Ann Inst Stat Math (2019) 71:439–471
https://doi.org/10.1007/s10463-018-0647-z

Wishart exponential families on cones related to tridiagonal matrices

Piotr Graczyk[1] · **Hideyuki Ishi**[2] · **Salha Mamane**[3]

Received: 28 July 2016 / Revised: 5 January 2018 / Published online: 23 February 2018

Abstract Let G be the graph corresponding to the graphical model of nearest neighbor interaction in a Gaussian character. We study Natural Exponential Families (NEF) of Wishart distributions on convex cones Q_G and P_G, where P_G is the cone of tridiagonal positive definite real symmetric matrices, and Q_G is the dual cone of P_G. The Wishart NEF that we construct include Wishart distributions considered earlier for models based on decomposable(chordal) graphs. Our approach is, however, different and allows us to study the basic objects of Wishart NEF on the cones Q_G and P_G. We determine Riesz measures generating Wishart exponential families on Q_G and P_G, and we give the quadratic construction of these Riesz measures and exponential families. The mean, inverse-mean, covariance and variance functions, as well as moments of higher order, are studied and their explicit formulas are given.

Keywords Wishart distribution · Graphical model · Nearest neighbor interaction

✉ Piotr Graczyk
piotr.graczyk@univ-angers.fr

Hideyuki Ishi
hideyuki@math.nagoya-u.ac.jp

Salha Mamane
Salha.Mamane@wits.ac.za

1 LAREMA, University of Angers, 2 Boulevard Lavoisier, 49045 Angers Cedex 01, France

2 Graduate School of Mathematics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan

3 School of Statistics and Actuarial Science, University of the Witwatersrand, Private Bag X3, Wits, Johannesburg 2050, South Africa

1 Introduction

The classical Wishart distribution was first derived by Wishart (1928) as the distribution of the maximum likelihood estimator of the covariance matrix of the multivariate normal distribution. In the framework of graphical Gaussian models, the distribution of the maximum likelihood estimator of $\pi(\Sigma)$, where π denotes the canonical projection onto Q_G, was derived by Dawid and Lauritzen (1993), who called it the hyper Wishart distribution. Dawid and Lauritzen (1993) also considered the hyper-inverse Wishart distribution which is defined on Q_G as the Diaconis-Ylvisaker conjugate prior distribution for $\pi(\Sigma)$, and Roverato (2000) derived the so-called G-Wishart distribution on P_G, that is, the distribution of the concentration matrix $K = \Sigma^{-1}$ when $\pi(\Sigma)$ follows the hyper-inverse Wishart distribution. Letac and Massam (2007) constructed two classes of multi-parameter Wishart distributions on the cones Q_G and P_G associated with a decomposable(chordal) graph G and called them type I and type II Wishart distributions, respectively. They are more flexible because they have multiple shape parameters. In fact, the type I and type II Wishart distributions generalize the hyper Wishart distribution and the G-Wishart distribution, respectively.

The Wishart exponential families introduced and studied in this paper include the type I and type II Wishart distributions of Letac–Massam on the cones Q_G and P_G associated with the so-called path graphs $\overset{v_1}{\bullet} - \overset{v_2}{\bullet} - \cdots - \overset{v_n}{\bullet}$. Although path graphs are often denoted by P_n in the literature (e.g. Bondy and Murty 2008), we shall use the Dynkin diagram notation A_n for the path graph. This notation is well known in other fields of mathematics. In mathematical statistics, this notation is used by Letac and Massam (2007).

Our methods, which are new and different from methods of articles cited above, simplify in a significant way the Wishart theory for graphical models.

In Graczyk and Ishi (2014) and Ishi (2014), the theory of Wishart distributions on general convex cones was developed, with a strong accent on the quadratic constructions and on applications to homogeneous cones. In this article, we apply for the first time the ideas and results of Graczyk and Ishi (2014) to study important families of non-homogeneous cones.

Applications in estimation and other practical aspects of Wishart distributions are intensely studied, cf. Sugiura and Konno (1988), Tsukuma and Konno (2006), Konno (2007, 2009), Kuriki and Numata (2010).

The focus of this work is on non-homogeneous cones Q_{A_n} and P_{A_n} appearing in the statistical theory of graphical models, corresponding to the practical model of nearest neighbor interactions. In the Gaussian character $(X_1, X_2, \ldots, X_n)$, non-neighbors X_i, X_j, $|i-j| > 1$ are conditionally independent with respect to other variables. This family of decomposable graphical models presents many advantages: it encompasses the univariate case (A_1), a complete graph (A_2), a non-complete homogeneous graph (A_3) and an infinite number of non-homogeneous graphs $(A_n, n \geq 4)$.

Some of the results of our research may be extended to cones related to all chordal graphs (work in progress), e.g., the definitions of the power functions, which are described (see Definition 3) in general graph terminology. However, there exist difficulties to extend numerous results of this paper to all chordal graphs, since a part of the proposed methods does not work in the general setting.

Our results on the indexation of Riesz and Wishart measures by $M = 1, \dots, n$, on the inverse mean map and on the variance function are specific for the cones Q_{A_n} and P_{A_n}. Thanks to the indexation result, the Laplace transforms attributed to all the possible eliminating orders are treated in a simple and uniform way. A striking illustration of the importance and of the special role of path graphs A_n is the fact that the methods and results given in this article make it possible to proceed and to solve the Letac–Massam Conjecture (Letac and Massam 2007) for the cones Q_{A_n} (Graczyk et al. 2017) and they give much hope to prove the Letac–Massam Conjecture for the cones P_{A_n} (work in progress). Together with the results of this article, we achieve in this way the complete study of all classical objects of an exponential family for the Wishart NEF on the cones Q_{A_n}.

For all these reasons, we decided to deal in this article exclusively with the path graphs A_n. It is clear that a complete comprehension of this important class of graphs will contribute greatly to the research in mathematical statistics for chordal graphical models.

Plan of the article. Sections 2, 3 and 4 provide the main tools in order to define and to study the Wishart NEF on the cones Q_{A_n} and P_{A_n}. In Sect. 2, useful notions of eliminating orders $\prec$ on A_n and of generalized power functions $\delta_{\underline{s}}^{\prec}$ and $\Delta_{\underline{s}}^{\prec}$, $\underline{s} \in \mathbb{R}^n$ will be introduced on the cones Q_{A_n} and P_{A_n} respectively. In Theorem 1, a classical relation between the power functions $\delta_{\underline{s}}^{\prec}$ and $\Delta_{-\underline{s}}^{\prec}$ is proved as well as the dependence of $\delta_{\underline{s}}^{\prec}$ and $\Delta_{\underline{s}}^{\prec}$ on the maximal element M of $\prec$ only. Thus, in the sequel of the paper, only generalized power functions $\delta_{\underline{s}}^{(M)}$ and $\Delta_{\underline{s}}^{(M)}$ appear. Next important tool of analysis of Wishart exponential families are recurrent construction of the cones P_G and Q_G and corresponding changes of variables. They are introduced and studied in Sect. 3, and are immediately applied in Sect. 4 in order to compute the Laplace transform of generalized power functions $\delta_{\underline{s}}^{(M)}$ and $\Delta_{\underline{s}}^{(M)}$ (Theorems 2 and 3).

In Sect. 5, Wishart natural exponential families on the cones Q_{A_n} are defined, and all their classical objects are explicitly determined, beginning with the Riesz generating measures, Wishart densities, Laplace transform, mean and covariance. In Theorem 4 and Corollary 3, an explicit formula for the inverse mean map is proved. A key formula is obtained in Proposition 6, whence Theorem 4 follows by standard argument in information geometry of exponential family (Amari and Nagaoka 2007; Brown 1986; Speed and Kiiveri 1986). It provides an infinite number of versions of Lauritzen formulas for bijections between the cones Q_G and P_G. In Sect. 5.3, two explicit formulas are given for the variance function of a Wishart family. The formula of Theorem 5 is surprisingly simple and similar to the case of the symmetric cone S_n^+. Sections 5.4 and 5.5 are devoted to the quadratic constructions of Wishart exponential families on Q_G and to the computation of their higher moments in Theorem 6.

Section 6 is on Wishart natural exponential families on the cones P_{A_n} and follows a similar scheme as Sect. 5; however, the inverse mean map and variance function are not available on the cones P_{A_n}. The analysis on these cones is more difficult.

In the last Sect. 7, we establish the relations of the Wishart NEF defined and studied in our paper with the type I and type II Wishart distributions from Letac and Massam

(2007). Our methods give a simple proof of the formulas for Laplace transforms of type I and type II Wishart distributions from Letac and Massam (2007).

2 Preliminaries on A_n graphs and related cones

In this section, we study properties of graphs A_n that will be important in the theory of Riesz measures and Wishart distributions on the cones related to these graphs. In particular, we characterize all the eliminating orders of vertices and we introduce generalized power functions related to such orders. We show that they only depend on the maximal element $M \in \{1, \dots, n\}$ of the order.

An undirected graph is a pair $G = (V, \mathcal{E})$, where V is a finite set and $\mathcal{E}$ is a subset of $\mathcal{P}_2(V)$, the set of all subsets of $\mathcal{E}$ with cardinality two. The elements of V are called nodes or vertices, and the elements of $\mathcal{E}$ are called edges. If $\{v, v'\} \in \mathcal{E}$, then v and v' are said to be adjacent and this is denoted by $v \sim v'$. Graphs are visualized by representing each node by a point and each edge $\{v, v'\}$ by a line with the nodes v and v' as endpoints. For convenience, we introduce a subset $E \subset V \times V$ defined by $E := \{(v, v') : v \sim v'\} \cup \{(v, v) : v \in V\}$.

The graph $G = (V, \mathcal{E})$ with set of vertices $V = \{v_1, v_2, \dots, v_n\}$ and set of edges $\mathcal{E} = \{\{v_j,\ v_{j+1}\} : 1 \le j \le n-1\}$ is denoted by A_n. In what follows, we often denote the vertex v_i by i and the graph A_n : $\overset{v_1}{\bullet} - \overset{v_2}{\bullet} - \cdots - \overset{v_n}{\bullet}$ is simply represented as $1-2-3-\ldots-n$. An n-dimensional Gaussian model $(X_v)_{v\in V}$ is said to be Markov with respect to a graph G if for any $(v,\ v') \notin E$, the random variables X_v and $X_{v'}$ are conditionally independent given all the other variables. The conditional independence relations encoded in A_n graph are of the form: $X_{v_i} \perp\!\!\!\perp X_{v_j} | (X_{v_k})_{k \neq i,j}$, for all $|i-j| > 1$. Thus, A_n graphs correspond to nearest neighbor interaction models.

Let S_n be the space of real symmetric matrices of order n and let $S_n^+ \subset S_n$ be the cone of positive definite matrices. The notation for a positive definite matrix y is $y > 0$. For a graph G, let $Z_G \subset S_n$ be the vector space consisting of $y \in S_n$ such that $y_{ij} = 0$ if $(i\ ,\ j) \notin E$. Let $I_G = Z_G^*$ be the dual vector space with respect to the scalar product $\langle y, \eta \rangle = \mathrm{tr}(y\eta) = \sum_{(i,j)\in E} y_{ij}\eta_{ij},\quad y \in Z_G,\ \eta \in I_G$. In the statistical literature, the vector space I_G is commonly realized as the space of $n \times n$ symmetric matrices η, in which only the coefficients η_{ij}, $(i, j) \in E$, are given. We adopt this realization of I_G in this paper.

If $I \subset V$, we denote by y_I the submatrix of $y \in Z_G$ obtained by extracting from y the lines and the columns indexed by I. The same notation is used for $\eta \in I_G$. Let P_G be the cone defined by $P_G = \{y \in Z_G : y > 0\}$, and $Q_G \subset I_G$ the dual cone of P_G, that is, $Q_G = \{\eta \in I_G :\ \ \forall y \in \overline{P_G}\backslash\{0\}\ \ \langle y, \eta\rangle > 0\}$. A Gaussian vector model $(X_v)_{v\in V}$ is Markov with respect to G if and only if the concentration matrix $K = \Sigma^{-1}$ belongs to P_G. When $G = A_n$, the cone Q_G is described as $Q_G = \{\eta \in I_G : \eta_{\{i,i+1\}} > 0,\ i = 1, \dots, n-1\}$. Let $\pi = \pi_{I_G}$ be the projection of S_n onto I_G, $x \mapsto \eta$ such that $\eta_{ij} = x_{ij}$ if $(i, j) \in E$. Then, it is known (cf. Letac and Massam 2007; Andersson and Klein 2010) that the mapping $P_G \longrightarrow Q_G$, $y \longmapsto \pi(y^{-1})$ is a bijection.

Example 1 Consider the graph A_4 : $\overset{1}{\bullet} - \overset{2}{\bullet} - \overset{3}{\bullet} - \overset{4}{\bullet}$.

The elements of P_{A_4} are positive definite tridiagonal matrices of the form $y = \begin{pmatrix} y_{11} & y_{12} & 0 & 0 \\ y_{12} & y_{22} & y_{23} & 0 \\ 0 & y_{23} & y_{33} & y_{34} \\ 0 & 0 & y_{34} & y_{44} \end{pmatrix}$. The elements of Q_{A_4} are incomplete matrices of the form $x = \begin{pmatrix} x_{11} & x_{12} & * & * \\ x_{12} & x_{22} & x_{23} & * \\ * & x_{23} & x_{33} & x_{34} \\ * & * & x_{34} & x_{44} \end{pmatrix}$, with $x_{\{1,2\}} = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix}$, $x_{2:3} = \begin{pmatrix} x_{22} & x_{23} \\ x_{23} & x_{33} \end{pmatrix}$ and $x_{3:4} = \begin{pmatrix} x_{33} & x_{34} \\ x_{34} & x_{44} \end{pmatrix}$ positive definite matrices.

The matrix $z = y^{-1} = \begin{pmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{12} & z_{22} & z_{23} & z_{24} \\ z_{13} & z_{23} & z_{33} & z_{34} \\ z_{14} & z_{24} & z_{34} & z_{44} \end{pmatrix}$ is a positive definite matrix that does not belong to P_{A_4}. The projection of z on Q_{A_4} is $\pi(z) = \begin{pmatrix} z_{11} & z_{12} & * & * \\ z_{12} & z_{22} & z_{23} & * \\ * & z_{23} & z_{33} & z_{34} \\ * & * & z_{34} & z_{44} \end{pmatrix}$. π can be understood as the operator to ignore the elements of z corresponding to non-adjacent vertices.

In the sequel, unless otherwise stated, $G = A_n$,

2.1 Eliminating orders

Different orders of vertices $v_1, v_2, \ldots, v_n$ should be considered in order to have a harmonious theory of Riesz and Wishart distributions on the cones related to A_n graphs. The orders that will be important in this work are called *eliminating orders of vertices* and will be presented now.

Definition 1 Consider a graph $G = (V, \mathcal{E})$ and an ordering $\prec$ of the vertices of G. The set of future neighbors of a vertex v is defined as $v^+ = \{w \in V : v \prec w \text{ and } v \sim w\}$. The set of all predecessors of a vertex $v \in V$ with respect to $\prec$ is defined as $v^- = \{u \in V : u \prec v\}$.

Definition 2 An ordering $\prec$ of the vertices of a graph G is said to be an eliminating order if v^+ is complete for all $v \in V$.

In this section, we present a characterization of the eliminating orders in the case of the graph A_n. An algorithm that generates all eliminating orders for a general graph is given by (Chandran et al. 2003).

Proposition 1 *Consider a graph $A_n : 1 - 2 - 3 - \cdots - n$. All eliminating orders are obtained by an intertwining of two sequences $1 \prec \ldots \prec M$ and $n \prec \cdots \prec M$ for an $M \in V$. There are 2^{n-1} eliminating orders on the graph A_n.*

The proof is easy and is omitted.

2.2 Generalized power functions

In this section, we define and study generalized power functions on the cones P_G and Q_G. Here, we introduce useful notation. For $1 \le i \le j \le n$, let $\{i : j\} \subset V$ be the set of $a \in V$ for which $i \le a \le j$. Then, for $y \in Z_G$ and $1 \le i \le n$, the matrix $y_{\{1:i\}}$ is the upper left submatrix of y of size i, and $y_{\{i:n\}}$ is the lower right submatrix of size $n - i + 1$. Recall that on the cone S_n^+, the generalized power functions are $\Delta_{\underline{s}}(y) = \prod_{i=1}^n |y_{\{1:i\}}|^{s_i - s_{i+1}}$ and $\delta_{\underline{s}}(y) = \prod_{i=1}^n |y_{\{i:n\}}|^{s_i - s_{i-1}}$, with $s_0 = s_{n+1} = 0$.

Definition 3 For $\underline{s} \in \mathbb{R}^V$, setting $\det y_\emptyset = 1 = \det \eta_\emptyset$, we define

$$\Delta_{\underline{s}}^{\prec}(y) := \prod_{v \in V} \left(\frac{\det y_{\{v\} \cup v^-}}{\det y_{v^-}} \right)^{s_v} \qquad (y \in P_G), \tag{1}$$

$$\delta_{\underline{s}}^{\prec}(\eta) := \prod_{v \in V} \left(\frac{\det \eta_{\{v\} \cup v^+}}{\det \eta_{v^+}} \right)^{s_v} \qquad (\eta \in Q_G). \tag{2}$$

Note that Definition 3 applied to the complete graph with the usual order $1 < \ldots < n$ gives $\Delta_{\underline{s}}$ and $\delta_{\underline{s}}$. For any $\underline{s}$, the following formula $\delta_{\underline{s}}(y^{-1}) = \Delta_{-\underline{s}}(y)$ is well known. In Theorem 1, we find an analogous formula in the case of the cones P_G and Q_G.

We will see in Theorem 1 that on the cones related to the graphs A_n, different order-depending power functions $\Delta_{\underline{s}}^{\prec}$ and $\delta_{\underline{s}}^{\prec}$ defined in Definition 3 may be expressed in terms of explicit "M-power functions" $\Delta_{\underline{s}}^{(M)}$ and $\delta_{\underline{s}}^{(M)}$ that will be defined below. They depend only on the choice of $M \in V$.

Definition 4 Let $M \in V$, $y \in P_G$ and $\eta \in Q_G$. We define the M-power functions $\Delta_{\underline{s}}^{(M)}(y)$ on P_G and $\delta_{\underline{s}}^{(M)}(x)$ on Q_G by the following formulas:

$$\Delta_{\underline{s}}^{(M)}(y) = \prod_{i=1}^{M-1} |y_{\{1:i\}}|^{s_i - s_{i+1}} |y|^{s_M} \prod_{i=M+1}^{n} |y_{\{i:n\}}|^{s_i - s_{i-1}}, \tag{3}$$

$$\delta_{\underline{s}}^{(M)}(\eta) = \frac{\prod_{i=1}^{M-1} |\eta_{\{i:i+1\}}|^{s_i} \prod_{i=M+1}^{n} |\eta_{\{i-1:i\}}|^{s_i}}{\prod_{i=2}^{M-1} \eta_{ii}^{s_{i-1}} \cdot \eta_{MM}^{s_{M-1} - s_M + s_{M+1}} \cdot \prod_{i=M+1}^{n-1} \eta_{ii}^{s_{i+1}}}. \tag{4}$$

Observe that for $M = 1, n$ there are $n - 1$ factors in the denominator of (4), and for $M = 2, \ldots n - 1$ there are $n - 2$ factors (powers of $\eta_{22} \ldots \eta_{n-1,n-1}$).

The main result of this section is the following theorem.

Theorem 1 *Consider a graph $G = A_n$ with an eliminating order $\prec$. Let M be the maximal element with respect to $\prec$. Then for all $y \in P_G$, we have*

$$\delta_{\underline{s}}^{\prec}(\pi(y^{-1})) = \Delta_{-\underline{s}}^{\prec}(y) = \Delta_{-\underline{s}}^{(M)}(y). \tag{5}$$

The proof of Theorem 1 is preceded by a series of elementary lemmas.

Lemma 1 *Let $y \in P_G$ and $i < j < j+1 < k < m$. The determinant of the submatrix $y_{\{i:j\}\cup\{k:m\}}$ can be factorized as $|y_{\{i:j\}\cup\{k:m\}}| = |y_{\{i:j\}}||y_{\{k:m\}}|$.*

Lemma 2 *Let $y \in P_G$ and $\eta = \pi(y^{-1})$. Then for all $i, i+1 \in V$, we have*

$$\left|\eta_{\{i,i+1\}}\right| = |y|^{-1}|y_{V\setminus\{i,i+1\}}|.$$

Proof (of Theorem 1) *Part 1: $\delta_{\underline{s}}^{\prec}(\pi(y^{-1})) = \Delta_{-\underline{s}}^{(M)}(y)$.* From Proposition 1, we have

$$i^+ = \begin{cases} \{i+1\} & \text{if } i \leq M-1, \\ \emptyset & \text{if } i = M, \\ \{i-1\} & \text{if } i \geq M+1. \end{cases}$$

Using $\eta_{ii} = |y|^{-1}|y_{V\setminus\{i\}}|$ with $\eta = \pi(y^{-1})$ and Lemmas 1 and 2, we get $\delta_{\underline{s}}^{\prec}(\pi(y^{-1})) = \Delta_{-\underline{s}}^{(M)}(y)$. *Part 2: $\Delta_{\underline{s}}^{\prec}(y) = \Delta_{\underline{s}}^{(M)}(y)$.* Let us first consider the eliminating order $\prec_M$ given by

$$1 \prec_M 2 \prec_M \ldots \prec_M M-1 \prec_M n \prec_M n-1 \prec_M \ldots \prec_M M+1 \prec_M M. \quad (6)$$

Using $\eta_{ii} = |y|^{-1}|y_{V\setminus\{i\}}|$, Lemmas 1 and 2 again, we get $\Delta_{\underline{s}}^{\prec_M}(y) = \Delta_{\underline{s}}^{(M)}(y)$.

It is easy to see using Proposition 1 and the factorization from Lemma 1 that for any other eliminating order $\prec$, the factors of $\Delta_{\underline{s}}^{\prec}(y)$ under the powers s_i are exactly the same as for $\prec_M$. Indeed, if $i \leq M-1$, let $n-j$ be the largest vertex greater than M such that $n-j \prec i$. Then, the factor under the power s_i is

$$\frac{|y_{\{i\}\cup i^-}|}{|y_{i^-}|} = \frac{|y_{\{1:i\}}||y_{\{n-j:n\}}|}{|y_{\{1:i-1\}}||y_{\{n-j:n\}}|} = \frac{|y_{\{1:i\}}|}{|y_{\{1:i-1\}}|}.$$

A similar argument shows that this is also true for $i = M$ and for $i > M$. □

Corollary 1 *Let $\prec_1$ and $\prec_2$ be two eliminating orders on G such that* $\max_{\prec_1} V = \max_{\prec_2} V$. *Then $\delta_{\underline{s}}^{\prec_1}(\eta) = \delta_{\underline{s}}^{\prec_2}(\eta)$ for all $\eta \in Q_G$. If* $\max_{\prec} V = M$ *then we have $\delta_{\underline{s}}^{\prec}(\eta) = \delta_{\underline{s}}^{(M)}(\eta)$.*

3 Recurrent construction of the cones P_G and Q_G and changes of variables

In this section, we introduce very useful recurrent constructions of the cones P_{A_n} and Q_{A_n} from the cones $P_{A_{n-1}}$ and $Q_{A_{n-1}}$. There are two variants of them for $A_{n-1} : 2 - \cdots - n$ and $A_{n-1} : 1 - \cdots - (n-1)$. Corresponding changes of variables for integration on P_{A_n} and Q_{A_n} are introduced.

Proposition 2 1. *For $n \geq 2$, let $\Phi_n : \mathbb{R}^+ \times \mathbb{R} \times P_{A_{n-1}} \longrightarrow P_{A_n}$, $(a, b, z) \longmapsto y$ with*

$$y = A(b) \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & z & \\ 0 & & & \end{pmatrix} A(b)^T, \qquad A(b) = \begin{pmatrix} 1 & & & \\ b & 1 & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & 1 \end{pmatrix},$$

and let $\Psi_n : \mathbb{R}^+ \times \mathbb{R} \times Q_{A_{n-1}} \longrightarrow Q_{A_n}$, $(\alpha, \beta, x) \longmapsto \eta$ with

$$\eta = \pi \left(A(\beta)^T \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & x & \\ 0 & & & \end{pmatrix} A(\beta) \right).$$

Then the maps Φ_n and Ψ_n are bijections.

2. *Let $\tilde{\Phi}_n : \mathbb{R}^+ \times \mathbb{R} \times P_{A_{n-1}} \longrightarrow P_{A_n}$, $(a, b, z) \longmapsto \tilde{y}$ with*

$$\tilde{y} = B(b)^T \begin{pmatrix} & & & 0 \\ & z & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & a \end{pmatrix} B(b), \quad B(b) = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & \dots & b & 1 \end{pmatrix},$$

and let $\tilde{\Psi}_n : \mathbb{R}^+ \times \mathbb{R} \times Q_{A_{n-1}} \longrightarrow Q_{A_n}$, $(\alpha, \beta, x) \longmapsto \tilde{\eta}$ with

$$\tilde{\eta} = \pi \left(B(\beta) \begin{pmatrix} & & & 0 \\ & x & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & \alpha \end{pmatrix} B(\beta)^T \right).$$

Then the maps $\tilde{\Phi}_n$ and $\tilde{\Psi}_n$ are bijections.

3. *The Jacobians of the changes of variables $y = \Phi_n(a, b, z)$ and $y = \tilde{\Phi}_n(a, b, z)$ are given by*

$$J_{\Phi_n}(a, b, z) = a, \quad J_{\tilde{\Phi}_n}(a, b, z) = a. \tag{7}$$

The Jacobians of the changes of variables $\eta = \Psi_n(\alpha, \beta, x)$ and $\eta = \tilde{\Psi}_n(\alpha, \beta, x)$ are given by

$$J_{\Psi_n}(\alpha, \beta, x) = x_{22}, \quad J_{\tilde{\Psi}_n}(\alpha, \beta, x) = x_{n-1,n-1}. \tag{8}$$

It should be noted that for $\Phi_n(a, b, z)$ and $\Psi_n(\alpha, \beta, x)$ the rows and columns of z and x are numbered $2, \dots, n$ while for $\tilde{\Phi}_n(a, b, z)$ and $\tilde{\Psi}_n(\alpha, \beta, x)$ they are numbered $1, \dots, n-1$.

Proof 1. Let $y' = \begin{pmatrix} a & 0 \dots 0 \\ 0 & \\ \vdots & z \\ 0 & \end{pmatrix}$ and $\eta' = \begin{pmatrix} \alpha & 0 \dots 0 \\ 0 & \\ \vdots & x \\ 0 & \end{pmatrix}$. Then

$$y_{ij} = \begin{cases} ab & \text{if } (i,j) = (1,2) \text{ or } (i,j) = (2,1), \\ ab^2 + z_{22} & \text{if } i = j = 2, \\ y'_{ij} & \text{otherwise.} \end{cases} \tag{9}$$

Thus, on the one hand, if $(a, b, z) \in \mathbb{R}^+ \times \mathbb{R} \times P_{A_{n-1}}$, then $y \in Z_{A_n}$. And $z > 0$ implies $y' > 0$ as every principal minor of y' equals a times a principal minor of z. From $y = Ty'T^T$ with $T = A(b)$, we get $y \in P_{A_n}$. On the other hand, if $y \in P_{A_n}$, we have $a = y_{11} > 0$, $b = \frac{y_{12}}{y_{11}}$, $z_{22} = y_{22} - \frac{y_{12}^2}{y_{11}}$ and $z_{ij} = y_{ij}$ for all $i \neq 2$ and $j \neq 2$. We use the notation $z = (z_{ij})_{2 \leq i,j \leq n}$. Now, let us show that $z \in P_{A_{n-1}}$. We have $y' = T^{-1} y \,(T^T)^{-1} > 0$. Hence, we have also $z > 0$ since each principal minor of z equals $1/a$ times a principal minor of y'. Therefore, the map Φ_n is indeed a bijection from $\mathbb{R}^+ \times \mathbb{R} \times P_{A_{n-1}}$ onto P_{A_n}.

Let us turn to Ψ_n. The relation between η and η' is given by

$$\eta_{ij} = \begin{cases} \alpha + \beta^2 x_{22} & \text{if } i = j = 1, \\ \beta x_{22} & \text{if } (i,j) = (1,2) \text{ or } (i,j) = (2,1), \\ \eta'_{ij} & \text{otherwise.} \end{cases} \tag{10}$$

First we show that if $(\alpha, \beta, x) \in \mathbb{R}^+ \times \mathbb{R} \times Q_{A_{n-1}}$, then $\eta \in I_{A_n}$. Actually, since $x_{\{2,3\}} > 0$, we have $\alpha + \beta^2 x_{22} > 0$ and $\eta_{\{1,2\}} = \begin{pmatrix} \alpha + \beta^2 x_{22} & \beta x_{22} \\ \beta x_{22} & x_{22} \end{pmatrix} > 0$. On the other hand, if $\eta \in Q_{A_n}$, we have $x_{ij} = \eta_{ij}$ for all $i, j = 2, \dots, n$. Thus, $\eta \in Q_{A_n}$ implies $x \in Q_{A_{n-1}}$.

2. Let $\tilde{y}' = \begin{pmatrix} & 0 \\ z & \vdots \\ & 0 \\ 0 \dots 0 & a \end{pmatrix}$ and $\tilde{\eta}' = \begin{pmatrix} & 0 \\ x & \vdots \\ & 0 \\ 0 \dots 0 & \alpha \end{pmatrix}$. Then we have

$$\tilde{y}_{ij} = \begin{cases} ab & \text{if } (i,j) = (n-1,n) \text{ or } (i,j) = (n,n-1), \\ ab^2 + z_{n-1,n-1} & \text{if } i = j = n-1, \\ \tilde{y}'_{ij} & \text{otherwise,} \end{cases} \tag{11}$$

and

$$\tilde{\eta}_{ij} = \begin{cases} \alpha + \beta^2 x_{n-1,n-1} & \text{if } i = j = n, \\ \beta x_{n-1,n-1} & \text{if } (i,j) = (n-1,n) \text{ or } (i,j) = (n,n-1), \\ \tilde{\eta}'_{ij} & \text{otherwise.} \end{cases} \tag{12}$$

Similar reasoning as above shows that $\tilde{\Phi}$ and $\tilde{\Psi}$ are indeed bijections.

3. The proof is by direct computation. □

Let us define $\varphi_{A_n} : Q_{A_n} \to \mathbb{R}_+$ by $\varphi_{A_1}(\eta) = \eta^{-1}$, and for $n \geq 2$

$$\varphi_{A_n}(\eta) = \prod_{i=1}^{n-1} |\eta_{\{i,i+1\}}|^{-3/2} \prod_{i \neq 1,n} \eta_{ii}. \tag{13}$$

Lemma 3 1. *Let $y = \Phi_n(a, b, z)$ and $\eta = \Psi_n(\alpha, \beta, x)$. Then, for all $M = 2, \ldots, n$,*

$$\Delta_{\underline{s}}^{(M)}(y) = a^{s_1} \Delta_{(s_2,\ldots,s_n)}^{(M)}(z), \tag{14}$$

$$\delta_{\underline{s}}^{(M)}(\eta) = \alpha^{s_1} \delta_{(s_2,\ldots,s_n)}^{(M)}(x). \tag{15}$$

Let $y = \tilde{\Phi}_n(a, b, z)$ and $\eta = \tilde{\Psi}_n(\alpha, \beta, x)$. Then, for all $M = 1, \ldots, n-1$,

$$\Delta_{\underline{s}}^{(M)}(y) = a^{s_n} \Delta_{(s_1,\ldots,s_{n-1})}^{(M)}(z), \tag{16}$$

$$\delta_{\underline{s}}^{(M)}(\eta) = \alpha^{s_n} \delta_{(s_1,\ldots,s_{n-1})}^{(M)}(x). \tag{17}$$

2. *Let $\eta = \Psi_n(\alpha, \beta, x)$ and $\tilde{\eta} = \tilde{\Psi}_n(\alpha, \beta, x)$. Then,*

$$\varphi_{A_n}(\eta) = x_{22}^{-1/2} \alpha^{-3/2} \varphi_{A_{n-1}}(x) \tag{18}$$

and

$$\varphi_{A_n}(\tilde{\eta}) = x_{n-1,n-1}^{-1/2} \alpha^{-3/2} \varphi_{A_{n-1}}(x). \tag{19}$$

3. *If $y = \Phi_n(a, b, z)$ and $\eta = \Psi_n(\alpha, \beta, x)$, then*

$$\operatorname{tr}(y\eta) = a\alpha + ax_{22}(b + \beta)^2 + \operatorname{tr}(zx). \tag{20}$$

If $y = \tilde{\Phi}_n(a, b, z)$ and $\eta = \tilde{\Psi}_n(\alpha, \beta, x)$, then

$$\operatorname{tr}(y\eta) = a\alpha + ax_{n-1,n-1}(b + \beta)^2 + \operatorname{tr}(zx). \tag{21}$$

Proof 1. For $M \geq 2$, we have

$$\frac{\Delta_{\underline{s}}^{(M)}(y)}{\Delta_{(s_2,\ldots,s_n)}^{(M)}(z)} = (y_{11})^{s_1 - s_2} \left[\prod_{i=2}^{M-1} \left(\frac{|y_{\{1:i\}}|}{|z_{\{2:i\}}|} \right)^{s_i - s_{i+1}} \right] \left(\frac{|y|}{|z|} \right)^{s_M}.$$

Using Lemma 7, we have $|y_{\{1:i\}}| = a|z_{\{2:i\}}|$. Thus,

$$\frac{\Delta_{\underline{s}}^{(M)}(y)}{\Delta_{(s_2,\ldots,s_n)}^{(M)}(z)} = a^{s_1}.$$

Noting that $a = y_{nn}$, we have for $M = 1, \dots, n-1$,

$$\Delta_{\underline{s}}^{(M)}(\tilde{y}) = |\tilde{y}|^{s_1} \prod_{i=2}^{n} |\tilde{y}_{\{i:n\}}|^{s_i - s_{i-1}} = a^{s_1} |z|^{s_1} \prod_{i=2}^{n-1} \left(a\, |z_{\{i:n\}}|^{s_i - s_{i-1}} \right) a^{s_n - s_{n-1}}$$
$$= a^{s_n} |z|^{s_1} \prod_{i=2}^{n-1} |z_{\{i:n\}}|^{s_i - s_{i-1}} = a^{s_n} \Delta^{(M)}_{(s_1,\dots,s_{n-1})}(z).$$

Similarly, we show that $\delta_{\underline{s}}^{(M)}(\eta) = \alpha^{s_1} \delta^{(M)}_{(s_2,\dots,s_n)}(x)$ for $M \geq 2$ and that $\delta_{\underline{s}}^{(M)}(\eta) = \alpha^{s_n} \delta_{\underline{s}}^{(M)}(x)$ for all $M \leq n-1$.
2. Let $\eta = \Psi(\alpha, \beta, x)$ and $\tilde{\eta} = \tilde{\Psi}(\alpha, \beta, x)$. For $n = 2$, we have

$$\varphi_{A_2}(\eta) = |\eta_{\{1,2\}}|^{-3/2} = \begin{vmatrix} \alpha + \beta^2 x & \beta x \\ \beta x & x \end{vmatrix}^{-3/2} = \alpha^{-3/2} x^{-3/2}$$
$$= x^{-1/2} \alpha^{-3/2} \varphi_{A_1}(x).$$

For $n > 2$, using (10), we have

$$\varphi_{A_n}(\eta) = \eta_{22}\, |\eta_{\{1,2\}}|^{-3/2} \frac{\prod_{i=2}^{n-1} |\eta_{\{i,i+1\}}|^{-3/2}}{\prod_{i=3}^{n-1} \eta_{ii}^{-1}} = x_{22}^{-1/2} \alpha^{-3/2} \varphi_{A_{n-1}}(x).$$

The proof of the second part is analogous.
3. The proof is by direct computation. □

4 Laplace transform of generalized power functions on Q_G and P_G

Theorem 2 *For all $n \geq 1$, for all $1 \leq M \leq n$ and for all $y \in P_{A_n}$, the integral $\int_{Q_{A_n}} \mathrm{e}^{-\operatorname{tr}(y\eta)} \delta_{\underline{s}}^{(M)}(\eta) \varphi_{A_n}(\eta) \mathrm{d}\eta$ converges if and only if $s_i > \frac{1}{2}$ for all $i \neq M$, and $s_M > 0$. In this case, we have*

$$\int_{Q_{A_n}} \mathrm{e}^{-\operatorname{tr}(y\eta)} \delta_{\underline{s}}^{(M)}(\eta) \varphi_{A_n}(\eta) \mathrm{d}\eta = \pi^{(n-1)/2} \left\{ \prod_{i \neq M} \Gamma(s_i - \frac{1}{2}) \right\} \Gamma(s_M) \Delta_{-\underline{s}}^{(M)}(y). \tag{22}$$

Proof We will proceed by induction on the number n of vertices. For $n = 1$, we have the gamma integral that converges if and only if $s > 0$, so that

$$\int_0^\infty \mathrm{e}^{-y\eta} \delta_s^{(1)}(\eta) \varphi_{A_1}(\eta) \mathrm{d}\eta = \int_0^\infty \mathrm{e}^{-y\eta} \eta^{s-1} \mathrm{d}\eta = \Gamma(s) y^{-s}.$$

Now assume that the assertion holds for a graph with $n-1$ vertices.
Case $M>1$. Let $y=\Phi_n(a,b,z)$ and let us make the change of variable $\eta=\Psi_n(\alpha,\beta,x)$. The induction hypothesis gives

$$\begin{aligned}\int_{Q_{A_{n-1}}} & \mathrm{e}^{-\operatorname{tr}(zx)}\delta^{(M)}_{(s_2,\dots,s_n)}(x)\varphi_{A_{n-1}}(x)\mathrm{d}x \\ &= \pi^{(n-2)/2}\Big\{\prod_{i\neq 1,M}\Gamma(s_i-\frac{1}{2})\Big\}\Gamma(s_M)\Delta^{(M)}_{-(s_2,\dots,s_n)}(z)\end{aligned} \tag{23}$$

if and only if $s_i>\frac{1}{2}$ for all $i\neq M$, and $s_M>0$. By Lemma 3, the change of variable $\eta=\Psi_n(\alpha,\beta,x)$ gives $\mathrm{d}\eta=x_{22}d\alpha d\beta\mathrm{d}x$. Thus, we have

$$\begin{aligned}\int_{Q_{A_n}} & \mathrm{e}^{-\operatorname{tr}(y\eta)}\delta^{(M)}_{\underline{s}}(\eta)\varphi_{A_n}(\eta)\mathrm{d}\eta \\ &= \int_0^\infty\int_{-\infty}^\infty\int_{Q_{A_{n-1}}}\mathrm{e}^{-(a\alpha+ax_{22}(b+\beta)^2+\operatorname{tr}(zx))}\alpha^{s_1-3/2} \\ &\quad\times\delta^{(M)}_{(s_2,\dots,s_n)}(x)\varphi_{A_{n-1}}(x)x_{22}^{1/2}\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}x,\end{aligned}$$

where we used parts 3 and 1 of Lemma 3. Now, using the Gaussian integral

$$\int_{-\infty}^\infty \mathrm{e}^{-ax_{22}(b+\beta)^2}\mathrm{d}\beta=\pi^{1/2}a^{-1/2}x_{22}^{-1/2}$$

and the gamma integral

$$\int_0^\infty \mathrm{e}^{-a\alpha}\alpha^{s_1-3/2}\mathrm{d}\alpha=a^{-s_1+1/2}\Gamma\left(s_1-\frac{1}{2}\right),$$

that is finite if and only if $s_1>\frac{1}{2}$, we get

$$\begin{aligned}\int_{Q_{A_n}} & \mathrm{e}^{-\operatorname{tr}(y\eta)}\delta^{(M)}_{\underline{s}}(\eta)\varphi_{A_n}(\eta)\mathrm{d}\eta \\ &= \pi^{1/2}a^{-s_1}\Gamma(s_1-\frac{1}{2})\int_{Q_{A_{n-1}}}\mathrm{e}^{-\operatorname{tr}(zx)}\delta^{(M)}_{(s_2,\dots,s_n)}(x)\varphi_{A_{n-1}}(x)\mathrm{d}x.\end{aligned} \tag{24}$$

Finally, using Formulas (23) and (14) completes the proof in the case $M>1$. *Case* $M=1$. Let $y=\tilde{\Phi}_n(a,b,z)$ and let us make the change of variable $\eta=\tilde{\Psi}_n(\alpha,\beta,x)$. The proof is similar. □

Theorem 3 *For all* $n\geq 1$*, for all* $1\leq M\leq n$ *and for all* $\eta\in Q_{A_n}$*, the integral* $\int_{P_{A_n}}\mathrm{e}^{-\operatorname{tr}(y\eta)}\Delta^{(M)}_{\underline{s}}(y)\mathrm{d}y$ *converges if and only if* $s_i>-\frac{3}{2}$ *for all* $i\neq M$*, and* $s_M>-1$.

In this case, we have

$$\int_{P_{A_n}} \mathrm{e}^{-\operatorname{tr}(y\eta)} \Delta_{\underline{s}}^{(M)}(y)\mathrm{d}y = \pi^{(n-1)/2}\left\{\prod_{i\neq M} \Gamma\left(s_i+\frac{3}{2}\right)\right\}\Gamma(s_M+1)\delta_{-\underline{s}}^{(M)}(\eta)\varphi_{A_n}(\eta). \tag{25}$$

Proof Similar to the proof of Theorem 2 using Proposition 2 and Lemma 3. □

The characteristic function φ_Ω of a convex cone Ω is defined as the Laplace transform of the Lebesgue measure of the dual cone: $\varphi_\Omega(x) = \int_{\Omega^*} \mathrm{e}^{-\langle y,x\rangle}\mathrm{d}y$, where Ω^* is the dual of Ω. The measure $\varphi_\Omega(x)\mathrm{d}x$ is called the canonical measure of Ω. It is invariant by linear automorphisms of Ω (Faraut and Korányi 1994).

Corollary 2 $\varphi_{Q_{A_n}} = \mathrm{const}\,.\,\varphi_{A_n}$.

Proof The result, $\left(\frac{4}{\pi^2}\right)^{\frac{n-1}{2}} \int_{P_{A_n}} \mathrm{e}^{-\operatorname{tr}(y\eta)}\mathrm{d}y = \varphi_{A_n}(\eta)$, is obtained by substituting $\underline{s} = (0,\dots,0)$ into Theorem 3. □

Remark 1 Formulas (22) and (25) may seem similar but in (25) the integrand does not contain the characteristic function of the cone P_{A_n}. This function is unknown except for A_4 when it is not a power function (Letac and Massam 2007, Prop.3.2).

5 Wishart exponential families on Q_G

Let us define the Riesz measure $R_{\underline{s}}^{(M)}$ on Q_G by

$$\mathrm{d}R_{\underline{s}}^{(M)}(x) = C_{\underline{s}}\delta_{\underline{s}}^{(M)}(x)\varphi_{A_n}(x)1_{Q_{A_n}}(x)\mathrm{d}x, \tag{26}$$

where $C_{\underline{s}}^{-1} = \pi^{(n-1)/2}\left(\prod_{i\neq M}\Gamma(s_i-\frac{1}{2})\right)\Gamma(s_M)$. Therefore, from Theorem 2, the Laplace transform of the measure $\mathrm{d}R_{\underline{s}}^{(M)}$ is given for all $s_i > \frac{1}{2}$, $i \neq M$ and $s_M > 0$ by

$$\mathcal{L}(R_{\underline{s}}^{(M)})(y) = \int_{Q_{A_n}} \mathrm{e}^{-\operatorname{tr}(y\eta)}\mathrm{d}R_{\underline{s}}^{(M)}(\eta) = \Delta_{-\underline{s}}^{(M)}(y), \quad y \in P_{A_n}. \tag{27}$$

Wishart natural exponential family $\gamma_{\underline{s},y}^{(M)}$ on Q_G is, by definition, generated by the Riesz measure $\mathrm{d}R_{\underline{s}}^{(M)}$. The density function of the Wishart distribution on Q_G is given by

$$\gamma_{\underline{s},y}^{(M)}(\mathrm{d}x) = C_{\underline{s}}\mathrm{e}^{-\operatorname{tr}(yx)}\Delta_{\underline{s}}^{(M)}(y)\delta_{\underline{s}}^{(M)}(x)\varphi_{A_n}(x)1_{Q_{A_n}}(x)\mathrm{d}x. \tag{28}$$

The Laplace transform of $\gamma^{(M)}_{\underline{s},y}(\mathrm{d}x)$ is

$$\mathcal{L}(\gamma^{(M)}_{\underline{s},y})(z)=\frac{\mathcal{L}(R^{(M)}_{\underline{s}})(z+y)}{\mathcal{L}(R^{(M)}_{\underline{s}})(y)}=\frac{\Delta^{(M)}_{-\underline{s}}(z+y)}{\Delta^{(M)}_{-\underline{s}}(y)}.$$

The family $\gamma^{(M)}_{\underline{s},y}$ does not depend on the normalization of the Riesz measure.

5.1 Mean and covariance of the Wishart distributions on Q_G

In this subsection, we derive a formula for the mean of the Wishart exponential family on the cones Q_G. It is known from the general theory of exponential families of distributions (Brown 1986) that the mean of $\gamma^{(M)}_{\underline{s},y}$ is obtained by differentiation with respect to y of the cumulant generating function of the Riesz measure:

$$m^{(M)}_{\underline{s}}(y)=-\mathrm{grad}_y \log \Delta^{(M)}_{-\underline{s}}(y)\in Q_G. \tag{29}$$

For all matrix A in Z_G and a subset $B\subset V$ of the set of vertices V of G, we note $(A_B)^0$ the matrix in Z_G such that $(A_B)^0_{ij}=\begin{cases}A_{ij} & \text{if } i,j\in B,\\ 0 & \text{otherwise.}\end{cases}$

Proposition 3 *The mean function of the Wishart family $\gamma^{(M)}_{\underline{s},y}$ on Q_G is equal to*

$$m^{(M)}_{\underline{s}}(y)=\pi\left(\sum_{i=1}^{M-1}(s_i-s_{i+1})[(y_{\{1:i\}})^{-1}]^0+s_My^{-1}+\sum_{i=M+1}^{n}(s_i-s_{i-1})[(y_{\{i:n\}})^{-1}]^0\right). \tag{30}$$

Proof Use formulas (3), (29) and $\mathrm{grad}\log|y_A|=\left((y_A)^{-1}\right)^0$. □

Proposition 4 *For all $y\in P_G$, we have*

$$\langle m^{(M)}_{\underline{s}}(y),y\rangle=\kappa(\underline{s}),$$

where the constant $\kappa(\underline{s})$ is $\sum_{i=1}^n s_i-(n-M)s_M$.

Proof Observe that by (3), for any $c>0$, $\Delta^{(M)}_{-\underline{s}}(cy)=c^{-\kappa(\underline{s})}\Delta^{(M)}_{-\underline{s}}(y)$. By (29), $\langle m^{(M)}_{\underline{s}}(y),y\rangle=-\langle \mathrm{grad}_y\log\Delta^{(M)}_{-\underline{s}}(y),y\rangle$. Set $F(y)=\log\Delta^{(M)}_{-\underline{s}}(y)$. By the chain rule, $\langle\mathrm{grad}_y F(y),y\rangle=\frac{\mathrm{d}}{\mathrm{d}t}F(ty)\big|_{t=1}$. The map $t\to F(ty)=\log\varphi(t),\mathbb{R}^+\to\mathbb{R}$, where $\varphi(t)=\Delta^{(M)}_{-\underline{s}}(ty)$, satisfies $\varphi(ct)=c^{-\kappa(\underline{s})}\varphi(t)$. Hence $\varphi(c)=c^{-\kappa(\underline{s})}\varphi(1)$ and $\frac{\mathrm{d}}{\mathrm{d}t}F(ty)\big|_{t=1}=\frac{\varphi'(1)}{\varphi(1)}=-\kappa(\underline{s})$. Thus $\langle\mathrm{grad}_yF(y),y\rangle=-\kappa(\underline{s})$ and the result follows. □

Differentiating the mean function gives the covariance function. For $A \in S_n$, let $\mathbb{P}(A) : Z_G \to I_G$ be the quadratic operator defined by $\mathbb{P}(A)u = \pi(AuA), \quad u \in Z_G$.

Proposition 5 *The covariance function of the Wishart family $\gamma_{\underline{s},y}^{(M)}$ on Q_G is equal to*

$$v(y) = -\frac{\mathrm{d}m_{\underline{s}}^{(M)}(y)}{\mathrm{d}y} = \sum_{i=1}^{M-1}(s_i - s_{i+1})\,\mathbb{P}\left[\left((y_{\{1:i\}})^{-1}\right)^0\right] + s_M\,\mathbb{P}(y^{-1}) \tag{31}$$

$$+ \sum_{i=M+1}^{n}(s_i - s_{i-1})\,\mathbb{P}\left[\left((y_{\{i:n\}})^{-1}\right)^0\right].$$

5.2 Inverse mean map

In the study of the exponential family $(\gamma_{\underline{s},y}^{(M)})_{y\in P_G}$, it is important to determine explicitly the inverse of the mean map $\psi_{\underline{s}}^{(M)} : m = m_{\underline{s}}^{(M)}(y) \mapsto y$, which we refer to as the inverse mean map in the sequel. The following theorem is known for Wishart exponential families on homogeneous cones (Ishi 2014). Surprisingly, it is also true on Q_G.

Theorem 4 *The inverse mean map $\psi_{\underline{s}}^{(M)}$ is given by the formula*

$$\psi_{\underline{s}}^{(M)}(m) = \mathrm{grad}_m \log \delta_{\underline{s}}^{(M)}(m), \; m \in Q_G. \tag{32}$$

The proof consists in following steps:

1. One shows that there exists a constant $c_{\underline{s}}$ depending only on $\underline{s}$ such that for any $y \in P_G$

$$\delta_{\underline{s}}^{(M)}(m_{\underline{s}}^{(M)}(y)) = c_{\underline{s}}\Delta_{-\underline{s}}^{(M)}(y) = c_{\underline{s}}\delta_{\underline{s}}^{(M)}(\pi(y^{-1})).$$

   This is done in Proposition 6 below.
2. One uses the Fenchel–Legendre duality, following a standard argument in information geometry of exponential family.

Proposition 6 *The following formula holds for any $y \in P_G$ and $\underline{s} \in \mathbb{R}^n$:*

$$\delta_{\underline{s}}^{(M)}(m_{\underline{s}}^{(M)}(y)) = \left(\prod_{i=1}^{n} s_i^{s_i}\right)\Delta_{-\underline{s}}^{(M)}(y) = \left(\prod_{i=1}^{n} s_i^{s_i}\right)\delta_{\underline{s}}^{(M)}(\pi(y^{-1})).$$

The proof of Proposition 6 will need a generalization of Lemma 2, where coefficients of inverse matrices of principal submatrices $y_{\{1:k\}}$ (or of $y_{\{k:n\}}$) are simultaneously considered. Define for $y \in P_G$, $\eta^{(k)} = (y_{\{1:k\}})^{-1}, \quad \eta^{[k]} = (y_{\{k:n\}})^{-1}$. The rows and the columns of the matrix $\eta^{(k)}$ are numbered by $i = 1, \ldots, k$ and the rows and the columns of the matrix $\eta^{[k]}$ are numbered by $i = k, \ldots, n$.

Lemma 4 *Let $y \in P_G$.*

1. *For all $i \in V$ and $k, m \geq i+1$ we have*

$$D_i^{k,m} := \begin{vmatrix} \eta_{ii}^{(k)} & \eta_{i,i+1}^{(m)} \\ \eta_{i,i+1}^{(k)} & \eta_{i+1,i+1}^{(m)} \end{vmatrix} = |y_{\{1:m\}}|^{-1} |y_{\{1:m\}\setminus\{i,i+1\}}|. \tag{33}$$

2. *For all $i \in V$ and $k, m \leq i < n$ we have*

$$D_i^{[k,m]} := \begin{vmatrix} \eta_{ii}^{[k]} & \eta_{i,i+1}^{[m]} \\ \eta_{i,i+1}^{[k]} & \eta_{i+1,i+1}^{[m]} \end{vmatrix} = |y_{\{k:n\}}|^{-1} |y_{\{k:n\}\setminus\{i,i+1\}}|. \tag{34}$$

Proof (of Proposition 6) We will deal with $\delta_{\underline{s}}^{(M)}(m_{\underline{s}}^{(M)}(y)) = \delta_{\underline{s}}^{\prec_M}(m_{\underline{s}}^{(M)}(y))$ where the order $\prec_M$ was defined in (6). By formula (30) and by the definition of $\delta_{\underline{s}}^{\prec_M}$ we obtain that $\delta_{\underline{s}}^{\prec_M}(m_{\underline{s}}(y))$ equals

$$\prod_{i=1}^{M-1} \left(\frac{1}{c_i} \begin{vmatrix} x_i + a_i & b_i \\ b_i & c_i \end{vmatrix} \right)^{s_i} (s_M \eta_{MM}^{(n)})^{s_M} \prod_{i=M+1}^{n} \left(\frac{1}{c_i'} \begin{vmatrix} x_i' + a_i' & b_i' \\ b_i' & c_i' \end{vmatrix} \right)^{s_i},$$

where $x_i = (s_i - s_{i+1})\eta_{ii}^{(i)}$, $a_i = \sum_{k=i+1}^{M-1} (s_k - s_{k+1})\eta_{ii}^{(k)} + s_M \eta_{ii}^{(n)}$,

$$\begin{aligned}
b_i &= \sum_{k=i+1}^{M-1} (s_k - s_{k+1})\eta_{i,i+1}^{(k)} + s_M \eta_{i,i+1}^{(n)}, \\
c_i &= \sum_{k=i+1}^{M-1} (s_k - s_{k+1})\eta_{i+1,i+1}^{(k)} + s_M \eta_{i+1,i+1}^{(n)}, \\
a_i' &= \sum_{k=M+1}^{i-1} (s_k - s_{k-1})\eta_{ii}^{[k]} + s_M \eta_{ii}^{[1]}, \\
b_i' &= \sum_{k=M+1}^{i-1} (s_k - s_{k-1})\eta_{i,i-1}^{[k]} + s_M \eta_{i,i-1}^{[1]}, \\
c_i' &= \sum_{k=M+1}^{i-1} (s_k - s_{k-1})\eta_{i-1,i-1}^{[k]} + s_M \eta_{i-1,i-1}^{[1]},
\end{aligned}$$

and $x_i' = (s_i - s_{i-1})\eta_{ii}^{[i]}$. Let us first compute the factors $\begin{vmatrix} x_i + a_i & b_i \\ b_i & c_i \end{vmatrix} / c_i$ for $i = 1, \ldots, M-1$. We will show that

$$\frac{1}{c_i} \begin{vmatrix} x_i + a_i & b_i \\ b_i & c_i \end{vmatrix} = s_i \eta_{ii}^{(i)}, \quad i = 1, \ldots, M-1. \tag{35}$$

We have $\frac{1}{c_i}\begin{vmatrix} x_i+a_i & b_i \\ b_i & c_i \end{vmatrix} = x_i + \frac{1}{c_i}\begin{vmatrix} a_i & b_i \\ b_i & c_i \end{vmatrix}$, so in order to prove (35), it is sufficient to prove that

$$\frac{1}{c_i}\begin{vmatrix} a_i & b_i \\ b_i & c_i \end{vmatrix} = s_{i+1}\eta_{ii}^{(i)}. \tag{36}$$

In order to prove (36), we first use the multilinearity of the determinant with respect to its columns and we write, using the notation $D_i^{k,m}$ from Lemma 4,

$$\begin{aligned}\begin{vmatrix} a_i & b_i \\ b_i & c_i \end{vmatrix} &= \sum_{k,m=i+1}^{M-1}(s_k - s_{k+1})(s_m - s_{m+1})D_i^{k,m} + s_M\sum_{k=i+1}^{M-1}(s_k - s_{k+1})D_i^{k,n} \\ &\quad + s_M\sum_{m=i+1}^{M-1}(s_m - s_{m+1})D_i^{n,m} + s_M^2 D_i^{n,n}.\end{aligned}$$

By Part 1 of Lemma 4 we have $D_i^{k,m} = |y_{\{1:m\}}|^{-1}|y_{\{1:m\}\setminus\{i,i+1\}}|$, which is independent of the left index k. The last fact allows to write

$$\begin{aligned}\begin{vmatrix} a_i & b_i \\ b_i & c_i \end{vmatrix} &= s_{i+1}\sum_{m=i+1}^{M-1}(s_m - s_{m+1})D_i^{n,m} + s_{i+1}s_M D_i^{n,n} \\ &= s_{i+1}\left(\sum_{m=i+1}^{M-1}(s_m - s_{m+1})\frac{|y_{\{1:m\}\setminus\{i,i+1\}}|}{|y_{\{1:m\}}|} + s_M\frac{|y_{\{1:n\}\setminus\{i,i+1\}}|}{|y|}\right).\end{aligned}$$

We factorize the determinants $|y_{\{1:m\}\setminus\{i,i+1\}}|$ and $|y_{\{1:n\}\setminus\{i,i+1\}}|$ in the last sum according to Lemma 1 and we write this sum as

$$\frac{|y_{\{1:i-1\}}|}{|y_{\{1:i\}}|}\left(\sum_{m=i+1}^{M-1}(s_m - s_{m+1})\frac{|y_{\{1:i\}}||y_{\{i+2:m\}}|}{|y_{\{1:m\}}|} + s_M\frac{|y_{\{1:i\}}||y_{\{i+2:n\}}|}{|y|}\right).$$

We have $|y_{\{1:m\}}|^{-1}|y_{\{1:i\}}||y_{\{i+2:m\}}| = \eta_{i+1,i+1}^{(m)}$. By definition of c_i we finally obtain

$$\begin{vmatrix} a_i & b_i \\ b_i & c_i \end{vmatrix} = s_{i+1}\frac{|y_{\{1:i-1\}}|}{|y_{\{1:i\}}|}c_i = s_{i+1}\eta_{ii}^{(i)}c_i$$

and formulas (36) and (35) are proved.

A "mirror" proof based on Part 2 of Lemma 4 shows that

$$\frac{1}{c_i'}\begin{vmatrix} x_i'+a_i' & b_i' \\ b_i' & c_i' \end{vmatrix} = s_i\eta_{ii}^{[i]}, \quad i = M+1,\ldots,n \tag{37}$$

and that $\delta_{\underline{s}}^{(M)}(m_{\underline{s}}^{(M)}(y)) = \prod_{i=1}^n s_i^{s_i} \prod_{i=1}^{M-1}(\eta_{ii}^{(i)})^{s_i}(\eta_{MM}^{(n)})^{s_M} \prod_{i=M+1}^n(\eta_{ii}^{[i]})^{s_i}$. Recall that

$$\eta_{ii}^{(i)} = \frac{|y_{\{1:i-1\}}|}{|y_{\{1:i\}}|}, \quad \eta_{ii}^{[i]} = \frac{|y_{\{i+1:n\}}|}{|y_{\{i:n\}}|}, \quad \eta_{MM}^{(n)} = \frac{|y_{\{1:M-1\}}||y_{\{M+1:n\}}|}{|y|},$$

so we deduce, using formula (3) that

$$\prod_{i=1}^{M-1}(\eta_{ii}^{(i)})^{s_i}(\eta_{MM}^{(n)})^{s_M} \prod_{i=M+1}^{n}(\eta_{ii}^{[i]})^{s_i} = \Delta_{-\underline{s}}^{(M)}(y).$$

Applying Theorem 1, we see that $\delta_{\underline{s}}^{(M)}(m_{\underline{s}}^{(M)}(y)) = \prod_{i=1}^n s_i^{s_i} \delta_{\underline{s}}^{(M)}(\pi(y^{-1}))$. $\square$

Proof (of Theorem 4). In order to apply techniques in information geometry of exponential family of Amari and Nagaoka (2007), see also Brown (1986), we introduce a variable $\theta = -y \in (-P_G)$. By (29), we have

$$m_{\underline{s}}^{(M)}(y) = -\mathrm{grad}_y \log \Delta_{-\underline{s}}^{(M)}(y) = \mathrm{grad}_\theta \log \Delta_{-\underline{s}}^{(M)}(-\theta).$$

Let $f(\theta) := \log \Delta_{-\underline{s}}^{(M)}(-\theta)$, which is the cumulant generating function of the Riesz measure by (27). Then, f is a convex function on the domain $-P_G$. Let $g(m)$ be the Fenchel–Legendre transform of $f(\theta)$, that is to say, $g(m) = \langle m, \theta\rangle - f(\theta)$, where

$$m = \mathrm{grad}_\theta f(\theta) = m_{\underline{s}}^{(M)}(-\theta), \qquad \theta \in (-P_G).$$

Thanks to the Fenchel–Legendre duality, the inverse map of $\mathrm{grad}_\theta f : (-P_G) \ni \theta \mapsto m = m_{\underline{s}}^{(M)}(-\theta) \in Q_G$ is given by $\mathrm{grad}_m g : Q_G \ni m \mapsto \theta \in (-P_G)$. Now, by Proposition 4, we have

$$\langle m, \theta\rangle = \langle m_{\underline{s}}^{(M)}(y), -y\rangle = -\kappa(\underline{s}).$$

On the other hand, Proposition 6 implies that

$$f(\theta) = \log \Delta_{-\underline{s}}^{(M)}(-\theta) = \log\left(c_{\underline{s}}^{-1}\delta_{\underline{s}}^{(M)}(m)\right),$$

where $c_{\underline{s}} = \prod_{i=1}^r s_i^{s_i}$. Therefore, we have

$$\begin{aligned}\theta = \mathrm{grad}_m g(m) &= \mathrm{grad}_m\big\{-\kappa(\underline{s}) + \log c_{\underline{s}} - \log \delta_{\underline{s}}^{(M)}(m)\big\}\\ &= -\mathrm{grad}_m \log \delta_{\underline{s}}^{(M)}(m),\end{aligned}$$

which leads us to $y = -\theta = \mathrm{grad}_m \log \delta_{\underline{s}}^{(M)}(m)$. $\square$

Corollary 3 *The inverse mean map $\psi_{\underline{s}}^{(M)} : Q_G \to P_G$ is given by*

$$\psi_{\underline{s}}^{(M)}(m) = \sum_{k=1}^{M-1} s_k \left((m_{\{k:k+1\}})^{-1}\right)^0 + \sum_{k=M+1}^{n} s_k \left((m_{\{k-1:k\}})^{-1}\right)^0$$
$$- \sum_{k=2}^{M-1} s_{k-1} \left((m_{\{kk\}})^{-1}\right)^0 - (s_{M-1} - s_M + s_{M+1}) \left((m_{\{MM\}})^{-1}\right)^0$$
$$- \sum_{k=M+1}^{n-1} s_{k+1} \left((m_{\{kk\}})^{-1}\right)^0. \tag{38}$$

Proof The result is obtained by computing the gradient of $\log \delta_{\underline{s}}^{(M)}(m)$, as indicated in (32). We use the formula (4). □

The Lauritzen formula (Lauritzen 1996) is an explicit formula for a bijection between Q_G and P_G. It states that for all $x \in Q_G$, the unique $y \in P_G$ such that $\pi(y^{-1}) = x$ is given by

$$y = \sum_{i=1}^{n-1} (x_{\{i:i+1\}}^{-1})^0 - \sum_{i=2}^{n-1} (x_{ii}^{-1})^0. \tag{39}$$

Setting $s_1 = \ldots = s_n = 1$ in formula (30) for the mean function, we get

$$m_{(1,\ldots,1)}^{(M)}(y) = \pi(y^{-1}) = x. \tag{40}$$

Thus,

$$\psi_{(1,\ldots,1)}^{(M)}(x) = y \tag{41}$$

is the Lauritzen formula. Indeed, for $s_1 = \ldots = s_n = 1$, formula (38) gives

$$\psi_{(1,\ldots,1)}^{(M)}(m) = \sum_{i=1}^{n-1} (m_{\{i:i+1\}}^{-1})^0 - \sum_{i=2}^{n-1} (m_{ii}^{-1})^0. \tag{42}$$

Thus we found a new proof of the Lauritzen formula, based on the observation that the Lauritzen map is the inverse mean map for $\underline{s} = \mathbf{1} = (1, 1, \ldots, 1)$. At the same time, we find an infinite number of explicit isomorphisms from Q_G onto P_G, given by the inverse mean maps $\psi_{\underline{s}}^{(M)}$. It is an essential generalization of the Lauritzen formula. Each map $\psi_{\underline{s}}^{(M)}$ is a generalized Lauritzen map.

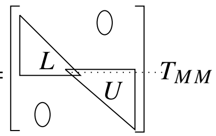
5.3 Variance function

5.3.1 Properties of lower-upper M-triangular matrices

Here, we define and prove basic properties of lower-upper M-triangular matrices, that we will denote by LU(M). They are very important in proofs of this section.

Definition 5 A matrix T is said to be an LU(M) triangular matrix if for all $i < M$, $T_{ij} = 0$ if $j > i$ and for all $i > M$, $T_{ij} = 0$ if $i > j$.

In particular, T is an LU(n) triangular matrix if and only if it is lower triangular, and T is an LU(1) triangular matrix if and only if it is upper triangular. An LU(M) triangular matrix T is a succession of an $M \times M$ lower triangular matrix $L = T_{\{1:M\}}$ and an $(N - M) \times (N - M)$ upper triangular matrix $U = T_{\{M:n\}}$ with diagonal term T_{MM} in common. We write $T = s(L, U)$.

$T =$ 

Proposition 7 1. *$s(L, U)s(L', U') = s(LL', UU')$.*
2. *If $s(L, U)$ is invertible, then $(s(L, U))^{-1}$ is also an LU(M) triangular matrix and $(s(L, U))^{-1} = s(L^{-1}, U^{-1})$.*
3. *The set of $LU(M)$ triangular matrices is a group.*

Lemma 5 *Let S and T be LU(M) triangular $n \times n$ matrices.*

1. (a) *Let $A = K^0$ with $K = A_{\{1:k\}}$. If $k \leq M - 1$, then $S^T AT = \left(S^T_{\{1:k\}} K T_{\{1:k\}}\right)^0$.*
   (b) *Let $B = K^0$ with $K = B_{\{k:n\}}$. If $k \geq M + 1$, then $S^T BT = \left(S^T_{\{k:n\}} K T_{\{k:n\}}\right)^0$.*
2. *Let A be an $n \times n$ matrix. Then $(TAS^T)_{\{1:i\}} = T_{\{1:i\}} A_{\{1:i\}} S^T_{\{1:i\}}$ for $i \leq M - 1$, and $(TAS^T)_{\{i:n\}} = T_{\{i:n\}} A_{\{i:n\}} S^T_{\{i:n\}}$ for $i \geq M + 1$.*
3. *If T is invertible, then*
   (a) *$(T_{\{1:k\}})^{-1} = (T^{-1})_{\{1:k\}}$ for all $k \leq M - 1$;*
   (b) *$(T_{\{k:n\}})^{-1} = (T^{-1})_{\{k:n\}}$ for all $k \geq M + 1$.*

Proposition 8 *For all $y \in P_{A_n}$, for all $1 \leq M \leq n$, there exists an LU(M) triangular matrix T satisfying $T_{ij} = 0$ if $i \nsim j$ and such that $y = T\,T^T$.*

Proof We will proceed by induction on n. The statement is obviously true for $n = 1$. Let us assume that the statement is true for $n - 1$. Let $y \in P_{A_n}$ and $M \neq 1$. Let us write $y = \Phi_n(a, b, z)$ with $z \in P_{A_{n-1}}$. The induction assumption implies that there exists V an $(n-1) \times (n-1)$ LU(M) triangular matrix such that $V_{ij} = 0$ if $i \nsim j$ and such that $z = V\,V^T$. Let us write
$$T = \begin{pmatrix} 1 & & & & \\ b & 1 & & & \\ 0 & & \ddots & & \\ \vdots & & & \ddots & \\ 0 & \dots & \dots & \dots\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{a} & 0 \dots 0 \\ 0 & \\ \vdots & V \\ 0 & \end{pmatrix} = \begin{pmatrix} \sqrt{a} & 0 \dots 0 \\ \sqrt{ab} & \\ 0 & \\ \vdots & V \\ 0 & \end{pmatrix}.$$

T is LU(M) triangular satisfying $T_{ij} = 0$ if $i \nsim j$ and $y = T\,T^T$.

For $M = 1$, we use $y = \tilde{\Phi}_n(a, b, z)$ with $z \in P_{A_{n-1}}$. □

5.3.2 Two formulas for the variance function

Let $m \in Q_G$. We note $\hat{m} \in S_n^+$ the unique symmetric positive definite matrix verifying $\pi(\hat{m}) = m,\ \ \hat{m}^{-1} \in P_G$. Define $y = \psi_{\underline{s}}^{(M)}(m) \in P_G$. Decompose $y = T\,T^T$, with T an LU(M) triangular matrix such that $T_{ij} = 0$ when $i \nsim j$.

Lemma 6 *We have*

$$\hat{m} = (T^{-1})^T \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix} T^{-1}. \tag{43}$$

Proof Note that $y = \psi_{\underline{s}}^{(M)}(m)$ is equivalent to $m = m_{\underline{s}}^{(M)}(y)$. The formula of the mean function (30) gives $m = \pi(Z)$, where

$$Z = \sum_{i=1}^{M-1} (s_i - s_{i+1})[(y_{\{1:i\}})^{-1}]^0 + s_M y^{-1} + \sum_{i=M+1}^{n} (s_i - s_{i-1})[(y_{\{i:n\}})^{-1}]^0. \tag{44}$$

Using Part 2 of Lemma 5, we have $y_{\{1:i\}} = T_{\{1:i\}} I_{\{1:i\}} (T_{\{1:i\}})^T$ for $i \le M-1$. Here $I_{\{1:i\}}$ is the $i \times i$ identity matrix. By Part 3 of Lemma 5, we get $(y_{\{1:i\}})^{-1} = (T^{-1})^T_{\{1:i\}} I_{\{1:i\}} (T^{-1})_{\{1:i\}}$. Finally, using Part 1 of Lemma 5, we obtain

$$[(y_{\{1:i\}})^{-1}]^0 = (T^{-1})^T (I_{\{1:i\}})^0 T^{-1}, \quad i \le M-1. \tag{45}$$

Similarly, we have

$$[(y_{\{i:n\}})^{-1}]^0 = (T^{-1})^T (I_{\{i:n\}})^0 T^{-1}, \quad i \ge M+1. \tag{46}$$

Thus,

$$\begin{aligned} Z &= (T^{-1})^T \left(\sum_{i=1}^{M-1} (s_i - s_{i+1})(I_{\{1:i\}})^0 + s_M I + \sum_{i=M+1}^{n} (s_i - s_{i-1})(I_{\{i:n\}})^0 \right) T^{-1} \\ &= (T^{-1})^T \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix} T^{-1}. \end{aligned}$$

Therefore, Z is positive definite and $Z^{-1} = T \begin{pmatrix} s_1^{-1} & & 0 \\ & \ddots & \\ 0 & & s_n^{-1} \end{pmatrix} T^T \in P_{A_n}$. Indeed, for all $i < i+1 < j$, we have $(Z^{-1})_{ij} = \sum_{k=1}^n T_{ik} T_{jk} s_k^{-1}$. Since $T_{ik} = 0$ for $|k-i| > 1$, $(Z^{-1})_{ij} = T_{i,i-1} T_{j,i-1} s_{i-1}^{-1} + T_{ii} T_{ji} s_i^{-1} + T_{i,i+1} T_{j,i+1} s_{i+1}^{-1}$. But since $|j-i| > 1$, we have $T_{j,i-1} = 0 = T_{ji}$ and $(Z^{-1})_{ij} = T_{i,i+1} T_{j,i+1} s_{i+1}^{-1}$. Now since T is LU(M), we have $T_{i,i+1}T_{j,i+1} = 0$. In fact, $T_{i,i+1} = 0$ for $i \le M-1$ and $T_{j,i+1} = 0$ for $i \ge M$. In conclusion, we have shown that $m = \pi(Z)$ with $Z^{-1} \in P_{A_n}$, which implies $Z = \hat{m}$. □

The following Proposition derives the formula for the variance function $V(m)$ which, for each fixed $m \in Q_G$ is a continuous operator $V(m) : Z_G \to I_G$ (Casalis and Letac 1996). Recall that $\mathbb{P}(A) : Z_G \to I_G$ is the quadratic operator defined by $\mathbb{P}(A)u = \pi(AuA)$. For $A, B \in S_n$, let $\mathbb{P}(A, B)u = \frac{1}{2}\pi(AuB + BuA)$. For all $m \in Q_G$ and $I \subset V$, we note

$$M_I = [((\hat{m}^{-1})_I)^{-1}]^0. \tag{47}$$

Proposition 9 *The variance function $V(m)$ of a Wishart NEF on Q_G is equal to*

$$\begin{aligned}
&\sum_{i=1}^{M-1}(s_i - s_{i+1})\,\mathbb{P}\left(\sum_{j=1}^{i-1}\left(\frac{1}{s_j} - \frac{1}{s_{j+1}}\right)M_{\{1:j\}} + \frac{1}{s_i}M_{\{1:i\}}\right)\\
&\quad + s_M\,\mathbb{P}\left(\frac{\hat{m}}{s_M} + \sum_{j=1}^{M-1}\left(\frac{1}{s_j} - \frac{1}{s_{j+1}}\right)M_{\{1:j\}} + \sum_{k=M+1}^{n}\left(\frac{1}{s_k} - \frac{1}{s_{k-1}}\right)M_{\{k:n\}}\right)\\
&\quad + \sum_{i=M+1}^{n}(s_i - s_{i-1})\,\mathbb{P}\left(\frac{1}{s_i}M_{\{i:n\}} + \sum_{j=i+1}^{n}\left(\frac{1}{s_j} - \frac{1}{s_{j-1}}\right)M_{\{j:n\}}\right).
\end{aligned} \tag{48}$$

Proof The variance function is given for all $m \in Q_{A_n}$ by $V(m) = v(\psi_{\underline{s}}^{(M)}(m))$, where $v(y)$ is given by (31). Let $y = \psi_{\underline{s}}^{(M)}(m) = TT^T$, where T is LU(M). From Lemma 6, we have

$$\hat{m}^{-1} = T\begin{pmatrix} s_1^{-1} & & 0\\ & \ddots & \\ 0 & & s_n^{-1}\end{pmatrix}T^T.$$

Using Lemma 5, we get

$$M_{\{1:i\}} = (T^{-1})^T\,(diag(s_1, \ldots, s_i))^0\,T^{-1}, \quad i \le M-1 \tag{49}$$

and

$$M_{\{i:n\}} = (T^{-1})^T\,(diag(s_i, \ldots, s_n))^0\,T^{-1}, \quad i \ge M+1. \tag{50}$$

Thus, for all $2 \le i \le M-1$, we have

$$\frac{1}{s_1}M_1 = (T^{-1})^T e_1 T^{-1}, \quad \frac{1}{s_i}(M_{\{1:i\}} - M_{\{1:i-1\}}) = (T^{-1})^T e_i T^{-1}, \tag{51}$$

and for all $n-1 \geq i \geq M+1$, we have

$$\frac{1}{s_n} M_n = (T^{-1})^T e_n T^{-1}, \quad \frac{1}{s_i}(M_{\{i:n\}} - M_{\{i+1:n\}}) = (T^{-1})^T e_i T^{-1}, \tag{52}$$

where e_i is the matrix with $e_{ii} = 1$ and $e_{ij} = 0$ for all $i \neq j$. Observing that $(I_{\{1:i\}})^0 = \sum_{k=1}^{i} e_k$ and $(I_{\{i:n\}})^0 = \sum_{k=i}^{n} e_k$, and using (45) and (51), we obtain for $i \leq M-1$

$$\begin{aligned}
[(y_{\{1:i\}})^{-1}]^0 &= (T^{-1})^T (I_{\{1:i\}})^0 T^{-1} = (T^{-1})^T \left(\sum_{k=1}^{i} e_k \right) T^{-1} = \sum_{k=1}^{i} \left((T^{-1})^T e_k T^{-1} \right) \\
&= \frac{1}{s_1} M_{\{1\}} + \frac{1}{s_2}(M_{\{1:2\}} - M_{\{1\}}) + \ldots + \frac{1}{s_i}(M_{\{1:i\}} - M_{\{1:i-1\}}) \\
&= \left(\frac{1}{s_1} - \frac{1}{s_2} \right) M_{\{1\}} + \ldots + \left(\frac{1}{s_{i-1}} - \frac{1}{s_i} \right) M_{\{1:i-1\}} + \frac{1}{s_i} M_{\{1:i\}}.
\end{aligned}$$

Similarly, using (46) and (52), we obtain for $i \geq M+1$,

$$[(y_{\{i:n\}})^{-1}]^0 = \frac{1}{s_i} M_{\{i:n\}} + \left(\frac{1}{s_{i+1}} - \frac{1}{s_i} \right) M_{\{i+1:n\}} + \ldots + \left(\frac{1}{s_n} - \frac{1}{s_{n-1}} \right) M_{\{n\}}.$$

We also observe that

$$(T^{-1})^T e_M T^{-1} = \frac{1}{s_M} \left(\hat{m} - M_{\{1:M-1\}} - M_{\{M+1:n\}} \right). \tag{53}$$

Thus, by (51), (52) and (53), we get

$$\begin{aligned}
y^{-1} &= \sum_{i=1}^{n} (T^{-1})^T e_i T^{-1} = \sum_{i=1}^{M-1} (T^{-1})^T e_i T^{-1} + (T^{-1})^T e_M T^{-1} \\
&\quad + \sum_{i=M+1}^{n} (T^{-1})^T e_i T^{-1} \\
&= \frac{\hat{m}}{s_M} + \sum_{j=1}^{M-1} \left(\frac{1}{s_j} - \frac{1}{s_{j+1}} \right) M_{\{1:j\}} + \sum_{j=M+1}^{n} \left(\frac{1}{s_j} - \frac{1}{s_{j-1}} \right) M_{\{j:n\}}.
\end{aligned} \tag{54}$$

Substituting these expressions of $[(y_{\{1:i\}})^{-1}]^0$, y^{-1} and $[(y_{\{i:n\}})^{-1}]^0$ into $v(y)$ given by (31), we obtain the stated result. □

We prove now a much simpler formula for the variance function on Q_G, surprisingly similar to the variance function on a homogeneous cone, in particular on the symmetric cone S_n^+ (cf. Graczyk et al. 2016).

Theorem 5 *The variance function of the Wishart exponential family* $\gamma^{(M)}_{\underline{s},y}$ *is*

$$V(m) = \left(\frac{1}{s_1}+\frac{1}{s_n}-\frac{1}{s_M}\right)\mathbb{P}(\hat{m}) \\ + \sum_{i=1}^{M-1}\left(\frac{1}{s_{i+1}}-\frac{1}{s_i}\right)\mathbb{P}(\hat{m}-M_{\{1:i\}}) + \sum_{i=M+1}^{n}\left(\frac{1}{s_{i-1}}-\frac{1}{s_i}\right)\mathbb{P}(\hat{m}-M_{\{i:n\}}), \tag{55}$$

where $M_{\{1:i\}}$ *and* $M_{\{i:n\}}$ *are defined in* (47).

Proof Using $\mathbb{P}(a-b)=\mathbb{P}(a)+\mathbb{P}(b)-2\,\mathbb{P}(a,b)$, we see that (55) is equivalent to

$$\begin{aligned} &V(m) \\ &= \frac{1}{s_M}\mathbb{P}(\hat{m}) + \sum_{i=1}^{M-1}\left(\frac{1}{s_{i+1}}-\frac{1}{s_i}\right)\mathbb{P}(M_{\{1:i\}}) + \sum_{i=M+1}^{n}\left(\frac{1}{s_{i-1}}-\frac{1}{s_i}\right)\mathbb{P}(M_{\{i:n\}}) \\ &\quad - 2\left(\sum_{i=1}^{M-1}\left(\frac{1}{s_{i+1}}-\frac{1}{s_i}\right)\mathbb{P}(\hat{m},M_{\{1:i\}}) + \sum_{i=M+1}^{n}\left(\frac{1}{s_{i-1}}-\frac{1}{s_i}\right)\mathbb{P}(\hat{m},M_{\{i:n\}})\right). \end{aligned} \tag{56}$$

We show that the right-hand sides of (48) and (56) are the same. Below, we expand (48) using $\mathbb{P}(a+b)=\mathbb{P}(a)+\mathbb{P}(b)+2\,\mathbb{P}(a,b)$ and compute the coefficients in the expanded formula. Note that for all $Z\in Z_G$, $\mathbb{P}(M_{\{1:i\}},M_{\{k:n\}})Z=0$ for all $i\le M-1$ and $k\ge M+1$, since $Z_{\{1:i\},\{k:n\}}=0$.

For a fixed $r\le M-1$, the coefficient of $\mathbb{P}(M_{\{1:r\}})$ is

$$\frac{s_r-s_{r+1}}{s_r^2}+\sum_{i=r+1}^{M-1}(s_i-s_{i+1})\left(\frac{1}{s_r}-\frac{1}{s_{r+1}}\right)^2+s_M\left(\frac{1}{s_r}-\frac{1}{s_{r+1}}\right)^2=\frac{1}{s_{r+1}}-\frac{1}{s_r}.$$

By a mirror argument, for a fixed $r\ge M+1$, the coefficient of $\mathbb{P}(M_{\{r:n\}})$ is $\frac{1}{s_{r-1}}-\frac{1}{s_r}$. On the other hand, the coefficient of $\mathbb{P}(\hat{m})$ is $\frac{1}{s_M}$.

For a fixed r, the coefficient of $\mathbb{P}(\hat{m},M_{\{1:r\}})$ is $\frac{1}{s_r}-\frac{1}{s_{r+1}}$ if $r\le M-1$, and the coefficient of $\mathbb{P}(\hat{m},M_{\{r:n\}})$ is $\frac{1}{s_r}-\frac{1}{s_{r-1}}$ if $r\ge M+1$. Moreover, if $k<r\le M-1$, the coefficient of $\mathbb{P}(M_{\{1:r\}},\ M_{\{1:k\}})$ is

$$\begin{aligned} &(s_r-s_{r+1})\frac{1}{s_r}\left(\frac{1}{s_k}-\frac{1}{s_{k+1}}\right)+\sum_{i=r+1}^{M-1}(s_i-s_{i+1})\left(\frac{1}{s_r}-\frac{1}{s_{r+1}}\right)\left(\frac{1}{s_k}-\frac{1}{s_{k+1}}\right) \\ &\quad + s_M\left(\frac{1}{s_r}-\frac{1}{s_{r+1}}\right)\left(\frac{1}{s_k}-\frac{1}{s_{k+1}}\right) \\ &= \left(\frac{1}{s_k}-\frac{1}{s_{k+1}}\right)\left(1-\frac{s_{r+1}}{s_r}+s_{r+1}\left(\frac{1}{s_r}-\frac{1}{s_{r+1}}\right)\right)=0. \end{aligned}$$

By a mirror argument, for a fixed $M+1 \le k < r$, the coefficient of $\mathbb{P}(M_{\{k:n\}}, M_{\{r:n\}})$ is 0. $\square$

Remark 2 $\hat{m}$ is easy to compute, using, for non-adjacent i and j the formula $\hat{m}_{ij} = m_{i,V\setminus\{i,j\}}(\hat{m}^{-1}_{V\setminus\{i,j\},V\setminus\{i,j\}})m_{V\setminus\{i,j\},j}$. (Letac and Massam 2007, p.1279).

In the next Corollary, we consider $\underline{s} = p\mathbf{1}$, $p > 1/2$. We note that $\delta^{(M)}_{p\mathbf{1}}$ and $\gamma^{(M)}_{p\mathbf{1},y} := \gamma_{p,y}$ do not depend on M.

Corollary 4 *The variance function of the Wishart exponential family $\gamma_{p,y}$ is*

$$V(m) = \frac{1}{p}\,\mathbb{P}(\hat{m}).$$

5.3.3 A relation between the inverse mean map and $m_{\frac{1}{\underline{s}}}$

Recall that for the classical Wishart exponential families $W_{s\mathbf{1},y}$ on the symmetric cone Sym_n^+ the bijection between the cone Q_G and P_G is given by $L(m) = m^{-1}$. The mean map is $m_s(y) = sy^{-1}$ and the inverse mean map $\psi_s(m) = sm^{-1}$. It follows that

$$\psi_s = L \circ m_{\frac{1}{s}} \circ L,$$

that is, the maps ψ_s and $m_{\frac{1}{s}}$ are intertwined by the bijection L.

An analogous property holds on the cone Q_{A_n}, with the intertwiner given by the Lauritzen map. The bijection $L : Q_{A_n} \to P_{A_n}$ is the Lauritzen map $L(m) = (\hat{m})^{-1}$. Its inverse $L^{-1} : P_{A_n} \to Q_{A_n}$ is $L^{-1}(y) = \pi(y^{-1})$.

Proposition 10 *The inverse mean map $\psi^{(M)}_{\underline{s}} : Q_G \to P_G$ satisfies*

$$\psi^{(M)}_{\underline{s}} = L \circ m^{(M)}_{\frac{1}{\underline{s}}} \circ L.$$

Equivalently, for any $m \in Q_G$, $\pi(\psi^{(M)}_{\underline{s}}(m)^{-1}) = m^{(M)}_{\frac{1}{\underline{s}}}(\hat{m}^{-1})$.

$$\begin{tikzcd}
Q_{A_n} \arrow[r, "\psi^{(M)}_{\underline{s}}"] \arrow[d, "L"'] & P_{A_n} \\
P_{A_n} \arrow[r, "m^{(M)}_{\frac{1}{\underline{s}}}"'] & Q_{A_n} \arrow[u, "L"']
\end{tikzcd}$$

Proof Using formula (30) of the mean function and definition (47) of $M_{\{1:i\}}$ and $M_{\{i:n\}}$, we see that $m^{(M)}_{\frac{1}{\underline{s}}}(\hat{m}^{-1})$ equals

$$\pi\left(\sum_{j=1}^{M-1}\left(\frac{1}{s_j}-\frac{1}{s_{j+1}}\right)M_{\{1:j\}}+\frac{\hat{m}}{s_M}+\sum_{j=M+1}^{n}\left(\frac{1}{s_j}-\frac{1}{s_{j-1}}\right)M_{\{j:n\}}\right).$$

Confronting this result with (54), we obtain $m^{(M)}_{\frac{1}{\underline{s}}}(\hat{m}^{-1})=\pi\left(\psi^{(M)}_{\underline{s}}(m)^{-1}\right)$. $\square$

5.4 Quadratic construction of Riesz measures and Wishart distributions on Q_G

Let $I\subset\{1,\dots,n\}$. We define $|I|$-dimensional subspaces W_I of $\mathbb{R}^n$ by

$$W_I=\{x\in\mathbb{R}^n \mid x_i=0,\ i\notin I\}.$$

Denote by q^I the quadratic map $q^I(x)=xx^T$ from W_I into $\mathrm{Sym}(n,\mathbb{R})$ and by q^I_* its projection onto I_G, i.e. $q^I_*=\pi\circ q^I$. The maps q^I_* are clearly Q_G-positive (submatrices y_I of a positive definite matrix y are positive definite for any $I\subset\{1,\dots,n\}$). In Graczyk and Ishi (2014), p.322, Riesz measures μ_q associated to a quadratic map q were defined and their Laplace transform computed. Recall that the measure $\mu_{q^I_*}$ is the image of the Lebesgue measure on W_I by q^I_* and that its Laplace transform equals

$$\mathcal{L}(\mu_{q^I_*})(y)=\pi^{|I|/2}|y_I|^{-1/2},\quad y\in P_G. \tag{57}$$

When $I=\{1,\dots,k\}$, we write $q^I_*=q^k_*$. When $I=\{k,\dots,n\}$, we write $q^I_*=\tilde{q}^k_*$.

Fix $M\in\{1,\dots,n\}$. We define the set B_M of *basic quadratic maps* for the Riesz $R^{(M)}_{\underline{s}}$ and Wishart $\gamma^{(M)}_{\underline{s},y}$ families on Q_G by $B_M=\{q^1_*,\dots,q^{M-1}_*,q^n_*,\tilde{q}^{M+1}_*,\dots,\tilde{q}^n_*\}$. Note that the basic quadratic maps with values in Q_G are different for each fixed $M=1,\dots,n$.

Proposition 11 *Let $\sigma_i\in\mathbb{R}$, $i=1,\dots,m$. A virtual quadratic map $q^{\underline{\sigma}}_*=\sum^{\oplus}_{i<M}(q^i_*)^{\oplus\sigma_i}\oplus(q^n_*)^{\oplus\sigma_M}\oplus\sum^{\oplus}_{i>M}(\tilde{q}^i_*)^{\oplus\sigma_i}$. exists if there exists $\underline{s}$ satisfying $s_i>\frac{1}{2}$, $i\neq M$, $s_M>0$ and*

$$\frac{\sigma_i}{2}=s_i-s_{i+1},\ 1\le i<M,\quad \frac{\sigma_M}{2}=s_M,\quad \frac{\sigma_i}{2}=s_i-s_{i-1},\ M<i\le n. \tag{58}$$

Proof We compare the Laplace transform of $\mu_{q^{\underline{\sigma}}_*}$, computed thanks to (57), with (27). As a result, we see that there exists a constant $c>0$ such that $R^{(M)}_{\underline{s}}=c\mu_{q^{\underline{\sigma}}_*}$. $\square$

Thus, all the Riesz $R^{(M)}_{\underline{s}}$ measures on Q_G defined in this paper are obtained as virtual or true (i.e. for $\sigma_i\in\mathbb{N}$) quadratic Riesz families, with basic maps from B_M.

Observe that by the quadratic construction, we can obtain absolutely continuous Riesz measures on Q_G not belonging to $\cup_M\{R_{\underline{s}}^{(M)}\}$, e.g. when $n = 3$, consider μ_q associated with the quadratic map $q = q_*^2 \oplus (q_*^3)^{\oplus 2} \oplus \tilde{q}_*^2$.

5.5 Higher-order moments of Wishart families on Q_{A_n}

Thanks to the identification of Wishart families $\gamma_{\underline{s},y}^{(M)}$ with quadratically constructed Wishart distributions $\gamma_{q_*^{\underline{\sigma}}}$ in Sect. 5.4, we can compute moments of any order N of a Wishart random variable X on Q_{A_n}.

Theorem 6 *Let X be a Q_{A_n}-valued random variable with the Wishart law $\gamma_{\underline{s},y}^{(M)}$. Let $z^{(1)}, z^{(2)}, \ldots, z^{(N)} \in Z_G$. Then, denoting by $C(\pi)$ the set of cycles of a permutation $\pi \in S_N$, the N-th moment $\mathbf{E}(\langle X, z^{(1)}\rangle \ldots \langle X, z^{(N)}\rangle)$ equals*

$$\sum_{\pi\in S_N} \prod_{c\in C(\pi)} \left\{ \sum_{i=1}^{M-1} (s_i - s_{i+1}) \operatorname{tr} \prod_{j\in c} (y_{\{1:i\}})^{-1} z_{\{1:i\}}^{(j)} \right.$$
$$\left. + s_M \operatorname{tr} \prod_{j\in c} y^{-1} z^{(j)} + \sum_{i=M+1}^{n} (s_i - s_{i-1}) \operatorname{tr} \prod_{j\in c} (y_{\{i:n\}})^{-1} z_{\{i:n\}}^{(j)} \right\}.$$

Proof We apply Theorem 2.13 from Graczyk and Ishi (2014) and formula (58). □

Corollary 5 *If $\underline{s} = s\mathbf{1}$, $s > \frac{1}{2}$, then $\gamma_{\underline{s},y}^{(M)} = \gamma_{s,y}$ does not depend on M. Moreover, for X with law $\gamma_{s,y}$, we have*

$$\mathbf{E}(\langle X, z^{(1)}\rangle \ldots \langle X, z^{(N)}\rangle) = \sum_{\pi\in S_N} s^{|C(\pi)|} \prod_{c\in C(\pi)} \operatorname{tr} \prod_{j\in c} y^{-1} z^{(j)}.$$

Example. For any graph A_n and $N = 3$ we get for X with law $\gamma_{s,y}$:

$$\mathbf{E}(\langle X, z^{(1)}\rangle\langle X, z^{(2)}\rangle\langle X, z^{(3)}\rangle) = s^3 \prod_{j=1}^{3} \operatorname{tr}(y^{-1}z^{(j)})$$
$$+s^2 \Big[\operatorname{tr} y^{-1}z^{(1)}y^{-1}z^{(2)} \operatorname{tr} y^{-1}z^{(3)} + \operatorname{tr} y^{-1}z^{(1)}y^{-1}z^{(3)} \operatorname{tr} y^{-1}z^{(2)}$$
$$+ \operatorname{tr} y^{-1}z^{(2)}y^{-1}z^{(3)} \operatorname{tr} y^{-1}z^{(1)}\Big] + s \left[\operatorname{tr} \prod_{j=1}^{3} y^{-1}z^{(j)} + \operatorname{tr} y^{-1}z^{(1)}y^{-1}z^{(3)}y^{-1}z^{(2)}\right].$$

6 Wishart exponential families on the cone P_G

A measure $\tilde{R}$ on P_G is said to be a Riesz measure if, for some $1 \leq M \leq n$, $s_M > -1$ and $s_i > -3/2, i \neq M,$ its Laplace transform is given by

$$L_{\tilde{R}}(x) = \int_{P_G} \mathrm{e}^{-\langle x,y\rangle} \tilde{R}(\mathrm{d}y) = \delta^{(M)}_{-\underline{s}}(x)\varphi_{Q_{A_n}}(x). \tag{59}$$

From formula (22), the measure $\tilde{R}^{(M)}_{\underline{s}}(\mathrm{d}y) = C_{\underline{s}}\Delta^{(M)}_{\underline{s}}(y)\mathrm{d}y$, where

$$C_{\underline{s}}^{-1} = \pi^{(n-1)/2}\left\{\prod_{i\neq M}\Gamma(s_i+\frac{3}{2})\right\}\Gamma(s_M+1),$$

is a Riesz measure. The exponential family of generated by $\tilde{R}^{(M)}_{\underline{s}}$ will be called the exponential family of Wishart distributions on P_G. Its density function is

$$\tilde{\gamma}^{(M)}_{\underline{s}}(y) = \frac{1}{\delta^{(M)}_{-\underline{s}}(x)\varphi_{Q_{A_n}}(x)}\mathrm{e}^{-\langle x,y\rangle}\tilde{R}^{(M)}_{\underline{s}}(\mathrm{d}y). \tag{60}$$

Its Laplace transform is

$$L_{\tilde{\gamma}^{(M)}_{\underline{s}}}(\theta) = \int_{P_G} \mathrm{e}^{-\langle\theta,y\rangle}\tilde{\gamma}^{(M)}_{\underline{s}}(y) = \frac{L_{\mu^{(M)}_{\underline{s}}}(\theta+x)}{L_{\mu^{(M)}_{\underline{s}}}(x)} = \frac{\delta^{(M)}_{-\underline{s}}(\theta+x)\varphi_{Q_{A_n}}(\theta+x)}{\delta^{(M)}_{-\underline{s}}(x)\varphi_{Q_{A_n}}(x)}. \tag{61}$$

6.1 Mean and covariance

Theorem 7 *The mean function of the Wishart exponential family on P_G is for all $s_i > -\frac{3}{2}$ and $x \in Q_G$,*

$$\begin{aligned}\tilde{m}^{(M)}_{\underline{s}}(x) = &\sum_{i=1}^{M-1}(s_i+\tfrac{3}{2})(x^{-1}_{\{i:i+1\}})^0 + \sum_{i=M+1}^{n}(s_i+\tfrac{3}{2})(x^{-1}_{\{i-1:i\}})^0\\ &-\sum_{i=2}^{M-1}(s_{i-1}+1)(x^{-1}_{ii})^0 - (s_{M-1}-s_M+s_{M+1}+1)(x^{-1}_{MM})^0\\ &-\sum_{i=M+1}^{n-1}(s_{i+1}+1)(x^{-1}_{ii})^0.\end{aligned} \tag{62}$$

The covariance function $\tilde{v}(x) : I_G \to Z_G$ *of the Wishart exponential family on* P_G *equals*

$$\begin{aligned}\tilde{v}(x) = &\sum_{i=1}^{M-1}(s_i+\tfrac{3}{2})\mathbb{P}\left[(x^{-1}_{\{i:i+1\}})^0\right] + \sum_{i=M+1}^{n}(s_i+\tfrac{3}{2})\mathbb{P}\left[(x^{-1}_{\{i-1:i\}})^0\right]\\ &-\sum_{i=2}^{M-1}(s_{i-1}+1)\mathbb{P}\left[(x^{-1}_{ii})^0\right] - (s_{M-1}-s_M+s_{M+1}+1)\mathbb{P}\left[(x^{-1}_{MM})^0\right]\\ &-\sum_{i=M+1}^{n-1}(s_{i+1}+1)\mathbb{P}\left[(x^{-1}_{ii})^0\right],\end{aligned}$$

where we identify I_G *with* Z_G *by the trace inner product.*

Proof We have $\tilde{m}^{(M)}_{\underline{s}}(x) = -\operatorname{grad}\log L_{\mu^{(M)}_{\underline{s}}}(x) = -\operatorname{grad}\log \delta^{(M)}_{-\underline{s}}(x)\varphi_{Q_{A_n}}(x)$. The covariance operator is obtained by differentiation of (62). □

6.2 Quadratic construction of Riesz measures and Wishart distributions on P_G

Let $M \in \{1,\dots,n\}$. Suppose $s_i > -\frac{3}{2}$, for all $i \neq M$ and $s_M > -1$. Let $\theta \in Q_G$. In order to establish a relation between quadratically constructed Riesz measures $\tilde{\mu}_q$ on P_G and the measures $\tilde{R}^{(M)}_{\underline{s}}$ we consider the sets $J_k = \{k, k+1\}$ and $J'_k = \{k\}$. As *basic quadratic maps* we choose the quadratic P_G-positive maps q^{J_k} and $q^{J'_k}$. For $\alpha = (\alpha_1,\dots,\alpha_{n-1})$ and $\beta = (\beta_1,\dots,\beta_n)$ with $\alpha_i, \beta_j \in \mathbb{N}$ define $q^{\alpha,\beta} = \sum^{\oplus}_{k<n}(q^{J_k})^{\oplus\alpha_k} \oplus \sum^{\oplus}_{k\le n}(q^{J'_k})^{\oplus\beta_k}$. The following proposition is easy to prove by comparing $\delta^{(M)}_{-\underline{s}}(\eta)\varphi_{Q_{A_n}}(\eta)$ with the Laplace transform

$$L_{\tilde{\mu}_{q^{\alpha,\beta}}}(\eta) = \pi^{(\sum_{k<n}\alpha_k+\sum_{k\le n}\beta_k)/2}\prod_{i<n}|\eta_{\{i,i+1\}}|^{-\alpha_i/2}\prod_{j\le n}|\eta_{jj}|^{-\beta_j/2}.$$

Proposition 12 *Let* $M \in \{2,\dots,n-1\}$. *Then there exists a constant* $c > 0$ *such that* $c\tilde{R}^{(M)}_{\underline{s}} = \tilde{\mu}_{q^{\alpha,\beta}}$ *if and only if* $\alpha_i/2 = s_i+3/2$, $i \le M-1$, $\alpha_i/2 = s_{i+1}+3/2$, $i \ge M$, $\beta_1 = 0$, $\beta_i/2 = -s_{i-1}-1$, $2 \le i \le M-1$, $\beta_M/2 = -s_{M-1}+s_M-s_{M+1}-1$, $\beta_i/2 = -s_{i+1}-1$, $M+1 \le i < n$, $\beta_n = 0$. *For* $M = 1, n$ *the condition* $\beta_M = 0$ *is suppressed.*

Proposition 12 implies easily two following facts.

Corollary 6 1. *All Riesz measures* $\tilde{R}^{(M)}_{\underline{s}}$ *are equal (up to a factor) to a virtual quadratic Riesz measure* $\tilde{\mu}_{q^{\alpha,\beta}}$.
2. *For* $n \ge 4$, *no true quadratic Riesz measure* $\tilde{\mu}_{q^{\alpha,\beta}}$, $\alpha_i, \beta_j \in \mathbb{N}$, *belongs (up to a factor) to the set of Riesz measures* $\tilde{R}^{(M)}_{\underline{s}}$.

Proof To prove Part 2, we have conditions $s_i + \frac{3}{2} = \alpha_{i'}/2$ and $s_j + 1 = -\beta_{j'}/2$, so all (except at most one) $s_i \ge -1$, and all (except at most one) $s_j \le -3/2$ simultaneously. □

6.3 Higher-order moments of Wishart families on P_{A_n}

Thanks to Part 1 of Corollary 6, all the moments of the Wishart Exponential Families $\tilde{\gamma}^{(M)}_{\underline{s},\theta}$ can be computed, using Theorem 2.13 from Graczyk and Ishi (2014) and Proposition 12.

Theorem 8 *Let Y be a P_{A_n}-valued random variable with the Wishart law $\tilde{\gamma}^{(M)}_{\underline{s},\theta}$. Let $x^{(1)}, x^{(2)}, \ldots, x^{(N)} \in I_G$. Then, denoting by $C(\pi)$ the set of cycles of a permutation $\pi \in S_N$, the N-th moment $\mathbf{E}(\langle Y, x^{(1)}\rangle \ldots \langle Y, x^{(N)}\rangle)$ equals*

$$\sum_{\pi \in S_N} \prod_{c \in C(\pi)} \Big\{ \sum_{i=1}^{M-1} (s_i + \frac{3}{2}) \operatorname{tr} \prod_{j \in c} (\theta_{\{i:i+1\}})^{-1} x^{(j)}_{\{i:i+1\}}$$
$$+ \sum_{i=M}^{n-1} (s_{i+1} + \frac{3}{2}) \operatorname{tr} \prod_{j \in c} (\theta_{\{i:i+1\}})^{-1} x^{(j)}_{\{i:i+1\}} - \sum_{i=2}^{M-1} (s_{i-1} + 1)\theta_{ii}^{-|c|} \prod_{j \in c} x^{(j)}_{ii}$$
$$- (s_{M-1} - s_M + s_{M+1} + 1)\theta_{MM}^{-|c|} \prod_{j \in c} x^{(j)}_{MM} - \sum_{i=M+1}^{n-1} (s_{i+1} + 1)\theta_{ii}^{-|c|} \prod_{j \in c} x^{(j)}_{ii} \Big\}.$$

7 Relations with the type I and type II Wishart distributions of Letac and Massam (2007)

In this section, we will explain the relation between our work and type 1 and type 2 Wishart distributions constructed by Letac and Massam (2007).

Letac and Massam (2007) introduced, studied and used the function $H(\alpha, \beta, x)$ on Q_G as a generalized power function for constructing type I and type II Wishart distributions. The reader is referred to the cited paper for the general definition of the function $H(\alpha, \beta, x)$ as well as for graphical theoretic notions such as cliques, separators and perfect order of cliques (see also Lauritzen 1996). For our purpose, it is sufficient to recall that for $\alpha \in \mathbb{R}^{n-1}$ and $\beta \in \mathbb{R}^{n-2}$

$$H(\alpha, \beta; x) = \frac{\prod_{i=1}^{n-1} |x_{\{i,i+1\}}|^{\alpha_i}}{\prod_{i=2}^{n-1} x_{ii}^{\beta_i}}, \quad x \in Q_{A_n}, \tag{63}$$

that the cliques(i.e. the sets of vertices of maximal complete subgraphs) are $\{1, 2\}, \ldots, \{n-1, n\}$ and the separators $\{2\}, \ldots, \{n-1\}$. The definition of the function $H(\alpha, \beta; x)$ does not include any restrictions on the values of the parameter (α, β) of dimension $2n - 3$.

However, the existence of type I Wishart distributions on Q_G is only showed for (α, β) belonging to some set A_P dependent on a perfect order of cliques P, i.e. for $(\alpha, \beta) \in \mathcal{A}_0 = \cup_P A_P$, where the union is on all perfect order of cliques. Proposition 14 describes this set for A_n graphs. It also makes clear a phenomenon observed by Letac and Massam (2007) for the graph A_4, where there are only two different sets A_P although there are 4 perfect orders of cliques. To prove Proposition 14, we use the following explicit relation between two concepts: perfect orders of cliques used by Letac and Massam (2007) and eliminating orders of vertices used in this work.

Proposition 13 *Let $G = A_n : 1-2-3-\cdots-n$. A clique ordering $C'_1 < \cdots < C'_{n-1}$ is perfect if and only if $C'_{n-1} \prec \ldots \prec C'_1$ is an eliminating order on the A_{n-1} graph $G' : C_1 - C_2 \ldots - C_{n-1}$. There are 2^{n-2} perfect orders of cliques on A_n.*

The proof is easy and is omitted.

Proposition 14 *Let $P' : C'_1 < C'_2 < \ldots < C'_{n-1}$ and $P'' : C''_1 < C''_2 < \ldots < C''_{n-1}$ be two perfect orders of cliques on $G = A_n$. Let S'_2 and S''_2 be the first separators of P' and P''. If $S'_2 = S''_2$ then $A_{P'} = A_{P''}$, i.e. the parameter set A_P depends only on the first separator S_2 with respect to the clique order P. If $S_2 = \{M\}$ then the set A_P is described by the conditions:*

$$\begin{cases} \alpha_j = \beta_{j+1} & \text{if } 1 \leq j \leq M-2, \\ \alpha_j = \beta_j & \text{if } M+1 \leq j \leq n-1, \end{cases} \tag{64}$$

and

$$\alpha_j > \frac{1}{2} \text{ for all } 1 \leq j \leq n-1;\ \alpha_{M-1} + \alpha_M - \beta_M > 0. \tag{65}$$

Thus $\mathcal{A}_0 = \cup_P A_P$ is the set of (α, β) such that there exists $2 \leq M \leq n-1$ for which (64) and (65) are satisfied.

Proof We use Propositions 1 and 13. □

The reference measure μ_G used by Letac and Massam (2007) is, on the cone Q_{A_n},

$$\mu_{A_n}(x)(\mathrm{d}x) = H_{A_n}(-\frac{3}{2}\mathbb{1}, -\mathbb{1}; x) 1_{Q_{A_n}}(x)\mathrm{d}x. \tag{66}$$

By (13), we observe that $\mu_{A_n}(x)(\mathrm{d}x) = \varphi_{Q_{A_n}}(x) 1_{Q_{A_n}}(x)\mathrm{d}x$. Namely, the reference measure μ_G is the characteristic measure of the cone $G = Q_{A_n}$.

Theorem 9 [Letac and Massam (2007) Theorem 3.3] *If $(\alpha, \beta) \in \mathcal{A}_0$, then, for a constant $\Gamma_{1(\alpha,\beta)}$, and for all $y \in P_{A_n}$*

$$\int_{Q_{A_n}} \mathrm{e}^{-\operatorname{tr}(xy)} H(\alpha, \beta; x) \mu_{A_n}(x)(\mathrm{d}x) = \Gamma_{1(\alpha,\beta)} H(\alpha, \beta; \pi(y^{-1})).$$

The methods of our article give a new simple proof of Theorem 9, see the proof of Corollary 7 below.

Let us compare now the functions $H(\alpha, \beta; x)$ and $H(\alpha, \beta; \pi(y^{-1}))$ with the generalized power functions $\delta_{\underline{s}}^{(M)}$ and $\Delta_{\underline{s}}^{(M)}$.

Proposition 15 1. *Let* $\alpha \in \mathbb{R}^{n-1}$ *and* $\beta \in \mathbb{R}^{n-2}$. *There exists* $\underline{s} \in \mathbb{R}^n$ *such that* $H(\alpha, \beta; x) = \delta_{\underline{s}}^{(M)}(x)$ *if and only if* (64) *holds for some* $2 \le M \le n-1$.
Then $s_j = \alpha_j$ *if* $1 \le j \le M-1$, $s_M = \alpha_{M-1} + \alpha_M - \beta_M$ *and* $s_j = \alpha_{j-1}$ *if* $M+1 \le j \le n$.

2. *Moreover, under the hypothesis of Part 1, we have* $H(\alpha, \beta; \pi(y^{-1})) = \Delta_{-\underline{s}}^{(M)}(y)$.

The proof is easy and is omitted.

Corollary 7 *The type I Wishart distributions indexed by the set* $\mathcal{A}_0$ *are equal to the subset* $\bigcup_{M=2}^{n-1} (\gamma_{\underline{s},y}^{(M)})_{y \in P_G}$ *of Wishart NEF families defined in Sect.* 5. *Thus, they are strictly contained in the set of all Wishart NEF families on* Q_G, *equal to* $\bigcup_{M=1}^{n} (\gamma_{\underline{s},y}^{(M)})_{y \in P_G}$.

The proof is easy and is omitted.

Note that Theorem 2 implies Theorem 9 of Letac and Massam (2007).

The family of functions $H(\alpha, \beta, x)$ does not contain the power functions $\delta_{\underline{s}}^{(1)}$ or $\delta_{\underline{s}}^{(n)}$. In fact, the last functions contain powers of $n-1$ diagonal elements x_{ii}, whereas the function $H(\alpha, \beta, x)$ contains powers of $n-2$ such elements. Similar comparisons can be done on the cones P_G. In this case, Letac and Massam (2007) define type II Wishart distributions on P_G indexed by a set $\mathcal{B}_0$, analogous to the set $\mathcal{A}_0$ for Q_G. Similar arguments as on the cone Q_G lead to

Corollary 8 *The type II Wishart distributions on* P_G *indexed by the set* $\mathcal{B}_0$ *are equal to the subset* $\bigcup_{M=2}^{n-1} (\tilde{\gamma}_{\underline{s},x}^{(M)})_{x \in Q_G}$ *of Wishart NEF families defined in Sect.* 6. *Thus they are strictly contained in the set of all Wishart NEF families on* P_G, *equal to* $\bigcup_{M=1}^{n} (\tilde{\gamma}_{\underline{s},x}^{(M)})_{x \in Q_G}$.

Acknowledgements The authors would like to thank Gérard Letac and two anonymous referees for their insightful suggestions.

8 Appendix

We list here some properties of triangular matrices, used in proofs.

Lemma 7 1. *Let* $A = K^0$, *where* $K = A_{\{1:k\}}$ *and let* L *be lower triangular and* U *upper triangular* $n \times n$ *matrices. Then* $UAL = \left(U_{\{1:k\}} K L_{\{1:k\}}\right)^0$.

2. *Let* M, L, U *be matrices* $n \times n$, *with* L *lower triangular and* U *upper triangular. Then, for all* $i = 1, \ldots, n$, $(LMU)_{\{1:i\}} = L_{\{1:i\}} M_{\{1:i\}} U_{\{1:i\}}$ *and* $(UML)_{\{i:n\}} = U_{\{i:n\}} M_{\{i:n\}} L_{\{i:n\}}$.
3. *If* T *is an invertible triangular matrix then* $(T_{\{1:k\}})^{-1} = (T^{-1})_{\{1:k\}}$ *for all* $k = 1, \ldots, n$.

References

Amari, S., Nagaoka, H. (2007). *Methods of information geometry. Translations of mathematical monographs* (Vol. 191). New York: American Mathematical Society.

Andersson, S. A., Klein, T. (2010). On Riesz and Wishart distributions associated with decomposable undirected graphs. *Journal of Multivariate Analysis*, *101*(4), 789–810.

Bondy, J. A., Murty, U. S. (2008). *Graph theory. Graduate texts in mathematics* (Vol. 244). London: Springer.

Brown, L. D. (1986). *Fundamentals of statistical exponential families with applications in statistical decision theory*. Lecture Notes-Monograph Series, 9. Hayward, California: Institute of Mathematical Statistics.

Casalis, M., Letac, G. (1996). The Lukacs–Olkin–Rubin characterization of Wishart distributions on symmetric cones. *The Annals of Statistics*, *24*(2), 763–786.

Chandran, L. S., Ibarra, L., Ruskey, F., Sawada, J. (2003). Generating and characterizing the perfect elimination orderings of a chordal graph. *Theoretical Computer Science*, *307*(2), 303–317.

Dawid, A. P., Lauritzen, S. L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. *The Annals of Statistics*, *21*(3), 1272–1317.

Faraut, J., Korányi, A. (1994). *Analysis on symmetric cones. Oxford mathematical monographs*. New York: Oxford University Press.

Graczyk, P., Ishi, H. (2014). Riesz measures and Wishart laws associated to quadratic maps. *Journal of the Mathematical Society of Japan*, *66*(1), 317–348.

Graczyk, P., Ishi, H., Kołodziejek, B. (2016). Wishart laws and variance function on homogeneous cones. arXiv:1802.02352

Graczyk, P., Ishi, H., Mamane, S., Ochiai, H. (2017). On the Letac–Massam conjecture on cones Q_{A_n}. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, *93*(3), 16–21.

Ishi, H. (2014). Homogeneous cones and their applications to statistics. In P. Graczyk, A. Hassairi (Eds.) *Modern Methods of Multivariate Statistics*, vol. 82, pp. 135–154. Paris: Hermann.

Konno, Y. (2007). Estimation of normal covariance matrices parametrized by irreducible symmetric cones under Stein's loss. *Journal of Multivariate Analysis*, *98*(2), 295–316.

Konno, Y. (2009). Shrinkage estimators for large covariance matrices in multivariate real and complex normal distributions under an invariant quadratic loss. *Journal of Multivariate Analysis*, *100*(10), 2237–2253.

Kuriki, S., Numata, Y. (2010). Graph presentations for moments of noncentral Wishart distributions and their applications. *Annals of the Institute of Statistical Mathematics*, *62*(4), 645–672.

Lauritzen, S. L. (1996). *Graphical models. Oxford statistical series* (Vol. 17). New York: Oxford University Press.

Letac, G., Massam, H. (2007). Wishart distributions for decomposable graphs. *The Annals of Statistics*, *35*(3), 1278–1323.

Roverato, A. (2000). Cholesky decomposition of a hyper inverse Wishart matrix. *Biometrika*, *87*(1), 99–112.

Speed, T., Kiiveri, H. (1986). Gaussian Markov distributions over finite graphs. *The Annals of Statistics*, *14*(1), 138–150.

Sugiura, N., Konno, Y. (1988). Entropy loss and risk of improved estimators for the generalized variance and precision. *Annals of the Institute of Statistical Mathematics*, *40*(2), 329–341.

Tsukuma, H., Konno, Y. (2006). On improved estimation of normal precision matrix and discriminant coefficients. *Journal of Multivariate Analysis*, *97*(7), 1477–1500.

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