{A'}$, $A \in \mathcal{A}$. Note that the center $\mathcal{A}_c = \mathcal{A}$ and in particular that K is a one-point set. Then $n_{\cdot 1} = 16$, $n_{\cdot 2} = 8$, $n_{\cdot 3} = 0$, $n_{1\cdot} = 0$, $n_{2\cdot} = 8$, $n_{3\cdot} = 16$, $n_i = 9$, $i = 1, 2, 3$, and $n_{\cdot} = 27$.

The subspaces $\mathcal{T}_l$, $\mathcal{T}_u$, and $\mathcal{H}$ become the lower triangular 3×3 octonion matrices with real-valued diagonal entries, denoted by $\mathcal{T}_{l,\mathbb{R}}(3, \mathbb{O})$, the upper triangular 3×3 octonion matrices with real-valued diagonal entries, denoted by $\mathcal{T}_{u,\mathbb{R}}(3, \mathbb{O})$, and the Hermitian 3×3 octonion matrices denoted by $\mathcal{H}(3, \mathbb{O})$, respectively. With the standard (nonassociative) multiplication and inner product $q \mapsto \|q\|^2$, $\mathcal{M}_{\mathbb{R}}(3, \mathbb{O})$ becomes a Vinberg algebra. Since we only consider 3×3 matrices the "associative" condition (A) in Section 4 becomes vacuous and the nonassociativity of the multiplication in $\mathbb{O}$ does not come into play. The Vinberg multiplication becomes standard matrix multiplication of 3×3 octonion matrices followed by taking the real part of diagonal elements.

The groups $\mathcal{T}_l^+$ and $\mathcal{T}_u^+$ are the groups $\mathcal{T}_l^+(3, \mathbb{O})$ and $\mathcal{T}_u^+(3, \mathbb{O})$ of lower and upper triangular 3×3 octonion matrices with positive diagonal entries, respectively. The homogeneous cone in this Vinberg algebra is

$$\mathcal{P} \equiv \mathcal{P}(3, \mathbb{O}) := \{T\bar{T}' | T \in \mathcal{T}_l^+(3, \mathbb{O})\}, \tag{6.17}$$

and $\mathcal{P} = \mathcal{P}^* = \mathcal{P}_c$. (The definition (6.17) provides a definition of a positive definite 3×3 octonion matrix.)

The Lie algebra of the connected component G is $\mathfrak{g} = \mathfrak{g}_c = L \oplus \mathbb{R}$, where $L = \mathfrak{e}_{6(26)}$ is one of the exceptional simple Lie algebras, cf. for example Faraut and Korányi (1994), page 97. Thus by our now standard argument, every multiplier $\chi : G \to \mathbb{R}_+$ has the form $\chi(g) = |\det(g)|^{\alpha}$, so $\lambda_1 = \lambda_2 = \lambda_3 = 9\alpha =: \lambda$ does not depend on $i \in \{1, 2, 3\}$ and λ can then replace χ whenever χ is used as an index or a parameter. The three conditions (5.6) reduce to the single condition $\lambda > 8$. From (5.19) it follows that $\Sigma^{-\chi} = \lambda\Sigma^{-1}$, where $\Sigma^{-1} = (\overline{T'})^{-1}T^{-1}$, when $\Sigma = T\overline{T'}$. The Wishart distribution (5.15) then takes the form

$$dW_{\Sigma,\lambda}(S) = \frac{\lambda^{3\lambda}\det(S)^{\lambda-9}\exp\{-\lambda\mathrm{tr}(\Sigma^{-1}S)\}}{\pi^{12}\Gamma(\lambda)\Gamma(\lambda-4)\Gamma(\lambda-8)\det(\Sigma)^{\lambda}}\,dS, \tag{6.18}$$

where

$$\det(S) := s_{[1]\bullet}s_{[2]\bullet}s_{[3]\bullet} \tag{6.19}$$
$$= s_{11}s_{22}s_{33} - s_{11}\|s_{32}\|^2 - s_{22}\|s_{31}\|^2 - s_{33}\|s_{21}\|^2 + 2\mathbf{Re}(s_{32}s_{21}s_{13})$$

could be called the determinant of $S \equiv (s_{ij}|i,j = 1,2,3) \in \mathcal{P}(3,\mathbb{O})$. Note that $\det(S)$ does not depend of the choice of the total ordering of I. This distribution is called the *octonion Wishart distribution* on $\mathcal{P}(3,\mathbb{O})$ with shape parameter $\lambda > 8$ and expectation Σ. The octonion Wishart distribution is also defined in Casalis and Letac (1996), page 772. The r^{th} moment $(r \geq 0)$ of the normalized generalized variance is given by, cf. (5.18),

$$\lambda^{-3r\lambda}\frac{\Gamma(\lambda+r)\Gamma(\lambda+4+r)\Gamma(\lambda+8+r)}{\Gamma(\lambda)\Gamma(\lambda+4)\Gamma(\lambda+8)}. \tag{6.20}$$

Example 6.9. Let $V \neq \{0\}$ be a Euclidean space with inner product $(\cdot,\cdot)$. Let $C \subset V$ be a symmetric cone, i.e., a homogeneous cone with $C = C^*$ under the identification of V with V^* through $(\cdot,\cdot)$. It is established in Faraut and Korányi (1994) that a symmetric cone is isomorphic to the cone $\mathcal{J}^+$ of positive elements in a formally real Jordan algebra $\mathcal{J}$, cf. (x) in Section 1. Through the well-known classical structure theorem for such Jordan algebras one obtains the structure theorem for symmetric cones, or equivalently, for the homogeneous cones consisting of the positive elements in a Jordan algebra. Every symmetric cone decomposes into a product of indecomposable cones each being one of the six non-overlapping types from the Examples 6.1, 6.2 $(p > 2)$, 6.3 $(p > 2)$, 6.4 $(p > 2)$, 6.6, and 6.8. Note that the last type consists of *only* one cone. Thus again the Wishart distributions on symmetric cones are completely described. The reason for the extra exceptional type in Example 6.8 compared to the result by Tolver Jensen (1988) is that he also require that the Jordan algebra should be represented as a Jordan algebra of symmetric linear mappings on a vector space. This excludes the exceptional cone. Tolver Jensen thus uses the structure theorem for formally real Jordan algebras *together* with the structure of their representations as symmetric linear mappings in an Euclidean space.

Example 6.10. Let $I = \{a,b,1\}$ with partial ordering given by $a < 1$ and $b < 1$. Set $E_{1a} = E_{1b} = \mathbb{R}$ and $f_{1a} = f_{1b} = \mathrm{Id}_{\mathbb{R}}$. The vector space $\mathcal{A}$ is thus the vector space of all $I \times I$ real matrices with zeroes at the (a,b) and (b,a) entries. The involution is given by $A^* = A'$, $A \in \mathcal{A}$. Note that $\mathcal{A}_c$ consists of all the diagonal matrices in $\mathcal{A}$ and in particular that $K = I$. Then $n_{\cdot a} = 1$, $n_{\cdot b} = 1$, $n_{\cdot 1} = 0$, $n_{a\cdot} = 0$, $n_{b\cdot} = 0$, $n_{1\cdot} = 2$, $n_a = n_b = \frac{3}{2}$,

$n_1 = 2$, and $n. = 5$. With the standard multiplication and the standard inner product on $\mathbb{R}$, the vector space $\mathcal{A}$ becomes a Vinberg algebra. The Vinberg multiplication is given by the standard matrix multiplication except that the entries corresponding to (a, b) and (b, a) are by definition forced to be zero. The homogeneous cone $\mathcal{P}$ becomes

$$\mathcal{P} := \left\{ \begin{pmatrix} t_{aa} & 0 & 0 \\ 0 & t_{bb} & 0 \\ t_{1a} & t_{1b} & t_{11} \end{pmatrix} \begin{pmatrix} t_{aa} & 0 & t_{1a} \\ 0 & t_{bb} & t_{1b} \\ 0 & 0 & t_{11} \end{pmatrix} \right. \tag{6.21}$$

$$\left. \in \mathcal{A} | t_{aa}, t_{bb}, t_{11} > 0,\ t_{1a}, t_{1b} \in \mathbb{R} \right\}.$$

Since TT', $T \in \mathcal{T}_l^+$, in this case corresponds to standard multiplication for lower triangular matrices we obtain

$$\mathcal{P} := \{S \in \mathcal{A} | S \text{ is positive definite in the classical sense}\}. \tag{6.22}$$

The mapping $\pi : \mathcal{T}_l^+ \to G$ is in this case a group isomorphism , cf. Vinberg (1962). In fact its Lie algebra is $\mathfrak{g} = \mathfrak{g}_c \oplus \mathcal{T}_l^0$ with $\mathfrak{g}_c = \mathbb{R}^{\{a,b,1\}}$, where $\mathcal{T}_l^0$ is the Lie algebra of lower triangular matrices in $\mathcal{A}$ with zeros in the diagonal. Therefore every multiplier $\chi : G \to \mathbb{R}_+$ is given by $\chi(\pi(T)) = t_{aa}^{2\lambda_a} t_{bb}^{2\lambda_b} t_{11}^{2\lambda_1}$, $T \equiv (t_{ij} | (i, j) \in I \times I) \in \mathcal{T}_l^+$, for some $(\lambda_a, \lambda_b, \lambda_1) \in \mathbb{R}^{\{a,b,1\}}$, so $(\lambda_a, \lambda_b, \lambda_1)$ can replace χ whenever χ is used as an index or a parameter. The three conditions (5.6) become

$$\lambda_a > 0, \quad \lambda_b > 0, \quad \lambda_1 > 1. \tag{6.23}$$

The generalized Wishart distribution (5.15) takes the form

$$dW_{\Sigma,(\lambda_a,\lambda_b,\lambda_1)}(S) \tag{6.24}$$

$$= \frac{\lambda_a^{\lambda_a} \lambda_b^{\lambda_b} \lambda_c^{\lambda_c} s_{aa}^{\lambda_a - \frac{3}{2}} s_{bb}^{\lambda_b - \frac{3}{2}} s_{[1]\bullet}^{\lambda_1 - 2}}{\pi \Gamma(\lambda_a) \Gamma(\lambda_b) \Gamma(\lambda_c - 1) \sigma_{aa}^{\lambda_a} \sigma_{bb}^{\lambda_b} \sigma_{[1]\bullet}^{\lambda_1}} \exp\{-\mathrm{tr}(\Sigma^{-(\lambda_a,\lambda_b,\lambda_1)} S)\} dS,$$

where $S \equiv (s_{ij} | i, j = a, b, 1) \in \mathcal{P}$, $\Sigma \equiv (\sigma_{ij} | i, j = a, b, 1) \in \mathcal{P}$, $s_{[1]\bullet} = s_{11} - s_{1a} s_{aa}^{-1} s_{a1} - s_{1b} s_{bb}^{-1} s_{b1}$, and $\sigma_{[1]\bullet} = \sigma_{11} - \sigma_{1a} \sigma_{aa}^{-1} \sigma_{a1} - \sigma_{1b} \sigma_{bb}^{-1} \sigma_{b1}$, cf. Remark 4.2. This distribution is new and has expectation Σ. The r^{th} moment $(r \geq 0)$ of the normalized generalized variance is given by, cf. (5.18),

$$\lambda_a^{-r\lambda_a} \lambda_b^{-r\lambda_b} \lambda_1^{-r\lambda_1} \frac{\Gamma(\lambda_a + r)\Gamma(\lambda_b + r)\Gamma(\lambda_1 - 1 + r)}{\Gamma(\lambda_a)\Gamma(\lambda_b)\Gamma(\lambda_1 - 1)}. \tag{6.25}$$

From (5.19) it follows that $\Sigma^{-\chi} \equiv \Sigma^{-(\lambda_a,\lambda_b,\lambda_1)} \in \mathcal{P}^*$ can be obtained explicitly, but as with standard matrix inversion the formula is complicated.

For the special case of $\lambda_a = \lambda_b = \lambda_1 =: \lambda > 1$ we obtain $\Sigma^{-(\lambda,\lambda,\lambda)} = \lambda\Sigma^{-1}$, where $\Sigma^{-1} = (T')^{-1}T^{-1}$ when $\Sigma = TT'$, $T \in \mathcal{T}_l^+$.

Example 6.11. The dual cone $\mathcal{P}^*$ to $\mathcal{P}$ from Example 6.10 is obtained by reversing the ordering of the index set I. In the light of the reverse ordering we will also rename the index 1 to 0, and consider $I = \{a, b, 0\}$, with partial ordering given by $0 < a$ and $0 < b$. This yields a new Vinberg algebra with the same underlying algebra, with the same involution, with the same Vinberg multiplication, but whose lower triangular matrices are the upper triangular matrices from the previous Example 6.10. If the elements in the index set I are placed in a never decreasing sequence, for example $0, a, b$, the lower triangular matrices will again appear to be lower triangular in the traditional sense. Nevertheless since the appearance does not matter we will continue to list indices in the same order as before, namely $a, b, 0$, in order to make comparisons with the previous example easier. Thus

$$\mathcal{P}^* :=$$

$$\left\{ \begin{pmatrix} t_{aa} & 0 & t_{0a} \\ 0 & t_{bb} & t_{0b} \\ 0 & 0 & t_{00} \end{pmatrix} \begin{pmatrix} t_{aa} & 0 & 0 \\ 0 & t_{bb} & 0 \\ t_{0a} & t_{0b} & t_{00} \end{pmatrix} \in \mathcal{A} | t_{aa}, t_{bb}, t_{cc} > 0,\ t_{0a}, t_{0b} \in \mathbb{R} \right\}$$

$$= \left\{ \begin{pmatrix} s_{aa} & 0 & s_{a0} \\ 0 & s_{bb} & s_{b0} \\ s_{a0} & s_{b0} & s_{00} \end{pmatrix} \in \mathcal{A} | s_{00} > 0,\ s_{[a]\bullet} > 0,\ s_{[b]\bullet} > 0 \right\}, \tag{6.26}$$

where $s_{[a]\bullet} := s_{aa} - s_{a0}s_{00}^{-1}s_{0a}$ and $s_{[b]\bullet} := s_{bb} - s_{b0}s_{00}^{-1}s_{0b}$, cf. Remark 4.2. The defining conditions in (6.26) express that the $\{a, 0\} \times \{a, 0\}$ submatrix and the $\{b, 0\} \times \{b, 0\}$ submatrix are positive definite in the traditional sense.

The group G in this "dual" example is isomorphic to G in the previous Example 6.10, in particular, the mapping $\pi : \mathcal{T}_l^+ \to G$ is again a group isomorphism. Therefore every multiplier $\chi : G \to \mathbb{R}_+$ is given by $\chi(\pi(T)) = t_{aa}^{2\mu_a}t_{bb}^{2\mu_b}t_{00}^{2\mu_0}$, $T \equiv (t_{ij}|i, j \in \{a, b, 0\}) \in \mathcal{T}_l^+$, for some $(\mu_a, \mu_b, \mu_0) \in \mathbb{R}^{\{a,b,0\}}$. Thus (μ_a, μ_b, μ_0) can replace χ whenever χ is used as an index or a parameter. The three conditions (5.6) become

$$\mu_a > \frac{1}{2}, \quad \mu_b > \frac{1}{2}, \quad \mu_0 > 0. \tag{6.27}$$

The general Wishart distribution (5.15) on this dual homogeneous cone $\mathcal{P}^*$ takes the form

$$dW_{\Sigma,(\mu_a,\mu_b,\mu_0)}(S) \tag{6.28}$$

$$= \frac{\mu_a^{\mu_a}\mu_b^{\mu_b}\mu_0^{\mu_0}s_{[a]\bullet}^{\mu_a-\frac{3}{2}}s_{[b]\bullet}^{\mu_b-\frac{3}{2}}s_{00}^{\mu_0-1}}{\pi\Gamma(\mu_a-\frac{1}{2})\Gamma(\mu_b-\frac{1}{2})\Gamma(\mu_0)\sigma_{[a]\bullet}^{\mu_a}\sigma_{[b]\bullet}^{\mu_b}\sigma_{00}^{\mu_0}} \exp\{-\mathrm{tr}(\Sigma^{-(\mu_a,\mu_b,\mu_0)}S)\}dS$$

where $S \equiv (s_{ij}|i,j = a,b,0) \in \mathcal{P}^*$, $\Sigma \equiv (\sigma_{ij}|i,j = a,b,0) \in \mathcal{P}^*$, and where $\sigma_{[a]\bullet}$ and $\sigma_{[b]\bullet}$ are defined similar to $s_{[a]\bullet}$ and $s_{[b]\bullet}$. This distribution is new and has expectation Σ. The r^{th} moment ($r \geq 0$) of the normalized generalized variance is then given by, cf. (5.18),

$$\mu_a^{-r\mu_a}\mu_b^{-r\mu_b}\mu_0^{-r\mu_0}\frac{\Gamma(\mu_a - \frac{1}{2} + r)\Gamma(\mu_b - \frac{1}{2} + r)\Gamma(\mu_0 + r)}{\Gamma(\mu_a - \frac{1}{2})\Gamma(\mu_b - \frac{1}{2})\Gamma(\mu_0)} \tag{6.29}$$

From (5.19) it follows that $\Sigma^{-\chi} \equiv \Sigma^{-(\mu_a,\mu_b,\mu_0)} \in \mathcal{P}$ can be obtained explicitly. The special case $\mu_a = \mu_b = \mu_0 > \frac{1}{2}$ reduces in a way similar to the case $\lambda_a = \lambda_b = \lambda_1 > 1$ in the previous Example 6.10.

Example 6.12. This example generalizes Example 6.10, cf. Remark 6.2 below. Let I_1 be a set and $(I_\kappa|\kappa \in K)$ be a family of sets indexed by the set K. Let $I = \dot\cup(I_\kappa|\kappa \in K)\dot\cup I_1$ with a partial ordering given by $i_\kappa < i_1$ for any $i_\kappa \in I_\kappa$, $\kappa \in K$ and any $i_1 \in I_1$, together with any total ordering of the elements *within* each I_κ, $\kappa \in K$, and I_1. Set $E_{ij} = \mathbb{R}$ and $f_{ij} = \mathrm{Id}_{\mathbb{R}}$, $i,j \in I$ with $i > j$. The vector space $\mathcal{A}$ is thus the vector space of all real $I \times I$ matrices A blocked according to the decomposition of I and having the form

$$A = \begin{pmatrix} \operatorname{diag}(A_{\kappa\kappa}|\kappa \in K) & (A_{\kappa 1}|\kappa \in K) \\ (A_{1\kappa}|\kappa \in K) & A_{11} \end{pmatrix}, \tag{6.30}$$

where $A_{\kappa\kappa} \in \mathcal{M}(I_\kappa \times I_\kappa, \mathbb{R})$, $A_{\kappa 1} \in \mathcal{M}(I_\kappa \times I_1, \mathbb{R})$, $A_{1\kappa} \in \mathcal{M}(I_1 \times I_\kappa, \mathbb{R})$, and $A_{11} \in \mathcal{M}(I_1 \times I_1, \mathbb{R})$. The involution is $A^* = A'$, $A \in \mathcal{A}$. The decomposition of I into equivalence classes, cf. Remark 4.4, is the above decomposition and the center $\mathcal{A}_c$ is the set of all matrices A of the form (6.30) with $A_{1\kappa} = 0$ and $A_{\kappa 1} = 0$, $\kappa \in K$. For any $i \in I_\kappa$ define $I_{\kappa<i>} := \{i' \in I_\kappa | i' < i\}$, $\kappa \in K$ and for any $i \in I_1$ define $I_{1<i>} := \{i' \in I_1 | i' < i\}$. Then $n_{\cdot i} = I_1 + I_\kappa - I_{\kappa<i>} - 1$, $i \in I_\kappa$, $\kappa \in K$, $n_{\cdot i} = I_1 - I_{1<i>} - 1$, $i \in I_1$, $n_{i\cdot} = I_{\kappa<i>}$, $i \in I_\kappa$, $\kappa \in K$, $n_{i\cdot} = \sum(I_\kappa|\kappa \in K) + I_{1<i>}$, $i \in I_1$, $n_i = \frac{1}{2}(I_1 + I_\kappa + 1)$, $i \in I_\kappa$, $\kappa \in K$, and $n_i = \frac{1}{2}(I_1 + \sum(I_\kappa|\kappa \in K) + 1)$, $i \in I_1$.

With the standard multiplication and inner product on $\mathbb{R}$ the vector space $\mathcal{A}$ becomes a Vinberg algebra. Since the Vinberg multiplication TT', $T \in \mathcal{T}_l^+$ in this case is standard matrix multiplication, the homogeneous cone is

$$\mathcal{P} := \{S \in \mathcal{A} | S \text{ is positive definite in the classical sense}\}. \tag{6.31}$$

The center cone is

$$\mathcal{P}_c = \times(\mathcal{P}(I_\kappa, \mathbb{R})|\kappa \in K) \times \mathcal{P}(I_1, \mathbb{R}).$$

The Lie algebra of G is $\mathfrak{g} = \mathfrak{g}_c \oplus \mathcal{T}_{lB}^0$, where $\mathcal{T}_{lB}^0$ is the Lie algebra all matrices of the form (6.30) with $A_{\kappa\kappa} = 0$, $A_{\kappa 1} = 0$, $\kappa \in K$, and $A_{11} = 0$.

If $(\lambda_i|i \in I)$ corresponds to a multiplier $\chi : G \to \mathbb{R}_+$ then there exists $((\lambda_\kappa|\kappa \in K), \lambda_1) \in \mathbb{R}^K \times \mathbb{R}$, such that $\lambda_i = \lambda_\kappa$, $i \in I_\kappa$, $\kappa \in K$, and $\lambda_i = \lambda_1$,

$i \in I_1$, cf. Remark 5.3. Thus $((\lambda_\kappa|\kappa \in K), \lambda_1)$ can replace χ whenever χ is used as an index or a parameter. The I conditions (5.6) become the $K+1$ conditions

$$\lambda_\kappa > \frac{I_\kappa - 1}{2}, \quad \kappa \in K, \quad \lambda_1 > \frac{\sum(I_\kappa|\kappa \in K) + I_1 - 1}{2}. \tag{6.32}$$

The generalized Wishart distribution (5.15) takes the form

$$\begin{aligned} &dW_{\Sigma,((\lambda_\kappa|\kappa\in K),\lambda_1)}(S) \qquad (6.33)\\ &= \frac{\prod(\lambda_\kappa^{I_\kappa\lambda_\kappa}|\kappa\in K)\lambda_1^{I_1\lambda_1}\prod(|S_{\kappa\kappa}|^{\lambda_\kappa-\frac{1}{2}(I_1+I_\kappa+1)}|\kappa\in K)|S_{[1]\bullet}|^{\lambda_1-\frac{1}{2}(I_1+\sum(I_\kappa|\kappa\in K))}}{\pi^{n_{\cdot}-I}\prod(\prod(\Gamma(\lambda_\kappa-\frac{n_{i\cdot}}{2})|i\in I_\kappa)|\kappa\in K)\prod(\Gamma(\lambda_1-\frac{n_{i\cdot}}{2})|i\in I_1)\prod(|\Sigma_{\kappa\kappa}|^{\lambda_\kappa}|\kappa\in K)|\Sigma_{[1]\bullet}|^{\lambda_1}} \\ &\times\exp\{-\mathrm{tr}(\Sigma^{-((\lambda_\kappa|\kappa\in K),\lambda_1)}S)\}dS, \end{aligned}$$

where $S \equiv (S_{bb'}|b, b' \in K\dot{\cup}\{1\}) \in \mathcal{P}$, $\Sigma \equiv (\Sigma_{bb'}|b, b' \in K\dot{\cup}\{1\}) \in \mathcal{P}$, $S_{[1]\bullet} = S_{11} - \sum(S_{1\kappa}S_{\kappa\kappa}^{-1}S_{\kappa 1}|\kappa \in K)$, and $\Sigma_{[1]\bullet}$ is defined in a similar way. Thus $|\cdot| = \det(\cdot)$, cf. Remark 4.4. This distribution is new and has expectation Σ. The r^{th} moment $(r \geq 0)$ of the normalized generalized variance is then given by, cf. (5.18),

$$\begin{aligned} &\prod(\lambda_\kappa^{-rI_\kappa\lambda_\kappa}\prod(\tfrac{\Gamma(\lambda_\kappa-\frac{n_{i\cdot}}{2}+r)}{\Gamma(\lambda_\kappa-\frac{n_{i\cdot}}{2})}|i \in I_\kappa)|\kappa \in K) \\ &\times\lambda_1^{-rI_1\lambda_1}\prod(\tfrac{\Gamma(\lambda_1-\frac{n_{i\cdot}}{2}+r)}{\Gamma(\lambda_1-\frac{n_{i\cdot}}{2})}|i \in I_1). \end{aligned} \tag{6.34}$$

For the special case of $\lambda_\kappa = \lambda_1 =: \lambda > \frac{\sum(I_\kappa|\kappa\in K)+I_1-1}{2}$, $\kappa \in K$, we obtain $\Sigma^{-((\lambda|\kappa\in K),\lambda)} = \lambda\Sigma^{-1}$, where $\Sigma^{-1} = (T')^{-1}T^{-1}$, for $\Sigma = TT'$, $T \in \mathcal{T}_l^+$.

Remark 6.2. Two special cases of Examples 6.12 should be pointed out. One case is $K = 2$ and the other case is $I_\kappa = I_1 = 1$, $\kappa \in K$. If both $K = 2$, e.g., $K = \{a, b\}$, and $I_a = I_b = I_1 = 1$ Example 6.12 reduces to Example 6.10. □

Example 6.13. This example generalizes Example 6.11, cf. Remark 6.3 below. The dual cone $\mathcal{P}^*$ to $\mathcal{P}$ in (6.31) is obtained by reversing the ordering of the index set in the previous Example 6.12. In the light of the reverse ordering we will again as in Example 6.11 rename the index 1 to 0, i.e., consider $I = \dot{\cup}(I_\kappa|\kappa \in K)\dot{\cup}I_0$. Then $n_{\cdot i} = I_\kappa - I_{\kappa<i>} - 1$, $i \in I_\kappa$, $\kappa \in K$, $n_{\cdot i} = \sum(I_\kappa|\kappa \in K) + I_0 - I_{0<i>} - 1$, $i \in I_0$, $n_{i\cdot} = I_0 + I_{\kappa<i>}$, $i \in I_\kappa$, $\kappa \in K$, $n_{i\cdot} = I_{0<i>}$, $i \in I_0$, $n_i = \frac{1}{2}(I_0 + I_\kappa + 1)$, $i \in I_\kappa$, $\kappa \in K$, and $n_i = \frac{1}{2}(I_0 + \sum(I_\kappa|\kappa \in K) + 1)$, $i \in I_0$. A direct investigation using the characterization of the positive definite matrices in a Vinberg algebra in Remark 4.2 shows that

$$\mathcal{P}^* = \left\{ S \equiv \begin{pmatrix} \mathrm{Diag}(S_{\kappa\kappa}|\kappa \in K) & (S_{\kappa 0}|\kappa \in K) \\ (S_{0\kappa}|\kappa \in K) & S_{00} \end{pmatrix} \right.$$

$$\left. \in \mathcal{H} | S_{00} > 0,\ S_{[\kappa]\bullet} > 0, \kappa \in K \right\}, \tag{6.35}$$

where $\mathcal{H}$ denotes the Hermitian matrices in $\mathcal{A}$, $S_{[\kappa]\bullet} = S_{\kappa\kappa} - S_{\kappa 0} S_{00}^{-1} S_{0\kappa}$, $\kappa \in K$, and "> 0" means that the matrix is positive definite in the classical sense. The defining conditions express that the $(I_\kappa \dot\cup I_0) \times (I_\kappa \dot\cup I_0)$ submatrices of S, $\kappa \in K$, are positive definite in the classical sense. Again $((\mu_\kappa|\kappa \in K), \mu_0) \in \mathbb{R}^K \times \mathbb{R}$ may replace multipliers $\chi : G \to \mathbb{R}_+$. The I conditions (5.6) become the $K+1$ conditions

$$\mu_\kappa > \frac{I_0 + I_\kappa - 1}{2}, \quad \kappa \in K, \quad \mu_0 > \frac{I_0 - 1}{2}. \tag{6.36}$$

The generalized Wishart distribution (5.15) takes the form

$$dW_{\Sigma,((\mu_\kappa|\kappa\in K),\mu_0)}(S) \tag{6.37}$$

$$= \frac{\prod(\mu_\kappa^{I_\kappa \mu_\kappa}|\kappa\in K)\mu_0^{I_0\mu_0} \prod(|S_{[\kappa]\bullet}|^{\mu_\kappa - \frac{1}{2}(I_0+I_\kappa+1)}|\kappa\in K)|S_{00}|^{\mu_0 - \frac{1}{2}(I_0+\sum(I_\kappa|\kappa\in K)+1)}}{\pi^{n.-I}\prod(\prod(\Gamma(\mu_\kappa - \frac{n_{i.}}{2})|i\in I_\kappa)|\kappa\in K)\prod(\Gamma(\mu_0 - \frac{n_{i.}}{2})|i\in I_0)\prod(|\Sigma_{[\kappa]\bullet}|^{\mu_\kappa}|\kappa\in K)|\Sigma_{00}|^{\mu_0}}$$

$$\times \exp\{-\mathrm{tr}(\Sigma^{-((\mu_\kappa|\kappa\in K),\mu_0)} S)\} dS$$

where $\Sigma \in \mathcal{P}^*$ is partitioned similarly to $S \in \mathcal{P}^*$ and $\Sigma_{[\kappa]\bullet}$ is defined similarly to $S_{[\kappa]\bullet}$, $\kappa \in K$. Again $|\cdot| = \det(\cdot)$, cf. Remark 4.4. This distribution is new and has expectation Σ. The r^{th} moment ($r \geq 0$) of the normalized generalized variance may be obtained using (5.18).

For the special case $\mu_\kappa = \mu_0 =: \mu > \max(\frac{I_0+I_\kappa-1}{2}|\kappa \in K)$, $\kappa \in K$, the χ-inverse $\Sigma^{-((\mu_\kappa|\kappa\in K),\mu_0)}$ reduces in a way similar to the previous Example 6.12.

Remark 6.3. Two special cases of Examples 6.13 should be pointed out. One case is $K = 2$ and the other case is $I_\kappa = I_0 = 1$, $\kappa \in K$. If both $K = 2$, e.g., $K = \{a, b\}$, and $I_a = I_b = I_0 = 1$ Example 6.13 reduces to Example 6.11. □

Example 6.14. Let I be a partial ordered set and let $E_{ij} = \mathbb{R}$, $i, j \in I$ with $i > j$. The vector space $\mathcal{A}$ is the vector space of all $A \equiv (a_{ij}|(i,j) \in I \times I) \in \mathcal{M}(I, \mathbb{R})$ with $a_{ij} = 0$ when i and j are *not* related. The multiplication $C = A \cdot B$, $A, B \in \mathcal{A}$, defined as standard multiplication of $I \times I$ matrices except that entries (i,j) in C, where i and j are not related (or equal) in the partial ordering, are forced to be zero, will *not* define a Vinberg algebra in general. Nevertheless, if I has the *four element property*: i.e., there do not exist 4 different elements $i_0, i_a, i_b, i_1 \in I$ with the property that $i_0 < i_a < i_1$, $i_0 < i_b < i_1$, $i_0 < i_1$, with i_a and i_b being unrelated then the multiplication defines a Vinberg algebra. On the other hand if the multiplication defines a Vinberg algebra then I has the four element property. We now have the one to one correspondences

$$\begin{array}{ccccc} \mathcal{P} & \leftrightarrow & \mathcal{T}_l^+ & \leftrightarrow & \mathbf{P}(I, \leq) \\ T \cdot T^t & \leftarrow\!\shortmid & T & \mapsto & TT^t, \end{array} \tag{6.38}$$

where " $T \cdot T^t$ " in this example indicates the Vinberg multiplication, TT^t is the standard multiplication of $I \times I$ matrices, and $\mathbf{P}(I, \leq) := \{TT^t | T \in \mathcal{T}_l^+\}$. Note that $\mathbf{P}(I, \leq)$ is the set of $I \times I$ covariance matrices Σ for a normal distribution on $\mathbb{R}^I$ satisfying the set of conditional independences (CI) given by the poset I or equivalently by the *transitive* acyclic directed graph (ADG) I, cf. for example Andersson and Perlman (1994), §4, with $V = I$ and $I_v = \{v\}$, together with Proposition 11.2. In fact it is well-known and trivial that all (CI) restrictions in §4 of this reference can be reformulated such that $V = I$ and $I_v = \{v\}$. Normal models of this type were first introduced and analyzed by Andersson and Perlman (1993) as *lattice conditional independence* (LCI) models. Thus the set of unknown covariance matrices $\mathbf{P}(I, \leq)$ for a normal distribution with CI restrictions given by a transitive ADG with the four element property can be parametrized in a one-to-one fashion by a homogeneous cone $\mathcal{P} \subset \mathcal{A}$. The inverse parametrization replaces with a zero any entry σ_{ij} in $\Sigma \equiv (\sigma_{ij} | i, j \in I) \in \mathbf{P}(I, \leq)$ where i and j are not comparable in the partial ordering. The parametrization itself is given explicitly by the Reconstruction Algorithm given by Andersson and Perlman (1994), §5.

Let $I = \dot{\cup}(I_\kappa | \kappa \in K)$ be the decomposition given in Remark 4.4. Note that the decomposition in this case only depends on the poset I (all non-zero n_{ij} equal to 1) and thus its construction and description is a purely graphical problem. This description is addressed in the doctoral thesis by Wojnar (2000). The ordering of I can always be described in terms of an ordering[‖] of K, total orderings of all elements in I_κ, $\kappa \in K$, and in addition $i < j$ if $i \in I_\kappa$, $j \in I_{\kappa'}$, with $\kappa < \kappa'$.

Set $I_{\kappa<i>} := \{i' \in I_\kappa | i' < i\}$, $i \in I_\kappa$, $\kappa \in K$. Then $n_{i\cdot} = \sum(I_{\kappa'} | \kappa' < \kappa) + I_{\kappa<i>}$, $i \in I_\kappa$, $n_{\cdot i} = \sum(I_{\kappa'} | \kappa' > \kappa) + I_\kappa - I_{\kappa<i>} - 1$, $i \in I_\kappa$, $n_i = \frac{1}{2}(\sum(I_{\kappa'} | \kappa' < \kappa) + \sum(I_{\kappa'} | \kappa' > \kappa) + I_\kappa + 1) =: n_\kappa$, $i \in I_\kappa$.

Let the multiplier χ correspond to $(\lambda_i | i \in I) \in \mathbb{R}^I$. Then there exists $(\lambda_\kappa \in K) \in \mathbb{R}^K$ such that $\lambda_i = \lambda_\kappa$, $i \in I_\kappa$, cf. Remark 5.3. Thus $(\lambda_\kappa \in K)$ can replace $(\lambda_i | i \in I)$ as an index. The conditions (5.6) are

[‖]The possible kinds of orderings of K in the present example are described in the doctoral thesis by Wojnar (2000).

$$\lambda_\kappa > \frac{\sum(I_{\kappa'}|\kappa' < \kappa) + I_\kappa - 1}{2}, \quad \kappa \in K. \tag{6.39}$$

The generalized Wishart distribution takes the form

$$\begin{aligned} & dW_{\Sigma,(\lambda_\kappa|\kappa\in K)} \\ &= \frac{\prod(\lambda_\kappa^{I_\kappa\lambda_\kappa}|S_{[\kappa]\bullet}|^{\lambda_\kappa - n_\kappa}|\kappa\in K)}{\pi^{n.-I}\prod(|\Sigma_{[\kappa]\bullet}|^{\lambda_\kappa}\prod(\Gamma(\lambda_\kappa - \frac{1}{2}n_{i\cdot})|i=1,\cdots,I_\kappa)|\kappa\in K)} \\ &\quad \times \exp\{-\mathrm{tr}(\Sigma^{-(\lambda_\kappa|\kappa\in K)}S)\}dS, \end{aligned} \tag{6.40}$$

where $|S_{[\kappa]\bullet}| := \prod(s_{[i]\bullet}|i \in I_\kappa)$, $\kappa \in K$. The distribution on $\mathcal{P}$ is new and has expectation $\Sigma \in \mathcal{P}$. The r^{th} moments ($r \geq 0$) of the normalized generalized variance are given by

$$\prod\left(\lambda_\kappa^{I_\kappa\lambda_\kappa}\prod\left(\frac{\Gamma(\lambda_\kappa - \frac{1}{2}n_{i\cdot} + r)}{\Gamma(\lambda_\kappa - \frac{1}{2}n_{i\cdot})}\Big|i \in I_\kappa\right)\Bigg|\kappa \in K\right). \tag{6.41}$$

For the special case $\lambda_\kappa =: \lambda > \frac{h(I)}{2}$, $\kappa \in K$, where $h(I)$ denotes the *height*** of the poset I, we obtain $\Sigma^{(\lambda|\kappa\in K)} = \lambda\Sigma^{-1}$, where $\Sigma^{-1} = (T^t)^{-1}\cdot T^{-1}$, when $\Sigma = T\cdot T^t \in \mathcal{P}$, $T \in \mathcal{T}_l^+$.

The distribution (6.40) can be transformed to a distribution on $\mathbf{P}(I, \leq)$. The dual cone $\mathcal{P}^*$ and its Wishart distributions are obtained by reversing the order in I. The Examples 6.12 and 6.13 are both special cases of this example.

Example 6.15. This example contains 3 subexamples, denoted (i), (ii), and (iii). The details will be omitted since they are trivial and by now routine. In all three cases, let $I = \{1, 2, 3\}$ with the usual total ordering. Consider

Case (i): $E_{12} = E_{23} = \mathbb{R}$ and $E_{13} = \mathbb{C}$

Case (ii): $E_{12} = \mathbb{R}$ and $E_{23} = E_{13} = \mathbb{C}$

Case (iii): $E_{23} = \mathbb{R}$ and $E_{12} = E_{13} = \mathbb{C}$.

In all three cases, standard conjugations, standard multiplications, and standard inner products on $\mathbb{R}$ and $\mathbb{C}$ are used. The Vinberg multiplication in $\mathcal{A}$ becomes standard multiplication of 3×3 matrices followed by taking the real part of entries whenever needed to have a well-defined multiplication. All three cases constitute Vinberg algebras.

**The maximal obtainable length n of a sequence of the form $i_1 < \cdots < i_n$.

Let

$$\Sigma \equiv \begin{pmatrix} \sigma_{11} & \bar{\sigma}_{21} & \bar{\sigma}_{31} \\ \sigma_{21} & \sigma_{22} & \bar{\sigma}_{32} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \in \mathcal{H}.$$

The homogeneous cones $\mathcal{P}$ are given by the condition that $\Sigma \in \mathcal{H}$ belongs to $\mathcal{P} \subset \mathcal{H}$ if and only if, cf. Remark 4.2:

Case (i): $\sigma_{11} > 0$, $\sigma_{[2]\bullet} := \sigma_{22} - \frac{\sigma_{21}^2}{\sigma_{11}} > 0$ and

$$\sigma_{[3]\bullet} := \sigma_{33} - (\mathbf{Re}(\sigma_{31}), \sigma_{32}) \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Re}(\bar{\sigma}_{31}) \\ \sigma_{32} \end{pmatrix} - \frac{\mathbf{Im}(\sigma_{31})^2}{\sigma_{11}}$$

$$= \sigma_{33} - (\sigma_{31}, \sigma_{32}) \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \bar{\sigma}_{31} \\ \sigma_{32} \end{pmatrix} + \frac{\mathbf{Im}(\sigma_{31})^2 \sigma_{21}^2}{\sigma_{11}(\sigma_{11}\sigma_{22} - \sigma_{21}^2)} > 0,$$

where $\mathbf{Re}(\cdot)$ and $\mathbf{Im}(\cdot)$ denote the real and imaginary part, respectively. The two expressions for $\sigma_{[3]\bullet}$ show that the real part of Σ, $\mathbf{Re}(\Sigma)$, is positive definite in the classical sense, and that the positive definite complex 3×3 matrices Σ (in the classical sense) with $\sigma_{21}, \sigma_{32} \in \mathbb{R}$ form a proper subset of $\mathcal{P}$.

Case (ii): $\sigma_{11} > 0$, $\sigma_{[2]\bullet} := \sigma_{22} - \frac{\sigma_{21}^2}{\sigma_{11}} > 0$ and

$$\sigma_{[3]\bullet} := \sigma_{33} - (\sigma_{31}, \sigma_{32}) \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \bar{\sigma}_{31} \\ \sigma_{32} \end{pmatrix} > 0.$$

This shows that $\mathcal{P}$ is the cone of positive definite 3×3 complex matrices Σ (in the the classical sense) with $\sigma_{12} \in \mathbb{R}$.

Case (iii): $\sigma_{11} > 0$, $\sigma_{[2]\bullet} := \sigma_{22} - \frac{\|\sigma_{21}\|^2}{\sigma_{11}} > 0$ and

$$\sigma_{[3]\bullet} := \sigma_{33} - (\sigma_{31}, \sigma_{32}) \begin{pmatrix} \sigma_{11} & \bar{\sigma}_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \bar{\sigma}_{31} \\ \sigma_{32} \end{pmatrix} + \frac{\mathbf{Im}(\sigma_{31}\bar{\sigma}_{21})^2}{\sigma_{11}(\sigma_{11}\sigma_{22} - \|\sigma_{21}\|^2)} > 0.$$

The constants are:

Case (i): $n_{1\cdot} = 0$, $n_{2\cdot} = 1$, $n_{3\cdot} = 3$, $n_{\cdot 1} = 3$, $n_{\cdot 2} = 1$, $n_{\cdot 3} = 0$, $n_1 = \frac{5}{2}$, $n_2 = 2$, $n_3 = \frac{5}{2}$, and $n_{\cdot} = 7$

Case (ii): $n_{1\cdot} = 0$, $n_{2\cdot} = 1$, $n_{3\cdot} = 4$, $n_{\cdot 1} = 3$, $n_{\cdot 2} = 2$, $n_{\cdot 3} = 0$, $n_1 = \frac{5}{2}$, $n_2 = \frac{5}{2}$, $n_3 = 3$, and $n_{\cdot} = 8$

Case (iii): $n_{1\cdot} = 0$, $n_{2\cdot} = 2$, $n_{3\cdot} = 3$, $n_{\cdot 1} = 4$, $n_{\cdot 2} = 1$, $n_{\cdot 3} = 0$, $n_1 = 3$, $n_2 = \frac{5}{2}$, $n_3 = \frac{5}{2}$, and $n_{\cdot} = 8$

The decompositions and the center cone $\mathcal{P}_c$ become:
Case (i): $I = \{1\}\dot{\cup}\{2\}\dot{\cup}\{3\}$ and $\mathcal{P}_c = \mathbb{R}_+^3$
Case (ii): $I = \{1,2\}\dot{\cup}\{3\}$ and $\mathcal{P}_c = \mathcal{P}(\{1,2\},\mathbb{R}) \times \mathbb{R}_+$
Case (iii) $I = \{1\}\dot{\cup}\{2,3\}$ and $\mathcal{P}_c = \mathbb{R}_+ \times \mathcal{P}(\{2,3\},\mathbb{R})$.

The conditions (5.6) are:
Case (i): $\lambda_1 > 0, \quad \lambda_2 > \frac{1}{2}, \quad \lambda_3 > \frac{3}{2}$
Case (ii): $\lambda_{12} > \frac{1}{2}, \quad \lambda_3 > 2$
Case (iii): $\lambda_1 > 0, \quad \lambda_{23} > \frac{3}{2}$.

The Wishart distributions are:

Case(i) : $$dW_{\Sigma,(\lambda_1,\lambda_2,\lambda_3)}(S)$$

$$= \frac{\lambda_1^{\lambda_1}\lambda_2^{\lambda_2}\lambda_3^{\lambda_3} s_{11}^{\lambda_1-\frac{5}{2}} s_{[2]\bullet}^{\lambda_2-2} s_{[3]\bullet}^{\lambda_3-\frac{5}{2}}}{\pi^2\Gamma(\lambda_1)\Gamma(\lambda_2-\frac{1}{2})\Gamma(\lambda_3-\frac{3}{2})\sigma_{11}^{\lambda_1}\sigma_{[2]\bullet}^{\lambda_2}\sigma_{[3]\bullet}^{\lambda_3}} \exp\{-\mathrm{tr}(\Sigma^{-(\lambda_1,\lambda_2,\lambda_3)}S)\}$$

Case(ii) : $$dW_{\Sigma,(\lambda_{12},\lambda_3)}(S)$$

$$= \frac{\lambda_{12}^{2\lambda_{12}}\lambda_3^{\lambda_3}|S_{\{12\}}|^{\lambda_{12}-\frac{5}{2}} s_{[3]\bullet}^{\lambda_3-3}}{\pi^{\frac{5}{2}}\Gamma(\lambda_{12})\Gamma(\lambda_{12}-\frac{1}{2})\Gamma(\lambda_3-2)|\Sigma_{\{12\}}|^{\lambda_{12}}\sigma_{[3]\bullet}^{\lambda_3}} \exp\{-\mathrm{tr}(\Sigma^{-(\lambda_{12},\lambda_3)}S)\}$$

Case(iii) : $$dW_{\Sigma,(\lambda_1,\lambda_{23})}(S)$$

$$= \frac{\lambda_1^{\lambda_1}\lambda_{23}^{2\lambda_{23}} s_{11}^{\lambda_1-3} S_{[23]\bullet}^{\lambda_{23}-\frac{5}{2}}}{\pi^{\frac{5}{2}}\Gamma(\lambda_1)\Gamma(\lambda_{23}-1)\Gamma(\lambda_{23}-\frac{3}{2})\sigma_{11}^{\lambda_1}\Sigma_{[23]\bullet}^{\lambda_{23}}} \exp\{-\mathrm{tr}(\Sigma^{-(\lambda_1,\lambda_{23})}S)\},$$

where $|S_{\{12\}}|$ is the determinant of the $\{1,2\} \times \{1,2\}$ submatrix of $S \in \mathcal{P}$, $|S_{[23]\bullet}| := s_{[2]\bullet}s_{[3]\bullet}$, and similarly for $\Sigma \in \mathcal{P}$.

7. Conclusion and further research

The basic theory of the general Wishart distributions on homogeneous cones presented in the present paper opens several interesting new areas in multivariate statistical analysis.

a) The general solution to testing problems arising within the class of Wishart distribution on homogeneous cones is under preparation, cf. Andersson (2001). The results contain a classification of the test problems, the abstract central distribution of the maximal invariant statistics, the explicit likelihood ratio (LR) statistics, and the central distribution of the LR statistics in terms of its moments.

The representation of the maximal invariant statistics in several special cases leads to new eigenvalue problems, the interpretations of the eigenvalues, and their central distribution.

The non-central distribution of the maximal invariant statistics in general expressed as an integral will probably in several special cases lead to new hypergeometric functions and their expansion into sums of new zonal polynomials.

b) Further properties of the general Wishart distributions are under investigation. One result is a generalization of the following well-known result for the classical Wishart distribution: Let S follow the Wishart distribution (1.3). Let $I = I_1 \dot{\cup} I_2$ be a decomposition of the set I, and let $S = (S_{ij} | i, j = 1, 2)$ be the corresponding partitions of S. Then it follows that $S_{22\bullet1} := S_{22} - S_{21}S_{11}^{-1}S_{12}$ and $(S_{21}S_{11}^{-1}, S_{11})$ are independent, $S_{22\bullet1}$ follows a Wishart distributions, the conditional distribution of $S_{21}S_{11}^{-1}$ given S_{11} follows a normal distribution, and S_{11} follows a Wishart distribution. The parameters in these distributions are simple expressions in Σ, λ, I_1, and S_{11}.

c) The result in b) leads to a new and natural unique representation of homogeneous cones and their Wishart distributions. This representations arises from a generalization of (4.5) into block matrices. In particular $\mathcal{D}^+$ is replaced by block diagonal matrices where each diagonal block is an indecomposable symmetric cone, i.e., one of the five types in Examples 6.2, 6.3, 6.4, 6.6, and 6.8. The generalization of (4.7) then gives the new representation of a homogeneous cone. This induces a natural decomposition of any generalized Wishart distribution into Wishart distributions on symmetric cones (only five different types), and conditional normal distributions.

d) An interesting subclass of Vinberg algebras occurs under the added requirement that all non-trivial entries in the arrays are real numbers. As mentioned in Example 6.14 the corresponding homogeneous cones arise from the set of covariance matrices in the lattice models introduced by Andersson and Perlman (1993). These special Wishart distributions are under investigation.

e) The decomposition of the Wishart distribution mentioned in c) and the restriction of the theory mentioned in d) together suggest another generalization of Wishart distributions to cones arising from the set of unknown covariance matrices given by conditional independence restrictions induced by acyclic directed graphs or more generally by chain graphs, cf. Andersson and Perlman (2001) and Lauritzen (1996). These topics are also under investigation.

f) From Remark 3.2 it is seen that the role of the classical inverse Wishart distribution as a prior distribution can be generalized to the Wishart models in the present paper. Thus it is interesting to evaluate the solution to

Bayesian inference with conjugate prior distribution in the many new examples. Furthermore the generalized inverse Wishart distribution may give new contributions to the so-called Bayesian networks. The latter requires that the extension mentioned in e) be used.

Acknowledgments. The first author wishes to thank Gerard Letac and Helene Massam for insightful discussions, in particular during 3 months at the Fields Institute, Toronto, Canada in 1999. He also wishes to thank Hans Brøns and Søren Tolver Jensen for the inspiration for the present work, originating from the countless hours they worked together on "an algebraic theory for normal models" at the University of Copenhagen (1970-1989). Homogeneous cones were already then a part of their discussions. Additionally, he wishes to thank Friedrich Pukelsheim for the hospitality and support he provided while visiting Augsburg Universität in the first 3 months of 2001. The present work was finished in that period. Both authors also thank Michael Perlman for his careful reading of the manuscript and his many suggestions for improving the presentation.

References

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis,* 2nd ed. Wiley, New York.

Anderson, T. W. (1969). Statistical inference for covariance matrices with linear structure. In: *Proc. Second Internat. Sympos. Multivariate Analysis*, Ed: P. R. Krishnaiah, Academic Press, New York, 55–66.

Anderson, T. W. (1970). Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices. In: *Essays in Probability and Statistics*, Eds: R. C. Bose, I. M. Chakravati, P. C. Mahalanobis, C. R. Rao, and K. J. C. Smith, Univ. North Carolina Press, Chapel Hill, 1–24.

Anderson, T. W. (1973). Asymptotically efficient estimation of covariances with linear structure. *Ann. Statist.* **1**, 135–141.

Andersson, S. A. (1975a). *Forelæsningsnoter i flerdimensional statististisk analyse. Lecture notes.* Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

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

Andersson, S. A. (1976). *Forelæsningsnoter i flerdimensional statististisk analyse. Lecture notes.* Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

Andersson, S. A. (1978). *On the mathematical foundation of multivariate statistical analysis. Lecture Notes.* Inst. Mathematical Statistics, Univ. Copenhagen.

Andersson, S. A. (1992). *Normal linear models given by group symmetry. Lecture Notes.* Inst. Mathematical Statistics, Univ. Copenhagen.

Andersson, S. A. (2001). Hypothesis testing for the general Wishart distribution on homogeneous cones. (In preparation)

Andersson, S. A., Brøns, H. K. and Tolver Jensen, S. (1975). En algebraisk teori for normale statistiske modeller. Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

Andersson, S. A., Brøns, H. K. and Tolver Jensen, S. (1983). Distribution of Eigenvalues in multivariate statistical analysis. *Ann. Statist.* **11**, 392–415.

Andersson, S. A., Madigan, D., Perlman, M. D. and Triggs, C. M. (1995). On the relation between conditional independence models determined by finite distributive lattice and directed acyclic graphs. *J. Statist. Plann. Inference* **48**, 25–46.

Andersson, S. A., Madigan, D., Perlman, M. D. and Triggs, C. M. (1997). A graphical characterization of lattice conditional independence models. *Artificial Intelligence* **21**, 27–50.

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

Andersson, S. A. and Perlman, M. D. (1993). Lattice models for conditional indepence in a multivariate normal distribution. *Ann. Statist.* **21**, 1318–1358.

Andersson, S. A. and Perlman, M. D. (1994). Normal linear models with lattice conditional independence restrictions; In: *Multivariate Analysis and Its Applications* **24**, 97–110.

Andersson, S. A., Madigan, D. and Perlman, M. D. (2001). Alternative Markov properties for chain graphs. *Scand. J. Statist.* **28**, 33–85.

Andersson, S. A., Letac, G. and Massam, H. (2001). The variance of the Wishart distributions on homogeneous cones. (In preparation)

Arnold, S. F. (1973). Application of the theory of products of certain patterned covariance matrices. *Ann. Statist.* **1**, 682–699.

Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory.* Wiley, New York.

Bourbaki, N. (1963). *Éléments de Mathématique.* Intégration, Ch. 7 à 8. Herman, Paris.

Bourbaki, N. (1971). *Éléments de Mathématique.* Topologie Générale, Ch. 1 à 4. Herman, Paris.

Braun, H. and Koecher, M. (1966). *Jordan-Algebren.* Springer, Berlin.

Brøns, H. K. (1969). Forelæsninger i den normale fordelings teori. Inst. Mathematical Statistics, Univ. Copenhagen. (Personal communication, in Danish)

Brøns, H. K. (1998). *Flerdimensional statistisk analyse. Manuscript.* Dept. of Mathematical Statistics, Univ. Copenhagen. (In Danish)

Casalis, M. and Letac, G. (1996). The Lukacs-Olkin-Rubin characterization of the Wishart distribution on symmetric cones. *Ann. Statist.* **24**, 763–786.

Faraut, J. and Korányi, A. (1994). *Analysis on Symmetric Cones.* Oxford University Press, New York.

Goodman, N. R. (1963). Statistical analysis based on certain multivariate complex distributions. *Ann. Math. Statist.* **34**, 152–176.

Jacobson, N. (1968). *Structure and Representations of Jordan Algebras.* Amer. Math. Soc., Providence, R.I.

Jensen, J. L. (1991). A large deviation-type approximation for the "Box class" of likelihood ratio criteria. *J. Amer. Statist. Assoc.* **86**, 437–440.

Khatri, C. G. (1965a). Classical statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Statist.* **36**, 98–114.

Khatri, C. G. (1965b). A test of reality of a covariance matrix in certain complex Gaussian distributions. *Ann. Math. Statist.* **36**, 115–119.

Lauritzen, S. L. (1996). *Graphical Models.* Oxford University Press, Oxford.

Letac, G. and Massam, H. (1998). Quadratic and inverse regression for the Wishart distribution. *Ann. Statist.* **26**, 573–595.

Letac, G. and Massam, H. (2000). Representations of the Wishart distributions. *Contemp. Math.* **261**, 121–142.

Massam, H. and Neher, E. (1997). On transformations and determinants of Wishart variables on symmetric cones. *J. Theoret. Probab.* **10**, 867–902.

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

Olkin, I. (1973). Testing and estimation for structures which are circularly symmetric in blocks. In: *Proc. Res. Sem. Multivariate Statistical Inference. Dalhousie, Nova Scotia*, North Holland, Amsterdam, 183–195.

Perlman, M. D. (1987). Group symmetry covariance models. *Statistical Science* **2**, 421–425.

Seely, J. (1971). Quadratic subspaces and completeness. *Ann. Math. Statist.* **42**, 710–721.

Seely, J. (1972). Completeness for a family of multivariate normal distributions. *Ann. Math. Statist.* **43**, 1644–1647.

Tolver Jensen, S. (1973). *Forelæsninger i matematisk statistik. Lecture notes.* Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

Tolver Jensen, S. (1974). *Forelæsningsnoter flerdimensional statistisk analyse. Lecture notes.* Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

Tolver Jensen, S. (1977). *Flerdimensional statistisk analyse. Lecture notes.* Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

Tolver Jensen, S. (1983). *Symmetrimodeller. Lecture notes.* Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish)

Tolver Jensen, S. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist.* **16**, 302–322.

Vinberg, E. B. (1960). Homogeneous cones. *Dokl. Akad. Nauk SSSR* **133**, 9–12; *Soviet Math. Dokl.* **1**, 787–790.

Vinberg, E. B. (1962). Automorphisms of homogeneous convex cones. *Dokl. Akad. Nauk SSSR* **143**, 265–268; *Soviet Math. Dokl.* **3**, 371–374.

Vinberg, E. B. (1963). The theory of convex homogeneous cones. *Trudy Moskov. Mat. Obshch.* **12**, 303–358; *Trans. Moscow Math. Soc.* **12**, 340–403.

Vinberg, E. B. (1965). The structure of the group of automorphisms of a homogeneous convex cone. *Trudy Moskov. Mat. Obshch.* **13**, 56–83; *Trans. Moscow Math. Soc.* **13**, 63–93.

Votaw, D. F. (1948). Testing compound symmetry in a normal multivariate distribution. *Ann. Math. Statist.* **19**, 447–473.

Wilks, S. S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Ann. Math. Statist.* **17**, 257–281.

Wishart, J. (1928). The generalized product moment distribution in a sample from a normal multivariate population. *Biometrika* **20A**, 32–52.

Wojnar, G. G. (2000). *Generalized Wishart Models on Convex Homogeneous Cones.* Ph.D. Thesis, Indiana University.

Department of Mathematics, Indiana University, Bloomington, IN 47405, USA

E-mail address: `standers@indiana.edu`

Department of Mathematics, Frostburg State University, Frostburg, MD 21532, USA

E-mail address: `gwojnar@frostburg.edu`