The Annals of Statistics
1998, Vol. 26, No. 3, 1051–1082

ESTIMATION AND TESTING FOR LATTICE CONDITIONAL INDEPENDENCE MODELS ON EUCLIDEAN JORDAN ALGEBRAS

BY HÉLÈNE MASSAM[1] AND ERHARD NEHER[2]

York University and University of Virginia, and University of Ottawa

In this paper we generalize the major results of Andersson and Perlman on LCI models to the setting of symmetric cones and give an explicit closed form formula for the estimate of the covariance matrix in the generalized LCI models that we define.

To this end, we replace the cone $H_I^+(\mathbb{R})$ sitting inside the Jordan algebra of symmetric real $I \times I$-matrices by the symmetric cone Ω of an Euclidean Jordan algebra V. We introduce the *Andersson-Perlman cone* $\Omega(\mathscr{K}) \subseteq \Omega$ which generalizes $\mathscr{P}(\mathscr{K}) \subseteq H_I^+(\mathbb{R})$. We prove several characterizations and properties of $\Omega(\mathscr{K})$ which allows us to recover, though with different proofs, the main results of Andersson and Perlman regarding $\mathscr{P}(\mathscr{K})$. The new lattice conditional independence models are defined, assuming that the Euclidean Jordan algebra V has a symmetric representation. Using standard results from the theory of Jordan algebras, we can reduce the general model to the case where V is the Jordan algebra of Hermitian matrices over the real, complex or quaternionic numbers, and Ω is the corresponding cone of positive-definite matrices. Our main statistical result is a closed-form formula for the estimate of the covariance matrix in the generalized LCI model. We also give the likelihood ratio test for testing a given model versus another one, nested within the first.

1. Introduction.

1.1. *Background.* Consider a finite index set $I = \{1, \dots, n\}$. Let $\mathscr{K}$ be a *lattice* on I, that is, a family of subsets of I which is closed under union and intersection and contains $\varnothing$ and I. Andersson and Perlman (abbreviated henceforth as AP) (1993) introduced a class of normal $N(0, \Sigma)$ models for which conditional independence was determined by a lattice $\mathscr{K}$. These models are called *lattice conditional independence models*, abbreviated "LCI models," and are denoted by $N(\mathscr{K})$. The set of covariance matrices of the $N(0, \Sigma)$ distributions in $N(\mathscr{K})$ form a (in general nonconvex) cone $\mathscr{P}(\mathscr{K})$, contained in the cone $H_I^+(\mathbb{R})$ of all real positive-definite $I \times I$-matrices, so that

$$N(\mathscr{K}) = \{N(0, \Sigma): \Sigma \in \mathscr{P}(\mathscr{K})\}.$$

AP (1993) gives several characterizations and properties of the cone $\mathscr{P}(\mathscr{K})$ and also an algorithm to obtain the maximum likelihood estimate of the covariance matrix $\Sigma \in \mathscr{P}(\mathscr{K})$.

Received July 1996; revised January 1998.

[1]Supported in part by NSERC Grant A8947.

[2]Supported by NSERC Grant A8836.

AMS 1991 *subject classification*. 62E15.

Key words and phrases. Lattice, conditional independence, covariance selection, Wishart distributions, symmetric cones, Euclidean Jordan algebras.

1.2. *A detailed description.* We now give a detailed description of the contents of this paper. The AP cone $\Omega(\mathscr{K})$ which we study in Section 2 is defined in terms of a complete orthogonal system $\mathscr{E} = (e_1, \dots, e_n)$ of idempotents of V and a lattice $\mathscr{K}$ of subsets of $I = \{1, \dots, n\}$. For its definition we need the following notation. Let P be the quadratic representation of V, defined in terms of the Jordan algebra product by $P(x)y = 2x(xy) - x^2 y$ for $x, y \in V$. For $K \subseteq I = \{1, \dots, n\}$ and $y \in V$, let $e_K = \sum_{i \in K} e_i$, $y_K = P(e_K)y$ and $V_K = P(e_K)V$. If $x \in \Omega$ and $K \neq \varnothing$, one knows that $x_K \in \Omega_K = P(e_K)\Omega$, the cone of the Euclidean Jordan algebra V_K, and hence x_K is invertible in V_K. We denote by x_K^{-1} the inverse of x_K in V_K (not to be confused with the V_K-component of x^{-1}). For easier formulation in the following, it is useful to define a sum over an empty set as 0 (it will be clear from the context where the 0 lies) so that $e_\varnothing = 0$ and $V_\varnothing = \{0\}$. Also, we put $0^{-1} = 0$. Let tr be the trace form of V. The *AP cone* $\Omega(\mathscr{K})$ is then defined as the set of $x \in \Omega$ satisfying for all $L, M \in \mathscr{K}$ and $y \in V$,

$$\operatorname{tr}(x_{L\cup M}^{-1}\, y_{L\cup M}) + \operatorname{tr}(x_{L\cap M}^{-1}\, y_{L\cap M}) = \operatorname{tr}(x_L^{-1}\, y_L) + \operatorname{tr}(x_M^{-1}\, y_M).$$

It is easily seen that this condition generalizes the cone $\mathscr{P}(\mathscr{K})$ of AP (1993). As in that paper, we then give several other descriptions of $\Omega(\mathscr{K})$; see Theorem 1. First, since the trace form is nondegenerate, the trace condition is equivalent to

$$x_{L\cup M}^{-1} + x_{L\cap M}^{-1} = x_L^{-1} + x_M^{-1}$$

for all $L, M \in \mathscr{K}$. It is remarkable that $\Omega(\mathscr{K})$ can in fact be defined by a single equation, involving the set $\mathscr{J}(\mathscr{K})$ of all join-irreducible sets in $\mathscr{K}$. Here, $\varnothing \neq K \in \mathscr{K}$ is called *join-irreducible* if it is not a union of proper nonempty subsets belonging to $\mathscr{K}$. If we define $\langle K\rangle := \cup\{K' \in \mathscr{K};\ K' \subseteq K, K' \neq K\} \in \mathscr{K}$ then K is join-irreducible if and only if $K \neq \langle K\rangle$. We show that $x \in \Omega(\mathscr{K})$ if and only if

$$x^{-1} = \sum_{K \in \mathscr{J}(\mathscr{K})} \big(x_K^{-1} - x_{\langle K\rangle}^{-1}\big).$$

The different characterizations of $\Omega(\mathscr{K})$ are used to introduce the *Frobenius coordinates* of $x \in \Omega(\mathscr{K})$ and the so-called *$\mathscr{K}$-parametrization* of $\Omega(\mathscr{K})$; letting

$$x_{[K\rangle} = P\big(e_{\langle K\rangle} + e_{[K]}\big)x - P\big(e_{\langle K\rangle}\big)x - P(e_{[K]})x,$$

the map

$$\Omega(\mathscr{K}) \ni x \mapsto \prod_{K \in \mathscr{J}(\mathscr{K})} \big(x_{\langle K\rangle}^{-1} x_{[K\rangle},\, x_{[K]} - P(x_{[K\rangle})x_{\langle K\rangle}^{-1}\big)$$

is injective, and its image can be precisely described; see Section 2.7.

The reader who is not familiar with Jordan algebras can easily translate the expressions above into standard matrix notation. If we use the notation

$$x = x_1 + x_{12} + x_0 \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

for the *same* matrix where x is the Jordan algebra notation and Σ is the standard block-matrix notation in the space of Hermitian matrices on $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, the quaternions, we have

$$x_1 = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & 0 \end{pmatrix}, \qquad x_{12} = \begin{pmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{pmatrix},$$

$$x_0 = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{22} \end{pmatrix}, \qquad P(x_{12})x_1^{-1} = \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

and

$$x_0 - P(x_{12})x_1^{-1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \Sigma_{22\bullet 1}.$$

If $\Sigma = (\Sigma_{ij})_{1\le i,\, j\le n}$ and $K \subseteq \{1, \dots, n\}$, we have $x_K = (x_{ij})_{i,\, j\in K}$. For easier notation, say $K = \{1, \dots, m\}$ and $\langle K\rangle = \{1, \dots, k\}$, so that $[K] = \{k + 1, \dots, m\}$. Then $x_K = x_{\langle K\rangle} + x_{[K\rangle} + x_{[K]}$ with

$$x_{\langle K\rangle} = \begin{pmatrix} \Sigma_{\langle K\rangle} & 0 \\ 0 & 0 \end{pmatrix},$$

$$x_{[K\rangle} = \begin{pmatrix} 0 & \Sigma_{\langle K]} \\ \Sigma_{[K\rangle} & 0 \end{pmatrix}$$

and

$$x_{[K]} = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{[K]} \end{pmatrix}.$$

Such a matrix is positive definite if and only if Σ_{11} and $\Sigma_{22\bullet 1}$ are positive definite. Indeed, for any $(m-k) \times k$ matrix y the *Frobenius transformation* $\tau(y)$ defined by

$$\tau(y)(\Sigma) = \begin{pmatrix} I & 0 \\ y & I \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} I & y^t \\ 0 & I \end{pmatrix}$$

leaves the cone of positive-definite matrices invariant, and for $y = -\Sigma_{21}\Sigma_{11}^{-1}$ we obtain

$$\tau(-\Sigma_{21}\Sigma_{11}^{-1}) \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22\cdot 1} \end{pmatrix}.$$

The Frobenius transformation is what statisticians have traditionally called the "sweep operator." We observe that $x = \Sigma$ can be completely recovered once we know its Frobenius coordinates $(\Sigma_{11}, \Sigma_{21}\Sigma_{11}^{-1}, \Sigma_{22\bullet 1})$ in matrix notation, or

$$\left(x_1, x_1^{-1}x_{12}, x_2 - P(x_{21})x^{-1}\right) = \left(x_{\langle K\rangle}, x_{\langle K\rangle}^{-1}x_{[K\rangle}, x_{[K]} - P(x_{[K\rangle})x_{\langle K\rangle}^{-1}\right)$$

in Jordan algebra notation. The $\mathscr{K}$-parametrization of $\Omega(\mathscr{K})$ for Hermitian matrices over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ is

$$\Omega(\mathscr{K}) \ni \Sigma \mapsto \prod_{K\in\mathscr{J}(\mathscr{K})} \left(\Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1}, \Sigma_{[K]} - P(\Sigma_{[K\rangle})\Sigma_{\langle K\rangle}^{-1}\right)$$

as given in AP (1993), for the real case. We denote

$$\Sigma_{[K]\bullet} = \Sigma_{[K]} - P(\Sigma_{[K\rangle})\Sigma_{\langle K\rangle}^{-1} \quad \text{and} \quad P(\Sigma_{[K\rangle})\Sigma_{\langle K\rangle}^{-1} = \Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1}\Sigma_{\langle K]}.$$

For our main statistical result, Theorem 3, giving the closed-form expression of $\widehat{\Sigma}$ in $\Omega(\mathscr{K})$, we do not need the full $\mathscr{K}$-parametrization of $\Omega(\mathscr{K})$. Indeed, the proof of Theorem 2 giving the generalized $\mathscr{K}$-parametrization is based on Proposition 3, which, in a special case, can be reformulated as follows. For $L \in \mathscr{K}$, let $\mathscr{K}_L = \{K \in \mathscr{K} | K \subset L\}$. Given a regular decomposition of L, that is M, K in $\mathscr{K}$ such that $L = M \dot\cup [K]$ and $[K] \neq \varnothing$, then $\langle K\rangle \subset M$ and we can prove that for a given $(\Sigma_M, \Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1}, \Sigma_{[K]\bullet})$ there exists a unique $\Sigma_L \in \Omega(\mathscr{K}_L)$ such that

$$(\Sigma_L)_M = \Sigma_M, \tag{1.1}$$

$$\Sigma_L = \tau(\Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1})\begin{pmatrix} \Sigma_M & 0 \\ 0 & \Sigma_{[K]\bullet} \end{pmatrix}. \tag{1.2}$$

From (1.1) and (1.2), it is clear that the Frobenius transformation $\tau(\Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1})$ leaves $(\Sigma_L)_M = \Sigma_M$ unchanged. We do not need the uniqueness of Σ with a given $\mathscr{K}$-parametrization. It is just the uniqueness of Σ_L satisfying (1.1) and (1.2) and the fact that Σ_M is unchanged by $\tau(\Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1})$ that allows us to show that, if S denotes the sample covariance matrix and $\widehat{\Sigma_L}$ is the maximum likelihood estimate (abbreviated mle) of Σ in the L-marginal models, then

$$\widehat{\Sigma_L^{-1}} - \widehat{\Sigma_M^{-1}} = \widehat{S_K^{-1}} - \widehat{S_{\langle K\rangle}^{-1}}. \tag{1.3}$$

That is, the difference between the mle of the concentration matrices of the L- and M-marginal models, with the given independences, is the same as the differences between the mle of the concentration matrices in the saturated K- and $\langle K\rangle$-marginal models. The usage of some trace and determinant formulas to be given below, and an induction argument yields Theorem 3.

1.3. *Advantages.* All the main results of Section 2 of this paper are proved in AP (1988) and (1993), Section 2, for the special case of real symmetric matrices, but the proofs there do not seem to be adjustable to the setting of symmetric cones. We follow a different approach here which emphasizes Peirce decompositions and Frobenius transformations, while the main techniques of AP (1993) are transformation groups. We give more details in Theorem 1 and Section 2.3(d). Our proofs are new, even for the case of real symmetric matrices. We use the framework of Euclidean Jordan algebras because this provides us with an easy and, in our opinion, elegant notational scheme which allows us to avoid lengthy and cumbersome matrix calculations. Throughout the paper, we translate the most important results into matrix notation for the convenience of the reader.

1.4. *Relation to other work.* In an LCI model, the conditional independences between the different variables are determined by a lattice $\mathscr{K}$ on the index set I. These independences can also be represented by a transitive acyclic digraph. Indeed, Andersson, Madigan, Perlman and Triggs (1997) have proved that LCI models coincide with a subclass of the class of graphical Markov models determined by acyclic digraphs (ADG), namely the subclass of transitive ADG models. Our work is therefore on a special class of graphical models and is related to a number of other papers on graphical models apart from AP (1988, 1993), mentionned above. The main ones are Lauritzen (1989, 1996), Anderson, Højbjerre, Sørensen and Eriksen (1995), AP (1995a, b, 1997) and Andersson and Madsen (1998). We give more details in Section 3.3.2. The models considered in Lauritzen (1989, 1996) are ADG models, not necessarily transitive ones. We show in Section 3.3.1 that the results in our paper also hold for the more general class of ADG Markov models.

2. The AP cone. In this section we will develop the mathematical background for lattice conditional independence models on Euclidean Jordan algebras.

2.1. *Rings of sets.* We begin by reviewing some combinatorics which will be needed in the definition of the main object of this section.

Let I be a finite nonempty set. We denote by $\mathscr{D}(I)$ the set of all subsets of I. A *unital subring of* $\mathscr{D}(I)$ is a set of subsets of I which is closed under union and intersection and which contains $\varnothing$ and I. A *ring of sets* is a unital subring of some $\mathscr{D}(I)$. In the following, $\mathscr{K} \subseteq \mathscr{D}(I)$ will always be a ring of sets. In particular, $\mathscr{K}$ is a finite distributive lattice with respect to the partial order $\subseteq$. One calls $\varnothing \neq K \in \mathscr{K}$ *join-irreducible* if $K = L \cup M$, where $L, M \in \mathscr{K}$, implies $K = L$ or $K = M$. To describe the set $\mathscr{J}(\mathscr{K})$ of join-irreducible elements of $\mathscr{K}$, let us introduce, for $K \in \mathscr{K}$, $K \neq \varnothing$, the subsets

$$\langle K \rangle := \cup\{K' \in \mathscr{K};\ K' \subseteq K,\ K' \neq K\} \quad \text{and} \quad [K] := K \setminus \langle K \rangle.$$

Hence, any $K \in \mathscr{K}$, $K \neq \varnothing$ has a decomposition $K = \langle K \rangle \uplus [K]$ where $\uplus$ indicates a disjoint union, and

$$J \in \mathscr{J}(\mathscr{K}) \quad \Leftrightarrow \quad J \neq \langle J \rangle \quad \Leftrightarrow \quad [J] \neq \varnothing.$$

For $K \in \mathscr{K}$ put $\mathscr{K}_K = \{L \in \mathscr{K};\ L \subseteq K\}$. We note that $K = \varnothing$ is allowed, in which case we have $\mathscr{K}_\varnothing = \{\varnothing\}$ and put $\mathscr{J}(\mathscr{K}_\varnothing) = \varnothing$. Then [AP (1993), Section 2.1] for any $K, L \in \mathscr{K}$,

$$\mathscr{J}(\mathscr{K}_K) = \mathscr{J}(\mathscr{K}) \cap \mathscr{K}_K, \tag{2.1}$$

$$\mathscr{J}(\mathscr{K}_{K\cup L}) = \mathscr{J}(\mathscr{K}_K) \cup \mathscr{J}(\mathscr{K}_L), \qquad \mathscr{J}(\mathscr{K}_{K\cap L}) = \mathscr{J}(\mathscr{K}_K) \cap \mathscr{J}(\mathscr{K}_L), \tag{2.2}$$

$$([J];\ J \in \mathscr{J}(\mathscr{K}_K)) \text{ is a partition of } K \in \mathscr{K},\ K \neq \varnothing. \tag{2.3}$$

In particular, $K = \uplus([K'];\ K' \subseteq K)$. By AP (1993), Section 2.7, one can always find a *never-decreasing listing* of the poset $(\mathscr{J}(\mathscr{K}), \subseteq)$, that is, an enumeration

$\mathcal{J}(\mathcal{K}) = (K_1, \dots, K_q)$ with the property $i < j \Rightarrow K_j \not\subseteq K_i$. For such a listing one has

$$K_1 \cup K_2 \cup \cdots \cup K_k = [K_1] \uplus [K_2] \uplus \cdots \uplus [K_k] \ (1 \le k \le q). \tag{2.4}$$

A decomposition of $\varnothing \ne L \in \mathcal{K}$ of the form

$$L = [K] \uplus M,\ \ K \in \mathcal{J}(\mathcal{K}),\ \ M \in \mathcal{K} \tag{2.5}$$

will be called a *regular decomposition of* L. Regular decompositions always exist. Indeed, if $\mathcal{J}(\mathcal{K}_L) = (K_1, \dots, K_l)$ is a never-decreasing listing then $K = K_l$ and $M = K_1 \cup \cdots \cup K_{l-1}$ satisfy the conditions of (2.5).

LEMMA 1. *Let R be an Abelian group written additively, and let t: $\mathcal{K} \to R$ be a function with $t(\varnothing) = 0$. Then the following are equivalent*:

$$t(L \cup M) + t(L \cap M) = t(L) + t(M) \quad \textit{for all } L, M \in \mathcal{K}, \tag{2.6}$$

$$t(L) = \sum_{K \in \mathcal{J}(\mathcal{K}_L)} t(K) - t(\langle K \rangle) \quad \textit{for all } L \in \mathcal{K}, \tag{2.7}$$

$$t(L) = t(K) - t(\langle K \rangle) + t(M) \quad \textit{for any } \varnothing \ne L \in \mathcal{K} \setminus \mathcal{J}(\mathcal{K}) \tag{2.8}$$

and any regular decomposition of L as in (2.5),

$$t(L) = t(K) - t(\langle K \rangle) + t(M) \quad \textit{for any } \varnothing \ne L \in \mathcal{K} \setminus \mathcal{J}(\mathcal{K}) \tag{2.9}$$

and some regular decomposition of L as in (2.5).

PROOF [Inspired by the proof of AP (1993), Theorem 2.1]. The direction (2.7) $\Rightarrow$ (2.6) follows immediately from (2.2) and (2.7) for $L \cup M$, $L \cap M$, L and M. For the proof of (2.6) $\Rightarrow$ (2.7) we first observe that (2.7) always holds for $L = \varnothing$ due to the assumption $t(\varnothing) = 0$ and our convention that the sum over an empty set is 0. We can therefore assume $L \ne \varnothing$. We use induction on $|\mathcal{J}(\mathcal{K})| = n$. If $n = 1$ we have $\mathcal{K} = \{\varnothing, I\}$ by (2.3), and so $\mathcal{J}(\mathcal{K}) = \{I\}$ with $\langle I \rangle = \varnothing$ and (2.7) holds. Hence assume $n > 1$ and let $L \in \mathcal{K}$, $L \ne \varnothing$. For easier notation, put $\delta(K) = t(K) - t(\langle K \rangle)$.

CASE 1. $L \in \mathcal{J}(\mathcal{K})$: then $\langle L \rangle$ is a proper subset of L, and hence $\mathcal{J}(\mathcal{K}_L) = \{L\} \uplus \mathcal{J}(\mathcal{K}_{\langle L \rangle})$ by (2.1). In particular, $|\mathcal{J}(\mathcal{K}_{\langle L \rangle})| < n$ and hence $t(\langle L \rangle) = \sum_{K \in \mathcal{J}(\mathcal{K}_{\langle L \rangle})} \delta(K)$ by induction. But then (2.7) follows from $t(L) = (t(L) - t(\langle L \rangle)) + t(\langle L \rangle)$.

CASE 2. $L \notin \mathcal{J}(\mathcal{K})$: then $L = K \cup M$ with $K \ne L \ne M$. In particular, $|\mathcal{J}(\mathcal{K}_K)| < n$ by (2.3), and so $t(K) = \sum_{N \in \mathcal{J}(\mathcal{K}_K)} \delta(N)$. Similarly, $t(M) = \sum_{N \in \mathcal{J}(\mathcal{K}_M)} \delta(N)$ and $t(K \cap M) = \sum_{N \in \mathcal{J}(\mathcal{K}_{K \cap M})} \delta(N)$ so that (2.7) follows from (2.6) and (2.2).

(2.6) $\Rightarrow$ (2.8): for any regular decomposition $L = [K] \uplus M$, where $K \in \mathcal{J}(\mathcal{K})$ and $M \in \mathcal{K}$, we have $K \cap M = \langle K \rangle$ and hence (2.8) follows from (2.6) applied to $L = K \cup M$.

The implication (2.8) $\Rightarrow$ (2.9) is trivial, so it suffices to show (2.9) $\Rightarrow$ (2.7) which we will do by induction on $|\mathscr{J}(L)|$. As in the proof above, the case $|\mathscr{J}(L)| \le 1$ is clear. So we can assume $|\mathscr{J}(L)| \ge 2$. If $L \notin \mathscr{J}(\mathscr{K})$, let $L = [K] \uplus M$ be the given regular decomposition; if $L \in \mathscr{J}(\mathscr{K})$ we take the regular decomposition $L = [L] \uplus M$ with $M = \langle L \rangle$. In both cases $\mathscr{J}(\mathscr{K}) = \{K\} \uplus \mathscr{J}(\mathscr{K}_M)$. Hence (2.7) holds for M by induction. But then (2.7) holds for L. Indeed, in the first case this follows from (2.9), while in the second we observe that the term $t(L)$ appears on both sides of (2.7), and hence (2.7) becomes an equation for $t(\langle L \rangle)$. $\square$

2.2. *Euclidean Jordan algebras.* The other fundamental concept needed in the definition of the main object of this paper is the notion of a Euclidean Jordan algebra. All results needed from the theory of Euclidean Jordan algebras are contained in the recent monograph by Faraut and Koranyi (1994) and will be used without further reference. The reader is also referred to Massam (1994) and Massam and Neher (1997) for a review of this theory with a special emphasis on the connection to statistics. In this subsection we set up our notation which will be used throughout the paper.

Always, V denotes an Euclidean Jordan algebra with quadratic representation P defined in terms of the algebra product of V by $P(u)v = 2u(uv) - u^2 v$. For $u, v, w \in V$ the linearization of P and the Jordan triple product $\{u\,v\,w\}$ are given by

$$P(u, w)v := \{u\,v\,w\} := 2u(vw) + 2w(uv) - 2(uw)v.$$

We denote by det the determinant function (also called reduced norm) of V. For an endomorphism φ of V, φ^* is the adjoint of φ with respect to the positive definite trace form tr. We denote the symmetric cone of V by $\Omega = \Omega(V)$.

For an idempotent c, that is, an element $c \in V$ satisfying $c^2 = c$, we denote the Peirce spaces of c by $V(c, i) = \{v \in V;\ cv = iv\},\ i \in \{0, \frac{1}{2}, 1\}$. An arbitrary $y \in V$ can then be uniquely written in the form $y = y_1 + y_{12} + y_0$ where $y_i \in V(c, i)$ for $i = 0, 1$ and $y_{12} \in V(c, \frac{1}{2})$. We will refer to this as the Peirce decomposition of y. If y_1 is invertible in $V(c, 1)$, its inverse is denoted y_1^{-1}. Observe that in general $y_1^{-1} \ne (y^{-1})_1$. The symmetric cone and the determinant of the Euclidean Jordan algebra $V(c, 1)$ are denoted Ω_c, respectively, $\det_c$ (we put $\Omega_0 = \{0\}$ and $\det_0 = 1$).

We assume that we are given a complete orthogonal system $\mathscr{E} = (e_1, \dots, e_n)$ of primitive idempotents of V. For any $L \subseteq I = \{1, \dots, n\}$, we put

$$\begin{aligned} e_L &:= \sum_{i \in L} e_i, \qquad V_L := V(e_L, 1),\\ y_L &:= P(e_L)y = V_L\text{-component of } y \in V,\\ \mathscr{E}_L &:= P(e_L)\mathscr{E} = (e_i;\ i \in L), \qquad \Omega_L := P(e_L)\Omega = \Omega(V_L),\\ \det_L &:= \text{determinant of } V_L. \end{aligned}$$

[The equality $P(e_L)\Omega = \Omega(V_L)$ is, e.g., proved in Massam and Neher (1997), Section 3.2.] The reader is reminded of our conventions: x_L^{-1} is the inverse of

x_L in V_L, $e_\varnothing = 0$, $V_\varnothing = \{0\} = \Omega_\varnothing$, $0^{-1} = 0$. For $L, M \subseteq I$ and $y \in V$ we put

$$y_{L\setminus M,\, M\setminus L} := P(e_{L\setminus M}, e_{M\setminus L})y.$$

Note that e_K is orthogonal to e_J if $K \cap J = \varnothing$. Hence, the element $y_{L\setminus M,\, M\setminus L}$ is the V_{12}-component of y with respect to the orthogonal system $(e_{L\setminus M}, e_{M\setminus L}, e - e_{L\cup M})$.

THEOREM 1. *In the setting defined above, let $\mathscr{K} \subseteq \mathscr{D}(I)$ be a unital subring and let $x \in \Omega$. Then the following conditions* (i)–(vii) *are equivalent*:

(i) *For all $L, M \in \mathscr{K}$ and $y \in V$*: $\operatorname{tr}(x_{L\cup M}^{-1}\, y_{L\cup M}) + \operatorname{tr}(x_{L\cap M}^{-1}\, y_{L\cap M}) = \operatorname{tr}(x_L^{-1}\, y_L) + \operatorname{tr}(x_M^{-1}\, y_M)$.

(ii) *For all $L, M \in \mathscr{K}$*: $x_{L\cup M}^{-1} + x_{L\cap M}^{-1} = x_L^{-1} + x_M^{-1}$.

(iii) *For all $L, M \in \mathscr{K}$*: $(x_{L\cup M}^{-1})_{L\setminus M,\, M\setminus L} = 0$.

(iv) *For all $L \in \mathscr{K}$ and $y \in V$*: $\operatorname{tr}(x_L^{-1}\, y_L) = \sum_{K\in\mathscr{J}(\mathscr{K}_L)} \operatorname{tr}(x_K^{-1}\, y_K) - \operatorname{tr}(x_{\langle K\rangle}^{-1}\, y_{\langle K\rangle})$.

(v) *For all $L \in \mathscr{K}$*: $x_L^{-1} = \sum_{K\in\mathscr{J}(\mathscr{K}_L)} (x_K^{-1} - x_{\langle K\rangle}^{-1})$.

(vi) *For every $\varnothing \neq L \in \mathscr{K} \setminus \mathscr{J}(\mathscr{K})$ and every regular decomposition $L = [K] \uplus M$*: $x_L^{-1} = x_K^{-1} - x_{\langle K\rangle}^{-1} + x_M^{-1}$.

(vi′) *For every $\varnothing \neq L \in \mathscr{K} \setminus \mathscr{J}(\mathscr{K})$ and some regular decomposition $L = [K] \uplus M$*: $x_L^{-1} = x_K^{-1} - x_{\langle K\rangle}^{-1} + x_M^{-1}$.

(vii) $x^{-1} = \sum_{K\in\mathscr{J}(\mathscr{K})} (x_K^{-1} - x_{\langle K\rangle}^{-1})$.

In this case, we have for $L, M \in \mathscr{K}$:

(viii) $\det_{L\cup M}(x_{L\cup M}) \det_{L\cap M}(x_{L\cap M}) = \det_L(x_L) \det_M(x_M)$;

(ix) $\det_L(x_L) = \prod_{K\in\mathscr{J}(\mathscr{K}_L)} \det_K(x_K) \det_{\langle K\rangle}(x_{\langle K\rangle}^{-1})$.

The proof of this theorem will be given later in this section. It will require new results in the theory of Euclidean Jordan algebras which will be established in the following subsections and Propositions 1–3.

As we will explain in Section 2.3(d) below, for the case of real symmetric matrices these conditions appear—sometimes explicitly, sometimes less explicitly—in AP (1988, 1993). It should be noted that, like our results, the results in AP (1988) are presented in a coordinate invariant fashion. However, our proof of the equivalence of (i)–(vii) is different from the one given by AP (1988). We have not been able to follow their proof in the setting of Euclidean Jordan algebras.

2.3. *The Andersson–Perlman cone $\Omega(\mathscr{K})$: definition and elementary properties.* The set of elements in $\Omega = \Omega(V)$ satisfying the seven equivalent conditions (i)–(vii) above will be denoted $\Omega(\mathscr{K})$. In the special case where V is the Euclidean Jordan algebra of real symmetric matrices, $\Omega(\mathscr{K})$ coincides with the cone $\mathbf{P}(\mathscr{K})$ studied in AP (1988, 1993); see (d) below. We therefore call it the *AP cone*. We use the notation $\Omega(\mathscr{K})$ since the AP cone is a generalization of the cone Ω of V; see (c) below. For examples of the AP cone, the reader is referred to AP (1993), 2.8 or to the examples below. In the following we collect some of its elementary properties.

(a) It is obvious that $\Omega(\mathscr{K})$ is a cone: $x \in \Omega(\mathscr{K})$, $s \in \mathbf{R}$, $s > 0 \Rightarrow sx \in \Omega(\mathscr{K})$. It is also clear, for example from Theorem 1(ii), that $\Omega(\mathscr{K})$ is closed in Ω. It is, however, in general not a convex cone. For example, let $I = \{1, 2, 3\}$, $L = \{1, 2\}$ and $M = \{1, 3\}$, and consider the ring $\mathscr{K}$ generated by L and M. We have $\mathscr{K} = \{\varnothing, \{1\}, L, M, I\}$, $\{1\} = \langle L\rangle = L \cap M = \langle M\rangle \in \mathscr{J}(\mathscr{K})$, $[L] = \{2\} \in \mathscr{J}(\mathscr{K})$ and $[M] = \{3\} \in \mathscr{J}(\mathscr{K})$. Then

$$x \in \Omega(\mathscr{K}) \Leftrightarrow x^{-1} + x_{L\cap M}^{-1} = x_L^{-1} + x_M^{-1} \Leftrightarrow (x^{-1})_{23} = 0 \Leftrightarrow x_{23} = \{x_{12}\, x_{11}^{-1}\, x_{13}\}.$$

Any "diagonal" element $a = a_1 \oplus a_2 \oplus a_3 \in \Omega_1 \oplus \Omega_2 \oplus \Omega_3$ lies in $\Omega(\mathscr{K})$, but for a nondiagonal $x \in \Omega(\mathscr{K})$, $a + x$ does not in general satisfy the defining condition above.

(b) The conditions on $\Omega(\mathscr{K})$ involve the family of idempotents $(e_L;\ L \in \mathscr{K})$ which in turn depend on the orthogonal system $\mathscr{E}$. However, if $\tilde{\mathscr{E}}$ is another complete orthogonal system of primitive idempotents, there exists an automorphism φ of the Euclidean Jordan algebra V such that $\varphi(\mathscr{E}) = \tilde{\mathscr{E}}$. One then easily sees that the cone $\Omega(\mathscr{K}, \mathscr{E})$ defined with respect to $\mathscr{E}$ and the cone $\Omega(\mathscr{K}, \tilde{\mathscr{E}})$ defined with respect to the orthogonal system $\tilde{\mathscr{E}}$ are mapped onto each other by φ. We have therefore left out $\mathscr{E}$ in the notation $\Omega(\mathscr{K})$.

(c) For any $\mathscr{K}$ we have

$$\Omega_{e_1} \oplus \cdots \oplus \Omega_{e_n} \subseteq \Omega(\mathscr{K}) \subseteq \Omega. \tag{2.10}$$

Indeed, any "diagonal" element $x = x_1 \oplus \cdots \oplus x_n \in \Omega_1 \oplus \cdots \oplus \Omega_n$ has $x_{L\cup M}^{-1} = \sum_{i\in L\cup M} x_i^{-1}$; in particular x^{-1} is diagonal too, and hence x fulfills condition (ii) in Theorem 1. We point out that the lower bound $\Omega_{e_1} \times \cdots \times \Omega_{e_n}$ in (2.10) is attained. Namely, if $\mathscr{K}$ contains disjoint subsets L, M such that $I = L \dot\cup M$, condition (iii) in Theorem 1 yields $(x^{-1})_{LM} = 0$, that is, $x^{-1} = (x^{-1})_L \oplus (x^{-1})_M$, which implies $x = x_L \oplus x_M$. Thus, $\Omega(\mathscr{K}) \subseteq \Omega_L \oplus \Omega_M$. Hence, if $\mathscr{K} = \mathscr{D}(I)$, then, by (2.10), $\Omega(\mathscr{K}, \mathscr{E}) = \Omega_{e_1} \times \cdots \times \Omega_{e_n}$. Also the upper bound in (2.10) is obtained for a suitable $\mathscr{K}$. Namely, observe that conditions (i)–(iii) in Theorem 1 are trivially fulfilled for all pairs (L, M) with $L \subseteq M$ or $M \subseteq L$. Hence, they only need to be checked for all $L, M \in \mathscr{K}$ with $L \not\subseteq M$ and $M \not\subseteq L$. In particular, if the poset $(\mathscr{K}, \subseteq)$ is a chain, that is, $L \subseteq M$ or $M \subseteq L$ for all $L, M \in \mathscr{K}$, we have $\Omega(\mathscr{K}) = \Omega$.

(d) For the Euclidean Jordan algebra V of real symmetric $n \times n$ matrices, the set $\mathbf{P}(\mathscr{K})$ defined in AP (1993) coincides with our $\Omega(\mathscr{K})$ where $\mathscr{E} = (e_1, \dots, e_n)$ is the standard Jordan frame; that is, $e_i = E_{ii}$ is the matrix which has 1 at the position (ii) and 0's elsewhere. To see this, note that by AP (1993), Lemma 2.1, a positive-definite matrix Σ lies in $\mathbf{P}(\mathscr{K})$ if and only if for all $L, M \in \mathscr{K}$ and $x \in \mathbf{R}^n$ we have

$$\begin{aligned} &\operatorname{tr}(\Sigma_{L\cup M}^{-1} x_{L\cup M}\, x_{L\cup M}^t) + \operatorname{tr}(\Sigma_{L\cap M}^{-1}\, x_{L\cap M}\, x_{L\cap M}^t) \\ &\qquad = \operatorname{tr}(\Sigma_L^{-1}\, x_L\, x_L^t) + \operatorname{tr}(\Sigma_M^{-1}\, x_M\, x_M^t). \end{aligned} \tag{2.11}$$

Since $x_L x_L^t = (xx^t)_L$ for any $L \subseteq I = \{1, \dots, n\}$, (2.11) is equivalent to

$$\begin{aligned}\operatorname{tr}\big(\Sigma_{L\cup M}^{-1}(xx^t)_{L\cup M}\big) + \operatorname{tr}\big(\Sigma_{L\cap M}^{-1}(xx^t)_{L\cap M}\big)\\ = \operatorname{tr}\big(\Sigma_L^{-1}(xx^t)_L\big) + \operatorname{tr}\big(\Sigma_M^{-1}(xx^t)_M\big).\end{aligned} \tag{2.12}$$

By standard linear algebra, every real symmetric $n \times n$ matrix is a linear combination of matrices of the form xx^t, $x \in \mathbf{R}^n$ (rank-1 matrices). Since (2.12) is linear in xx^t, a positive-definite matrix Σ lies in $\mathbf{P}(\mathscr{K})$ if and only if for all $L, M \in \mathscr{K}$ and every real symmetric $n \times n$ matrix A,

$$\operatorname{tr}\big(\Sigma_{L\cup M}^{-1} A_{L\cup M}\big) + \operatorname{tr}\big(\Sigma_{L\cap M}^{-1} A_{L\cap M}\big) = \operatorname{tr}\big(\Sigma_L^{-1} A_L\big) + \operatorname{tr}\big(\Sigma_M^{-1} A_M\big),$$

which is exactly our condition in Theorem 1(i).

2.4. *Peirce formulas.* The two main techniques to prove Theorem 1 are Peirce identities and Frobenius coordinates. For convenient reference, we list in this subsection the Peirce formulas which will be needed several times.

Throughout, indices indicate to which Peirce space of an idempotent $c \in V$ the elements belong; see Section 2.2. The following formulas hold:

$$y_1(y_{12}\, y_0) = (y_1 y_{12}) y_0, \tag{2.13}$$

$$\{y_1\ y_{12}\ y_0\} = 4y_1(y_{12}\, y_0), \tag{2.14}$$

$$P(\{y_1\ y_{12}\ y_0\})x_0 = P(y_1)\, P(y_{12})\, P(y_0) x_0, \tag{2.15}$$

$$4P(y_1\, y_{12}) y_0 = P(y_1) P(y_{12}) y_0, \tag{2.16}$$

$$P(y_{12}) y_0 = 2c(y_{12}(y_{12} y_0)), \tag{2.17}$$

$$x_1(y_1\, y_{12}) + y_1(x_1 y_{12}) = (x_1\, y_1) y_{12}, \tag{2.18}$$

$$\begin{aligned}4x_1^{-1}(x_1\, x_{12}) &= 4x_1(x_1^{-1}\, x_{12})\\ &= x_{12} \quad \text{for invertible } x_1 \in V(c, 1),\end{aligned} \tag{2.19}$$

$$\det(x) = \det_c(x_1)\det_{e-c}(x_0 - P(x_{12})x_1^{-1}) \quad \text{for } x \in \Omega. \tag{2.20}$$

Formula (2.13) follows from Faraut and Koranyi (1994), Proposition II.1.1(ii) for $x = y_0$, $y = y_1$, $z = c$, and the Peirce multiplication rules [see Faraut and Koranyi (1994), Proposition IV.1.1]. These are also used to derive (2.14) from (2.13). Formula (2.15) is proven in Massam and Neher (1997), Section 3.9. Specializing $y_0 = e - c$ in (2.14) shows $2y_1 y_{12} = \{y_1\ y_{12}(e - c)\}$ and this implies (2.16) in view of (2.15). Letting $y_1 = c$ in (2.16) then yields (2.17) by using the Peirce multiplication rules. Formulas (2.18)–(2.20) are proved in Massam and Neher (1997), Section 3.3.

2.5. *Frobenius coordinates.* Frobenius coordinates have been used before, for example, in Faraut and Koranyi (1994), Chapter VI.3, and in Massam and Neher (1997), Section 3.3, but not in the generality needed here. Frobenius coordinates rather than the $\mathscr{K}$-parametrization, presented in Section 2.7 below, will be fundamental for Section 3.

Let us first recall that for any $g \in G(\Omega) = \{g \in GL(V);\ g\Omega = \Omega\}$ and $x \in \Omega$, one knows [Faraut and Koranyi (1994), VIII.2.5 and VIII.2.8] that gx is invertible with inverse

$$(gx)^{-1} = g^{*-1}x^{-1}. \tag{2.21}$$

For an idempotent c and $z \in V(c, \frac{1}{2})$ the *Frobenius transformation* on V is defined as $\tau_c(z) = \exp(L(z, c)) \in G(\Omega)$. It is straightforward to check that τ_c: $V(c, \frac{1}{2}) \to G(\Omega)$ is a homomorphism, thus

$$\tau_c(z + z') = \tau_c(z)\tau_c(z'), \qquad \tau_c(-z) = \tau_c(z)^{-1}. \tag{2.22}$$

If $x = x_1 + x_{12} + x_0$ is the Peirce decomposition of $x \in V$ with respect to c then, by Faraut and Koranyi (1994), VI.3.1, and (2.17) for the idempotent $e-c$ instead of c,

$$\begin{aligned} \tau_c(z)x &= x_1 \oplus 2zx_1 + x_{12} \oplus 2(e-c)[z(zx_1) + zx_{12}] + x_0 \\ &= x_1 \oplus 2zx_1 + x_{12} \oplus P(z)x_1 + 2(e-c)(zx_{12}) + x_0. \end{aligned} \tag{2.23}$$

For any $x \in \Omega$ there exists a unique $z \in V(c, \frac{1}{2})$ such that $\tau_c(z)^{-1}x \in V(c, 1) \oplus V(c, 0)$, namely $z = 2x_1^{-1}x_{12}$ and we have

$$x = \tau_c(2x_1^{-1}x_{12})(x_1 \oplus (x_0 - P(x_{12})x_1^{-1})). \tag{2.24}$$

Note that $x_1 \in \Omega_c$ and

$$x_{0\bullet} := x_0 - P(x_{12})x_1^{-1} \in \Omega_{e-c}. \tag{2.25}$$

We call $(2x_1^{-1}x_{12}, x_1, x_{0\bullet}) \in V(c, \frac{1}{2}) \times \Omega_c \times \Omega_{e-c}$ the *Frobenius coordinates of* $x \in \Omega$ *with respect to* c. We also will need the operation of the adjoint map $\tau(z)^*$. By Massam and Neher (1997), (2.6.5), and (2.17) here we have

$$\tau_c(z)^*x = (x_1 + 2c(zx_{12}) + P(z)x_0) \oplus (x_{12} + 2zx_0) \oplus x_0. \tag{2.26}$$

Therefore, (2.21) implies for x as in (2.24):

$$\begin{aligned} x^{-1} &= \tau_c(-2x_1^{-1}x_{12})^*(x_1^{-1} \oplus x_{0\bullet}^{-1}) \\ &= (x_1^{-1} + 4P(x_1^{-1}x_{12})x_{0\bullet}^{-1}) \oplus -4(x_1^{-1}x_{12})x_{0\bullet}^{-1} \oplus x_{0\bullet}^{-1}. \end{aligned} \tag{2.27}$$

Here $x_{0\bullet}^{-1}$ denotes the inverse of $x_{0\bullet}$ in $V(c, 0)$. Taking the trace with an arbitrary $y = y_1 + y_{12} + y_0 \in V$ yields

$$\begin{aligned} \operatorname{tr}(x^{-1}y) = \operatorname{tr}([x_1^{-1} &+ 4P(x_1^{-1}x_{12})x_{0\bullet}^{-1}]y_1) \\ &- 4\operatorname{tr}([(x_1^{-1}x_{12})x_{0\bullet}^{-1}]y_{12}) + \operatorname{tr}(x_{0\bullet}^{-1}y_0), \end{aligned} \tag{2.28}$$

and using the associativity of tr, this can be rewritten in the form

$$\operatorname{tr}(x^{-1}y) = \operatorname{tr}(x_1^{-1}y_1) + \operatorname{tr}(x_{0\bullet}^{-1}[4P(x_1^{-1}x_{12})y_1 - 4(x_1^{-1}x_{12})y_{12} + y_0]). \tag{2.29}$$

Now let $(c_1, \dots, c_n) \subseteq V$ be a complete orthogonal system of arbitrary idempotents. We denote by V_{ij}, $1 \le i, j \le n$, the Peirce spaces of the orthogonal system $(c_1, \dots, c_n)$ [Faraut and Koranyi (1994), IV.2] and define, for $1 \le i < n$, subspaces

$$V^{(i)} := \bigoplus_{k=i+1}^{n} V_{ik} = [V(c_i + \cdots + c_n, 1)](c_i, \tfrac{1}{2}).$$

For $x \in V$ we let $x = \sum_{i \le j} x_{ij}$, $x_{ij} \in V_{ij}$, be the Peirce decomposition of $x \in V$. We abbreviate $\tau_i = \tau_{c_i}$ and $\Omega_i = \Omega_{c_i} = \Omega(V_{ii})$, $1 \le i \le n$. For $x \in \Omega$ there exists a unique $z_1 \in V^{(1)}$ such that

$$\tau_1(z_1)^{-1} x = y_1 \oplus w \in \Omega_1 \oplus \Omega_{c_2 + \cdots + c_n},$$

where $y_1 = x_{11}$, and hence by (2.23),

$$x = \tau_1(z_1)(y_1 \oplus w) = (\tau_1(z_1) y_1) + w.$$

We can repeat this process with the Euclidean Jordan algebra $V(c_1, 0) = V(c_2 + \cdots + c_n, 1)$, the idempotent $c_2 \in V(c_1, 0)$ and $w \in \Omega_{c_2 + \cdots + c_n} \subseteq V(c_1, 0)$. We obtain a unique $z_2 \in V^{(2)}$ such that

$$\tau_2(z_2)^{-1} w = y_2 \oplus v \in \Omega_2 \oplus \Omega_{c_3 + \cdots + c_n}.$$

Again by (2.23), $\tau_2(z_2)$ fixes every element of $V_{11} \oplus V(c_3 + \cdots + c_n)$. Hence,

$$x = \tau_1(z_1)\tau_2(z_2)(y_1 \oplus y_2 \oplus v) = \tau_1(z_1) y_1 + \tau_2(z_2) y_2 + v.$$

Continuing in this manner, we obtain the first part of the following proposition. The second part is an immediate application of (2.21) and (2.22).

Proposition 1. *The map* $F\colon V^{(1)} \times \cdots \times V^{(n-1)} \times \Omega_1 \times \cdots \times \Omega_n \to \Omega$ *given by*

$$\begin{aligned} &F(z_1, \dots, z_{n-1}, y_1, \dots, y_n) \\ &\quad := \tau_1(z_1) \cdots \tau_{n-1}(z_{n-1})(y_1 \oplus \cdots \oplus y_n) \\ &\quad = \tau_1(z_1) y_1 + \tau_2(z_2) y_2 + \cdots + \tau_{n-1}(z_{n-1}) y_{n-1} + y_n \end{aligned} \tag{2.30}$$

is a bijection. For $x = F(z_1, \dots, z_{n-1}, y_1, \dots, y_n) \in \Omega$, *the inverse of* x *is given by*

$$x^{-1} = \tau_1(-z_1)^* \cdots \tau_{n-1}(-z_{n-1})^* (y_1^{-1} \oplus \cdots \oplus y_n^{-1}). \tag{2.31}$$

We will call $F^{-1}(x) = (z_1, \dots, z_{n-1}, y_1, \dots, y_n)$ or sometimes $((z_{jk}), y_1, \dots, y_n)$ for $z_j = \sum_{k>j} z_{jk}$ the *Frobenius coordinates of* $x \in \Omega$ *with respect to the orthogonal system* $(c_1, \dots, c_n)$. We note that the proposition generalizes Faraut and Koranyi (1994), VI.3.5.

2.6. *An example for Frobenius coordinates.* In Proposition 2 we will need the precise formulas for the Frobenius coordinates in the case $n = 3$. Thus, let $x = \sum_{i \le j} x_{ij} \in \Omega$ with Peirce components $x_{ij} \in V_{ij}$. We put

$$y_1 = x_{11}, \qquad z_{12} = 2x_{11}^{-1}x_{12}, \qquad z_{13} = 2x_{11}^{-1}x_{13}. \tag{2.32}$$

Then

$$\begin{aligned} x &= \tau_1(z_{12} + z_{13})(y_1 + x_{22} + x_{23} + x_{33} - P(x_{12} + x_{13})y_1^{-1}) \\ &= \tau_1(z_{12} + z_{13})(y_1 \oplus (x_{22} - P(x_{12})y_1^{-1}) \\ &\quad \oplus (x_{23} - \{x_{12}\, y_1^{-1}\, x_{13}\}) \oplus (x_{33} - P(x_{13})y_1^{-1})). \end{aligned}$$

Now we consider $w = y_2 \oplus w_{23} \oplus w_3 \in V(c_1, 0) = V(c_2 + c_3, 1)$ where

$$\begin{aligned} &y_2 = x_{22} - P(x_{12})x_{11}^{-1}, \qquad w_{23} = x_{23} - \{x_{12}\, x_{11}^{-1}\, x_{13}\}, \\ &w_3 = x_{33} - P(x_{13})x_{11}^{-1}. \end{aligned} \tag{2.33}$$

with

$$z_{23} = 2y_2^{-1}w_{23} = 2(x_{22} - P(x_{12})x_{11}^{-1})^{-1}(x_{23} - \{x_{12}\, x_{11}^{-1}\, x_{13}\}). \tag{2.34}$$

We then obtain $w = \tau_2(z_{23})(y_2 \oplus w_3 - P(w_{23})y_2^{-1})$. Thus, the Frobenius coordinates of x with respect to (c_1, c_2, c_3) are $(z_{12}, z_{13}, z_{23}, y_1, y_2, y_3)$ where z_{jk} and y_1, y_2 are defined above and

$$\begin{aligned} y_3 &= w_3 - P(w_{23})y_2^{-1} \\ &= x_{33} - P(x_{13})x_{11}^{-1} - P(x_{23} - \{x_{12}\, x_{11}^{-1}\, x_{13}\})(x_{22} - P(x_{12})x_{11}^{-1})^{-1}. \end{aligned} \tag{2.35}$$

As an application, let us compute the (23)-component of x^{-1}. We have, by (2.31),

$$x^{-1} = \tau(-z_{12})^* \tau(-z_{13})^* \tau(-z_{23})^* (y_1^{-1} \oplus y_2^{-1} \oplus y_3^{-1})$$

and hence, by (2.26),

$$[x^{-1}]_{23} = [\tau(-z_{23})^*(y_1^{-1} \oplus y_2^{-1} \oplus y_3^{-1})]_{23} = -2z_{23}y_3^{-1}. \tag{2.36}$$

We also note that the $V(c_1+c_2, 1)$-component of x, that is $P(c_1+c_2)x$, can be expressed in terms of the Frobenius coordinates of x with respect to (c_1, c_2, c_3) as

$$P(c_1 + c_2)x = \tau_1(z_{12})(y_1 \oplus y_2). \tag{2.37}$$

Indeed, working in the Euclidean Jordan algebra $V(c_1 + c_2, 1)$, the Frobenius coordinates of

$$P(c_1 + c_2)x = x_{11} \oplus x_{12} \oplus x_{22} \in \Omega(V(c_1 + c_2, 1))$$

with respect to c_1 are

$$\left(2x_{11}^{-1}x_{12}, x_{11}, x_{22} - P(x_{12})x_{11}^{-1}\right) = (z_{12}, y_1, y_2).$$

PROPOSITION 2. *Let $\mathscr{O} = (c_1, c_2, c_3)$ be a complete orthogonal system of idempotents in V. Put $M = \{1,2\}$, $c_M = c_1 + c_2$, $V_M = V(c_M, 1)$ and for $u \in V$ let $u_M = P(c_M)u = u_{11} + u_{12} + u_{22}$ be the V_M-component of u. For $K = \{1,3\}$ define c_K, V_K and u_K analogously.*

(a) *Given arbitrary $v_{13} \in V_{13}$, $w_M \in \Omega_M = \Omega(V_M)$ and $w_3 \in \Omega_3$, there exists a unique $x \in \Omega$ such that x has V_M-component w_M and Frobenius coordinates $z_{13} = v_{13}$, $z_{23} = 0$ and $y_3 = w_3$ with respect to $\mathscr{O}$, namely $x = \tau_1(v_{13})(w_M \oplus w_3)$.*

(b) *Let $x \in \Omega$ with Frobenius coordinates $(z_{12}, z_{13}, z_{23}, y_1, y_2, y_3)$ and Peirce components x_{ij} with respect to $\mathscr{O}$. Denote by x_{11}^{-1} the inverse of $x_{11} \in V_{11}$ calculated in V_{11} and, similarly, by x_M^{-1} and x_K^{-1} the inverses of $x_M \in V(c_M, 1)$, respectively, $x_K \in V(c_K, 1)$. Then the following statements* (i)–(vi) *are equivalent*:

(i) $\operatorname{tr}(x^{-1}u) + \operatorname{tr}(x_{11}^{-1}u_{11}) = \operatorname{tr}(x_M^{-1}u_M) + \operatorname{tr}(x_K^{-1}u_K)$ *for all* $u \in V$;
(ii) $x^{-1} + x_{11}^{-1} = x_M^{-1} + x_K^{-1}$;
(iii) $(x^{-1})_{23} = 0$;
(iv) $z_{23} = 0$;
(v) $x_{23} = \{x_{12}\, x_{11}^{-1}\, x_{13}\}(= 0$ *if* $c_1 = 0)$;
(vi) $x = \tau_1(z_{13})(x_M \oplus y_3)$.

In this case,

(vii) $y_3 = x_{33} - P(x_{13})x_{11}^{-1}$;
(viii) $x = \tau_1(z_{12})(x_K \oplus y_2)$;
(ix) $\det(x)\det_{11}(x_{11}) = \det_M(x_M)\det_K(x_K)$

where $\det_{11}$, $\det_M$ *and* $\det_K$ *are the determinant functions of* V_{11}, V_M *and* V_K.

PROOF. (a) We first prove for arbitrary $v_{ij} \in V_{ij}$,

$$\{v_{12}\, v_{11}\, v_{13}\} = 4v_{12}(v_{13}\, v_{11}) = 4v_{13}(v_{12}\, v_{11}). \tag{2.38}$$

By the Peirce multiplication rules, we have $V_{12}V_{13} \subseteq V_{23}$ and $V_{11}V_{23} = 0$. Hence

$$\begin{aligned}\{v_{12}\, v_{11}\, v_{13}\} &= 2[v_{12}(v_{13}\, v_{11}) + v_{13}(v_{12}v_{11}) - v_{11}(v_{12}\, v_{13})]\\ &= 2[v_{12}(v_{13}\, v_{11}) + v_{13}(v_{12}v_{11})].\end{aligned}$$

We apply (2.18) for the idempotent c_1+c_2 and $x_1 = v_{11}$, $y_1 = v_{12}$, $y_{12} = v_{13}$, and obtain $v_{11}(v_{12}v_{13}) + v_{12}(v_{11}v_{13}) = (v_{11}\, v_{12})v_{13}$. Since $v_{11}(v_{12}v_{13}) = 0$ this shows $v_{12}(v_{11}v_{13}) = (v_{11}\, v_{12})v_{13}$, and then (2.38) follows from the formula above.

We next prove the existence part of (a). Given arbitrary $v_{13} \in V_{13}$, $w_M = w_1 \oplus w_{12} \oplus w_2 \in \Omega_M = \Omega(V_M)$ and $w_3 \in \Omega_3$, we define $x = \tau_1(v_{13})(w_M \oplus w_3)$. Then $x \in \Omega$ since $w_M \oplus w_3 \in \Omega_M \oplus \Omega_3 \subseteq \Omega$ and $\tau_1(v_{13}) \in G(\Omega)$ leaves the cone Ω invariant. By the Peirce multiplication rules, $(e-c_1)v_{23} = c_2v_{23} + c_3v_{23} = v_{23}$ for arbitrary $v_{23} \in V_{23}$. Hence, using (2.23),

$$\begin{aligned}x &= \tau_1(v_{13})(w_1 \oplus w_{12} \oplus (w_2 \oplus w_3))\\ &= w_1 \oplus w_{12} \oplus 2w_1v_{13} \oplus w_2 \oplus 2v_{13}\, w_{12} \oplus (P(v_{13})w_1 + w_3).\end{aligned}$$

In particular, $x_M = w$ and $x_{13} = 2w_1 v_{13}$. If (z_{ij}, y_k) are the Frobenius coordinates of x we obtain, by (2.19),

$$z_{13} = 2w_1^{-1} x_{13} = 4w_1^{-1}(w_1\, v_{13}) = v_{13}.$$

Using (2.38), this implies

$$\{x_{12}\, x_{11}^{-1}\, x_{13}\} = 4w_{12}(w_1^{-1} x_{13}) = 8w_{12}[w_1^{-1}(w_1 v_{13})] = 2w_{12}\, v_{13} = x_{23}.$$

Thus $z_{23} = 0$ by (2.34) and then, by (2.35),

$$y_3 = x_{33} - P(x_{13})x_{11}^{-1} = w_3 + P(v_{13})w_1 - P(2w_1 v_{13})w_1^{-1}.$$

Since $P(c_3)|V_{33} = Id$ and $P(w_1)w_1^{-1} = w_1$ it follows, using (2.14) and (2.15),

$$P(2w_1 v_{13})w_1^{-1} = P(\{w_1\, v_{13}\, c_3\})w_1^{-1} = P(c_3)P(v_{13})P(w_1)w_1^{-1} = P(v_{13})w_1,$$

which proves $y_3 = w_3$.

It remains to show the uniqueness part of (a). Since the Frobenius coordinate $z_{23} = 0$ we have, using (2.22), (2.37) and $\tau_1(z_{12})w_3 = w_3$,

$$\begin{aligned} x &= \tau_1(z_{12} + z_{13})(y_1 \oplus y_2 \oplus y_3) = \tau_1(v_{13})\tau_1(z_{12})(y_1 \oplus y_2 \oplus w_3) \\ &= \tau_1(v_{13})(x_M \oplus w_3). \end{aligned}$$

(b)(iv) $\Leftrightarrow$ (v) follows from (2.34) and (2.19), and (iv) $\Leftrightarrow$ (vi) follows from (a). The equivalence of (i)–(vi) will therefore hold if we can show (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (ii).

(i) $\Leftrightarrow$ (ii): since $\operatorname{tr}(V_{ij}) = 0$ for $i \neq j$, the Peirce decomposition is an orthogonal decomposition of V with respect to the positive definite form given by $\operatorname{tr}(uv)$. Thus, in (i) we may replace u_{11}, u_M and u_K by u, and (i) $\Leftrightarrow$ (ii) follows from the positive-definiteness of tr.

(ii) $\Rightarrow$ (iii) is immediate since $x_{11}^{-1} \in V_{11}$, $x_M^{-1} \in V_{11} \oplus V_{12} \oplus V_{22}$ and $x_K^{-1} \in V_{11} \oplus V_{13} \oplus V_{33}$.

(iii) $\Rightarrow$ (iv): by (2.34) we have $0 = z_{23} y_3^{-1}$ whence $z_{23} = 0$ by (2.19).

(iv) $\Rightarrow$ (ii): we know $x = \tau_1(z_{12} + z_{13})(y_1 \oplus y_2 \oplus y_3)$, and hence, by (2.27) and the Peirce multiplication rules,

$$\begin{aligned} x^{-1} &= y_1^{-1} + P(z_{12} + z_{13})(y_2^{-1} + y_3^{-1}) \oplus -2(z_{12} + z_{13})(y_2^{-1} + y_3^{-1}) \oplus (y_2^{-1} + y_3^{-1}) \\ &= (x_{11}^{-1} + P(z_{12})y_2^{-1} + P(z_{13})y_3^{-1}) \oplus -2z_{12} y_2^{-1} \oplus -2z_{13} y_3^{-1} \oplus y_2^{-1} \oplus y_3^{-1}. \end{aligned}$$

By (2.37), $x_M = \tau_1(z_{12})(y_1 \oplus y_2)$, which implies

$$x_M^{-1} = (x_{11}^{-1} + P(z_{12})y_2^{-1}) \oplus -2z_{12} y_2^{-1} \oplus y_2^{-1}.$$

Since $z_{23} = 0$, (2.35) implies (vii):

$$y_3 = x_{33} - P(x_{13})x_{11}^{-1}.$$

Using (2.24), we now find $x_K = x_{11} \oplus x_{13} \oplus x_{33} = \tau_1(z_{13})(y_1 \oplus y_3)$. By symmetry we therefore have

$$x_K^{-1} = (x_{11}^{-1} + P(z_{13})y_3^{-1}) \oplus -2z_{13}y_3^{-1} \oplus y_3^{-1}.$$

A comparison of the formulas for x^{-1}, x_M^{-1} and x_K^{-1} then shows (ii).

We have now shown the equivalence of (i)–(vi). Since (i)–(v) are symmetric in 2 and 3, so must be (vi), and hence we have (viii). Finally, (ix) is a consequence of

$$\det(\tau_c(z)y) = \det(y),$$

which, by Massam and Neher (1996), 3.1, holds for any $y \in V$ and $z \in V(c, \frac{1}{2})$, and the formulas above:

$$\det(x) = \det_{11}(x_{11}) \det_{22}(y_2) \det_{33}(y_3),$$
$$\det(x_M) = \det_{11}(x_{11}) \det_{22}(y_2), \qquad \det(x_K) = \det_{11}(x_{11}) \det_{33}(y_3). \qquad \square$$

PROOF OF THEOREM 1 (First part). For further development, it will be useful to prove that conditions (i)–(vi) in Theorem 1 are equivalent and that they imply the conditions (vii)–(ix). After the preparation in the previous subsections this can now be done as follows.

We have (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Rightarrow$ (viii) by Proposition 2 applied to $x_{L\cup M}$ and (i) $\Leftrightarrow$ (iv) by Lemma 1 for $R = (\mathbb{R}, +)$ and $t(L) = \operatorname{tr}(x_L^{-1} y_L)$. In (iv) we may replace y_L, y_K and $y_{\langle K\rangle}$ by y and then the nondegeneracy of tr shows (iv) $\Rightarrow$ (v). The reverse implication is obvious. The same argument also shows (i) $\Leftrightarrow$ (vi) $\Leftrightarrow$ (vi′). Since $I \in \mathscr{K}$, we have (v) $\Rightarrow$ (vii). Finally, (viii) $\Leftrightarrow$ (ix) by Lemma 1 for the Abelian group $R = (\mathbb{R} \setminus \{0\}, \cdot)$, now written multiplicatively, and $t(L) = \det_L(x_L)$.

Thus, all that remains to be shown for the proof of Theorem 1 is that condition (vii) implies (v). This will be done in the second part of the proof, using the following result in which $\Omega(\mathscr{K})$ is defined by the first six (equivalent) conditions of Theorem 1. $\square$

PROPOSITION 3. *As in Sections* 2.1 *and* 2.2, *let* $I = [K] \uplus M$ *be a regular decomposition with* $K \in \mathscr{J}(\mathscr{K})$ *and* $M \in \mathscr{K}$. *Define* $c_1 = e_{\langle K\rangle}$, $c_2 = e_{M\setminus K}$, $c_3 = e_{[K]}$, *and denote by* V_{ij} *the Peirce spaces with respect to the orthogonal system* $\mathscr{O} = (c_1, c_2, c_3)$. *Also, we let* $\Omega(\mathscr{K}_M) \subseteq V_M$ *be the AP cone constructed with respect to* $\mathscr{K}_M$ *and the orthogonal system* $\mathscr{E}_M$.

(a) *Let* $v_{13} \in V_{13}$, $w_3 \in \Omega(V_{33}) = \Omega_{c_3}$ *and* $w_M \in \Omega(\mathscr{K}_M)$. *Then there exists a unique* $x \in \Omega(\mathscr{K})$ *with* $x_M = w_M$ *and Frobenius coordinates* $(z_{12}, z_{13} = v_{13},\ z_{23} = 0,\ y_1, y_2, y_3 = w_3)$ *with respect to* $\mathscr{O}$, *namely* $x = \tau_{c_1}(v_{13})(w_M \oplus w_3)$. *The coordinates* z_{12}, y_1, y_2 *are uniquely determined by* $x_M = w_M$ *and we have*

$$z_{13} = 2x_{\langle K\rangle}^{-1}\, x_{\langle K\rangle,[K]}, \qquad w_3 = x_{[K]} - P(x_{[K\rangle})x_{\langle K\rangle}^{-1}. \tag{2.39}$$

(b) *An* $x \in \Omega$ *lies in* $\Omega(\mathscr{K})$ *if and only if* $x_M \in \Omega(\mathscr{K}_M)$ *and* $x^{-1} = (x_K^{-1} - x_{\langle K\rangle}^{-1}) + x_M^{-1}$.

PROOF. (a) We observe that $e_K = c_1 + c_3$ and $e_M = c_1 + c_2$. By Proposition 2, $x = \tau_{c_1}(v_{13})(w_M \oplus w_3)$ is the unique element of Ω satisfying $x_M = w_M$ and the requirements on the Frobenius coordinates. Since $x_{\langle K\rangle} = x_{11}$, $x_{\langle K\rangle,[K]} = x_{13}$ and $x_{[K]} - P(x_{[K\rangle})x_{\langle K\rangle}^{-1} = x_{33} - P(x_{13})x_{11}^{-1}$, (2.39) holds in view of (2.32) and Proposition 2(vii). Thus, it remains to show that the extra condition $w_M \in \Omega(\mathscr{K}_M)$ implies $x \in \Omega(\mathscr{K})$. We will verify condition (v) of Theorem 1.

So let $L \in \mathscr{K}$. If $L \subseteq M$, then (v) holds because $x_L = (x_M)_L = (w_M)_L$ and $w_M \in \Omega(\mathscr{K}_M)$. Thus, we can assume $L \not\subseteq M$. We claim that then $K \subseteq L$. Indeed, otherwise $L \cap K \neq K$ and then $L \cap K \subseteq \langle K\rangle \subseteq M$ which gives the contradiction $L = (L \cap M) \cup (L \cap K) \subseteq M$. We now know $L = \langle K\rangle \uplus (L \setminus K) \uplus [K]$ with $L \setminus K \subseteq M \setminus K$. Define

$$c_{2'} = e_{L\setminus K}, \qquad c_{2''} = e_{M\setminus L}.$$

Then $c_2 = c_{2'} + c_{2''}$ and $\mathscr{O}' = (c_1, c_{2'}, c_{2''}, c_3)$ is a complete orthogonal system of V whose Peirce spaces we will denote by U_{ij}, $i, j \in \{1, 2', 2'', 3\}$. Observe that $V_{ij} = U_{ij}$ for $i, j \in \{1, 3\}$ while

$$V_{12} = U_{12'} \oplus U_{12''}, \qquad V_{23} = U_{2'3} \oplus U_{2''3}, \qquad V_{22} = U_{2'2'} \oplus U_{2'2''} \oplus U_{2''2''}.$$

Let $x = \sum x_{ij}$ be the Peirce decomposition of x with respect to $\mathscr{O}$ and let $x = \sum u_{ij}$ be the Peirce decomposition of x with respect to $\mathscr{O}'$; thus $x_{ij} \in V_{ij}$ and $u_{ij} \in U_{ij}$. By the above $x_{ij} = u_{ij}$ for $i, j \in \{1, 3\}$ and $x_{12} = u_{12'} + u_{12''}$, $x_{23} = u_{2'3} + u_{2''3}$. Hence, by Proposition 2(v),

$$x_{23} = \{x_{12}\, x_{11}^{-1}\, x_{13}\} = \{u_{12'}\, u_{11}^{-1}\, u_{13}\} + \{u_{12''}\, u_{11}^{-1}\, u_{13}\}.$$

A comparison of Peirce components then shows $u_{2'3} = \{u_{12'}\, u_{11}^{-1}\, u_{13}\}$. We can therefore apply Proposition 2(b) to $x_L = u_{11} + u_{12'} + u_{13} + u_{2'2'} + u_{2'3} + u_{33} \in \Omega_L \subseteq V_L$ and the complete orthogonal system $(c_1, c_{2'}, c_3)$ of V_L. We obtain

$$x_L^{-1} + x_{\langle K\rangle}^{-1} = x_{L\cap M}^{-1} + x_K^{-1}. \tag{2.40}$$

Since $L \cap M \in \mathscr{K}_M$ and $x_N = (w_M)_N$ for all $N \subseteq M$, we know from Theorem 1(v) for w_M that

$$x_{L\cap M}^{-1} = \sum_{N \in \mathscr{J}(\mathscr{K}_{L\cap M})} (x_N^{-1} - x_{\langle N\rangle}^{-1}). \tag{2.41}$$

But $\mathscr{J}(\mathscr{K}_L) = \{K\} \uplus \mathscr{J}(\mathscr{K}_{L\cap M})$ by (2.3) so that (2.40) and (2.41) imply condition (v) in Theorem 1.

(b) If $x \in \Omega(\mathscr{K})$ then $x_M \in \Omega(\mathscr{K}_M)$, since for any $N \subseteq M$ we have $(x_M)_N = x_N$. Moreover, since $M \cap K = \langle K\rangle$, we also know $x^{-1} = (x_K^{-1} - x_{\langle K\rangle}^{-1}) + x_M^{-1}$ by Theorem 1(ii). Conversely, if these two conditions are fulfilled then, by Proposition 2(b), x is of the form $x = \tau_1(z_{13})(x_M \oplus y_3)$ where (z_{ij}, y_k) are the Frobenius coordinates of x and $z_{23} = 0$. However, then $x \in \Omega(\mathscr{K})$ by (a). □

CONCLUSION OF THE PROOF OF THEOREM 1. From the first part of the proof, it remains to show that if $x \in \Omega$ satisfies condition (vii),

$$x^{-1} = \sum_{K \in \mathscr{J}(\mathscr{K})} \left(x_K^{-1} - x_{\langle K \rangle}^{-1} \right),$$

then $x \in \Omega(\mathscr{K})$. We use induction on $q = |\mathscr{J}(\mathscr{K})|$. If $q = 1$ then $\mathscr{K} = \{\varnothing, I\}$ by (2.3), and hence $\Omega(\mathscr{K}) = \Omega$ by Section 2.3(c). So we can assume $q \geq 2$. Let $I = [K] \uplus M$ be a regular deomposition. We have $|\mathscr{J}(\mathscr{K}_M)| = q - 1$. In particular, we are in the situation of Proposition 3. Let $\mathscr{C} = (c_1, c_2, c_3)$ be as in Proposition 3, and denote by V_{ij} and v_{ij} the Peirce spaces, respectively, the Peirce components with respect to $\mathscr{C}$. By assumption,

$$x^{-1} = x_K^{-1} - x_{\langle K \rangle}^{-1} + \sum_{K' \in \mathscr{J}(\mathscr{K}_M)} \left(x_{K'}^{-1} - x_{\langle K' \rangle}^{-1} \right). \tag{2.42}$$

As $-x_{\langle K \rangle}^{-1} + \sum_{K' \in \mathscr{J}(\mathscr{K}_M)} (x_{K'}^{-1} - x_{\langle K' \rangle}^{-1}) \in V_M = V_{11} \oplus V_{12} \oplus V_{22}$ and $x_K^{-1} \in V_K = V_{11} \oplus V_{13} \oplus V_{33}$, we have $(x^{-1})_{23} = 0$. Hence, by Proposition 2, $x^{-1} = x_K^{-1} - x_{\langle K \rangle}^{-1} + x_M^{-1}$ and by comparison with (2.42),

$$x_M^{-1} = \sum_{K' \in \mathscr{J}(\mathscr{K}_M)} \left(x_{K'}^{-1} - x_{\langle K' \rangle}^{-1} \right).$$

However, then $x_M \in \Omega(\mathscr{K}_M)$ by induction, and hence $x \in \Omega(\mathscr{K})$ by Proposition 2(b). $\square$

2.7. *$\mathscr{K}$-parametrization.* For the case where V is the Euclidean Jordan algebra of real symmetric matrices, AP (1993), Theorem 2.2, proved a coordinatization theorem for $\Omega(\mathscr{K})$. In this section we will generalize their result to the setting of arbitray Euclidean Jordan algebras. Our proof uses a different method. Transformation groups of $\Omega(\mathscr{K})$, the method used by AP, will be studied in Neher (1997). To state our result, the following notation will be needed.

As in Section 2.2, we assume that we have a unital subring $\mathscr{K} \subseteq \mathscr{D}(\{1, \ldots, n\})$ and a complete orthogonal system $\mathscr{E} = (e_1, \ldots, e_n) \subseteq V$ of primitive idempotents. We put

$$V_{[K\rangle} := V\left(e_{[K]}, \tfrac{1}{2}\right) \cap V\left(e_{\langle K \rangle}, \tfrac{1}{2}\right), \qquad y_{[K\rangle} := y_{\langle K \rangle, [K]} = P\left(e_{\langle K \rangle}, e_{[K]}\right) y \in V_{[K\rangle},$$

so that $y_K = y_{[K]} + y_{[K\rangle} + y_{\langle K \rangle}$ for any $y \in V$. For $x \in \Omega$ we have $x_K \in \Omega_K$ and, using a notation from AP (1993) [see also (2.25)],

$$x_{[K]\bullet} := x_{[K]} - P(x_{[K\rangle}) x_{\langle K \rangle}^{-1} \in \Omega_{[K]}.$$

We also fix a never-decreasing listing of $\mathscr{J}(\mathscr{K}) = (K_1, \ldots, K_q)$ (cf. Section 2.1) and define for $1 \leq i \leq q$,

$$e_{[i]} = e_{[K_i]} = \sum_{j \in [K_i]} e_j, \qquad \Omega_{[i]} = \Omega_{[K_i]},$$

$$v_{[i]} = P(e_{[i]}) v = V_{[K_i]}\text{-component of } v \in V,$$

$$e_{\langle i\rangle} = e_{\langle K_i\rangle} = \sum_{j\in\langle K_i\rangle} e_j,$$

$$\tau_{\langle i\rangle} = \text{Frobenius transformation of the idempotent } e_{\langle i\rangle},$$

$$v_{[i\rangle} = v_{[K_i\rangle} = P(e_{\langle i\rangle}, e_{[i]})v = V_{[K_i\rangle}\text{-component of } v \in V.$$

THEOREM 2. *The map*

$$\Omega(\mathscr{K}) \to \prod_{K\in\mathscr{J}(\mathscr{K})} (V_{[K\rangle} \times \Omega_{[K]}): x \mapsto \prod_{K\in\mathscr{J}(\mathscr{K})} \big(x^{-1}_{\langle K\rangle} x_{[K\rangle}, x_{[K]\bullet}\big)$$

is a bijection. Its inverse is given by

$$F_{\mathscr{K}}: \prod_{j=1}^{q} (v_{[j\rangle}, w_{[j]}) \mapsto \tau_{\langle q\rangle}(v_{[q\rangle}) \cdots \tau_{\langle 2\rangle}(v_{[2\rangle})(w_{[1]} \oplus \cdots \oplus w_{[q]}).$$

Following AP (1993), *we call* $F^{-1}_{\mathscr{K}}(x) = \prod_{K\in\mathscr{J}(\mathscr{K})} (x^{-1}_{\langle K\rangle} x_{[K\rangle}, x_{[K]\bullet})$ *the* $\mathscr{K}$-*parameters of* $x \in \Omega(\mathscr{K})$.

PROOF. We define

$$M_i = K_1 \cup \cdots \cup K_i \in \mathscr{K},$$

$$\Omega_i = \Omega_{M_i},\ \ \mathscr{K}_i = \mathscr{K}_{M_i}$$

and

$$\mathscr{E}_i = \mathscr{E}_{M_i} = (e_j; j \in M_i).$$

By (2.4), $M_i = M_{i-1} \uplus [K_i]$. Hence we can apply Proposition 3 to $I = M_q$, $M_{q-1}, \dots, M_1 = K_1$; for any $x \in \Omega(\mathscr{K})$ there exist unique

$$v_{[q\rangle} \in V_{[K_q\rangle}, \qquad x_{q-1} = x_{M_{q-1}} \in \Omega_{q-1}(\mathscr{K}_{q-1})$$

and

$$w_{[q]} \in \Omega_{[q]}$$

such that $x = \tau_{\langle q\rangle}(v_{[q\rangle})(x_{q-1} \oplus w_{[q]})$. Then x_{q-1} can be uniquely written in the form

$$x_{q-1} = \tau_{\langle q-1\rangle}(v_{[q-1\rangle})(x_{q-2} \oplus w_{[q-1]})$$

with $x_{q-2} = x_{M_{q-2}} \in \Omega_{q-2}(\mathscr{K}_{q-2})$ and $w_{[q-1]} \in \Omega_{[q-1]}$. Since $V_{[q]} \subseteq V(e_{\langle q-1\rangle}, 0)$ it follows from (2.23) that $\tau_{\langle q-1\rangle}(v_{[q-1\rangle})$ leaves every element of $V_{[q]}$ invariant. Hence

$$x = \tau_{\langle q\rangle}(v_{[q\rangle})\tau_{\langle q-1\rangle}(v_{[q-1\rangle})\big(x_{q-2} \oplus w_{[q-1]} \oplus w_{[q]}\big).$$

Continuing in this manner, we see that $F_{\mathscr{K}}$ is a bijection. (In the last step, observe that $K_1 = [K_1]$, $\langle K_1\rangle = \varnothing$, $\mathscr{K}_{K_1} = \{\varnothing, K_1\}$ and hence $\Omega_1(\mathscr{K}_1) = \Omega_{[1]}$ by Section 2.3(c). That $F^{-1}_{\mathscr{K}}$ is given as stated in the theorem follows by a repeated application of (2.39). $\square$

3. The estimate of the covariance matrix. In this section, we define lattice conditional independence models $N(\mathscr{K})$ (abbreviated LCI models) with covariance matrices in the symmetric cone of a Euclidean Jordan algebra V. In fact, the cone of covariance matrices for such models, as we will see, is $\Omega(\mathscr{K})$. We then give a closed form formula of the maximum likelihood estimate $\hat{\Sigma}$ of the covariance matrix Σ in $\Omega(\mathscr{K})$ for such models.

3.1. *Models with covariance matrix in* $\Omega(\mathscr{K})$. AP (1993) defined LCI models, which are $N(0,\Sigma)$ models with conditional independences defined by a lattice $\mathscr{K}$. They proved that a real-valued random variable $x \sim N(0,\Sigma)$ satisfies these conditional independences if and only if the covariance matrix belongs to the subcone $\mathscr{P}(\mathscr{K})$ of the cone $H_I^+(\mathbb{R})$ of $I \times I$ positive definite real symmetric matrices. Since $H_I^+(\mathbb{R})$ is the symmetric cone of the particular Euclidean Jordan algebra of real symmetric matrices, it is natural to extend the LCI model with covariance in $H_I^+(\mathbb{R})$ to an LCI model with covariance in Ω as defined in Section 2. In Section 2, we also saw that $\Omega(\mathscr{K})$ is the natural generalization of $\mathscr{P}(\mathscr{K})$ when working with Euclidean Jordan algebras. We will therefore define below a new class of LCI models that are normal $N(0,\Sigma)$ models with Σ parametrized by $\Omega(\mathscr{K})$. We will assume V is a simple Euclidean Jordan algebra. If V is not simple, it decomposes as a direct product of simple algebras, and the cone $\Omega(\mathscr{K})$ decomposes correspondingly. It is easily seen that a model for a general V with a symmetric representation decomposes according to the same product. Let r be the rank of V. From now on, we assume $I = \{1,\dots,r\}$ and as usual $\mathscr{K}$ denotes a lattice on I.

To define an LCI model with covariance in Ω, we first need the definition of a normal model with covariance in Ω. This definition was given in Jensen (1988) and we refer the reader to this paper for details, but we recall the definition here for convenience. Let F denote a Euclidean space and $L_s(F)$ be the space of symmetric endomorphisms of F. Following Faraut and Koranyi (1994), we say that ϕ is a *symmetric representation* of the Euclidean Jordan algebra V if ϕ is a linear map from V to $L_s(F)$ such that $\phi(xy) = \frac{1}{2}[\phi(x)\phi(y) + \phi(y)\phi(x)]$ and $\phi(e)$ is the identity of $L_s(F)$.

DEFINITION 1. Let ϕ be a symmetric representation of V on $F = \mathbb{R}^n$ for an appropriate n. Then a random variable z in F is said to have the central Gaussian distribution with covariance $\phi(x)$ for some $x \in \Omega$ if its density with respect to the Lebesgue measure is

$$(2\pi)^{-n/2}(\det\phi(x))^{-1/2}\exp\left(-\tfrac{1}{2}\langle\phi(x)^{-1}z, z\rangle\right) \tag{3.1}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in F and where x belongs to the set $\{x \in \Omega \mid \phi(x) \text{ is positive definite}\}$. We write $z \sim N(0,\phi(x))$.

There are only four types of simple Euclidean Jordan algebras that admit a symmetric representation, namely, the space $H_r(\mathbb{D})$ of symmetric matrices over $\mathbb{D} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$, the quaternions and the Lorentz algebra defined in terms of an inner product $B(\cdot,\cdot)$ on a finite-dimensional real vector space W by $V = \mathbb{R} \times W$ and algebra product

$$(\lambda_1, w_1) \circ (\lambda_2, w_2) = (\lambda_1\lambda_2 + B(w_1, w_2), \lambda_1 w_2 + \lambda_2 w_1).$$

Since the covariance matrix of a normal variable with values in F can be viewed as a symmetric endomorphism of F, the covariance matrix of a normal random variable has to be of the form $\phi(x)$ where x belongs to one of the four simple Euclidean Jordan algebras given above. The standard representations for $H_r(\mathbb{D})$ are the following:

1. For $V = H_r(\mathbb{R})$, $F = \mathbb{R}^r$, ϕ is the identity.
2. For $V = H_r(\mathbb{C})$, $F = \mathbb{R}^{2r}$, if $x = (a_{st} + ib_{st})_{s,\,t=1,\dots,r} \in H_r(\mathbb{C})$, $\phi(x)$ is the $2r \times 2r$ matrix with (s,t) block
$$\begin{pmatrix} a_{st} & b_{st} \\ -b_{st} & a_{st} \end{pmatrix}.$$
3. For $V = H_r(\mathbb{H})$, $F = \mathbb{R}^{4r}$, if $x = (a_{st} + ib_{st} + jc_{st} + kd_{st})_{s,\,t=1,\dots,r} \in H_r(\mathbb{H})$, $\phi(x)$ is the $4r \times 4r$ matrix with (s,t) block
$$\begin{pmatrix} a_{st} & b_{st} & c_{st} & d_{st} \\ -b_{st} & a_{st} & d_{st} & -c_{st} \\ -c_{st} & -d_{st} & a_{st} & b_{st} \\ -d_{st} & c_{st} & -b_{st} & a_{st} \end{pmatrix}. \tag{3.2}$$
4. For $V = \mathbb{R} \times W$, the representations are much more intricate [see Jensen (1988), Theorem 6]. We do not give them here since, as we will see below, this case will not be of interest to us.

Let us now define an LCI model with covariance $\Sigma = \phi(x)$ where $x \in \Omega$. We fix a complete orthogonal system $\mathscr{E} = (e_1, \dots, e_r)$ of primitive idempotents of V. The cone $\Omega(\mathscr{K})$ has been defined in Section 2.3.

DEFINITION 2. Consider the set of $N(0,\Sigma)$ distributions with $\Sigma = \phi(x)$ for some x in $\Omega(\mathscr{K})$. We define the model $N(\mathscr{K})$ as the set

$$N(\mathscr{K}) = \{N(0, \phi(x)) \mid x \in \Omega(\mathscr{K})\}. \tag{3.3}$$

At this point, we need to make two important remarks. First, it is clear that the model is indeed a generalization of the LCI model defined in AP (1993). Indeed from Section 2.3(d), we know that for $V = H_r(\mathbb{R})$, $\Omega(\mathscr{K}) = \mathscr{P}(\mathscr{K})$ if $(e_1, \dots, e_r)$ is the system of idempotents that are diagonal and therefore $N(\mathscr{K})$ is as defined by AP (1993). In the case where $z \sim N(0,\Sigma)$, $\Sigma = \phi(x)$ with $x \in \Omega(\mathscr{K})$ and $V = H_r(\mathbb{C})$, $H_r(\mathbb{H})$ or the Lorentz algebra, we only need to show that for all $L, M \in \mathscr{K}$, z_L and z_M are conditionally independent given $z_{L\cap M}$. This is also immediate if we take a representation ϕ which maps diagonal idempotents to diagonal idempotents such as the representations given above, for $V = H_r(\mathbb{D})$. Indeed, by Theorem 1(ii), $x \in \Omega(\mathscr{K})$ if and only if $x_{L\cap M}^{-1} + x_{L\cup M}^{-1} = x_L^{-1} + x_M^{-1}$, which is equivalent then, to $\Sigma_{L\cap M}^{-1} + \Sigma_{L\cup M}^{-1} = \Sigma_L^{-1} + \Sigma_M^{-1}$. It follows immediately that $z_L \perp\!\!\!\perp z_M | z_{L\cap M}$ since we are then brought back to the real case.

Our second remark concerns the model $N(\mathscr{K})$ with $V = \mathbb{R} \times W$, the Lorentz algebra. The rank of $V = \mathbb{R} \times W$, the Lorentz algebra, is $r = 2$, and the only possible lattices in $I = \{1, 2\}$ are $\mathscr{K}_0 = \{\varnothing, \{1, 2\}\}$, $\mathscr{K}_1 = \{\varnothing, \{1\}, \{1, 2\}\}$ and $\mathscr{K}_2 = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$. Since $\mathscr{K}_0$ and $\mathscr{K}_1$ are chains, conditions in Theorem 1(ii) do not give any independence for $N(\mathscr{K})$. In $\mathscr{K}_2$, the only interesting relation is $x^{-1}_{\{1,2\}} = x^{-1}_{\{1\}} + x^{-1}_{\{2\}}$ since $\{1\}$ and $\{2\}$ are disjoint. This means of course that $z_{\{1\}}$ and $z_{\{2\}}$ are independent. We see that the $N(\mathscr{K})$ models for $r = 2$ present no new statistical interest. The only models that we are going to consider from now on are those $N(\mathscr{K})$ models with $V = H_r(\mathbb{D})$, $r \geq 3$, with representation in $F = \mathbb{R}^{rd}$ as given above, and where $d = 1, 2, 4$ is the Peirce constant of $H_r(\mathbb{D})$, $\mathbb{D} = \mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$, respectively.

3.2. *The estimate of the concentration matrix.* In this section, we give our main statistical result: the maximum likelihood estimate of the covariance matrix in an $N(\mathscr{K})$ model. Let $(\tilde{z}_1, \dots, \tilde{z}_n)$ be a sample from a normal $N(0, \Sigma)$ distribution on F, with $\Sigma = \phi(x/d)$, $x \in \Omega$. This sample corresponds to a sample $(z_1, \dots, z_n)$ from the real, complex or quaternionic Gaussian distribution with covariance x in $H_r(\mathbb{D})$, respectively. Let z be the matrix with columns $(z_1, \dots, z_n)$, let z^* be its conjugate transpose and let $y = zz^*$ be the unnormed sample covariance matrix. For the saturated model, that is, when $\Omega(\mathscr{K}) = \Omega$, it is well known [see Goodman (1963) and Andersson (1975)] that the mle of x is

$$s = \frac{y}{n}. \tag{3.4}$$

Then $S = \phi(s/d)$ is the maximum likelihood estimate of Σ corresponding to the sample $(\tilde{z}_1, \dots, \tilde{z}_n)$ from the $N(0, \Sigma)$ distribution on F.

Our aim in this section is to give an explicit closed form expression of the mle of x, when we have a sample $(z_1, \dots, z_n)$ from the $N(0, x)$ distribution on $\mathbb{D}$ and $\widehat{x} \in \Omega(\mathscr{K})$. Equivalently, we can say that we obtain the expression of the mle $\widehat{\Sigma}$ of Σ, given a sample $(\tilde{z}_1, \dots, \tilde{z}_n)$ from the $N(0, \Sigma)$ distribution, $\Sigma = \phi(x/d)$, $x \in \Omega(\mathscr{K})$.

In parallel to the two characterizations of $\Omega(\mathscr{K})$, given in Theorem 1(ii) and (vi), we give the mle of the concentration matrix $\delta = x^{-1}$ or $\Delta = \Sigma^{-1}$ in two forms. Theorem 3 gives $\widehat{\Delta}$ in closed form while Theorem 4 gives a recursive relation linking $\widehat{\Delta}$ with the mle of concentration matrices in smaller models. We consider a sample $(\tilde{z}_1, \dots, \tilde{z}_n)$ from a $N(0, \Sigma)$ distribution on $F = \mathbb{R}^{dr}$, $d = 1, 2$ or 4, with $\Sigma = \phi(x/d)$, $x \in \mathscr{P}(\mathscr{K})$ or equivalently, $(z_1, \dots, z_n)$ from a $N(0, x)$ distribution on $\mathbb{D}$, $x \in \Omega(\mathscr{K})$, and we let s be defined as in (3.4) with $S = \phi(s/d)$ the representation of s/d on F. We denote $\delta = x^{-1}$ the concentration matrix. We have the following results.

THEOREM 3. *Let $(z_1, \dots, z_n)$ be a sample as given above. Then the mle $\widehat{x}$ of x is unique and exists for a.e. $(z_1, \dots, z_n)$ if and only if $n \geq \max\{|K|: K \in J(\mathscr{K})\}$. When it exists, we have for all $L \in \mathscr{K}$,*

$$\widehat{\delta_L} = \Sigma\big(s^{-1}_K - s^{-1}_{\langle K\rangle} \,|\, K \in J(\mathscr{K}_L)\big) \tag{3.5}$$

and

$$\det_L \widehat{x_L} = \prod\left(\frac{\det_K s_K}{\det_{\langle K\rangle} s_{\langle K\rangle}} | K \in J(\mathscr{K}_L)\right) = \prod\left(\det_{[K]} s_{[K]\bullet} | K \in J(\mathscr{K}_L)\right). \tag{3.6}$$

In particular, for $L = I$, we have

$$\widehat{\delta} = \Sigma(s_K^{-1} - s_{\langle K\rangle}^{-1} | K \in J(K)) \quad and \quad \det \widehat{x} = \prod\left(\det_{[K]} s_{[K]\bullet} | K \in J(\mathscr{K})\right). \tag{3.7}$$

Equivalently,

$$\widehat{\Delta_L} = \Sigma\left(S_K^{-1} - S_{\langle K\rangle}^{-1} | K \in J(K_L)\right), \tag{3.5$'$}$$

$$\det_L \widehat{\Sigma_L} = \prod\left(\frac{\det_K S_K}{\det_{\langle K\rangle} S_{\langle K\rangle}} | K \in J(\mathscr{K}_L)\right) = \prod(\det_{[K]} S_{[K]\bullet} | K \in J(\mathscr{K}_L)), \tag{3.6$'$}$$

$$\widehat{\Delta} = \Sigma\left(S_K^{-1} - S_{\langle K\rangle}^{-1} | K \in J(\mathscr{K})\right) \tag{3.7$'$}$$

and

$$\det \widehat{\Sigma} = \prod\left(\det_{[K]} S_{[K]\bullet} | K \in J(\mathscr{K})\right). \tag{3.8}$$

THEOREM 4. *Let $(z_1, \dots, z_n)$ be a sample as given above. When it exists, the mle $\widehat{\delta}$ of δ is such that for all $L \notin J(\mathscr{K})$, $L \neq \varnothing$ and for $K_1, \dots, K_l$ a never-decreasing listing of the elements of $J(\mathscr{K}_L)$,*

$$\widehat{\delta_L} = \widehat{\delta_{K_l}} + \delta_{\widehat{\bigcup_{i=1}^{l-1} K_i}} - \widehat{\delta_{\langle K_l\rangle}} \tag{3.9}$$

or equivalently

$$\widehat{\Delta_L} = \widehat{\Delta_{K_l}} + \Delta_{\widehat{\bigcup_{i=1}^{l-1} K_i}} - \widehat{\Delta_{\langle K_l\rangle}}. \tag{3.10}$$

The proof of these two theorems is given in the following section.

We first illustrate Theorem 4, and consequently also Theorem 1(vi), with an example. We use the lattice given in Figure 1, where the empty dots mark elements of $J(\mathscr{K})$, and the full dots mark elements of $\mathscr{K}$ which are not in $J(\mathscr{K})$.

Here $J(\mathscr{K})^c = \{L \cup M, L', I\}$. Equation (3.9) becomes

$$\widehat{\delta_{L\cup M}} = \widehat{\delta_L} + \widehat{\delta_M} - \widehat{\delta_{L\cap M}},$$
$$\widehat{\delta_{L'}} = \widehat{\delta_{L''}} + \widehat{\delta_{L\cup M}} - \widehat{\delta}_L$$
$$\widehat{\delta} = \widehat{\delta_{M'}} + \widehat{\delta_{L'}} - \widehat{\delta_{L\cup M}}.$$

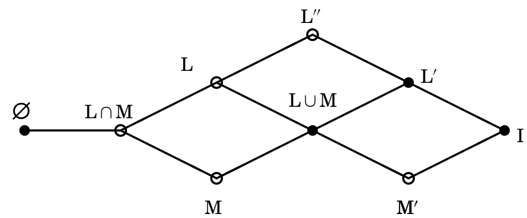

FIG. 1.

Using the first equality, we can rewrite the second and third one as

$$\widehat{\delta_{L'}} = \widehat{\delta_{L''}} + \widehat{\delta_M} - \widehat{\delta_{L\cap M}} = \widehat{\delta_{M'}} + \widehat{\delta_{L''}} - \widehat{\delta_L},$$

which shows that each one of the $\widehat{\delta}$ can be expressed in terms of $\widehat{\delta}_K$ for $K \in J(\mathscr{K})$.

PROOF. The proof of Theorem 3 is in three steps. The first step is to choose a regular decomposition for L, that is, $L, M \in \mathscr{K}$ and $K \in \mathscr{J}(\mathscr{K})$ such that $L = M\dot{\cup}[K]$ and to write the likelihood function for x_L in function of the Frobenius coordinates of x_L given by the complete orthogonal system ($c_1 = c_{\langle K\rangle}, c_2 = c_{M\setminus\langle K\rangle}, c_3 = c_{[K]}$). This uses the trace formula and the Frobenius parametrization and leads to the existence a.e. and uniqueness of the mle of the Frobenius coordinates given below in (3.11). The second step shows that for $\delta = x^{-1}$ the difference $\widehat{\delta_L} - \widehat{\delta_M}$ is equal to the difference $s_K^{-1} - s_{\langle K\rangle}^{-1}$ of the sample covariance submatrix s_K. It is this step that allows us to give a closed-form formula for $\widehat{\delta}$ rather than an algorithm. The third step is a simple induction on the previous step, that allows us to find our main result (3.5) and (3.6).

First step: the maximum likelihood estimate of the Frobenius coordinates. Consider the triplet L, M, K in $\mathscr{K}$ such that $K \in \mathscr{J}(\mathscr{K})$ and $L = M\dot{\cup}[K]$. By (2.3), $K \subseteq L$ and therefore $\langle K\rangle \subseteq M$. Now consider the complete orthogonal system ($c_1 = c_{\langle K\rangle},\ c_2 = c_{M\setminus\langle K\rangle},\ c_3 = c_{[K]}$). Since $z_{[K]} \perp\!\!\!\perp z_{M\setminus\langle K\rangle}|z_{\langle K\rangle}$, by Proposition 2(b)(vi),

$$x_L = \tau_{\langle K\rangle}\big(2x_{\langle K\rangle}^{-1}x_{[K\rangle}\big)(x_M + x_{[K]\bullet}), \tag{3.11}$$

which implies immediately, since the determinant of a Frobenius transformation is 1, that

$$\det_L x_L = \det_M x_M \det_{[K]} x_{[K]\bullet}. \tag{3.12}$$

The independence condition (3.11), combined with the trace formula (2.29) gives the following:

$$\begin{aligned}\operatorname{tr} x_L^{-1}s_L = \operatorname{tr} x_M^{-1}s_M + \operatorname{tr} x_{[K]\bullet}^{-1}\big(&4P(x_{\langle K\rangle}^{-1}x_{[K\rangle})s_{\langle K\rangle}\\ &- 4(x_{\langle K\rangle}^{-1}x_{[K\rangle})s_{\langle K]} + s_{[K]}\big).\end{aligned} \tag{3.13}$$

In matrix notation, (3.11), (3.12) and (3.13), for $x = \Sigma$ and $s = S$, read as

$$\Sigma_L = \begin{pmatrix} I_{\langle K\rangle} & 0_{\langle K\rangle,\, M\setminus\langle K\rangle} & 0_{\langle K]} \\ 0_{M\setminus\langle K\rangle,\, \langle K\rangle} & I_{M\setminus\langle K\rangle} & 0_{M\setminus\langle K\rangle,\, [K]} \\ \Sigma_{[K\rangle}\Sigma^{-1}_{\langle K\rangle} & 0_{[K],\, M\setminus\langle K\rangle} & I_{[K]} \end{pmatrix} \begin{pmatrix} & x_M & 0 \\ & & 0 \\ 0_{[K\rangle} & 0_{[K\rangle,\, M\setminus\langle K\rangle} & s_{[K]\bullet} \end{pmatrix}$$
$$\times \begin{pmatrix} I_{\langle K\rangle} & 0_{\langle K\rangle, M\setminus\langle K\rangle} & \Sigma^{-1}_{\langle K\rangle}\Sigma_{\langle K]} \\ 0_{M\setminus\langle K\rangle,\, \langle K\rangle} & I_{M\setminus\langle K\rangle} & 0_{M\setminus\langle K\rangle, [K]} \\ 0_{[K\rangle} & 0_{[K],\, M\setminus\langle K\rangle} & I_{[K]} \end{pmatrix},$$

$$\det_L \Sigma_L = \det_M \Sigma_M \det_{[K]} \Sigma_{[K]\bullet}$$

and

$$\operatorname{tr} \Sigma_L^{-1} S_L = \operatorname{tr} \Sigma_M^{-1} S_M + \operatorname{tr} \Sigma^{-1}_{[K]\bullet}\big(S_{[K]} - 2\Sigma_{[K\rangle}\Sigma^{-1}_{\langle K\rangle} S_{\langle K]} + \Sigma_{[K\rangle}\Sigma^{-1}_{\langle K\rangle} S_{\langle K\rangle}\Sigma^{-1}_{\langle K\rangle}\Sigma_{\langle K]}\big),$$

respectively. The reader will note that x_M, in the second matrix of the expression of Σ_L above belongs to $V(c_1 + c_2, 1)$.

From (3.12) and (3.13), we obtain the following factorization of the likelihood function $L(x_L)$:

$$\begin{aligned} L(x_L) &= C\, (\det_M x_M)^{-1/2} \exp\big[-\tfrac{1}{2} \operatorname{tr} x_M^{-1} s_M \big(\det_{[K]} x_{[K]\bullet}\big)^{-1/2}\big] \\ &\quad \times \exp\big[-\tfrac{1}{2} \operatorname{tr} x^{-1}_{[K]\bullet}\big(4P\big(x^{-1}_{\langle K\rangle} x_{[K\rangle}\big) s_{\langle K\rangle} - 4\big(x^{-1}_{\langle K\rangle} x_{[K\rangle}\big) s_{[K\rangle} + s_{[K]}\big)\big] \\ &= L_1(x_M) L_2\big(x^{-1}_{\langle K\rangle} x_{[K\rangle}, x_{[K]\bullet}\big), \end{aligned} \tag{3.14}$$

where C is a constant of proportionality. [This factorization is the parallel, for Frobenius coordinates, of Theorem 3.1 in AP (1993)]. Assuming that $\widehat{x_M}$ is known, we need only find the estimates of $(x^{-1}_{\langle K\rangle} x_{[K\rangle}, x_{[K]\bullet})$. Differentiating $\log L_2$ with respect to its two arguments leads to the classical $\widehat{x^{-1}_{\langle K\rangle} x_{[K\rangle}} = s^{-1}_{\langle K\rangle} s_{[K\rangle}$ and $\widehat{x_{[K]\bullet}} = s_{[K]} - P(s_{[K\rangle}) s^{-1}_{\langle K\rangle} = s_{[K]\bullet}$. By Proposition 3(a), there exists a unique $\widehat{x_L} \in \Omega(\mathscr{K}_L)$ with $(\widehat{x_L})_M = \widehat{x_M}$ and Frobenius coordinates $(\widehat{x_M}, s^{-1}_{\langle K\rangle} s_{[K\rangle}, 0, s_{[K]\bullet})$ and therefore, when it exists, the mle of x_L is equal to

$$\widehat{x_L} = \tau_{\langle K\rangle}(2 s^{-1}_{\langle K\rangle} s_{[K\rangle})(\widehat{x_M} + s_{[K]\bullet}). \tag{3.15}$$

Assuming $\widehat{x_M}$ exists, clearly $\widehat{x_L}$ exists if and only if $(s_{\langle K\rangle}, s^{-1}_{\langle K\rangle} s_{[K\rangle}, s_{[K]\bullet})$, the Frobenius coordinates of the mle of Σ_K in a saturated K-marginal model, exist. It is well known that this estimate exists, for almost all z, if and only if $n \geq |K|$. By the induction assumption given below in the third step, it follows immediately that the mle of x exists for a.e. $(z_1, \ldots, z_n)$ if and only if $n \geq \max\{|K|, K \in \mathscr{J}(\mathscr{K})\}$.

Second step: the estimate of $\delta_L - \delta_M$ as a function of the sample covariance matrix. We note that the z_{23}-coordinate of x_L is 0 and therefore $\tau_{\langle K\rangle}(2x_{\langle K\rangle}^{-1}x_{[K\rangle})(\widehat{x_M}+s_{[K]\bullet}) = \tau_M((2x_{\langle K\rangle}^{-1}x_{[K\rangle},0)(\widehat{x_M}+s_{[K]\bullet})$. Thus, we write

$$\widehat{x_L} = \tau_M(2s_{\langle K\rangle}^{-1}s_{[K\rangle})(\widehat{x_M}+s_{[K]\bullet}) \tag{3.16}$$

(see the section above for the corresponding matrix notation). Inverting both sides of (3.16) gives

$$\widehat{\delta_L} = \tau_M(-2s_{\langle K\rangle}^{-1}s_{[K\rangle})^*(\widehat{\delta_M}+s_{[K]\bullet}^{-1}) \tag{3.17}$$

From (2.26), we know that $\tau_M(-2s_{\langle K\rangle}^{-1}s_{[K\rangle})^*$ leaves the $V(c_{\langle K\rangle}+c_{M\setminus\langle K\rangle},1)$ component $\widehat{\delta_M}$ fixed and therefore, $\widehat{\delta_L}-\widehat{\delta_M}$ depends only on $s_{\langle K\rangle}^{-1}s_{[K\rangle}$. But using the Frobenius coordinates of s_K, we know that $s_K = \tau_{\langle K\rangle}(2s_{\langle K\rangle}^{-1}s_{[K\rangle})(s_{\langle K\rangle}+s_{[K]\bullet})$ or $s_K^{-1} = \tau_{\langle K\rangle}(-2s_{\langle K\rangle}^{-1}s_{[K\rangle})^*(s_{\langle K\rangle}^{-1}+s_{[K]\bullet}^{-1})$. Again, since $\tau_{\langle K\rangle}(-2s_{\langle K\rangle}^{-1}s_{[K\rangle})^*$ leaves $s_{\langle K\rangle}^{-1}$ fixed, we have that $s_K^{-1}-s_{\langle K\rangle}^{-1}$ depends on $s_{\langle K\rangle}^{-1}s_{[K\rangle}$ and $s_{[K]\bullet}^{-1}$ only and therefore $\widehat{\delta_L}-\widehat{\delta_M} = s_K^{-1}-s_{\langle K\rangle}^{-1}$.

Third step: induction on $|\mathscr{J}(\mathscr{K})|$. If $|\mathscr{J}(\mathscr{K})| = 1$, then clearly the single element K_1 of $\mathscr{J}(\mathscr{K})$ is such that $\langle K_1\rangle = \varnothing$. Otherwise, there would exist $K' \notin \mathscr{K}$ such that $K' \subset K_1$. Since K_1 is join-irreducible, this would imply the same for K' but this contradicts the fact that K_1 is the only element of a nondecreasing listing of $\mathscr{J}(\mathscr{K})$. Since $\langle K_1\rangle = \varnothing$ and since the independence conditions in $N(\mathscr{K})$ are $z_{[K_k]} \perp\!\!\!\perp z_{K_1\cup\cdots\cup K_{k-1}\setminus\langle K_k\rangle}|z_{\langle K_k\rangle}$, there is no independence condition in $N(\mathscr{K}_{K_1})$. The model is saturated and $\widehat{\delta_{K_1}} = s_{K_1}^{-1} = s_{K_1}^{-1} - s_{\langle K_1\rangle}^{-1}$, where we adopt the convention $s_{\varnothing}^{-1} = 0$. We also have $\det_{K_1}\widehat{x_{K_1}} = \det_{K_1}s_{K_1} = (\det_{K_1}s_{K_1}/\det_{\langle K_1\rangle}s_{\langle K_1\rangle})$ with the convention that $\det_{\varnothing}x = 1$. So (3.5) and (3.6) are verified for $\mathscr{K}_{K_1}$. Let us assume that (3.5) and (3.6) are true for M such that $|\mathscr{J}(\mathscr{K}_M)| = k-1$. Let $L \in \mathscr{K}$ such that $|\mathscr{J}(\mathscr{K}_L)| = k$. We choose a regular decomposition of L such that $L = M\dot{\cup}[K]$ for $K \in \mathscr{J}(\mathscr{K}_L)$. Then we have $|\mathscr{J}(\mathscr{K}_M)| = k-1$. By the previous step, we have $\widehat{\delta_L}-\widehat{\delta_M} = s_K^{-1}-s_{\langle K\rangle}^{-1}$ and by the induction assumption $\widehat{\delta_M} = \Sigma(s_K^{-1}-s_{\langle K\rangle}^{-1}|K \in \mathscr{J}(\mathscr{K}_M))$. Since $\mathscr{J}(\mathscr{K}_L) = \mathscr{J}(\mathscr{K}_M)\cup\{K\}$, this yields (3.5). Now, from (3.15), we have $\det_L\widehat{x_L} = \det_M\widehat{x_M}\det_{[K]}s_{[K]\bullet}$. By the induction assumption and the fact that $\det_{[K]}s_{[K]\bullet} = (\det_K s_K/\det_{\langle K\rangle}s_{\langle K\rangle})$, we immediately obtain (3.5).

Identities (3.5′), (3.6′) and (3.7′) follow immediately from (3.5), (3.6) and (3.7) using the representation ϕ.

PROOF OF THEOREM 4. The proof of Theorem 4 is immediate. If $L \in \mathscr{K}\setminus\mathscr{J}(\mathscr{K})$ and $\{K_1,\dots,K_l\}$ is a never-decreasing listing of the elements of $\mathscr{J}(\mathscr{K}_L)$, then for $M = K_1\cup\cdots\cup K_{l-1}$ and $K = K_l$, $L = M\dot{\cup}[K]$ is a regular decomposition of $L = \bigcup_{i=1}^l K_i$ with $L \neq K_l$. By (3.5),

$$\widehat{\delta_L} = \sum_{i=1}^{l} s_{K_i}^{-1} - s_{\langle K_i\rangle}^{-1} = \sum_{i=1}^{l-1} s_{K_i}^{-1} - s_{\langle K_i\rangle}^{-1} + s_{K_l}^{-1} - s_{K_{\langle K_l\rangle}}^{-1} \tag{3.18}$$

Since $L \neq K_l$, $\{K_1, \ldots, K_{l-1}\}$ is a never-decreasing listing of $\mathscr{J}(\mathscr{K}_M)$ and by (3.5) applied to M and K_l, respectively, we have $\widehat{\delta_L} = \widehat{\delta_M} + \widehat{\delta_{K_l}} - \widehat{\delta_{\langle K\rangle}}$, which completes the proof. $\square$

3.3. *Connection with various other results.*

3.3.1. *Our results for $N(0, \Sigma)$ models, Markov with respect to an acyclic digraph.* As mentioned in Section 1.4, Andersson, Perlman, Madigan and Triggs (1997) proved that the class of LCI models is identical to the class of transitive ADG models. It follows immediately that our results translate into corresponding results for transitive ADG models. In fact, we can be more general than this.

Our results (3.5) and (3.5′) for the estimate of the covariance matrix in an LCI model can be extended to the estimate of the covariance matrix of a $N(0, \Sigma)$ model Markov with respect to an acyclic digraph, not necessarily transitive. The proof follows the proof of Theorem 3 with the following correspondences. If $v(1), \ldots, v(T)$ denote the vertices of an acyclic digraph, $pa(v(i))$ the set of parents of $v(i)$ and $v_i = v(1) \cup \cdots \cup v(i)$, then L, M, K, $\langle K\rangle$ and $[K]$ in Sections 3.2 and 3.3 are replaced by v_T, v_{T-1}, $v(T) \cup pa(v(T))$, $pa(v(T))$ and $v(T)$, respectively. The estimate of the precision matrix is

$$\widehat{\delta_{v_T}} = \tau_{pa(v(T))}(-2s^{-1}_{pa(v(T))} s_{v(T),\, pa(v(T))})^* \big(\widehat{\delta_{v_{T-1}}} + s^{-1}_{v(T)\bullet}\big)$$

and we obtain that

$$\widehat{\delta_{v_T}} - \widehat{\delta_{v_{T-1}}} = s^{-1}_{v(T)\cup pa(v(T))} - s^{-1}_{pa(v(T))}.$$

An immediate induction argument yields

$$\hat{\delta} = \sum_{t=1}^{T} \big(s^{-1}_{v(t)\cup pa(v(t))} - s^{-1}_{pa(v(t))}\big). \tag{3.19}$$

Formula (3.19) has already been proved in Proposition 12 of Lauritzen (1989) for the real case. Our proof holds for a real, complex or quaternionic centered Gaussian model, Markov with respect to an acyclic digraph. Moreover the methods of proof are very different. Ours are purely algebraic.

3.3.2. *Relation to recent research.* Our paper has connections with Andersson and Madsen (1998) and AP (1998). Andersson and Madsen study the extension of normal LCI models to models combining group symmetry and conditional independence. AP (1998) discusses the structure and statistical properties of a real multivariate normal model with covariance structure determined by an acyclic digraph with the additional feature that compatible regression structures are also treated.

Our approach, however, is radically different from the group theoretic approach used in AP (1993) and the other two papers just mentioned. Our results lie in the framework of symmetric cones. Our paper is in the tradition of Massam (1994), Massam and Neher (1997), Casalis and Letac (1996) and Letac and Massam (1998).

Our paper is also connected, of course, to Lauritzen (1989) and Lauritzen (1996), Lemma 5.5, where we can find the fundamental result, in the real case, that $z_L \perp\!\!\!\perp z_M | z_{L\cap M}$, if and only if $x_{L\cup M}^{-1} + x_{L\cap M}^{-1} = x_L^{-1} + x_M^{-1}$. This is, of course, the characterization of $\Omega(\mathscr{K})$ where $\mathscr{K} = \{\varnothing, L\cap M, L, M, L\cup M = I\}$. This equivalence has been proved in the complex case by Anderson, Højbjerre, Sørensen and Eriksen (1995). Our results in Section 3 extend it to any LCI model on a symmetric cone.

4. Testing lattice conditional independence models. In this section, we consider the following testing problem. Given a sample $(z_1, \ldots, z_n)$ from the $N(0, x)$ distribution with $x \in H_r(\mathbb{D})$, $\mathbb{D} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, test

$$H_{\mathscr{K}}\colon x \in \Omega(\mathscr{K}) \text{ vs. } H_{\mathscr{M}}\colon x \in \Omega(\mathscr{M}), \tag{4.1}$$

where $\mathscr{M}$ is a proper sublattice of $\mathscr{K}$. This problem has been treated in the case $\mathbb{D} = \mathbb{R}$ by AP (1995a, b).

Following the same development as in AP (1995b), one can prove the equivalent of their Theorem 4.1 for $x \in H_r(\mathbb{D})$. This is done in Theorem 5 below, which gives the expression of the likelihood ratio statistic λ for testing (4.1), and its distribution. Before we state Theorem 5, we need to recall some concepts introduced in AP (1995b). We will then state Theorem 5 without proof since its proof follows the same development as the proof of Theorem 4.1 in AP (1995b) except for two points. The first is that we work in simple Euclidean Jordan algebras rather than just $H_r(\mathbb{R})$ and therefore we will need to prove the equivalent of their Lemma 2.1(iii) in the Jordan algebra framework. This will be Proposition 4 below. If w belongs to Ω and has the Wishart $W_r(p, x)$ distribution as defined below in (4.3), and if $w = w_1 + w_{12} + w_0$ and $x = x_1 + x_{12} + x_0$ are the Peirce decompositions of w and x with respect to an idempotent c with $tr(c) = k$, then Proposition 4 gives the well-known distributional and independence properties of w_1, $w_0 - P(w_{12})w_1^{-1}$ and w_{12}. It also states that if $x_{12} = 0$, then $P(w_{12})w_1^{-1}$ is also Wishart and $w_0 - P(w_{12})w_1^{-1}$, $P(w_{12})w_1^{-1}$ and w_1 are mutually independent. The second point of difference is that we view the different components ω_{mk} of λ, as given in (4.2) below, as determinants of Beta variables and use the generalized Beta distribution rather than use Wilks U-distribution.

4.1. *The likelihood ratio, its distribution.* To be able to compare the independence conditions given by the lattices $\mathscr{K}$ and $\mathscr{M}$, AP (1995b) introduced the following surjective poset homomorphism $\Psi\colon J(\mathscr{K}) \to J(\mathscr{M})$ defined by $\Psi(K) := \cap(M \in \mathscr{M} | M \supseteq K)$. Let $M_1, \ldots, M_s$ be a never decreasing listing of the elements of $J(\mathscr{M})$ and for each M_m, let $K_{m1}, \ldots, K_{mq_m}$ be a never decreasing listing of the elements of $\{K \in J(\mathscr{K}) | \Psi(K) = M_m\}$. Then the set $\{K_{mj},\ 1 \le m \le s,\ 1 \le j \le q_m\}$ form a never decreasing listing of $J(\mathscr{K})$. For ease of notation, we write $\langle m\rangle$ and $[m]$ for $K_{\langle m\rangle}$ and $K_{[m]}$ when those are used as indices. The sample covariance matrix can be divided into blocks $s_{K_{mj}}$. We will write $s_{[mk]\bullet} = s_{[mk]\bullet\langle mk\rangle} = s_{[mk]} - P(s_{[mk\rangle})s_{\langle mk\rangle}^{-1}$ and and in general, for any L, M, subsets of I, $s_{M\cdot L} = s_M - P(s_{M,L})s_L^{-1}$ (see Section 2).

THEOREM 5. *The likelihood ratio statistic* λ *for problem* (4.1) *can be factored as follows*:

$$\lambda^{2/n} = \prod(\omega_{mk} | m = 1, \dots, s, \ k = 1, \dots, q_m), \tag{4.2}$$

where

$$\omega_{mk} = \frac{\det s_{[mk]\bullet(\langle m\rangle \dot\cup [m1] \dot\cup \cdots \dot\cup [m(k-1)]}}{\det s_{[mk]\bullet\langle mk\rangle}}.$$

The factors ω_{mk} *are mutually independent and each* ω_{mk} *is distributed as the determinant of a Beta* $(n/2 - |\langle m\rangle \dot\cup [m1] \dot\cup \cdots \dot\cup [m(k-1)]|(d/2), |(rmk)|d/2)$ *variable where* d *is the Peirce constant of* $\mathbb{D}$, *that is*, $d = 1, 2$, *or* 4 *when* $\mathbb{D} = \mathbb{R}$, $\mathbb{C}$ *or* $\mathbb{H}$, *respectively, and* $(rmk) = \langle m\rangle \dot\cup [m1] \dot\cup \cdots \dot\cup [m(k-1)] \setminus \langle mk\rangle$. *The likelihood ratio statistic* λ *and the mle of* x *are mutually independent under* $H_{\mathscr{K}}$.

The proof of this theorem follows the same argument as the proof of Theorem 4.1 in AP (1995b) with their Lemma 2.1(iii) replaced by Proposition 4 below.

4.2. *Independence properties.* Let us recall the definition of the Wishart distribution on an irreducible symmetric cone [see Casalis (1990), Massam (1994) and Faraut and Koranyi (1994) for more details on these distributions]. We say that the random variable w on the irreducible symmetric cone Ω of rank r has the $W_r(p, x)$ Wishart distribution if its density with respect to the Lebesgue measure on Ω is

$$f(w) = \frac{1}{\Gamma_\Omega(p)} (\det x^{-1})^p (\det w)^{p - N/r} \exp -\frac{1}{2} \operatorname{tr} x^{-1} w \tag{4.3}$$

when N = dimension of Ω, r =rank(Ω), d is the Peirce constant for Ω and $\Gamma_\Omega(p) = (2\pi)^{(N-r)/2} \Gamma(p)\Gamma(p - d/2)\Gamma(p - 2d/2) \cdots \Gamma(p - (r-1)d/2)$ is the Γ-function of Ω.

PROPOSITION 4. *Let* $w \sim W_r(p, x)$. *Let* c *be an idempotent of* rank k, $w = w_1 + w_{12} + w_0$, $x = x_1 + x_{12} + x_0$ *the Peirce decomposition of* w *and* x *with respect to* c.

(i) *Then* $w_1 \sim W_k(p, x_1)$.

(ii) $w_0 - P(w_{12}) w_1^{-1} \sim W_{r-k}(p - kd/2, (x^{-1})_0^{-1}) = W_{r-k}(p - kd/2, x_0 - P(x_{12}) x_1^{-1})$. *Moreover* $w_0 - P(w_{12}) w_1^{-1}$ *is independent of* (w_{12}, w_1).

(iii) *If in addition* $x_{12} = 0$, *then* $w_0 - P(w_{12}) w_1^{-1} \sim W_{r-k}(p - k(d/2), x_0)$, $P(w_{12}) w_1^{-1} \sim W_{r-k}(k(d/2), x_0)$ *and* $w_0 - P(w_{12}) w_1^{-1}$, $P(w_{12}) w_1^{-1}$ *and* w_1 *are mutually independent*.

PROOF. The proof of (i) and (ii) for irreducible symmetric cones can be found in Massam and Neher (1997). Let us prove (iii). When $x_{12} = 0$, $x_0 - P(x_{12}) x_1^{-1} = x_0$ and therefore by (ii), $w_0 - P(w_{12}) w_1^{-1} \sim W_{r-k}(p - k(d/2), x_0)$.

Moreover by (ii), $w_0 - P(w_{12})w_1^{-1}$ is independent of (w_{12}, w_1) and therefore of $P(w_{12})w_1^{-1}$. By (i), we also know that $w_0 \sim W_{r-k}(p, x_0)$. Let us compute its moment generating function:

$$\begin{aligned} E(\exp(\theta w_0)) &= \det(e - x_0\theta)^{-p} \\ &= E(\exp(\theta(w_0 - P(w_{12})w_1^{-1}) + \theta P(w_{12})w_1^{-1})) \end{aligned} \tag{4.4}$$

and since $w_0 - P(w_{12})w_1^{-1}$ and $P(w_{12})w_1^{-1}$ are independent,

$$\begin{aligned} E(\exp(\theta w_0)) &= E(\exp(\theta(w_0 - P(w_{12})w_1^{-1})))E(\exp(\theta P(w_{12})w_1^{-1})) \\ &= \det(e - x_0\theta)^{-(p-k(d/2))}E(\exp(\theta P(w_{12})w_1^{-1})). \end{aligned} \tag{4.5}$$

From (4.4) and (4.5), it follows that $E(\exp(\theta P(w_{12})w_1^{-1})) = \det(e - x_0\theta)^{-k(d/2)}$ which proves that $P(w_{12})w_1^{-1} \sim W_{r-k}(k(d/2), x_0)$. It remains to show that $w_0 - P(w_{12})w_1^{-1}$, $P(w_{12})w_1^{-1}$ and w_1 are mutually independent. We know that $w_0 - P(w_{12})w^{-1}$ and w_1 are independent. Let us show that $P(w_{12})w_1^{-1}$ and w_1 are independent. Since w_0 and w_1 are independent and $w_0 - P(w_{12})w_1^{-1}$ and $P(w_{12})w_1^{-1}$ are also independent we have

$$\begin{aligned} E(\exp(\theta w_0)) &= E(\exp(\theta(w_0 - P(w_{12})w_1^{-1}) + \theta P(w_{12})w_1^{-1})) \\ &= E(\exp(\theta(w_0 - P(w_{12})w_1^{-1})))E(\exp(\theta P(w_{12})w_1^{-1})) \\ &= E(\exp(\theta w_0)|w_1) \\ &= E(\exp(\theta(w_0 - P(w_{12})w_1^{-1}))|w_1)E(\exp(\theta P(w_{12})w_1^{-1})|w_1) \\ &\qquad \text{(since independence implies conditional independence)} \\ &= E(\exp(\theta(w_0 - P(w_{12})w_1^{-1})))E(\exp(\theta P(w_{12})w_1^{-1})|w_1) \\ &\qquad \text{(since } w_0 - P(w_{12})w_1^{-1} \text{ and } w_1 \text{ are independent).} \end{aligned}$$

It follows that $E(\exp(\theta P(w_{12})w_1^{-1})|w_1) = E(\exp(\theta P(w_{12})w_1^{-1}))$ and the independence of $P(w_{12})w_1^{-1}$ and w_1 is proved. This completes the proof of (iii). □

4.3. *The distribution of ω_{mk}.* Let us now show that each ω_{mk} is of the form $\det(U_1/(U_1 + U_2))$ where U_1 and U_2 are independent Wishart variates on Ω with the same scale parameter. Here $U_1/(U_1 + U_2)$ is a symbolic notation for the variable $t_{U_1+U_2}^{-1}(U_1)$, where $t_{U_1+U_2}$ is the unique triangular transformation [see Theorem VI.3.6 in Faraut and Koranyi (1994)] in the group of automorphisms of Ω such that $t_{U_1+U_2}(e) = U_1 + U_2$ and where $t_{U_1+U_2}^{-1}$ is its inverse. This implies of course that $U_1/(U_1 + U_2)$ has the generalized Beta distribution on $\Omega \cap (e - \Omega)$ [see Massam (1994)]. Following the same argument as in AP (1995b) and using Proposition 4, we can prove that $\lambda^{2/n} = (\prod \omega_{mk}, m = 1, \dots, s, k = 1, \dots, q_m)$ where ω_{mk} is as in (4.2). Recall that $K_1, \dots, K_{mq_m}$ is a never-decreasing listing of the elements of $J(\mathscr{K}_{M_m})$. Recall that we use the notation $\langle m\rangle$, $[mk]$ and (m) for $K_{\langle m\rangle}$, $K_{[mk]}$ and K_m, respectively, and $(rmk) = \langle m\rangle \dot\cup [m1] \dot\cup \cdots \dot\cup [m(k-1)] \setminus \langle mk\rangle$ so that (m) can be decomposed into the disjoint union $(m) = (rmk)\dot\cup\langle mk\rangle\dot\cup[mk]$. Then let

$$U_1 = s_{[mk]\bullet\langle m\rangle\dot\cup[m1]\dot\cup\cdots\dot\cup[m(k-1)]} = s_{[mk]\bullet\langle mk\rangle} - P\big(s_{([mk],\,(rmk))\bullet\langle mk\rangle}\big)s^{-1}_{(rmk)\bullet\langle mk\rangle}$$

and let

$$U_2 = P(s_{([mk],\,(rmk))\bullet\langle mk\rangle})s^{-1}_{(rmk)\bullet\langle mk\rangle}$$

so that

$$s_{[mk]\langle mk\rangle} = U_1 + U_2.$$

Then clearly each factor ω_{mk} is of the form $\det U_1/\det(U_1+U_2)$ and it is easy to show that $\det t^{-1}_{U_1+U_2}(U_1) = \det U_1/\det U_1 + U_2$.

Let $p = n/2$. We know that ns, where s is the mle of x in the saturated $N(0,x)$ model, has the Wishart $W_r(p,x)$ distribution. Since $(m) = (rmk)\dot\cup\langle mk\rangle\dot\cup[mk]$, we have, by Proposition 4(ii),

$$ns_{((rmk)\cup[mk])\bullet\langle mk\rangle} \sim W\bigg(p - |\langle mk\rangle|\frac{d}{2},\ x_{((rmk)\cup[mk])\bullet\langle mk\rangle}\bigg)$$

and

$$nU_1 = ns_{[mk]\bullet(\langle mk\rangle\cup(rmk))} \sim W\bigg(p - |\langle mk\rangle\cup(rmk)|\frac{d}{2},\ x_{[mk]\bullet(\langle mk\rangle\cup(rmk))}\bigg).$$

Under $H_{\mathscr{K}}$, we have that [see AP (1995a), for $z\sim N(0,x)$, $x\in\Omega(\mathscr{K})$,

$$z_{[mk]} \perp\!\!\!\perp z_{\langle m\rangle\cup(m1)\cup\cdots\cup(m(k-1))}|z_{\langle mk\rangle}.$$

This implies, of course, that $x_{([mk],\,(rmk))\bullet\langle mk\rangle} = 0$ and by Proposition 4(iii), it follows that

$$nU_1 \sim W_{r-|\langle mk\rangle|-|(rmk)|}\bigg(p - |\langle mk\rangle|\frac{d}{2} - |(rmk)|\frac{d}{2},\ x_{[mk]\bullet\langle mk\rangle}\bigg),$$

$$nU_2 \sim W_{r-|\langle mk\rangle|-|(rmk)|}\bigg(|(rmk)|\frac{d}{2},\ x_{[mk]\bullet\langle mk\rangle}\bigg)$$

and nU_1 and nU_2 are independent, with the same scale parameter $x_{[mk]\bullet\langle mk\rangle}$. By Theorem 4.2 of Massam (1994), this implies that $U_1/(U_1+U_2) = t^{-1}_{U_1+U_2}(U_1)$ has the generalized Beta $B_\Omega(p_1,p_2)$ distribution with $p_1 = p - |\langle mk\rangle|(d/2) - |(rmk)|(d/2)$, $p_2 = |(rmk)|(d/2)$ and so

$$\begin{aligned}
E(\omega_{mk}) &= \frac{B_\Omega(p_1+1,\,p_2)}{B_\Omega(p_1,\,p_2)}\\
&= \frac{\Gamma_\Omega(p_1+1)\Gamma_\Omega(p_2)\,\Gamma_\Omega(p_1+p_2)}{\Gamma_\Omega(p_1+p_2+1)\,\Gamma_\Omega(p_1)\,\Gamma_\Omega(p_2)}\\
&= \frac{\prod_{j=0}^{r-1}\Gamma(p_1+1-j(d/2))\Gamma(p_1+p_2-j(d/2))}{\prod_{j=0}^{r-1}\Gamma(p_1-j(d/2))\Gamma(p_1+p_2+1-j(d/2))}.
\end{aligned}$$

Acknowledgments. We thank S. Andersson, G. Letac and M. Perlman for useful remarks on earlier versions of this paper. We also thank S. Lauritzen for several enlightening discussions.

REFERENCES

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

ANDERSON, H. H., HØJBJERRE, M., SØRENSEN, D. and ERIKSEN, P. S. (1995). *Linear and Graphical Models for the Multivariate Complex Normal Distribution*. Springer, Berlin.

ANDERSSON, S. A., MADIGAN, D., PERLMAN, M. D. and TRIGGS, C. M. (1997). A graphical representation of lattice conditional independence models. *Ann. of Math. 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. (1988). Lattice-models for conditional independence in a multivariate normal distribution. Technical report, Dept. Statistics, Univ. Washington.

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

ANDERSSON, S. A. and PERLMAN, M. D. (1995a). Testing lattice conditional independence models. *J. Multivariate Anal.* **53** 18–38.

ANDERSSON, S. A. and PERLMAN, M. D. (1995b). Unbiasedness for the likelihood ratio test for lattice conditional independence models. *J. Multivariate Anal.* **53** 1–17.

ANDERSSON, S. A. and PERLMAN, M. D. (1998). Normal linear regression models with recursive graphical Markov structure. *J. Multivariate Anal.* To appear.

CASALIS, M. and LETAC, G. (1996). The Lukacs–Olkin–Rubin characterization of Wishart distributions on symmetric cones. *Ann. Statist.* **24** 763–786.

FARAUT, J. and KORANYI, A. (1994). *Analysis on Symmetric Cones*. Clarendon Press, Oxford.

GOODMAN, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Statist.* **34** 152–176.

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

LAURITZEN, S. L. (1989). Mixed graphical association models. *Scand. J. Statist.* **16** 273–306.

LAURITZEN, S. L. (1996). *Graphical Models*. Oxford Univ. Press.

LETAC, G. and MASSAM, H. (1998). Quadratic and inverse regression for Wishart distributions. *Ann. Statist.* **26** 573–595.

MASSAM, H. (1994). An exact decomposition theorem and a unified view of some related distributions for a class of exponential transformation models on symmetric cones. *Ann. Statist.* **22** 369–394.

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

NEHER, E. (1997). Transformation groups of the Andersson–Perlman cone. Preprint.

DEPARTMENT OF STATISTICS
YORK UNIVERSITY
TORONTO M3J 1P3
CANADA
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22901
E-MAIL: hmm2n@virginia.edu

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF OTTAWA
OTTAWA K1N 6N5
CANADA
E-MAIL: neher@science.uottawa.ca