The Annals of Statistics
1998, Vol. 26, No. 2, 525–572

SYMMETRY AND LATTICE CONDITIONAL INDEPENDENCE IN A MULTIVARIATE NORMAL DISTRIBUTION

BY STEEN ANDERSSON[1] AND JESPER MADSEN[1,2]

Indiana University and University of Copenhagen

A class of multivariate normal models with symmetry restrictions given by a finite group and conditional independence restrictions given by a finite distributive lattice is defined and studied. The statistical properties of these models including maximum likelihood inference, invariance and hypothesis testing are discussed.

1. Introduction. Three of the most important concepts used in defining a statistical model are independence, conditional distribution and symmetry. (The assumption most often used in statistics is that of *i.i.d. observations*, that is, *independent and identical distributed observations*, which means *independence* between observations and *symmetry* under any permutation of the observations.) Statistical models given by a combination of two of these concepts, conditional distribution and independence, the so-called conditional independence (CI) models, have received increasing attention in recent years. The models are defined in terms of directed graphs, undirected graphs, or the combination of the two, the so-called chain graphs. See Whittaker (1990) or Lauritzen (1996) for an introduction to models of this type. The special connections between statistical models and graphs have been the subject of many of the contributions to this area, see, for example, Andersson and Perlman (1995b), Cox and Wermuth (1993), Lauritzen (1989, 1996), Lauritzen and Wermuth (1989), or Frydenberg (1990). The special class of CI models where all distributions are assumed to be multivariate normal is of special interest. Under this assumption, Andersson and Perlman (1993, 1995a) [hereafter abbreviated AP (1993) and AP (1995a), respectively] introduced the so-called lattice conditional independence (LCI) models and presented a complete solution to their estimation and testing problems. The relations between LCI models (without the assumption of normality) and other CI models are studied in Andersson, Madigan, Perlman and Triggs (1995a, b).

Received March 1996; revised January 1997.

[1]Supported in part by U.S. National Security Agency Grant MDA 904-92-H-3083 and by NSF Grant DMS-94-02714.

[2]Research carried out in part at the Department of Statistics, University of Washington.

AMS 1991 *subject classifications*. Primary 62H12, 62H15; secondary 62H10, 62H20, 62A05.

Key words and phrases. Group symmetry, invariance, orthogonal group representation, quotient space, conditional independence, distributive lattice, join-irreducible elements, maximum likelihood estimator, likelihood ratio test, multivariate normal distribution.

As part of a general development of the theory of the normal distribution, Brøns (1969) presented a general definition of group symmetry (GS) models to S. Andersson and S. T. Jensen. In the years 1972–1985, Andersson, Brøns, and Jensen together developed an algebraic theory for these models containing a complete solution to the likelihood inference problem. This basic theory, detailed in numerous Danish manuscripts [e.g., Andersson, Brøns and Jensen (1975), Andersson (1975a, 1976), Brøns (1969) and Jensen (1973, 1974, 1977, 1983)], has not yet been published. Several manuscripts in English summarize the theory [e.g., Andersson (1978, 1992)]. In Andersson (1975b) the structure of the models was explained and a solution to the estimation problem was given in a canonical form. Perlman (1987) reviews a small part of the theory. In Andersson, Brøns and Jensen (1983), the ten fundamental irreducible testing problems within this theory are discussed. Andersson and Perlman (1984) and Bertelsen (1989) treat the noncentral distributions connected to two of these ten testing problems. Since the present paper uses most of the basic theory of GS models, a summary is presented in Appendix A.

The present paper combines the lattice conditional independence restrictions with the group symmetry restrictions to obtain the *group symmetry lattice conditional independence* (*GS-LCI*) *models*. The GS models and the LCI models then become special cases of the GS-LCI models. In this paper we give necessary and sufficient conditions for the existence and uniqueness of the maximum likelihood (ML) estimator for an arbitrary observation, necessary and sufficient conditions for the existence and uniqueness of the ML estimator with probability 1, an explicit expression for the ML estimator, an explicit expression for the likelihood ratio statistic Q for testing one GS-LCI model against another and the central distribution of Q in terms of the moments $\mathrm{E}(Q^{\alpha})$, $\alpha > 0$.

Andersen, Højbjerre, Sørensen and Eriksen (1995) combine the symmetry given by the complex numbers, that is, the GS condition given by the group $\{\pm 1, \pm i\}$, with CI restrictions given by an undirected graph. In Hylleberg, Jensen and Ørnbøl (1993) a subgroup of the symmetric group is combined with CI restrictions given by an undirected graph. In both cases there is a nontrivial overlap with the models in the present paper. These arise from the overlap between LCI models and CI models given by undirected graphs, as explained in Andersson, Madigan, Perlman and Triggs (1995a, b). However, in the case of Hylleberg, Jensen and Ørnbøl (1993) the nontrivial overlap is also because the restriction of the interplay between the special group of permutations and the CI conditions is relaxed compared to the restriction between the general GS and LCI conditions in the present paper. Madsen (1996) discusses ML estimation in a class of models which extends both the GS-LCI models and those of Andersen, Højbjerre, Sørensen and Eriksen (1995) and Hylleberg, Jensen and Ørnbøl (1993).

We introduce the GS-LCI models by means of the following four simple examples.

EXAMPLE 1.1. Let $x_a = (x_{a1}, x_{a2})$, $x_b = (x_{b1}, x_{b2})$ and $x_c = (x_{c1}, x_{c2})$ be three pairs of random observations with a joint normal distribution with mean zero and covariance matrix $\Sigma = (\sigma_{l\nu, k\mu} | l, k = a, b, c;\ \nu, \mu = 1, 2)$, that is, $\sigma_{l\nu, k\mu}$ is the covariance between the two observations $x_{l\nu}$ and $x_{k\mu}$. For example (x_a, x_b, x_c) could be measurements of three different variables a, b, and c on two symmetric objects, for example, two plants within the same plot. Since the joint distribution should not depend on the (probably irrelevant) numbering of the two plants, it should remain invariant under the simple linear transformation that corresponds to permutation of plant indices. This implies that Σ has the restriction

$$\mathrm{H}_{\mathrm{GS}}\colon\ \sigma_{l\nu, k\mu} = \begin{cases} \gamma_{lk}, & \nu = \mu, \\ \omega_{lk}, & \nu \neq \mu \end{cases}$$

where $\gamma_{lk} = \gamma_{kl}$ and $\omega_{lk} = \omega_{kl}$ are real numbers, $l, k = a, b, c$. Thus, under H_{GS}, the six-dimensional variable $x = (x_{a1}, x_{b1}, x_{c1}, x_{a2}, x_{b2}, x_{c2})'$ has the 2×2 block covariance matrix

$$\Sigma = \begin{pmatrix} \Gamma & \Omega \\ \Omega & \Gamma \end{pmatrix}, \tag{1.1}$$

where $\Gamma = \Gamma' = (\gamma_{lk} | l, k = a, b, c)$ and $\Omega = \Omega' = (\omega_{lk} | l, k = a, b, c)$. This covariance structure is a special case of *multivariate complete symmetry*; compare Section A.6. Next, consider the assumption that x_a and x_c are conditionally independent given x_b, which we express in the familiar notation

$$\mathrm{H}_{\mathrm{LCI}}\colon\ x_a \perp x_c \mid x_b.$$

This restriction could occur if the three measured variables correspond to three "sites" on the plant where a is a neighbor to b, and b is a neighbor to c, in which case the dependence between the observations from "site" a and "site" c is indirect due only to their mutual dependence on the observations from "site" b. The lattice (ring) $\mathscr{K}$ of subsets of the index set $I = \{1a, 2a, 1b, 2b, 1c, 2c\}$, which defines this CI restriction, is given by

$$\mathscr{K} = \{\varnothing, \{1b, 2b\}, \{1b, 2b, 1a, 2a\}, \{1b, 2b, 1c, 2c\}, I\},$$

compare AP (1993), Example 2.5. The restriction imposed on Σ by both H_{GS} and $\mathrm{H}_{\mathrm{LCI}}$ can then be expressed as (1.1) together with the additional restriction

$$\mathrm{H}_{\text{GS-LCI}}\colon \begin{pmatrix} \gamma_{ac} & \omega_{ac} \\ \omega_{ac} & \gamma_{ac} \end{pmatrix} = \begin{pmatrix} \gamma_{ab} & \omega_{ab} \\ \omega_{ab} & \gamma_{ab} \end{pmatrix} \begin{pmatrix} \gamma_{bb} & \omega_{bb} \\ \omega_{bb} & \gamma_{bb} \end{pmatrix}^{-1} \begin{pmatrix} \gamma_{bc} & \omega_{bc} \\ \omega_{bc} & \gamma_{bc} \end{pmatrix}.$$

We thus have four hypotheses for the covariance matrix Σ, namely the unconstrained H, the two subhypotheses H_{GS} and $\mathrm{H}_{\mathrm{LCI}}$ and their intersection $\mathrm{H}_{\text{GS-LCI}}$.

Now consider N i.i.d. observations $x_1, \ldots, x_N$ of the six-dimensional random observation x. It is well known that under H the ML estimator exists and is unique with probability 1 if and only if $N \geq 6$. Moreover it is well known from classical multivariate analysis that in the models H_{GS} and $\mathrm{H}_{\mathrm{LCI}}$, the required conditions are $N \geq 3$ and $N \geq 4$, respectively. In all three cases, an explicit expression for the ML estimator is easily obtained. In the case of the model $\mathrm{H}_{\text{GS-LCI}}$, the results in the present paper applied to this simple case show that the condition for existence and uniqueness of the ML estimator with probability 1 is $N \geq 2$. The ML estimator can be found using a combination of the techniques applied for GS models and LCI models. First, one determines the ML estimator

$$\hat{\Sigma}_{\mathrm{GS}} = \begin{pmatrix} \hat{\Gamma} & \hat{\Omega} \\ \hat{\Omega} & \hat{\Gamma} \end{pmatrix},$$

for Σ under H_{GS}, where $\hat{\Gamma} = (\hat{\gamma}_{lk} | l, k = a, b, c)$ and $\hat{\Omega} = (\hat{\omega}_{lk} | l, k = a, b, c)$. Under $\mathrm{H}_{\mathrm{LCI}}$, the likelihood function (LF) factorizes into the product of the conditional LF of x_a given x_b, the conditional LF of x_c given x_b and the marginal LF of x_b. These factors then contain two 2×2 regression parameters R_a and R_c, two 2×2 conditional covariance matrices Λ_a and Λ_c and one 2×2 marginal covariance matrix Λ_b. The ML estimator $\hat{\Sigma}_{\text{GS-LCI}}$ for Σ under $\mathrm{H}_{\text{GS-LCI}}$, is then determined by

$$\hat{R}_l = \begin{pmatrix} \hat{\gamma}_{lb} & \hat{\omega}_{lb} \\ \hat{\omega}_{lb} & \hat{\gamma}_{lb} \end{pmatrix} \begin{pmatrix} \hat{\gamma}_{bb} & \hat{\omega}_{bb} \\ \hat{\omega}_{bb} & \hat{\gamma}_{bb} \end{pmatrix}^{-1},$$

$$\hat{\Lambda}_l = \begin{pmatrix} \hat{\gamma}_{ll} & \hat{\omega}_{ll} \\ \hat{\omega}_{ll} & \hat{\gamma}_{ll} \end{pmatrix} - \begin{pmatrix} \hat{\gamma}_{lb} & \hat{\omega}_{lb} \\ \hat{\omega}_{lb} & \hat{\gamma}_{lb} \end{pmatrix} \begin{pmatrix} \hat{\gamma}_{bb} & \hat{\omega}_{bb} \\ \hat{\omega}_{bb} & \hat{\gamma}_{bb} \end{pmatrix}^{-1} \begin{pmatrix} \hat{\gamma}_{bl} & \hat{\omega}_{bl} \\ \hat{\omega}_{bl} & \hat{\gamma}_{bl} \end{pmatrix},$$

for $l = a, c$, respectively, and

$$\hat{\Lambda}_b = \begin{pmatrix} \hat{\gamma}_{bb} & \hat{\omega}_{bb} \\ \hat{\omega}_{bb} & \hat{\gamma}_{bb} \end{pmatrix}.$$

Of the five possible testing problems within the design of the models given by H, H_{GS}, $\mathrm{H}_{\mathrm{LCI}}$ and $\mathrm{H}_{\text{GS-LCI}}$, the three involving $\mathrm{H}_{\text{GS-LCI}}$ seem to be new. The likelihood ratio statistic and its central distribution for these tests can easily be obtained from the general theory presented in this paper.

EXAMPLE 1.2. In Example 1.1, instead of $\mathrm{H}_{\mathrm{LCI}}$, consider the assumption

$$\mathrm{H}'_{\mathrm{LCI}}\colon\ x_a \perp x_c,$$

that is, x_a and x_c are *marginally independent*. The interpretation of this restriction is that the actual measurements of plants on "site" a do not contain any information about the measurements on "site" c and vice versa. The lattice (ring) $\mathscr{K}'$ of subsets of the index set $I = \{1a, 2a, 1b, 1c, 2c\}$ which defines this CI restriction is given by

$$\mathscr{K}' = \{\varnothing, \{1a, 2a\}, \{1c, 2c\}, \{1a, 2a, 1c, 2c\}, I\};$$

compare AP (1993), Example 2.4. The restriction imposed on Σ by both H_{GS} and $\mathrm{H}'_{\mathrm{LCI}}$ can then be expressed as (1.1) together with the additional restriction

$$\mathrm{H}'_{\text{GS-LCI}}\colon \gamma_{ac} = \omega_{ac} = 0.$$

Note that $\mathrm{H}_{\mathrm{LCI}}$ and $\mathrm{H}'_{\mathrm{LCI}}$ are nonnested and have a nontrivial intersection. We thus again consider four hypotheses for the covariance matrix Σ, namely the unconstrained H, the two subhypotheses H_{GS} and $\mathrm{H}'_{\mathrm{LCI}}$ and their intersection $\mathrm{H}'_{\text{GS-LCI}}$.

Consider N i.i.d. observations $x_1, \dots, x_N$ of the six-dimensional variable x. From Example 2.4 in AP (1993) it follows that under $\mathrm{H}'_{\mathrm{LCI}}$, the ML estimator exists and is unique with probability 1 if and only if $N \geq 6$. The results in the present paper shows that under $\mathrm{H}'_{\text{GS-LCI}}$, the required condition is $N \geq 3$. In this case, the ML estimator can be determined in the same way as in Example 1.1. Under $\mathrm{H}'_{\mathrm{LCI}}$, the likelihood function (LF) factorizes into the conditional LF of x_b given (x_a, x_c) and the marginal LFs of x_a and x_c, respectively. These factors then contain one 2×4 regression parameter R_b, one 2×2 conditional covariance matrix Λ_b and two 2×2 marginal covariance matrices Λ_a, Λ_c. The ML estimator $\hat{\Sigma}_{\text{GS-LCI}}$ for Σ under $\mathrm{H}'_{\text{GS-LCI}}$, is then determined by

$$\hat{R}_b = \begin{pmatrix} \hat{\gamma}_{ab} & \hat{\gamma}_{bc} & \hat{\omega}_{ab} & \hat{\omega}_{bc} \\ \hat{\omega}_{ab} & \hat{\omega}_{bc} & \hat{\gamma}_{ab} & \hat{\gamma}_{bc} \end{pmatrix} \begin{pmatrix} \hat{\gamma}_{aa} & \hat{\gamma}_{ac} & \hat{\omega}_{aa} & \hat{\omega}_{ac} \\ \hat{\gamma}_{ac} & \hat{\gamma}_{cc} & \hat{\omega}_{ac} & \hat{\omega}_{cc} \\ \hat{\omega}_{aa} & \hat{\omega}_{ac} & \hat{\gamma}_{aa} & \hat{\gamma}_{ac} \\ \hat{\omega}_{ac} & \hat{\omega}_{cc} & \hat{\gamma}_{ac} & \hat{\gamma}_{cc} \end{pmatrix}^{-1},$$

$$\hat{\Lambda}_b = \begin{pmatrix} \hat{\gamma}_{bb} & \hat{\omega}_{bb} \\ \hat{\omega}_{bb} & \hat{\gamma}_{bb} \end{pmatrix} - \hat{R}_b \begin{pmatrix} \hat{\gamma}_{ab} & \hat{\omega}_{ab} \\ \hat{\gamma}_{bc} & \hat{\omega}_{bc} \\ \hat{\omega}_{ab} & \hat{\gamma}_{ab} \\ \hat{\omega}_{bc} & \hat{\gamma}_{bc} \end{pmatrix}$$

and $\times$

$$\hat{\Lambda}_l = \begin{pmatrix} \hat{\gamma}_{ll} & \hat{\omega}_{ll} \\ \hat{\omega}_{ll} & \hat{\gamma}_{ll} \end{pmatrix},$$

for $l = a, c$, respectively. As in Example 1.1, the three testing problems involving $\mathrm{H}'_{\text{GS-LCI}}$ of the possible five within the design of the models given by H, H_{GS}, $\mathrm{H}'_{\mathrm{LCI}}$, and $\mathrm{H}'_{\text{GS-LCI}}$, seem to be new. The likelihood ratio statistic and its central distribution for these tests can be obtained from the general theory presented in this paper.

Example 1.3. Let $x_a = (x_{a1}, x_{a2}, \dots, x_{an_a})$, $x_b = (x_{b1}, x_{b2}, \dots, x_{bn_b})$ and $x_c = (x_{c1}, x_{c2}, \dots, x_{cn_c})$ be three families of n_a, n_b and n_c multivariate random observations, respectively. The dimensions of the multivariate observations within each of the families are p_a, p_b and p_c, respectively. The simultaneous distribution of these $n_a p_a + n_b p_b + n_c p_c$ real observations is assumed to be normal with mean vector zero and $(n_a + n_b + n_c) \times (n_a + n_b + n_c)$ block covariance matrix $\Sigma = (\Sigma_{l\nu, k\mu} | l, k = a, b, c;\ \nu = 1, \dots, n_l;\ \mu = 1, \dots, n_k)$; that is, $\Sigma_{l\nu, k\mu}$ is the $p_l \times p_k$ covariance matrix between the two multivariate observations $x_{l\nu}$ and $x_{k\mu}$. For example, x_a, x_b and x_c could be multivariate measurements on plants from three different varieties a, b and c, respectively. Since the joint distribution should not depend on the numbering of plants within a variety, it must remain invariant under any linear transformation of the sample space that corresponds to renumbering of plants within varieties. This implies that the covariance matrix Σ has the restrictions given by

$$\mathrm{H}_{\mathrm{GS}}\colon \Sigma_{l\nu, k\mu} = \begin{cases} \Gamma_l, & l = k,\ \nu = \mu, \\ \Omega_l, & l = k,\ \nu \neq \mu, \\ \Delta_{lk}, & l \neq k, \end{cases}$$

where $\Gamma_l = \Gamma_l'$ is a $p_l \times p_l$ matrix, $\Omega_l = \Omega_l'$ is a $p_l \times p_l$ matrix, and $\Delta_{lk} = \Delta_{kl}'$ is a $p_l \times p_k$ matrix, $l, k = a, b, c;\ l \neq k$. Thus, under H_{GS} the random vector

$$x = (x_{a1}', x_{a2}', \dots, x_{an_a}', x_{b1}', x_{b2}', \dots, x_{bn_b}', x_{c1}', x_{c2}', \dots, x_{cn_c}')'$$

of real dimension $n_a p_a + n_b p_b + n_c p_c$ has the block covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} & \Sigma_{ac} \\ \Sigma_{ba} & \Sigma_{bb} & \Sigma_{bc} \\ \Sigma_{ca} & \Sigma_{cb} & \Sigma_{cc} \end{pmatrix}, \tag{1.2}$$

where

$$\Sigma_{ll} = \begin{pmatrix} \Gamma_l & \Omega_l & \cdots & \cdots & \Omega_l \\ \Omega_l & \Gamma_l & \Omega_l & \cdots & \Omega_l \\ \vdots & \Omega_l & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \Omega_l & \Omega_l & \cdots & \Omega_l & \Gamma_l \end{pmatrix} \tag{1.3}$$

and

$$\Sigma_{lk} = \begin{pmatrix} \Delta_{lk} & \cdots & \Delta_{lk} \\ \vdots & \ddots & \vdots \\ \Delta_{lk} & \cdots & \Delta_{lk} \end{pmatrix}, \tag{1.4}$$

for $l, k = a, b, c;\ l \neq k$. This is an example of what we could call *multivariate compound symmetry*, first considered by Votaw (1948) in the univariate case; that is, $p_a = p_b = p_c = 1$ (see Section A.6). Next consider the assumption that the families x_a and x_c are conditionally independent given the family x_b, which we express in the familiar notation

$$\mathrm{H}_{\mathrm{LCI}}\colon x_a \perp x_c \mid x_b.$$

This restriction could occur if the three families of variables correspond to three plots a, b and c where a is neighbor to b, and b is a neighbor to c, in which case the dependence between the observations from plot a and plot c is due only to the observations from plot b. The lattice (ring) $\mathscr{K}$ of subsets of the index set

$$I = \{a1, a2, \dots, an_a, b1, b2, \dots, bn_b, c1, c2, \dots, cn_c\}$$

which defines this conditional independence is given by

$$\mathscr{K} = \left\{\varnothing, I_b, I_a \dot{\cup} I_b, I_b \dot{\cup} I_c, I\right\},$$

where $I_l = \{l1, l2, \dots, ln_l\}$, $l = a, b, c$; compare AP (1993), Example 2.5. The restriction imposed on Σ by both $\mathrm{H_{GS}}$ and $\mathrm{H_{LCI}}$ can then be expressed as (1.2), (1.3) and (1.4) together with the additional restriction

$\mathrm{H_{GS\text{-}LCI}}$:

$$\begin{pmatrix} \Delta_{ac} & \cdots & \Delta_{ac} \\ \vdots & \ddots & \vdots \\ \Delta_{ac} & \cdots & \Delta_{ac} \end{pmatrix}$$

$$= \begin{pmatrix} \Delta_{ab} & \cdots & \Delta_{ab} \\ \vdots & \ddots & \vdots \\ \Delta_{ab} & \cdots & \Delta_{ab} \end{pmatrix} \begin{pmatrix} \Gamma_b & \Omega_b & \cdots & \cdots & \Omega_b \\ \Omega_b & \Gamma_b & \Omega_b & \cdots & \Omega_b \\ \vdots & \Omega_b & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \Omega_b & \Omega_b & \cdots & \Omega_b & \Gamma_b \end{pmatrix}^{-1} \begin{pmatrix} \Delta_{bc} & \cdots & \Delta_{bc} \\ \vdots & \ddots & \vdots \\ \Delta_{bc} & \cdots & \Delta_{bc} \end{pmatrix},$$

or equivalently,

$$\mathrm{H_{GS\text{-}LCI}}\colon \Delta_{ac} = n_b \Delta_{ab} (\Gamma_b + (n_b - 1)\Omega_b)^{-1} \Delta_{bc}.$$

We thus again have four hypotheses in the covariance matrix Σ, namely the unconstrained H, the two subhypotheses $\mathrm{H_{GS}}$ and $\mathrm{H_{LCI}}$, and the intersection $\mathrm{H_{GS\text{-}LCI}}$.

Let $x_1, x_2, \dots, x_N$ be N i.i.d. observations of the $(n_a p_a + n_b p_b + n_c p_c)$-dimensional random vector x. It is well known that under H the ML estimator exists and is unique with probability 1 if and only if $N \geq n_a p_a + n_b p_b + n_c p_c$. The model given by $\mathrm{H_{GS}}$ is well known when $p_a = p_b = p_c = 1$; compare Votaw (1948). It follows from the theory of GS models presented in Appendix A that in the general case, the ML estimator for Σ exists and is unique with probability 1 if and only if $N \geq p_a + p_b + p_c$, $N(n_a - 1) \geq p_a$, $N(n_b - 1) \geq p_b$, and $N(n_c - 1) \geq p_c$; see Section A.4. In the familiar model given by $\mathrm{H_{LCI}}$, the condition is $N \geq \max\{n_a p_a + n_b p_b, n_c p_c + n_b p_b\}$; compare AP (1993), Example 2.5. For both models, the ML estimator is easily obtained. In the case of the model $\mathrm{H_{GS\text{-}LCI}}$, the theory presented in the present paper shows that the conditions for existence and uniqueness of the ML estimator with probability 1 become $N \geq p_a + p_b$, $N \geq p_c + p_b$, $N(n_a - 1) \geq p_a$, $N(n_b - 1) \geq p_b$ and $N(n_c - 1) \geq p_c$. The ML estimator can be found using a combination of the techniques from GS models and LCI models. First one finds the ML estimator $(\hat{\Gamma}_a, \hat{\Gamma}_b, \hat{\Gamma}_c, \hat{\Omega}_a, \hat{\Omega}_b, \hat{\Omega}_c, \hat{\Delta}_{ab}, \hat{\Delta}_{ac}, \hat{\Delta}_{bc})$ under H_{GS}. Let $y = (x_1, x_2, \dots, x_N)$ be the $I \times N$ observation matrix and let $y_{l\nu}$ be the $p_l \times N$ submatrix $((x_1)_{l\nu}, (x_2)_{l\nu}, \dots, (x_N)_{l\nu})$ of y, $\nu = 1, \dots, n_l$, $l = a, b, c$. We then obtain that

$$\hat{\Gamma}_l = \frac{1}{n_l} \sum (y_{l\nu} y'_{l\nu} \mid \nu = 1, \dots, n_l),$$

$$\hat{\Omega}_l = \frac{1}{n_l(n_l - 1)} \sum (y_{l\nu} y'_{l\mu} \mid \nu, \mu = 1, \dots, n_l,\ \nu \neq \mu),$$

$$\hat{\Delta}_{lk} = \frac{1}{n_l n_k} \sum (y_{l\nu} y'_{k\mu} \mid \nu = 1, \dots, n_l,\ \mu = 1, \dots, n_k),$$

where $l, k = a, b, c$, $l \neq k$.

Under $\mathrm{H}_{\mathrm{LCI}}$ the likelihood function (LF) can be factorized into the conditional LF of x_a given x_b, the conditional LF of x_c given x_b and the marginal LF of x_b. These factors then contain two multivariate regression parameters R_{ab} and R_{cb}, of dimensions $n_a p_a \times n_b p_b$ and $n_c p_c \times n_b p_b$, respectively; two multivariate conditional covariance matrices Λ_a and Λ_c of dimensions $n_a p_a \times n_a p_a$ and $n_c p_c \times n_c p_c$, respectively and one marginal covariance matrix Λ_b of dimension $n_b p_b \times n_b p_b$. Under $\mathrm{H}_{\mathrm{GS\text{-}LCI}}$, the regression parameters R_{lb}, $l = a, c$ and the variance parameters Λ_l, $l = a, b, c$ have the form (1.4) and (1.3), respectively. Thus,

$$R_{lb} = \begin{pmatrix} T_{lb} & \cdots & T_{lb} \\ \vdots & \ddots & \vdots \\ T_{lb} & \cdots & T_{lb} \end{pmatrix} \quad \text{and} \quad \Lambda_l = \begin{pmatrix} \Upsilon_l & \Phi_l & \cdots & \cdots & \Phi_l \\ \Phi_l & \Upsilon_l & \Phi_l & \cdots & \Phi_l \\ \vdots & \Phi_l & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \Phi_l & \Phi_l & \cdots & \Phi_l & \Upsilon_l \end{pmatrix},$$

where T_{lb} is a $p_l \times p_b$ matrix, $l = a, c$ and $\Upsilon_l = \Upsilon'_l$, $\Phi_l = \Phi'_l$ are $p_l \times p_l$ matrices, $l = a, b, c$. The ML estimator $\hat{\Sigma}$ for Σ under $\mathrm{H}_{\mathrm{GS\text{-}LCI}}$ is then determined by setting $\Upsilon_b = \hat{\Gamma}_b$, $\Phi_b = \hat{\Omega}_b$ and

$$T_{lb} = \hat{\Delta}_{lb}\big(\hat{\Gamma}_b + (n_b - 1)\hat{\Omega}_b\big)^{-1},$$

$$\Upsilon_l = \hat{\Gamma}_l - n_b \hat{\Delta}_{lb}\big(\hat{\Gamma}_b + (n_b - 1)\hat{\Omega}_b\big)^{-1}\hat{\Delta}_{bl},$$

$$\Phi_l = \hat{\Omega}_l - n_b \hat{\Delta}_{lb}\big(\hat{\Gamma}_b + (n_b - 1)\hat{\Omega}_b\big)^{-1}\hat{\Delta}_{bl},$$

for $l = a, c$, respectively.

Of the five possible testing problems within the design of the models given by H, H_{GS}, $\mathrm{H}_{\mathrm{LCI}}$, and $\mathrm{H}_{\mathrm{GS\text{-}LCI}}$, the problem of testing $\mathrm{H}_{\mathrm{LCI}}$ versus H is well known from the literature; compare AP (1995a), and the problem of testing H_{GS} versus H follows from the theory covered in Appendix A. The three tests involving the hypothesis $H_{\text{GS-LCI}}$ seem to be new. The likelihood ratio test statistics and a representation of the corresponding central distributions can easily be obtained from the general theory in the present paper.

EXAMPLE 1.4. Analogously to the construction of Example 1.2, consider the assumption

$$\mathrm{H}'_{\mathrm{LCI}}\colon\ x_a \perp x_c,$$

instead of $\mathrm{H}_{\mathrm{LCI}}$ in Example 1.3. This restriction could occur if the plants on plot a are assumed not to influence the plants on plot c and vice versa. The lattice (ring) $\mathscr{K}'$ of subsets of the index set I which defines this CI restriction is given by

$$\mathscr{K}' = \left\{\varnothing, I_a, I_c, I_a \dot{\cup} I_c, I\right\};$$

compare AP (1993), Example 2.4. The restriction imposed on Σ by both H_{GS} and $\mathrm{H}'_{\mathrm{LCI}}$ can then be expressed as (1.2), (1.3) and (1.4) together with the additional restriction

$$\mathrm{H}'_{\text{GS-LCI}}\colon\ \Delta_{ac} = 0.$$

As in Example 1.3, consider N i.i.d. observations $x_1, \dots, x_N$ of the $(n_a p_a + n_b p_b + n_c p_c)$-dimensional variable x. From Example 2.4 in AP (1993) it follows that under $\mathrm{H}'_{\mathrm{LCI}}$, the ML estimator exists and is unique with probability 1 if and only if $N \geq n_a p_a + n_b p_b + n_c p_c$, that is, the same as in the unconstrained case. The results in the present paper shows that under $\mathrm{H}'_{\text{GS-LCI}}$, the required conditions are $N \geq p_a + p_b + p_c$, $N(n_a - 1) \geq p_a$, $N(n_b - 1) \geq p_b$ and $N(n_c - 1) \geq p_c$, respectively, that is, the same as in the case of H_{GS}. In this case, the ML estimator can easily be determined in the same way as in the previous examples. Similarly, the three testing problems involving $\mathrm{H}'_{\text{GS-LCI}}$ of the possible five within the design of the models given by the unconstrained H, H_{GS}, $\mathrm{H}'_{\mathrm{LCI}}$, and $\mathrm{H}'_{\text{GS-LCI}}$, seem to be new, and the likelihood ratio statistic and its central distribution for these tests can be obtained from the general theory presented in this paper.

In general the observation space is $\mathbb{R}^I$ where I is a finite index set. The general definition of a GS-LCI model is stated in terms of an orthogonal group representation ρ of a finite group G on $\mathbb{R}^I$ together with a ring (lattice) $\mathscr{K}$ of subsets of the index set I. The GS-LCI model is then defined by imposing symmetry conditions given by ρ and conditional independence conditions given by $\mathscr{K}$. A condition on the interplay between the group representation and the ring is required to ensure the complete solution of the GS-LCI model. In Section 2 the GS-LCI models are defined (Section 2.4), the fundamental factorization of the parameter space $P_{G,\mathscr{K}}(I)$ of all $I \times I$ covariance matrices determined by the GS and LCI restrictions is obtained (Theorem 2.1) and the fundamental invariance group $\mathrm{GL}_{G,\mathscr{K}}(I)$ is defined together with its transitive action on $P_{G,\mathscr{K}}(I)$ (Theorem 2.2). The distribution results for the likelihood ratio statistics are greatly facilitated by this transitive action. The derivations of these distributions which generalize and improve the corresponding derivations for LCI models are presented in Appendix B; compare the Appendix of AP (1995a). In Section 3 a necessary and sufficient condition for the existence of the ML estimator and a necessary and sufficient condition for the uniqueness of the ML estimator for a fixed observation $x \in \mathbb{R}^I$ is obtained together with an almost explicit expression for the ML estimator (Theorem 3.1). In Proposition 3.2 the necessary and sufficient algebraic condition for the existence and uniqueness of the ML estimator with probability 1 is obtained. The structure constants for a GS-LCI model are then introduced. In terms of these, another, very useful, necessary and sufficient condition for the existence and uniqueness of the ML estimator with probability 1 is obtained (Proposition 3.3). Section 4 presents the general testing problem, and the likelihood ratio test statistic Q is derived. The central distribution of Q in terms of the moments $\mathrm{E}(Q^\alpha)$, $\alpha > 0$, is given as a function of the structure constants. In Section 5, it is established that independent repetitions (i.i.d.) of a GS-LCI model is again a GS-LCI model, except for a trivial reparametrization. Furthermore, it is shown how estimators and structure constants for the i.i.d. GS-LCI model are obtained in terms of the original GS-LCI model (Section 5.1). In Section 5.2 it is demonstrated how to construct new examples ad libitum based on well-known examples of GS models (cf. Section A.6) and the examples of LCI models in AP (1993). Finally, in Section 6, we indicate how the GS-LCI models can be extended in various ways.

2. Mathematical formulation. In this section we explain the mathematical set-up for the combined GS-LCI models to be investigated. Furthermore, we present some fundamental theorems describing the structure of the set $\mathrm{P}_{G,\mathscr{K}}(I)$ of covariance matrices that satisfy the GS-LCI restrictions. We have tried as much as possible to use the same type of notation as in AP (1993) and (1995a). In the following, let I and J denote finite index sets and let $|I|$ denote the number of elements in I.

2.1. *Notation.* Let $\mathbb{R}^I$ be the vector space of all families $x \equiv (x_i | i \in I)$ of real numbers indexed by I. For $K \subseteq I$, let $p_K\colon \mathbb{R}^I \to \mathbb{R}^K$ be the canonical projection and $u_K\colon \mathbb{R}^K \to \mathbb{R}^I$ the canonical imbedding; that is, $p_K((x_i | i \in I)) = (x_i | i \in K)$ and $u_K((x_i | i \in K)) = (x'_i | i \in I)$, where $x'_i = x_i$ for $i \in K$ and $x'_i = 0$ otherwise. For $x \in \mathbb{R}^I$, let x_K denote $p_K(x) \in \mathbb{R}^K$. Note that $\mathbb{R}^{\varnothing} = \{0\}$.

Let $\mathrm{M}(I \times J) \equiv \mathbb{R}^{I \times J}$ denote the vector space of all $I \times J$ matrices. The algebra $\mathrm{M}(I \times I)$ is denoted by $\mathrm{M}(I)$. For $A \in \mathrm{M}(I \times J)$ let $A' \in \mathrm{M}(J \times I)$ denote the transposed matrix. The group of all nonsingular $I \times I$ matrices, the group of all orthogonal $I \times I$ matrices, the cone of all positive semidefinite $I \times I$ matrices and the cone of all positive definite $I \times I$ matrices are denoted by $\mathrm{GL}(I)$, $\mathrm{O}(I)$, $\mathrm{PS}(I)$, and $\mathrm{P}(I)$, respectively. The action of the group $\mathrm{GL}(I)$ on $\mathrm{P}(I)$ given by

$$\begin{aligned} \mathrm{GL}(I) \times \mathrm{P}(I) &\to \mathrm{P}(I), \\ (A, \Sigma) &\mapsto A \Sigma A' \end{aligned} \tag{2.1}$$

is well known to be transitive and proper. The $I \times I$ identity matrix is denoted by 1_I.

For $A = (a_{ii'} \mid (i, i') \in I \times I) \in \mathrm{M}(I)$ and $K \subseteq I$, let A_K denote the $K \times K$ submatrix of A; that is, $A_K = (a_{ii'} \mid (i, i') \in K \times K) \in \mathrm{M}(K)$. If A_K is nonsingular, then A_K^{-1} denotes the inverse matrix $(A_K)^{-1}$.

For any subspace $U \subseteq \mathbb{R}^I$, let $U^{\perp}$ denote the orthogonal complement to U wrt the usual inner product in $\mathbb{R}^I$; that is, $U^{\perp} = \{x \in \mathbb{R}^I \mid \forall\ z \in U\colon\ x'z = 0\}$ and denote by $P_U \in \mathrm{M}(I)$ the corresponding orthogonal projection matrix.

For $\xi \in \mathbb{R}^I$ and $\Sigma \in \mathrm{P}(I)$ let $\mathrm{N}(\xi, \Sigma)$ denote the normal distribution on $\mathbb{R}^I$ with expectation ξ and covariance matrix Σ. Let $\mathrm{N}(\Sigma)$ denote $\mathrm{N}(0, \Sigma)$.

The overall normal model $(\mathrm{N}(\Sigma) \mid \Sigma \in \mathrm{P}(I))$ is invariant under the action of $\mathrm{GL}(I)$ on the observation space $\mathbb{R}^I$ given by

$$\begin{aligned} \mathrm{GL}(I) \times \mathbb{R}^I &\to \mathbb{R}^I, \\ (A, x) &\mapsto Ax, \end{aligned} \tag{2.2}$$

and the transitive action of $\mathrm{GL}(I)$ on the parameter space $\mathrm{P}(I)$ given by (2.1).

2.2. *The lattice conditional independence model.* Let $\mathscr{K}$ be a subring of the ring $\mathscr{D}(I)$ of all subsets of I; that is, $\mathscr{K}$ is closed under union and intersection. Since $\mathscr{K}$ is a distributive lattice wrt these operations, we usually refer to $\mathscr{K}$ as a *lattice of subsets of I*. Without loss of generality we assume that $I, \varnothing \in \mathscr{K}$.

A matrix $A \in \mathrm{M}(I)$ is called *$\mathscr{K}$-preserving* if for every $K \in \mathscr{K}$ and $x \in \mathbb{R}^I$, $(Ax)_K = A_K x_K$, or equivalently, if $Au_K(\mathbb{R}^K) \subseteq u_K(\mathbb{R}^K)$. Let $\mathrm{M}_{\mathscr{K}}(I)$ be the algebra of all $\mathscr{K}$-preserving matrices [in AP (1993), $\mathrm{M}_{\mathscr{K}}(I)$ was denoted $\mathrm{M}(\mathscr{K})$], and let $\mathrm{GL}_{\mathscr{K}}(I)$ be the group of all nonsingular $\mathscr{K}$-preserving matrices [in AP (1993), $\mathrm{GL}_{\mathscr{K}}(I)$ was denoted $\mathrm{GL}(\mathscr{K})$].

Define the subset $\mathrm{P}_{\mathscr{K}}(I) \subseteq \mathrm{P}(I)$ as follows: $\Sigma \in \mathrm{P}_{\mathscr{K}}(I)$ if and only if x_L and x_M are conditionally independent given $x_{L \cap M}$ for every $L, M \in \mathscr{K}$ whenever $x \in \mathbb{R}^I$ follows $\mathrm{N}(\Sigma)$ [in AP (1993), $\mathrm{P}_{\mathscr{K}}(I)$ was denoted by $\mathrm{P}(\mathscr{K})$]. The statistical model

$$(\mathrm{N}(\Sigma) \mid \Sigma \in \mathrm{P}_{\mathscr{K}}(I)) \tag{2.3}$$

with observation space $\mathbb{R}^I$ and parameter space $\mathrm{P}_{\mathscr{K}}(I)$ is called the *lattice conditional independence* (*LCI*) *model determined by* $\mathscr{K}$.

For $K \in \mathscr{K}$, define $\langle K \rangle = \bigcup(K' \in \mathscr{K} \mid K' \subset K)$ and $[K] = K \setminus \langle K \rangle$, so that

$$K = \langle K \rangle \mathbin{\dot\cup} [K], \tag{2.4}$$

where $\dot\cup$ indicates that the union is disjoint. Let $\mathscr{J}(\mathscr{K})$ denote the set of *join-irreducible elements* of $\mathscr{K}$, that is, $K \in \mathscr{J}(\mathscr{K})$ if and only if $\langle K \rangle \subset K$, or equivalently, if $[K] \neq \varnothing$. The subsets $[K]$ of I, $K \in \mathscr{J}(\mathscr{K})$, are all disjoint, and

$$K = \dot\bigcup([K'] \mid K' \in \mathscr{J}(\mathscr{K}),\ K' \subseteq K),$$

$K \in \mathscr{K}$. In particular,

$$I = \dot{\bigcup}([K] \mid K \in \mathscr{I}(\mathscr{K})) \tag{2.5}$$

[see AP (1993), Proposition 2.1].

For every $K \in \mathscr{I}(\mathscr{K})$ and $A \in \mathrm{M}(I)$, partition A_K according to the decomposition (2.4) as follows:

$$A_K = \begin{pmatrix} A_{\langle K \rangle} & A_{\langle K]} \\ A_{[K \rangle} & A_{[K]} \end{pmatrix},$$

so $A_{\langle K \rangle} \in \mathrm{M}(\langle K \rangle)$, $A_{\langle K]} \in \mathrm{M}(\langle K \rangle \times [K])$, $A_{[K \rangle} \in \mathrm{M}([K] \times \langle K \rangle)$ and $A_{[K]} \in \mathrm{M}([K])$. For $\Sigma \in \mathrm{P}(I)$ and $K \in \mathscr{I}(\mathscr{K})$, define $\Sigma_{[K]\cdot} = \Sigma_{[K]} - \Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1}\Sigma_{\langle K]}$. The following five results are the main tools in solving the estimation and testing problems for models of the type (2.3).

1. The mapping
$$\mathrm{P}_{\mathscr{K}}(I) \to \times(\mathrm{M}([K] \times \langle K \rangle) \times \mathrm{P}([K]) \mid K \in \mathscr{I}(\mathscr{K})),$$
$$\Sigma \mapsto \left(\left(\Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1}, \Sigma_{[K]\cdot}\right) \middle| K \in \mathscr{I}(\mathscr{K})\right),$$
is bijective [AP (1993), Theorem 2.2];
2. The covariance matrix $\Sigma \in \mathrm{P}_{\mathscr{K}}(I)$ if and only if
$$\mathrm{tr}(\Sigma^{-1}xx') = \sum\left(\mathrm{tr}\left(\Sigma_{[K]\cdot}^{-1}\left(x_{[K]} - \Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1}x_{\langle K \rangle}\right)(\cdots)'\right) \middle| K \in \mathscr{I}(\mathscr{K})\right), \tag{2.6}$$
for all $x \in \mathbb{R}^I$ [AP (1993), Theorem 2.1];
3. For $\Sigma \in \mathrm{P}_{\mathscr{K}}(I)$ and $L \in \mathscr{K}$,
$$\det(\Sigma_L) = \prod\left(\det\left(\Sigma_{[K]\cdot}\right) \middle| K \in \mathscr{I}(\mathscr{K}),\ K \subseteq L\right) \tag{2.7}$$
[AP (1993), Lemma 2.5]. In particular,
$$\det(\Sigma) = \prod\left(\det\left(\Sigma_{[K]\cdot}\right) \middle| K \in \mathscr{I}(\mathscr{K})\right); \tag{2.8}$$
4. The action of the group $\mathrm{GL}_{\mathscr{K}}(I)$ on $\mathrm{P}_{\mathscr{K}}(I)$ given by restriction of (2.1) is well defined, transitive and proper;
5. The model (2.3) is invariant under the action of $\mathrm{GL}_{\mathscr{K}}(I)$ on the observation space $\mathbb{R}^I$ given by the restriction of the action (2.2) and the transitive action of $\mathrm{GL}_{\mathscr{K}}(I)$ on the parameter space $\mathrm{P}_{\mathscr{K}}(I)$.

2.3. *The group symmetry model.* Let G be a finite group and ρ: $G \to \mathrm{O}(I)$ an *orthogonal group representation* of G on $\mathbb{R}^I$, that is, $\rho(1) = 1_I$ and $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$, for all $g_1, g_2 \in G$. Let $\mathrm{M}_G(I)$ denote the subalgebra of all matrices $A \in \mathrm{M}(I)$ that commute with $\rho(G)$, that is, $A\rho(g) = \rho(g)A$ for all $g \in G$. The group of all nonsingular matrices and the cone of all positive definite matrices in $\mathrm{M}_G(I)$ are denoted by $\mathrm{GL}_G(I)$ and $\mathrm{P}_G(I)$, respectively. Note that $\Sigma \in \mathrm{P}_G(I)$ if and only if $\Sigma \in \mathrm{P}(I)$ and Σ is G-invariant, that is, $\rho(g)\Sigma\rho(g)' = \Sigma$. Thus if $x \in \mathbb{R}^I$ follows the distribution $\mathrm{N}(\Sigma)$, where $\Sigma \in \mathrm{P}_G(I)$, then $\rho(g)x$ follows the same distribution for all $g \in G$. The statistical model

$$(\mathrm{N}(\Sigma) \mid \Sigma \in \mathrm{P}_G(I)) \tag{2.9}$$

with observation space $\mathbb{R}^I$ and parameter space $\mathrm{P}_G(I)$ is thus called the *group symmetry* (*GS*) *model given by G*. A summary of the basic theory of these models is presented in Appendix A; see the Introduction.

The smoothing ($\equiv$ averaging) mapping

$$\begin{aligned} \psi_I^G\colon \mathrm{PS}(I) &\to \mathrm{PS}_G(I), \\ S &\mapsto \frac{1}{|G|} \sum (\rho(g) S \rho(g)' \mid g \in G), \end{aligned} \tag{2.10}$$

where $\mathrm{PS}_G(I)$ denoting the cone of all positive semidefinite G-invariant $I \times I$-matrices is fundamental for likelihood inference for group symmetry models. When I, G or both are subsumed, we denote ψ_I^G by ψ^G, ψ_I and ψ, respectively.

Similar to (4) in Section 2.1, the action of the group $\mathrm{GL}_G(I)$ on $\mathrm{P}_G(I)$ given by restriction of (2.1) is well defined, transitive and proper (see Appendix A). The model (2.9) is invariant under the action of $\mathrm{GL}_G(I)$ on the observation space $\mathbb{R}^I$ given by the restriction of the action (2.2) and the transitive action of $\mathrm{GL}_G(I)$ on the parameter space $\mathrm{P}_G(I)$.

2.4. *Models having both GS and LCI restrictions*. Let $\mathscr{K} \subseteq \mathscr{D}(I)$ be a lattice of subsets of I and $\rho\colon G \to \mathrm{O}(I)$ an orthogonal group representation of G on $\mathbb{R}^I$. The intersection $\mathrm{P}_{\mathscr{K}}(I) \cap \mathrm{P}_G(I)$ is denoted by $\mathrm{P}_{G,\mathscr{K}}(I)$, that is, $\mathrm{P}_{G,\mathscr{K}}(I)$ is the set of covariance matrices having both symmetry restrictions w.r.t. G and conditional independence restrictions given by $\mathscr{K}$. The corresponding statistical model with observation space $\mathbb{R}^I$ and parameter space $\mathrm{P}_{G,\mathscr{K}}(I)$ is thus

$$(\mathrm{N}(\Sigma) \mid \Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)). \tag{2.11}$$

In the present paper we shall assume that all matrices $\rho(g)$, $g \in G$, are $\mathscr{K}$-preserving. Thus for $K \in \mathscr{K}$, the mapping $\rho_K\colon G \to \mathrm{O}(K)$ given by $\rho_K(g) = \rho(g)_K$, $g \in G$, is a well-defined orthogonal group representation of G on $\mathbb{R}^K$. Under this assumption, the model (2.11) is called the *group symmetry lattice conditional independence* (*GS-LCI*) *model determined by G and $\mathscr{K}$*.

The statistical interpretation of the above condition is that all the marginal distributions $\mathrm{N}(\Sigma_K)$, $K \in \mathscr{K}$ are G-invariant themselves. As a consequence (cf. Lemma 2.1), all the matrices $\rho(g)$, $g \in G$, are block diagonal w.r.t. to the decomposition (2.5). Symmetry conditions are thus only allowed to operate *within* each of the marginal variables $x_{[K]}$, $K \in \mathscr{J}(\mathscr{K})$, of $x \in \mathbb{R}^I$. Example 6.1 presents a model of type (2.11) where the matrices $\rho(g)$, $g \in G$, are not $\mathscr{K}$-preserving; that is, the model has GS and LCI restrictions but it is not a GS-LCI model.

LEMMA 2.1. *Let $g \in G$. The matrix $\rho(g)$ is $\mathscr{K}$-preserving if and only if $u_{[K]}(\mathbb{R}^{[K]})$ is a G-subspace, that is, $\rho(g)(u_{[K]}(\mathbb{R}^{[K]})) = u_{[K]}(\mathbb{R}^{[K]})$, $K \in \mathscr{J}(\mathscr{K})$.*

PROOF. First assume that $\rho(g)$ is $\mathscr{K}$-preserving. Then $\rho(g)' = \rho(g)^{-1} = \rho(g^{-1})$ is also $\mathscr{K}$-preserving. It then follows that $\rho(g)$ is block-diagonal w.r.t. the decomposition (2.5) [see AP (1993), Remark 2.1]. This establishes the "only if" claim. The converse is a consequence of (2.5). □

From Lemma 2.1, it follows that for $K \in \mathscr{K}$, the mapping $\rho_{[K]}: G \to \mathrm{O}([K])$ given by $\rho_{[K]}(g) = \rho(g)_{[K]}$, $g \in G$, is a well-defined orthogonal group representation of G on $\mathbb{R}^{[K]}$. Thus for $g \in G$, $\rho_K(g) = \mathrm{diag}(\rho_{[K']}(g) | K' \in \mathscr{I}(\mathscr{K}),\ K' \subseteq K)$, $K \in \mathscr{K}$. In particular, for $K \in \mathscr{I}(\mathscr{K})$,

$$\rho_K(g) = \mathrm{diag}\big(\rho_{\langle K \rangle}(g), \rho_{[K]}(g)\big). \tag{2.12}$$

For $K \in \mathscr{I}(\mathscr{K})$, denote by $\mathrm{M}_G([K] \times \langle K \rangle)$ the vector space of all $[K] \times \langle K \rangle$ matrices $R_{[K\rangle}$ that commute with G, that is, $\rho(g)_{[K]} R_{[K\rangle} = R_{[K\rangle} \rho(g)_{\langle K \rangle}$, $g \in G$. The following theorem is a generalization of Theorem 2.2 in AP (1993).

THEOREM 2.1. *The mapping*

$$\begin{aligned} \mathrm{P}_{G,\mathscr{K}}(I) &\to \times\big(\mathrm{M}_G([K] \times \langle K \rangle) \times \mathrm{P}_G([K]) | K \in \mathscr{I}(\mathscr{K})\big), \\ \Sigma &\mapsto \Big(\big(\Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1}, \Sigma_{[K]\cdot}\big)\Big| K \in \mathscr{I}(\mathscr{K})\Big), \end{aligned} \tag{2.13}$$

is well defined and bijective.

PROOF. Consider any $\Sigma \in \mathrm{P}_{\mathscr{K}}(I)$ and $g \in G$. From (2.12) it follows, that

$$(\rho(g)\Sigma\rho(g)')_{[K\rangle}(\rho(g)\Sigma\rho(g)')_{\langle K \rangle}^{-1} = \rho_{[K]}(g)\big(\Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1}\big)\rho_{\langle K \rangle}(g)^{-1}$$

and

$$(\rho(g)\Sigma\rho(g)')_{[K]\cdot} = \rho_{[K]}(g)\Sigma_{[K]\cdot}\rho_{[K]}(g)'.$$

On the other hand, by Theorem 2.2 and Proposition 2.3 of AP (1993), it follows that $\Sigma \in \mathrm{P}_G(I)$ if and only if

$$(\rho(g)\Sigma\rho(g)')_{[K\rangle}(\rho(g)\Sigma\rho(g)')_{\langle K \rangle}^{-1} = \Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1}$$

and

$$(\rho(g)\Sigma\rho(g)')_{[K]\cdot} = \Sigma_{[K]\cdot},$$

for all $K \in \mathscr{I}(\mathscr{K})$ and $g \in G$. Thus $\Sigma \in \mathrm{P}_G(I)$ if and only if $\Sigma_{[K\rangle}\Sigma_{\langle K \rangle}^{-1} \in \mathrm{M}_G([K] \times \langle K \rangle)$ and $\Sigma_{[K]\cdot} \in \mathrm{P}_G([K])$ for every $K \in \mathscr{I}(\mathscr{K})$. □

Now we discuss some invariance properties of $\mathrm{P}_{G,\mathscr{K}}(I)$ [compare to AP (1993), Section 2.4]. Let $\mathrm{M}_{G,\mathscr{K}}(I)$ denote the algebra $\mathrm{M}_{\mathscr{K}}(I) \cap \mathrm{M}_G(I)$ and $\mathrm{GL}_{G,\mathscr{K}}(I)$ the group of nonsingular elements in $\mathrm{M}_{G,\mathscr{K}}(I)$. Note that $\mathrm{GL}_{G,\mathscr{K}}(I) = \mathrm{GL}_{\mathscr{K}}(I) \cap \mathrm{GL}_G(I)$. The following lemma generalizes (2.19) in AP (1993).

LEMMA 2.2. *The mapping*

$$\begin{aligned} \mathrm{M}_{G,\mathscr{K}}(I) &\to \times\big(\mathrm{M}_G([K] \times \langle K \rangle) \times \mathrm{M}_G([K]) | K \in \mathscr{I}(\mathscr{K})\big), \\ A &\mapsto \Big(\big(A_{[K\rangle}, A_{[K]}\big)\Big| K \in \mathscr{I}(\mathscr{K})\Big), \end{aligned} \tag{2.14}$$

is well defined and bijective.

PROOF. The proof is similar to that of Theorem 2.1. Consider any $A \in \mathrm{M}_{\mathscr{K}}(I)$. For $K \in \mathscr{I}(\mathscr{K})$ and $g \in G$, it follows from (2.12) that

$$\left(\rho(g) A \rho(g)^{-1}\right)_{[K\rangle} = \rho_{[K]}(g) A_{[K\rangle} \rho_{\langle K\rangle}(g)^{-1}$$

and

$$\left(\rho(g) A \rho(g)^{-1}\right)_{[K]} = \rho_{[K]}(g) A_{[K]} \rho_{[K]}(g)^{-1}.$$

On the other hand, by (2.19) in AP (1993) together with the fact that $\rho(g) \in \mathrm{M}_{\mathscr{K}}(I)$, $g \in G$, it follows that $A \in \mathrm{M}_G(I)$ if and only if $(\rho(g) A \rho(g)^{-1})_{[K\rangle} = A_{[K\rangle}$ and $(\rho(g) A \rho(g)^{-1})_{[K]} = A_{[K]}$, for every $K \in \mathscr{I}(\mathscr{K})$. Thus $A \in \mathrm{M}_G(I)$ if and only if $A_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$ and $A_{[K]} \in \mathrm{M}_G([K])$ for every $K \in \mathscr{I}(\mathscr{K})$. □

REMARK 2.1. Under the correspondence (2.14), the subset $\mathrm{GL}_{G,\mathscr{K}}(I)$ corresponds to the subset

$$\times\left(\mathrm{M}_G([K] \times \langle K\rangle) \times \mathrm{GL}_G([K]) | K \in \mathscr{I}(\mathscr{K})\right).$$

The following lemma is a generalization of Lemma 2.4 in AP (1993).

LEMMA 2.3. *For any element*

$$\left((R_{[K\rangle}, \Lambda_{[K]}) \middle| K \in \mathscr{I}(\mathscr{K})\right) \in \times\left(\mathrm{M}_G([K] \times \langle K\rangle) \times \mathrm{P}_G([K]) | K \in \mathscr{I}(\mathscr{K})\right),$$

there exists a matrix $A \in \mathrm{GL}_{G,\mathscr{K}}(I)$ *such that for every* $K \in \mathscr{I}(\mathscr{K})$,

$$R_{[K\rangle} = A_{[K\rangle} A_{\langle K\rangle}^{-1}, \qquad \Lambda_{[K]} = A_{[K]} A'_{[K]}.$$

PROOF. We shall use induction on $q = |\mathscr{I}(\mathscr{K})|$. If $q = 1$ the assertion follows from the fact that $\mathrm{GL}_G(I)$ acts transitively on $\mathrm{P}_G(I)$ (see Proposition A.1).

Next, assume that the theorem holds for every ring of subsets $\mathscr{L}$ (of any index set) where $|\mathscr{I}(\mathscr{L})| < q$. Furthermore, let K denote a maximal element in $\mathscr{I}(\mathscr{K})$, and define the set $I_K = \cup(L \in \mathscr{I}(\mathscr{K}) | L \neq K)$. Finally, denote by $\mathscr{L}$ the sublattice $\{I_K \cap L | L \in \mathscr{K}\}$ of $\mathscr{K}$. Since K is maximal, it follows that $\mathscr{I}(\mathscr{L}) = \mathscr{I}(\mathscr{K}) \setminus \{K\}$, and hence, by assumption there exists a matrix $B \in \mathrm{GL}_{G,\mathscr{L}}(I_K)$ such that for every $L \in \mathscr{I}(\mathscr{K}) \setminus \{K\}$,

$$R_{[L\rangle} = B_{[L\rangle} B_{\langle L\rangle}^{-1}, \qquad \Lambda_{[L]} = B_{[L]} B'_{[L]}.$$

Furthermore, because $\mathrm{GL}_G([K])$ acts transitively on $\mathrm{P}_G([K])$, there exists a matrix $A_{[K]} \in \mathrm{GL}_G([K])$ such that $\Lambda_{[K]} = A_{[K]} A'_{[K]}$. Since $\langle K\rangle \subseteq I_K$, we can define

$$A_{[K\rangle} = R_{[K\rangle} B_{\langle K\rangle},$$

and it is straightforward to verify that $A_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$. By Remark 2.1, the family $((B_{[L\rangle}, B_{[L]}) | L \in \mathscr{I}(\mathscr{L}))$ together with $(A_{[K\rangle}, A_{[K]})$ uniquely determines a matrix $A \in \mathrm{GL}_{G,\mathscr{K}}(I)$, which satisfies the required conditions. □

The following theorem is a generalization of Theorem 2.3 of AP (1993).

THEOREM 2.2. *The action*

$$\text{(2.15)}\qquad \begin{aligned} \mathrm{GL}_{G,\mathscr{K}}(I) \times \mathrm{P}_{G,\mathscr{K}}(I) &\to \mathrm{P}_{G,\mathscr{K}}(I),\\ (A,\Sigma) &\mapsto A\Sigma A', \end{aligned}$$

is well defined, *transitive and proper*.

PROOF. Obviously, the action is well defined since $\mathrm{GL}_{\mathscr{K}}(I)$ acts on $\mathrm{P}_{\mathscr{K}}(I)$ [see AP (1993), Theorem 2.3] and $\mathrm{GL}_G(I)$ acts on $\mathrm{P}_G(I)$ (see Appendix A). The one-to-one correspondence (2.13) commutes with the action (2.15) and with the action given by the restriction of the action (2.26) in AP (1993) to the subset $\mathrm{GL}_{G,\mathscr{K}}(I) \times (\times(\mathrm{M}_G([K] \times \langle K\rangle) \times \mathrm{P}_G([K])|K \in \mathscr{I}(\mathscr{K})))$. Therefore, by Lemma 2.3, the action is transitive. It is proper since $\mathrm{GL}_{G,\mathscr{K}}(I)$ and $\mathrm{P}_{G,\mathscr{K}}(I)$ are closed subsets of $\mathrm{GL}(I)$ and $\mathrm{P}(I)$, respectively. □

The model (2.11) is invariant under the action of $\mathrm{GL}_{G,\mathscr{K}}(I)$ on the observation space $\mathbb{R}^I$ given by the restriction of the action (2.2) and the transitive action of $\mathrm{GL}_{G,\mathscr{K}}(I)$ on the parameter space $\mathrm{P}_{G,\mathscr{K}}(I)$.

3. Likelihood inference. Consider the GS-LCI model

$$\text{(3.1)}\qquad (\mathrm{N}(\Sigma)|\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I))$$

(cf. Section 2.3). Since $\mathrm{P}_{G,\mathscr{K}}(I) \subseteq \mathrm{P}_{\mathscr{K}}(I)$ it follows that the likelihood function $\mathbf{L}$: $\mathrm{P}_{G,\mathscr{K}}(I) \times \mathbb{R}^I \to]0,\infty[$ for the model (3.1) has the following factorization:

$$\text{(3.2)}\qquad \begin{aligned} \mathbf{L}(\Sigma, x) &= \det(\Sigma)^{-1/2}\exp\left(-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1}xx')\right)\\ &= \prod\Big(\det(\Sigma_{[K]\cdot})^{-1/2}\exp\Big(-\tfrac{1}{2}\mathrm{tr}\Big(\Sigma_{[K]\cdot}^{-1}\big(x_{[K]} - \Sigma_{[K\rangle}\Sigma_{\langle K\rangle}^{-1}x_{\langle K\rangle}\big)\\ &\qquad\qquad \times(\cdots)'\Big)\Big)\Big| K \in \mathscr{I}(\mathscr{K})\Big). \end{aligned}$$

Now consider the problem of existence and uniqueness of the maximum likelihood estimator $\hat{\Sigma}(x)$ of $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ based on an observation $x \in \mathbb{R}^I$. Because of (3.2) and the factorization of the parameter space $\mathrm{P}_{G,\mathscr{K}}(I)$ given in Theorem 2.1, it suffices for each $K \in \mathscr{I}(\mathscr{K})$ to consider the problem of maximizing

$$\text{(3.3)}\qquad \det(\Lambda_{[K]})^{-1/2}\exp\left(-\tfrac{1}{2}\mathrm{tr}\left(\Lambda_{[K]}^{-1}(x_{[K]} - R_{[K\rangle}x_{\langle K\rangle})(\cdots)'\right)\right)$$

for $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$ and $\Lambda_{[K]} \in \mathrm{P}_G([K])$. The subspace

$$\text{(3.4)}\qquad L_{[K]}(x_{\langle K\rangle}) = \left\{R_{[K\rangle}x_{\langle K\rangle}\middle| R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)\right\}$$

of $\mathbb{R}^{[K]}$ is preserved by $\mathrm{M}_G([K])$, that is, $\mathrm{M}_G([K])L_{[K]}(x_{\langle K\rangle}) = L_{[K]}(x_{\langle K\rangle})$. Thus, except for the parametrization of $L_{[K]}(x_{\langle K\rangle})$ by the regression parameter $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$, (3.3) is the likelihood function for a linear group symmetry (LGS) model determined by $L_{[K]}(x_{\langle K\rangle})$ and G, as defined in

Section A.8 with the sample space $\mathbb{R}^I$ replaced by $\mathbb{R}^{[K]}$. Let $P_{[K]} \in \mathrm{M}([K])$ be the unique orthogonal projection matrix of $\mathbb{R}^{[K]}$ onto $L_{[K]}(x_{\langle K \rangle})$ w.r.t. all $\Lambda_{[K]} \in \mathrm{P}_G([K])$; compare Lemma A.5. Thus for arbitrary $\Lambda_{[K]} \in \mathrm{P}_G([K])$, (3.3) is maximized for any element $\hat{R}_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K \rangle)$ that satisfies the equation

$$\hat{R}_{[K\rangle} x_{\langle K \rangle} = P_{[K]} x_{[K]}. \tag{3.5}$$

It follows from (2.12) that

$$\psi(S)_{\langle K \rangle} = \frac{1}{|G|} \sum \left(\rho_{\langle K \rangle}(g) S_{\langle K \rangle} \rho_{\langle K \rangle}(g)' | g \in G \right) = \psi_{\langle K \rangle}(S_{\langle K \rangle}),$$

$$\psi(S)_{[K]} = \frac{1}{|G|} \sum \left(\rho_{[K]}(g) S_{[K]} \rho_{[K]}(g)' | g \in G \right) = \psi_{[K]}(S_{[K]}),$$

$$\psi(S)_{[K\rangle} = \frac{1}{|G|} \sum \left(\rho_{[K]}(g) S_{[K\rangle} \rho_{\langle K \rangle}(g)' | g \in G \right)$$

and $\psi(S)_{\langle K]} = \psi(S)'_{[K\rangle}$, where $S = xx'$ is the overall empirical covariance matrix.

Since $1_{[K]} \in \mathrm{P}_G([K])$ and $\hat{R}_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K \rangle)$, it follows that for every $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K \rangle)$ and $g \in G$,

$$\begin{aligned}
0 &= (R_{[K\rangle} x_{\langle K \rangle})'(x_{[K]} - P_{[K]} x_{[K]}) \\
&= (R_{[K\rangle} x_{\langle K \rangle})'\left(x_{[K]} - \hat{R}_{[K\rangle} x_{\langle K \rangle}\right) \\
&= \mathrm{tr}\left(R'_{[K\rangle}\left(x_{[K]} - \hat{R}_{[K\rangle} x_{\langle K \rangle}\right) x'_{\langle K \rangle}\right) \\
&= \mathrm{tr}\left(R'_{[K\rangle}\left(S_{[K\rangle} - \hat{R}_{[K\rangle} S_{\langle K \rangle}\right)\right) \\
&= \mathrm{tr}\left(R'_{[K\rangle} \rho_{[K]}(g)' \rho_{[K]}(g) \left(S_{[K\rangle} - \hat{R}_{[K\rangle} S_{\langle K \rangle}\right)\right) \\
&= \mathrm{tr}\left(\rho_{\langle K \rangle}(g)' R'_{[K\rangle}\left(\rho_{[K]}(g) S_{[K\rangle} - \rho_{[K]}(g) \hat{R}_{[K\rangle} S_{\langle K \rangle}\right)\right) \\
&= \mathrm{tr}\left(R'_{[K\rangle}\left(\rho_{[K]}(g) S_{[K\rangle} \rho_{\langle K \rangle}(g)' - \hat{R}_{[K\rangle} \rho_{\langle K \rangle}(g) S_{\langle K \rangle} \rho_{\langle K \rangle}(g)'\right)\right).
\end{aligned}$$

Thus for every $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K \rangle)$,

$$\mathrm{tr}\left(R'_{[K\rangle}\left(\psi(S)_{[K\rangle} - \hat{R}_{[K\rangle} \psi_{\langle K \rangle}(S_{\langle K \rangle})\right)\right) = 0.$$

Because $\psi(S)_{[K\rangle} - \hat{R}_{[K\rangle} \psi_{\langle K \rangle}(S_{\langle K \rangle}) \in \mathrm{M}_G([K] \times \langle K \rangle)$, it then follows that

$$\psi(S)_{[K\rangle} = \hat{R}_{[K\rangle} \psi_{\langle K \rangle}(S_{\langle K \rangle}). \tag{3.6}$$

Conversely it follows that if $\hat{R}_{[K\rangle}$ is a solution to (3.6), then it will satisfy (3.5).

Now define

$$\psi(S)_{[K]^\circ} = \psi(S)_{[K]} - \hat{R}_{[K\rangle} \psi(S)_{\langle K]}. \tag{3.7}$$

From (3.5) and the fact that $\hat{R}_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K \rangle)$, it follows that

$$\begin{aligned}\psi_{[K]}&\big((x_{[K]} - P_{[K]}x_{[K]})(x_{[K]} - P_{[K]}x_{[K]})'\big)\\ &= \psi_{[K]}\Big(\Big(x_{[K]} - \hat{R}_{[K\rangle}x_{\langle K\rangle}\Big)\Big(x_{[K]} - \hat{R}_{[K\rangle}x_{\langle K\rangle}\Big)'\Big)\\ &= \psi(S)_{[K]} - \hat{R}_{[K\rangle}\psi(S)_{\langle K]} = \psi(S)_{[K]\circ}\end{aligned}$$

and thus $\psi(S)_{[K]\circ}$ does not depend on the solution $\hat{R}_{[K\rangle}$ to (3.6). It now follows from Theorem A.2 that the ML estimator $\hat{\Lambda}_{[K]}$ for $\Lambda_{[K]} \in \mathrm{P}_G([K])$ exists if and only if $\psi(S)_{[K]\circ}$ is nonsingular. In this case it is unique and given by

$$\hat{\Lambda}_{[K]} = \psi(S)_{[K]\circ}. \tag{3.8}$$

The maximum of the likelihood function (3.3) is then

$$\det\Big(\hat{\Lambda}_{[K]}\Big)^{-1/2}\exp\big(-\tfrac{1}{2}|[K]|\big). \tag{3.9}$$

We are now able to state the following theorem regarding ML estimation in the GS-LCI model given by G and $\mathscr{K}$ based on the observation $x \in \mathbb{R}^I$.

THEOREM 3.1. *In the model* (2.11), *the maximum likelihood estimator* $\hat{\Sigma} = \hat{\Sigma}(x)$ *of* $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ *for the observation* $x \in \mathbb{R}^I$ *exists if and only if the matrices* $\psi(xx')_{[K]\circ}$, $K \in \mathscr{I}(\mathscr{K})$, *all are positive definite.*

In this case, there is a one-to-one correspondence between all families $(\hat{R}_{[K\rangle} | K \in \mathscr{I}(\mathscr{K}))$ *of solutions to the equations*

$$\psi(xx')_{[K\rangle} = R_{[K\rangle}\psi(xx')_{\langle K\rangle}, \tag{3.10}$$

where $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$, $K \in \mathscr{I}(\mathscr{K})$, *and all ML estimators* $\hat{\Sigma}$ *for* $x \in \mathbb{R}^I$, *given by*

$$\hat{\Sigma}_{[K\rangle}\hat{\Sigma}^{-1}_{\langle K\rangle} = \hat{R}_{[K\rangle}, \qquad \hat{\Sigma}_{[K]\cdot} = \psi(xx')_{[K]\circ}, \tag{3.11}$$

$K \in \mathscr{I}(\mathscr{K})$ (*cf. Theorem* 2.1).

The maximum likelihood estimator $\hat{\Sigma}$ *is unique if and only if the equations*

$$R_{[K\rangle}\psi(xx')_{\langle K\rangle} = 0,$$

where $R_{[K\rangle} \in M_G([K] \times \langle K\rangle)$, $K \in \mathscr{I}(\mathscr{K})$, *only have the solutions* $R_{[K\rangle} = 0$, $K \in \mathscr{I}(\mathscr{K})$.

The proof follows from the above considerations and Remark 3.1.

REMARK 3.1. For any $K \in \mathscr{I}(\mathscr{K})$, the following statements are equivalent:

(a) The equation $\psi(S)_{[K\rangle} = R_{[K\rangle}\psi_{\langle K\rangle}(S_{\langle K\rangle})$ has a unique solution in $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$.

(b) The equation $R_{[K\rangle}\psi_{\langle K\rangle}(S_{\langle K\rangle}) = 0$ in $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$ only has the solution $R_{[K\rangle} = 0$.

(c) The parametrization mapping $R_{[K\rangle} \mapsto R_{[K\rangle}x_{\langle K\rangle}$ of $L_{[K]}(x_{\langle K\rangle})$ by $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$ is one-to-one.

REMARK 3.2. It follows from (3.2), (3.9) and (2.5), that the maximum of the likelihood function is

$$\det(\hat{\Sigma})^{-1/2}\exp\left(-\tfrac{1}{2}|I|\right). \tag{3.12}$$

REMARK 3.3. The explicit expression for $\hat{\Sigma}$ may be obtained from (3.11) by means of the reconstruction algorithm given in AP (1993), Section 2.7.

COROLLARY 3.1. *In the model* (2.11), *the maximum likelihood estimator* $\hat{\Sigma} = \hat{\Sigma}(x)$ *of* $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ *for the observation* $x \in \mathbb{R}^I$ *exists and is unique if the matrices* $\psi(xx')_K$, $K \in \mathscr{I}(\mathscr{K})$, *all are positive definite.*

In this case, $\hat{\Sigma}$ *is determined by*

$$\hat{\Sigma}_{[K\rangle}\hat{\Sigma}^{-1}_{\langle K\rangle} = \psi(xx')_{[K\rangle}\psi(xx')^{-1}_{\langle K\rangle}, \qquad \hat{\Sigma}_{[K]\cdot} = \psi(xx')_{[K]\cdot}, \tag{3.13}$$

$K \in \mathscr{I}(\mathscr{K})$ (*cf. Theorem* 2.1).

PROOF. Let $K \in \mathscr{I}(\mathscr{K})$. If $\psi(xx')_K$ is positive definite, then $\psi(xx')_{\langle K\rangle}$ is positive definite and thus the equation $R_{[K\rangle}\psi(xx')_{\langle K\rangle} = 0$ implies $R_{[K\rangle} = 0$. Furthermore $\psi(xx')_{[K]^\circ} = \psi(xx')_{[K]\cdot}$ is positive definite. □

REMARK 3.4. Note that the condition for existence and uniqueness of the ML estimator in Corollary 3.1 is equivalent to the condition that the matrices $\psi(xx')_K$ are nonsingular for all maximum elements K in the partially ordered set $\mathscr{I}(\mathscr{K})$.

We shall now establish a result for GS-LCI models which is similar to the result in Proposition A.2 for GS models. First we need to generalize the Lemmas A.2 and A.4. For $\Sigma \in \mathrm{PS}(I)$ the subspace

$$N_\Sigma = \{x \in \mathbb{R}^I \mid x'\Sigma x = 0\}$$

of $\mathbb{R}^I$ is called the *null-space* for Σ. Note that for any $B \in \mathrm{GL}(I)$, $N_{B'\Sigma B} = B^{-1}N_\Sigma$. As in Section A.2, let $\mathscr{A}_I(G)$ denote the subalgebra of $\mathrm{M}(I)$ generated by $\rho(G)$.

LEMMA 3.1. *Let* $N \subseteq \mathbb{R}^I$ *be a G-subspace. Then for* $x \in \mathbb{R}^I$, $\mathscr{A}_I(G)x \supseteq N^\perp$ *if and only if* $N_{\psi(xx')} \subseteq N$.

PROOF. Let $x \in \mathbb{R}^I$. First assume that $\mathscr{A}_I(G)x \supseteq N^\perp$ and let $z \in N_{\psi(xx')}$. From the proof of Lemma A.2 it then follows that $z'\rho(g)x = 0$ for all $g \in G$. This implies that z is orthogonal to $\mathscr{A}_I(G)x$ and therefore $z \in N$ by the assumption.

On the other hand, assume that $N_{\psi(xx')} \subseteq N$ and that $\mathscr{A}_I(G)x \not\supseteq N^\perp$. This implies that there exists a $z \in (\mathscr{A}_I(G)x)^\perp \setminus N$. But from the proof of Lemma A.2 it then follows that $z'\psi(xx')z = 0$; that is, $z \in N_{\psi(xx')}$, which is a contradiction. □

LEMMA 3.2. *Let $N \subseteq \mathbb{R}^I$ be a G-subspace and an $\mathrm{M}_G(I)$-subspace, that is, $\rho(G)N = N$ and $\mathrm{M}_G(I)N = N$, respectively. Then we have the following.*

(a) *The orthogonal complement $N^{\perp}$ to N (cf. Section 2.1) is a G-subspace and an $\mathrm{M}_G(I)$-subspace.*

(b) *For any $x \in \mathbb{R}^I$, $\mathscr{A}_I(G)x \supseteq N$ if and only if $\mathscr{A}_I(G)P_N x = N$, where $P_N \in \mathrm{M}(I)$ is the orthogonal projection matrix of $\mathbb{R}^I$ onto N (cf. Section 2.1). The latter condition states that $P_N x$ is a regular element in N w.r.t. G (cf. Section A.2).*

(c) *The set*

$$\Omega_{I,N} = \left\{x \in \mathbb{R}^I | \mathscr{A}_I(G)x \supseteq N\right\}$$

is open, and if it is nonempty, then the Lebesgue-measure is concentrated on $\Omega_{I,N}$; that is, $\mathbb{R}^I \setminus \Omega_{I,N}$ has Lebesgue-measure zero.

PROOF. It is easy to see that $N^{\perp}$ is both a G-subspace and an $\mathrm{M}_G(I)$-subspace. To show (b), observe that $P_N \in \mathrm{M}_G(I)$ and that P_N commutes with all matrices in $\mathrm{M}_G(I)$. From the bicommutant theorem [cf., e.g., Bourbaki (1958), Section 4, Number 2, Corollary 1] it then follows that $P_N \in \mathscr{A}_I(G)$. Now, if $\mathscr{A}_I(G)x \supseteq N$ then $\mathscr{A}_I(G)P_N x = P_N \mathscr{A}_I(G)x \supseteq P_N N = N$, since $P_N \in \mathrm{M}_G(I)$. On the other hand, if $\mathscr{A}_I(G)P_N x = N$, then $\mathscr{A}_I(G)x \supseteq \mathscr{A}_I(G)P_N x = N$, since $P_N \in \mathscr{A}_I(G)$.

To show (c), note that $\Omega_{I,N} = P_N^{-1}(\Omega_N)$, where Ω_N is the set of regular points in N w.r.t. G. The assertion now follows from Lemma A.4 with Ω and $\mathbb{R}^I$ replaced by Ω_N and N, respectively. □

Next, for $K \in \mathscr{K}$, let $\mathscr{A}_K(G)$ denote the subalgebra of $\mathrm{M}(K)$ generated by $\rho_K(G)$. When G is subsumed, we denote $\mathscr{A}_K(G)$ by $\mathscr{A}_K$. For $K \in \mathscr{I}(\mathscr{K})$ we define

$$N_{G,\langle K\rangle} = \left\{z_{\langle K\rangle} \in \mathbb{R}^{\langle K\rangle} | \forall\, R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle) \colon R_{[K\rangle} z_{\langle K\rangle} = 0\right\}$$

and

$$\Omega_{G,\langle K\rangle} = \left\{x_{\langle K\rangle} \in \mathbb{R}^{\langle K\rangle} \mid \mathscr{A}_{\langle K\rangle} x_{\langle K\rangle} \supseteq E_{G,\langle K\rangle}\right\}, \tag{3.14}$$

where $E_{G,\langle K\rangle} = (N_{G,\langle K\rangle})^{\perp}$ is the orthogonal complement to $N_{G,\langle K\rangle}$ in $\mathbb{R}^{\langle K\rangle}$. When G is subsumed we denote $N_{G,\langle K\rangle}$, $\Omega_{G,\langle K\rangle}$ and $E_{G,\langle K\rangle}$ by $N_{\langle K\rangle}$, $\Omega_{\langle K\rangle}$ and $E_{\langle K\rangle}$, respectively. It is easily verified that $N_{\langle K\rangle}$ is a G-subspace and an $\mathrm{M}_G(\langle K\rangle)$-subspace of $\mathbb{R}^{\langle K\rangle}$. Thus by Lemma 3.2, $E_{\langle K\rangle}$ is also a G-subspace and an $\mathrm{M}_G(\langle K\rangle)$-subspace of $\mathbb{R}^{\langle K\rangle}$.

LEMMA 3.3. *Let $K \in \mathscr{I}(\mathscr{K})$. For any $x_{\langle K\rangle} \in \mathbb{R}^{\langle K\rangle}$, the parametrization mapping $R_{[K\rangle} \mapsto R_{[K\rangle} x_{\langle K\rangle}$ of $L_{[K]}(x_{\langle K\rangle})$ by $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$ is one-to-one if and only if $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$.*

PROOF. Let $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$ and assume that $R_{[K\rangle} x_{\langle K\rangle} = 0$. Then

$$R_{[K\rangle} \mathscr{A}_{\langle K\rangle} x_{\langle K\rangle} = \mathscr{A}_{[K]} R_{[K\rangle} x_{\langle K\rangle} = \{0\}.$$

In particular $R_{[K\rangle} E_{\langle K\rangle} = \{0\}$. Thus for any $z_{\langle K\rangle} \in \mathbb{R}^{\langle K\rangle}$,

$$0 = R_{[K\rangle} P_{E_{\langle K\rangle}} z_{\langle K\rangle} = R_{[K\rangle} z_{\langle K\rangle} - R_{[K\rangle} P_{N_{\langle K\rangle}} z_{\langle K\rangle} = R_{[K\rangle} z_{\langle K\rangle},$$

and it follows that $R_{[K\rangle} = 0$. Here $P_{N_{\langle K\rangle}}$ and $P_{E_{\langle K\rangle}}$ denote the orthogonal projection matrices of $\mathbb{R}^{\langle K\rangle}$ onto $N_{\langle K\rangle}$ and $E_{\langle K\rangle}$, respectively (cf. Section 2.1).

On the other hand, assume that $x_{\langle K\rangle} \notin \Omega_{\langle K\rangle}$, that is, $\mathscr{A}_{\langle K\rangle} x_{\langle K\rangle} \cap E_{\langle K\rangle} \subset E_{\langle K\rangle}$. Then $U_{\langle K\rangle} = (\mathscr{A}_{\langle K\rangle} x_{\langle K\rangle})^{\perp} \cap E_{\langle K\rangle} \neq \{0\}$. Furthermore, $U_{\langle K\rangle}$ is a G-subspace orthogonal to $N_{\langle K\rangle}$ and $x_{\langle K\rangle} \in U_{\langle K\rangle}^{\perp}$. Let $y_{\langle K\rangle} \in U_{\langle K\rangle} \setminus \{0\}$. Since $y_{\langle K\rangle} \notin N_{\langle K\rangle}$, there exists $\tilde{R}_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$ such that $\tilde{R}_{[K\rangle} y_{\langle K\rangle} \neq 0$. Now let $R_{[K\rangle}$ be the matrix for the linear mapping $\mathbb{R}^{\langle K\rangle} \to \mathbb{R}^{[K]}$ defined by $R_{[K\rangle} z_{\langle K\rangle} = \tilde{R}_{[K\rangle} z_{\langle K\rangle}$ if $z_{\langle K\rangle} \in U_{\langle K\rangle}$ and $R_{[K\rangle} z_{\langle K\rangle} = 0$ if $z_{\langle K\rangle} \in U_{\langle K\rangle}^{\perp}$. Then $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$, $R_{[K\rangle} \neq 0$ and $R_{[K\rangle} x_{\langle K\rangle} = 0$. This shows that the parametrization mapping is not one-to-one. □

Now for $K \in \mathscr{I}(\mathscr{K})$ define

$$N_{G,K} = \{(x_{\langle K\rangle}, 0) \in \mathbb{R}^K | x_{\langle K\rangle} \in N_{\langle K\rangle}\}, \tag{3.15}$$

and

$$\Omega_{G,K} = \{x_K \in \mathbb{R}^K \,|\, \mathscr{A}_K x_K \supseteq E_K\},$$

where $E_{G,K} = (N_{G,K})^{\perp}$ is the orthogonal complement to $N_{G,K}$ in $\mathbb{R}^K$. When G is subsumed, we denote $N_{G,K}$, $\Omega_{G,K}$ and $E_{G,K}$ by N_K, Ω_K and E_K, respectively.

LEMMA 3.4. *Let $K \in \mathscr{I}(\mathscr{K})$ and $x_K \in \mathbb{R}^K$. Then $x_K \in \Omega_K$ if and only if $\psi_K(x_K x_K')_{[K]\circ}$ is nonsingular and the equation*

$$R_{[K\rangle} \psi_{\langle K\rangle}(x_{\langle K\rangle} x_{\langle K\rangle}') = 0, \tag{3.16}$$

only has the solution $R_{[K\rangle} = 0$ for $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$.

PROOF. Define the $K \times K$ matrix

$$B_K = \begin{pmatrix} 1_{\langle K\rangle} & 0 \\ \hat{R}_{[K\rangle} & 1_{[K]} \end{pmatrix},$$

where $\hat{R}_{[K\rangle}$ is any solution in $\mathrm{M}_G([K] \times \langle K\rangle)$ to the equation $\psi_K(x_K x_K')_{[K\rangle} = R_{[K\rangle} \psi_{\langle K\rangle}(x_{\langle K\rangle} x_{\langle K\rangle}')$. Then $B_K \in \mathrm{M}_G(K)$, B_K is nonsingular and

$$B_K \psi_K(x_K x_K')_{K\circ} B_K' = \psi_K(x_K x_K'), \tag{3.17}$$

where

$$\psi_K(x_K x_K')_{K\circ} = \begin{pmatrix} \psi_{\langle K\rangle}(x_{\langle K\rangle} x_{\langle K\rangle}') & 0 \\ 0 & \psi_K(x_K x_K')_{[K]\circ} \end{pmatrix}.$$

We now have that

$$\begin{aligned} x_K \in \Omega_K &\Leftrightarrow N_{\psi_K(x_K x_K')} \subseteq N_K \Leftrightarrow N_{\psi_K(x_K x_K')K^\circ} \subseteq N_K \\ &\Leftrightarrow \psi_K(x_K x_K')_{[K]^\circ} \text{ is nonsingular and } N_{\psi\langle K\rangle}(x_{\langle K\rangle} x'_{\langle K\rangle}) \subseteq N_{\langle K\rangle} \\ &\Leftrightarrow \psi_K(x_K x_K')_{[K]^\circ} \text{ is nonsingular and } \mathscr{A}_{\langle K\rangle} x_{\langle K\rangle} \supseteq E_{\langle K\rangle} \\ &\Leftrightarrow \psi_K(x_K x_K')_{[K]^\circ} \text{ is nonsingular and (3.16) only has the solution} \\ &\quad R_{[K\rangle} = 0, \end{aligned}$$

where the first biimplication follows from Lemma 3.1, the second from (3.17) and from the fact that $(B_K')^{-1} N_K = N_K$, the third from (3.15), the fourth from Lemma 3.1 and the fifth from Lemma 3.3 and from Remark 3.1. □

Now define

$$\Omega = \{x \in \mathbb{R}^I \mid \forall\ K \in \mathscr{I}(\mathscr{K}) \colon \mathscr{A}_K x_K \supseteq E_K\}. \tag{3.18}$$

PROPOSITION 3.1. *Let $x \in \mathbb{R}^I$. Then $x \in \Omega$ if and only if the maximum likelihood estimator $\hat{\Sigma}(x)$ of $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ in the model* (2.11) *exists and is unique.*

The proof follows from Lemma 3.4 and Theorem 3.1.

LEMMA 3.5. *The set Ω is an open subset of in $\mathbb{R}^I$. If $\Omega \neq \varnothing$, then the Lebesgue measure on $\mathbb{R}^I$ is concentrated on Ω, that is, $\mathbb{R}^I \setminus \Omega$ has Lebesgue measure zero.*

PROOF. By (3.18),

$$\Omega = \bigcap \left(p_K^{-1}(\Omega_K) | K \in \mathscr{I}(\mathscr{K}) \right).$$

We can assume that $\Omega \neq \varnothing$. It then follows from (c) in Lemma 3.2 that $p_K^{-1}(\Omega_K)$ is open and that the Lebesgue measure on $\mathbb{R}^I$ is concentrated on $p_K^{-1}(\Omega_K)$ for all $K \in \mathscr{I}(\mathscr{K})$. Hence the same holds for Ω. □

PROPOSITION 3.2. *The maximum likelihood estimator $\hat{\Sigma}$ of $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ in the model* (2.11) *exists and is unique with probability one w.r.t. all* $\mathrm{N}(\Sigma)$, *$\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$, if and only if $\Omega \neq \varnothing$.*

The proof follows from Lemma 3.5 and Proposition 3.1.

REMARK 3.5. Note that Proposition 3.1 and Proposition 3.2 imply that either the maximum likelihood estimator $\hat{\Sigma}(x)$ exists and is unique for almost all $x \in \mathbb{R}^I$ or else for any $x \in \mathbb{R}^I$ it will not exist or it will not be unique. The model (2.11) is called *regular* if the equivalent conditions in Proposition 3.2 hold.

REMARK 3.6. Consider the special case where the interplay between the representation of G on $\mathbb{R}^I$ and the lattice $\mathscr{K}$ yields that $N_{\langle K \rangle} = \{0\}$ for all $K \in \mathscr{I}(\mathscr{K})$. Then

$$\Omega_K = \left\{ x_K \in \mathbb{R}^K \mid \mathscr{A}_K x_K = \mathbb{R}^K \right\};$$

that is, Ω_K is the set of regular elements in $\mathbb{R}^K$ w.r.t. $G, K \in \mathscr{I}(\mathscr{K})$ (cf. Section A.2). From Lemma A.2 and Proposition 3.1 it then follows that in this case the condition in Corollary 3.1 is also *necessary*.

Let $((p_\mu^K, d_\mu, n_\mu) | \mu \in \mathfrak{M})$ be the structure constants for the representation ρ_K of G on $\mathbb{R}^K$, $K \in \mathscr{K}$ (cf. Section A.3). Let $K, K' \in \mathscr{K}$ with $K' \subseteq K$. Since $u_{K'}(\mathbb{R}^{K'}) \subseteq u_K(\mathbb{R}^K) \subseteq \mathbb{R}^I$ are G-subspaces, it follows that $p_\mu^{K'} \leq p_\mu^K$, $\mu \in \mathfrak{M}$, with equalities for all $\mu \in \mathfrak{M}$ if and only if $K = K'$. Furthermore, for $K \in \mathscr{I}(\mathscr{K})$, let $((p_\mu^{[K]}, d_\mu, n_\mu) \mid \mu \in \mathfrak{M})$ be the structure constants for the representation $\rho_{[K]}$ of G on $\mathbb{R}^{[K]}$. Let $K \in \mathscr{K}$. Since $u_K(\mathbb{R}^K)$ is the direct sum of the G-subspaces $u_{[K']}(\mathbb{R}^{[K']})$, $K' \in \mathscr{I}(\mathscr{K})$, $K' \subseteq K$, it follows that $p_\mu^K = \Sigma(p_\mu^{[K']} | K' \in \mathscr{I}(\mathscr{K}), K' \subseteq K)$, $\mu \in \mathfrak{M}$. In particular $p_\mu = \Sigma(p_\mu^{[K]} | K \in \mathscr{I}(\mathscr{K}))$, $\mu \in \mathfrak{M}$.

We shall call the family $(((p_\mu^{[K]} | K \in \mathscr{I}(\mathscr{K})), d_\mu, n_\mu) \mid \mu \in \mathfrak{M})$ *the structure constants given by* $\mathscr{K}$ *and* G for the model (3.1).

PROPOSITION 3.3. *The GS-LCI model* (3.1) *is regular if and only if*

$$\forall\ K \in \mathscr{I}(\mathscr{K}) \forall\ \mu \in \mathfrak{M}\colon\ p_\mu^{[K]} > 0 \Rightarrow n_\mu \geq p_\mu^K. \tag{3.19}$$

PROOF. For $K \in \mathscr{I}(\mathscr{K})$, let

$$\mathbb{R}^{\langle K \rangle} = \oplus \left(T_\mu^{\langle K \rangle} \mid \mu \in \mathfrak{M} \right)$$

and

$$\mathbb{R}^{[K]} = \oplus \left(T_\mu^{[K]} \mid \mu \in \mathfrak{M} \right)$$

be the unique decompositions of $\mathbb{R}^{\langle K \rangle}$ and $\mathbb{R}^{\langle K \rangle}$ into the orthogonal sums of their isotypic components w.r.t. G (cf. Section A.3). Since $N_{\langle K \rangle}$ and $E_{\langle K \rangle}$ are G-subspaces and $\mathrm{M}_G(\langle K \rangle)$-subspaces of $\mathbb{R}^{\langle K \rangle}$, they are both orthogonal sums of some of the isotypic components $T_\mu^{\langle K \rangle}$, $\mu \in \mathfrak{M}$, [cf., e.g., Bourbaki (1958), Section 3, Number 4, Proposition 11). It is easy to see that for $\mu \in \mathfrak{M}$, $T_\mu^{\langle K \rangle} \subseteq N_{\langle K \rangle}$ if and only if $T_\mu^{\langle K \rangle}$ and $T_\mu^{[K]}$ are disjoint, or equivalently, if $T_\mu^{\langle K \rangle} = 0$ or $T_\mu^{[K]} = 0$ (cf. Section A.3). From this it follows that

$$N_{\langle K \rangle} = \oplus \left(T_\mu^{\langle K \rangle} \mid \mu \in \mathfrak{M},\ p_\mu^{[K]} = 0 \right) \tag{3.20}$$

and

$$E_{\langle K \rangle} = \oplus \left(T_\mu^{\langle K \rangle} \mid \mu \in \mathfrak{M},\ p_\mu^{[K]} > 0 \right).$$

From the definition of N_K and E_K it then follows that

$$N_K = \oplus \left(T_\mu^K \mid \mu \in \mathfrak{M},\ p_\mu^{[K]} = 0 \right)$$

and

$$E_K = \oplus \left(T_\mu^K \mid \mu \in \mathfrak{M},\ p_\mu^{[K]} > 0 \right), \tag{3.21}$$

where

$$\mathbb{R}^K = \oplus\left(T_\mu^K \mid \mu \in \mathfrak{M}\right)$$

is the unique decomposition of $\mathbb{R}^K$ into the orthogonal sum of its isotypic components w.r.t. G. By Lemma 3.2, $\Omega_K \neq \emptyset$ if and only if the set of regular elements in E_K w.r.t. G is nonempty. The proposition now follows from (A.3) and (3.21). □

REMARK 3.7. It follows from (3.20) that $N_{\langle K \rangle} = 0$ if and only if for all $\mu \in \mathfrak{M}$, $p_\mu^{\langle K \rangle} > 0$ when $p_\mu^{[K]} > 0$. Thus in the special case where $N_{\langle K \rangle} = 0$ for all $K \in \mathscr{I}(\mathscr{K})$ (cf. Remark 3.6 and Corollary 3.1), (3.19) reduces to the condition that

$$n_\mu \geq p_\mu^K,$$

for all $\mu \in \mathfrak{M}$ and $K \in \mathscr{I}(\mathscr{K})$.

REMARK 3.8. For the model H_{GS} in Example 1.1, the family of structure constants becomes $(p_1, d_1, n_1) = (p_2, d_2, n_2) = (3, 1, 1)$, and the family of structure constants for the model $\mathrm{H}_{\mathrm{GS\text{-}LCI}}$ then simply becomes $(p_j^L, d_j, n_j) = (1, 1, 1)$, $L = \{a1, a2\}, \{b1, b2\}, \{c1, c2\}$, $j = 1, 2$. This family is essentially also the family of structure constants for the model $\mathrm{H}'_{\mathrm{GS\text{-}LCI}}$ in Example 1.2, but since $\mathscr{I}(\mathscr{K}) \neq \mathscr{I}(\mathscr{K}')$, the regularity conditions become different.

For the model H_{GS} in Example 1.3, the family of structure constants become $(p_{\mu_0}, d_{\mu_0}, n_{\mu_0}) = (p_a + p_b + p_c, 1, 1)$, $(p_{\mu_a}, d_{\mu_a}, n_{\mu_a}) = (p_a, 1, n_a - 1)$, $(p_{\mu_b}, d_{\mu_b}, n_{\mu_b}) = (p_b, 1, n_b - 1)$, and $(p_{\mu_c}, d_{\mu_c}, n_{\mu_c}) = (p_c, 1, n_c - 1)$, respectively. The family of structure constants for the model $\mathrm{H}_{\mathrm{GS\text{-}LCI}}$ is then given by $p_{\mu_0}^{I_a} = p_a$, $p_{\mu_0}^{I_b} = p_b$, $p_{\mu_0}^{I_c} = p_c$ and

$$p_{\mu_l}^{I_m} = \begin{cases} p_l, & l = m, \\ 0, & l \neq m, \end{cases}$$

for $l, m = a, b, c$. Similarly this family is essentially also the family of structure constants for the model $\mathrm{H}'_{\mathrm{GS\text{-}LCI}}$ in Example 1.4, but in this case too, the regularity conditions become different.

LEMMA 3.6. *Let $K \in \mathscr{I}(\mathscr{K})$, $x_{\langle K \rangle} \in \mathbb{R}^{\langle K \rangle}$, and let $(l_\mu(x_{\langle K \rangle}) \mid \mu \in \mathfrak{M})$ be the* $\mathrm{M}_G([K])$*-dimension of the* $\mathrm{M}_G([K])$*-subspace*

$$L_{[K]}(x_{\langle K \rangle}) = \left\{R_{[K\rangle} x_{\langle K \rangle} \mid R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K \rangle)\right\}$$

(*cf.* (3.4) *and Remark* A.6). *If $x_{\langle K \rangle} \in \Omega_{\langle K \rangle}$, then*

$$l_\mu(x_{\langle K \rangle}) = \begin{cases} p_\mu^{\langle K \rangle}, & p_\mu^{[K]} > 0, \\ 0, & p_\mu^{[K]} = 0, \end{cases}$$

$\mu \in \mathfrak{M}$.

PROOF. The vector spaces $\mathrm{M}_G([K] \times \langle K \rangle)$ and $L_{[K]}(x_{\langle K \rangle})$ are both $\mathrm{M}_G([K])$-modules under multiplication with the matrices in $\mathrm{M}_G([K])$ to the left, and the parametrization mapping

$$\begin{aligned} \mathrm{M}_G([K] \times \langle K\rangle) &\to L_{[K]}(x_{\langle K\rangle}), \\ R_{[K\rangle} &\mapsto R_{[K\rangle} x_{\langle K\rangle}, \end{aligned} \tag{3.22}$$

commutes with the module structures, that is,

$$(A_{[K]} R_{[K\rangle}) x_{\langle K\rangle} = A_{[K]}(R_{[K\rangle} x_{\langle K\rangle})$$

for all $A_{[K]} \in \mathrm{M}_G([K])$ and $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$. If $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$, then it follows from Lemma 3.3 that (3.22) is an isomorphism between the two $\mathrm{M}_G([K])$-modules, and hence they have the same $\mathrm{M}_G([K])$-dimension. In particular, $l_\mu = l_\mu(x_{\langle K\rangle})$ does not depend on $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$, $\mu \in \mathfrak{M}$.

Let $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$ and define $\mathfrak{M}^* = \{\mu \in \mathfrak{M} | p_\mu^{[K]} > 0\}$. It is easy to see that $l_\mu = 0$ when $p_\mu^{[K]} = 0$, $\mu \in \mathfrak{M}$, and thus we have that

$$\begin{aligned} &\forall\, \mu \in \mathfrak{M}^*\colon\ n_\mu \geq p_\mu^{[K]} + l_\mu \\ &\qquad \Leftrightarrow \forall\, \mu \in \mathfrak{M}\colon\ n_\mu \geq p_\mu^{[K]} + l_\mu. \end{aligned}$$

The latter condition is by Remark A.7 equivalent to the condition that there exists $x_{[K]} \in \mathbb{R}^{[K]}$ such that (3.3) with $x_K = (x_{\langle K\rangle}, x_{[K]})$ has a unique maximum for $(R_{[K]}, \Lambda_{[K]}) \in \mathrm{M}_G([K] \times \langle K\rangle) \times \mathrm{P}_G([K])$. By Lemma 3.4 and Theorem 3.1 this is equivalent to the condition that there exists $x_{[K]} \in \mathbb{R}^{[K]}$ such that $x_K = (x_{\langle K\rangle}, x_{[K]}) \in \Omega_K$. It then follows from Proposition 3.3 that

$$\begin{aligned} &\forall\, \mu \in \mathfrak{M}^*\colon\ n_\mu \geq p_\mu^{[K]} + l_\mu \\ &\qquad \Leftrightarrow \forall\, \mu \in \mathfrak{M}^*\colon\ n_\mu \geq p_\mu^{K}, \end{aligned}$$

and since $p_\mu^K = p_\mu^{\langle K\rangle} + p_\mu^{[K]}$, the lemma follows. □

Remark 3.9. Assume that the model (3.1) is regular and let $K \in \mathscr{I}(\mathscr{K})$. Then $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$ is regular with probability one (cf. (3.14) and Lemma 3.2). The function (3.3) is the likelihood function for the conditional model of $x_{[K]}$ given $x_{\langle K\rangle}$. Except for the parametrization of $L_{[K]}(x_{\langle K\rangle})$ by the regression parameter $R_{[K\rangle} \in \mathrm{M}_G([K] \times \langle K\rangle)$, this is the likelihood function for a LGS-model given by $L_{[K]}(x_{\langle K\rangle})$ and G (cf. Section A.8). By (A.13) and Lemma 3.6, it then follows that the distribution of $\det(\hat{\Sigma}_{[K]\cdot})/\det(\Sigma_{[K]\cdot})$ given $x_{\langle K\rangle}$ is the same as a product of independent variables

$$\prod \Big(\prod \Big(X_{j_\mu}^{d_\mu n_\mu} | j_\mu = 1, \dots, p_\mu^{[K]} \Big) \Big| \mu \in \mathfrak{M} \Big),$$

where X_{j_μ} follows a χ^2 distribution with $d_\mu(f_\mu^K - j_\mu + 1)$ degrees of freedom, scale $(d_\mu n_\mu)^{-1}$ and $f_\mu^K = n_\mu - p_\mu^{\langle K\rangle}$. Thus the distribution is independent of $x_{\langle K\rangle} \in \Omega_{\langle K\rangle}$ with probability 1 and is therefore also the distribution of $\det(\hat{\Sigma}_{[K]\cdot})/\det(\Sigma_{[K]\cdot})$.

4. Testing problems. In this section we consider the problem of testing additional conditional independence and/or symmetry restrictions.

Let $\mathscr{M}$ denote a subring of $\mathscr{K}$ and H a subgroup of G such that for all $h \in H$, $\rho(h) \in \mathrm{M}_{\mathscr{M}}(I)$. Since $\mathrm{P}_{G,\mathscr{K}}(I) \subseteq \mathrm{P}_{H,\mathscr{M}}(I)$, the problem of testing the GS-LCI model

$$(\mathrm{N}(\Sigma)|\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)) \tag{4.1}$$

versus the GS-LCI model

$$(\mathrm{N}(\Sigma)|\Sigma \in \mathrm{P}_{H,\mathscr{M}}(I)), \tag{4.2}$$

is well defined [compare AP (1995a)]. In the usual statistical language, this is the problem of testing the hypothesis

$$\mathrm{H}_{G,\mathscr{K}}\colon \Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$$

versus

$$\mathrm{H}_{H,\mathscr{M}}\colon \Sigma \in \mathrm{P}_{H,\mathscr{M}}(I),$$

based on a random observation $x \in \mathbb{R}^I$ from a normal distribution $\mathrm{N}(\Sigma)$, where $\Sigma \in \mathrm{P}_{H,\mathscr{M}}(I)$.

REMARK 4.1. Quantities such as $\langle K \rangle$ and $[K]$ depend not only on the subset K of I but also on the lattice of which K is considered a member. To alleviate this difficulty, the letter K shall denote a subset of I that is to be considered as a member of $\mathscr{K}$, while M shall denote a subset of I that is to be considered as a member of $\mathscr{M}$.

PROPOSITION 4.1. *Let $\Omega_{G,\mathscr{K}}$ and $\Omega_{H,\mathscr{M}}$ denote the set of regular elements for the GS-LCI models* (4.1) *and* (4.2), *respectively* [*cf*. (3.18)]. *Then $\Omega_{H,\mathscr{M}} \subseteq \Omega_{G,\mathscr{K}}$. In particular if the GS-LCI model* (4.2) *is regular*, *then the GS-LCI model* (4.1) *is regular*.

PROOF. Let $x \in \Omega_{H,\mathscr{M}}$. For $M \in \mathscr{J}(\mathscr{M})$ we have $\mathrm{M}_G([M] \times \langle M \rangle) \subseteq \mathrm{M}_H([M] \times \langle M \rangle)$ and thus $N_{G,\langle M \rangle} \supseteq N_{H,\langle M \rangle}$ which implies that $E_{G,M} \subseteq E_{H,M}$. Since $\mathscr{M}$ is a sublattice of $\mathscr{K}$, it follows [see Andersson (1990), Proposition 3.3] that there exists a surjective mapping $\varphi\colon \mathscr{J}(\mathscr{K}) \to \mathscr{J}(\mathscr{M})$ such that $K \subseteq \varphi(K)$ for all $K \in \mathscr{J}(\mathscr{K})$ and

$$[M] = \dot{\bigcup}([K]|K \in \mathscr{J}(\mathscr{K}), \varphi(K) = M), \tag{4.3}$$

for all $M \in \mathscr{J}(\mathscr{M})$, respectively. Let $K \in \mathscr{J}(\mathscr{K})$ and let $M = \varphi(K)$. Moreover let $p_{K,M}\colon \mathbb{R}^M \to \mathbb{R}^K$ denote the coordinate projection, that is, $p_{K,M}(x_M) = x_K$. It follows easily that

$$N_{G,K} = \{y \in \mathbb{R}^K \mid \forall\ A \in \mathrm{M}_G([K] \times K)\colon Ay = 0\},$$

where $\mathrm{M}_G([K], K)$ is the vector space of all $[K] \times K$ matrices that commute with G. Analogously

$$N_{G,M} = \{y \in \mathbb{R}^M \mid \forall\ A \in \mathrm{M}_G([M] \times M)\colon Ay = 0\}.$$

Since $p_{K,M}$ commutes with the group representations on $\mathbb{R}^M$ and $\mathbb{R}^K$, it then follows that $p_{K,M}(N_{G,M}) \subseteq N_{G,K}$, and thus $p_{K,M}(E_{G,M}) \supseteq E_{G,K}$. Since

$\mathscr{A}_I(H) \subseteq \mathscr{A}_I(G) \subseteq \mathrm{M}_{\mathscr{M}}(I) \subseteq \mathrm{M}_{\mathscr{K}}(I)$, it follows that for any $A \in \mathscr{A}_I(H)$,

$$A_K x_K = (Ax)_K = ((Ax)_M)_K = (A_M x_M)_K.$$

These considerations and the fact that $\mathscr{A}_K(H) \subseteq \mathscr{A}_K(G)$, shows that

$$\begin{aligned}\mathscr{A}_K(G) x_K \supseteq \mathscr{A}_K(H) x_K &= (\mathscr{A}_M(H) x_M)_K \\ &\supseteq (E_{H,M})_K \supseteq (E_{G,M})_K \supseteq E_{G,K}\end{aligned}$$

and hence $x \in \Omega_{G,\mathscr{K}}$. $\square$

Suppose that the normal model (4.2) is regular. It then follows from Proposition 4.1 and (3.12) that the LR test statistic Q for testing $\mathrm{H}_{G,\mathscr{K}}$ against $\mathrm{H}_{H,\mathscr{M}}$ exists with probability 1 and is given by

$$Q(x) = \left(\frac{\det(\tilde{\Sigma}(x))}{\det(\hat{\Sigma}(x))}\right)^{1/2}, \tag{4.4}$$

where $x \in \mathbb{R}^I$ and $\tilde{\Sigma}$ is the MLE for $\Sigma \in \mathrm{P}_{H,\mathscr{M}}(I)$ in the model (4.2).

Next we shall find the moments of the LR test statistic Q. First note that $\mathrm{GL}_{G,\mathscr{K}}(I) \subseteq \mathrm{GL}_{H,\mathscr{M}}(I)$. The testing problem is invariant under the action of the group $\mathrm{GL}_{G,\mathscr{K}}(I)$ on the sample space $\mathbb{R}^I$ given by the restriction of the action (2.2), and the action of $\mathrm{GL}_{G,\mathscr{K}}(I)$ on the parameter space $\mathrm{P}_{H,\mathscr{M}}(I)$ given by the restriction of the action of $\mathrm{GL}_{H,\mathscr{M}}(I)$ on $\mathrm{P}_{H,\mathscr{M}}(I)$ (cf. Theorem 2.2 with $\mathscr{K}$ replaced by $\mathscr{M}$ and G replaced by H).

Let

$$\pi: \mathbb{R}^I \to \mathbb{R}^I/\mathrm{GL}_{G,\mathscr{K}}(I)$$

denote the orbit projection (maximal invariant) of the action of $\mathrm{GL}_{G,\mathscr{K}}(I)$ on $\mathbb{R}^I$. The LR test statistic Q is invariant under this action and thus $Q(x)$ only depends on $x \in \mathbb{R}^I$ through $\pi(x)$. The central distribution of Q is then readily obtained from this fact and the theorem below.

THEOREM 4.1. *Suppose that the normal model* (4.2) *is regular. Under* $\mathrm{H}_{G,\mathscr{K}}$ *the orbit projection* π *and the ML estimators* $\hat{\Sigma}_{[K]\cdot}$ *of* $\Sigma_{[K]\cdot} \in \mathrm{P}_G([K])$, $K \in \mathscr{J}(\mathscr{K})$, *are mutually independent.*

For the proof, see Appendix B.

It follows from (4.4), (2.8) and Theorem 4.1 that for every $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ and $\alpha \in [0,\infty)$,

$$\mathrm{E}\Big(\det(\tilde{\Sigma})^{\alpha/2}\Big) = \mathrm{E}\Big(\det(\hat{\Sigma})^{\alpha/2} Q^{\alpha}\Big) = \mathrm{E}\Big(\det(\hat{\Sigma})^{\alpha/2}\Big)\mathrm{E}(Q^{\alpha}).$$

Hence

$$\mathrm{E}(Q^{\alpha}) = \frac{\mathrm{E}\Big(\det(\tilde{\Sigma})^{\alpha/2}\Big)}{\mathrm{E}\Big(\det(\hat{\Sigma})^{\alpha/2}\Big)}. \tag{4.5}$$

Thus it suffices to determine the moments $\mathrm{E}(\det(\hat{\Sigma})^{\alpha/2})$, where $\hat{\Sigma}$ is the MLE of $\Sigma \in \mathrm{P}_{G,\mathscr{K}}(I)$ in the GS-LCI model (4.1). Furthermore it follows from (2.8) and Theorem 4.1 that

$$\begin{aligned}\mathrm{E}\left(\det(\hat{\Sigma})^{\alpha/2}\right) &= \mathrm{E}\left(\prod\left(\det\left(\hat{\Sigma}_{[K]\cdot}\right)^{\alpha/2}\middle| K \in \mathscr{I}(\mathscr{K})\right)\right)\\ &= \prod\left(\mathrm{E}\left(\det\left(\hat{\Sigma}_{[K]\cdot}\right)^{\alpha/2}\right)\middle| K \in \mathscr{I}(\mathscr{K})\right),\end{aligned}$$

and hence it is enough to determine the moments $\mathrm{E}(\det(\hat{\Sigma}_{[K]\cdot})^{\alpha/2})$ for $K \in \mathscr{I}(\mathscr{K})$. These moments are obtained from Remark 3.9 as

$$\begin{aligned}&\mathrm{E}\left(\det\left(\hat{\Sigma}_{[K]\cdot}\right)^{\alpha/2}\right)\Big/\det\left(\Sigma_{[K]\cdot}\right)^{\alpha/2}\\ &\quad= \prod\left(\prod\left(\left(\frac{2}{d_\mu n_\mu}\right)^{d_\mu n_\mu(\alpha/2)} \frac{\Gamma\left(d_\mu\left(f_\mu^K - j_\mu + 1\right)/2 + d_\mu n_\mu(\alpha/2)\right)}{\Gamma\left(d_\mu\left(f_\mu^K - j_\mu + 1\right)/2\right)}\right|\right.\\ &\qquad\qquad\left.\left. j_\mu = 1,\dots, p_\mu^{[K]}\right)\middle| \mu \in \mathscr{M}\right),\end{aligned}$$

where $f_\mu^K = n_\mu - p_\mu^{\langle K\rangle}$, $\mu \in \mathfrak{M}$, $K \in \mathscr{I}(\mathscr{K})$ and $\alpha \in [0, \infty[$. Note that these moments are determined by the structure constants $(((p_\mu^{[K]}|K \in \mathscr{I}(\mathscr{K})), d_\mu, n_\mu)|\, \mu \in \mathscr{M})$ for the model (4.1). Thus by (4.5) and (2.8),

$$\begin{aligned}&\mathrm{E}(Q^\alpha)\\ &\quad= \left[\prod\left(\left(\frac{2}{d_\nu n_\nu}\right)^{d_\nu n_\nu(\alpha/2)} \frac{\Gamma\left(d_\nu\left(f_\nu^M - j_\nu + 1\right)/2 + d_\nu n_\nu(\alpha/2)\right)}{\Gamma\left(d_\nu\left(f_\nu^M - j_\nu + 1\right)/2\right)}\right|\right.\\ &\qquad\qquad\left.\left. M \in \mathscr{I}(\mathscr{M}),\ \nu \in \mathfrak{N},\ j_\nu = 1,\dots, p_\nu^{[M]}\right)\right]\\ &\qquad\times\left[\prod\left(\left(\frac{2}{d_\mu n_\mu}\right)^{d_\mu n_\mu(\alpha/2)} \frac{\Gamma\left(d_\mu\left(f_\mu^K - j_\mu + 1\right)/2 + d_\mu n_\mu(\alpha/2)\right)}{\Gamma\left(d_\mu\left(f_\mu^K - j_\mu + 1\right)/2\right)}\right|\right.\\ &\qquad\qquad\left.\left. K \in \mathscr{I}(\mathscr{K}),\ \mu \in \mathscr{M},\ j_\mu = 1,\dots, p_\mu^{[K]}\right)\right]^{-1},\end{aligned}$$

where $(((p_\nu^{[M]}|M \in \mathscr{I}(\mathscr{M})), d_\nu, n_\nu)|\nu \in \mathscr{N})$ are the structure constants for the model (4.2) and $f_\nu^M = n_\nu - p_\nu^{\langle M\rangle}$, $\nu \in \mathscr{N}$, $M \in \mathscr{I}(\mathscr{M})$. Since

$$\begin{aligned}&\sum\left(p_\mu^{[K]} d_\mu n_\mu |\, \mu \in \mathscr{M},\ K \in \mathscr{I}(\mathscr{K})\right) = |I|\\ &\qquad\qquad= \sum\left(p_\nu^{[M]} d_\nu n_\nu |\nu \in \mathscr{N},\ M \in \mathscr{I}(\mathscr{M})\right),\end{aligned}$$

an approximation for the central distribution of $-2\log Q$ may be obtained by the Box approximation as given in Anderson (1984), pages 311–316, or by means of the probably better saddle point approximation as given in, for example, Jensen (1991).

REMARK 4.2. If $G = H$, then $(((p_\mu^{[M]} | M \in \mathscr{I}(\mathscr{M})), d_\mu, n_\mu) \mid \mu \in \mathfrak{M})$ are the structure constants for the model (4.2), and thus

$$\begin{aligned} &\mathrm{E}(Q^\alpha) \\ &= \prod_{\mu \in \mathfrak{M}} \Bigg[\prod \Bigg(\left(\frac{2}{d_\mu n_\mu} \right)^{d_\mu n_\mu (\alpha/2)} \frac{\Gamma\big(d_\mu\big(f_\mu^M - j_\mu + 1\big)/2 + d_\mu n_\mu(\alpha/2)\big)}{\Gamma\big(d_\mu\big(f_\mu^M - j_\mu + 1\big)/2\big)} \Bigg| \\ &\qquad\qquad M \in \mathscr{I}(\mathscr{M}),\ j_\mu = 1, \dots, p_\mu^{[M]} \Bigg) \Bigg] \\ &\quad \times \Bigg[\prod \Bigg(\left(\frac{2}{d_\mu n_\mu} \right)^{d_\mu n_\mu (\alpha/2)} \frac{\Gamma\big(d_\mu\big(f_\mu^K - j_\mu + 1\big)/2 + d_\mu n_\mu(\alpha/2)\big)}{\Gamma\big(d_\mu\big(f_\mu^K - j_\mu + 1\big)/2\big)} \Bigg| \\ &\qquad\qquad K \in \mathscr{I}(\mathscr{K}), j_\mu = 1, \dots, p_\mu^{[K]} \Bigg) \Bigg]^{-1}; \end{aligned}$$

that is, the test statistic is a product of test statistics indexed by $\mathfrak{M}$. The factor associated with $\mu \in \mathfrak{M}$ corresponds to the problem of testing additional CI-restrictions based on n_μ independent repetitions from a D_μ-LCI model ($D_\mu = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$) on $D_\mu^{p^\mu}$ (cf. Section 5.2). In the special case where $G = H = \{e\}$ is the trivial group, this is the problem of testing the LCI model given by $\mathscr{K}$ against the one given by $\mathscr{M}$ based on $n = n_\mu$ independent repetitions [cf. AP (1995a)].

REMARK 4.3. If $\mathscr{K} = \mathscr{M}$, then $(((p_\nu^{[K]} | K \in \mathscr{I}(\mathscr{K})), d_\nu, n_\nu) | \mu \in \mathfrak{M})$ are the structure constants for the model (4.2), and thus

$$\begin{aligned} \mathrm{E}(Q^\alpha) &= \prod_{K \in \mathscr{I}(\mathscr{K})} \\ &= \Bigg[\prod \Bigg(\left(\frac{2}{d_\nu n_\nu} \right)^{d_\nu n_\nu (\alpha/2)} \frac{\Gamma\big(d_\nu\big(f_\nu^K - j_\nu + 1\big)/2 + d_\nu n_\nu(\alpha/2)\big)}{\Gamma\big(d_\nu\big(f_\nu^K - j_\nu + 1\big)/2\big)} \Bigg| \\ &\qquad\qquad \nu \in \mathfrak{N},\ j_\nu = 1, \dots, p_\nu^{[K]} \Bigg) \Bigg] \\ &\quad \times \Bigg[\prod \Bigg(\left(\frac{2}{d_\mu n_\mu} \right)^{d_\mu n_\mu (\alpha/2)} \frac{\Gamma\big(d_\mu\big(f_\mu^K - j_\mu + 1\big)/2 + d_\mu n_\mu(\alpha/2)\big)}{\Gamma\big(d_\mu\big(f_\mu^K - j_\mu + 1\big)/2\big)} \Bigg| \\ &\qquad\qquad \mu \in \mathfrak{M},\ j_\mu = 1, \dots, p_\mu^{[K]} \Bigg) \Bigg]^{-1}, \end{aligned}$$

that is, the test statistic is a product of test statistics indexed by $\mathscr{I}(\mathscr{K})$. The factor associated with $K \in \mathscr{I}(\mathscr{K})$ corresponds to the problem of testing additional symmetry restrictions determined by G in the generalized Wishart distribution (cf. Remark A.2) of $\hat{\Sigma}_{[K]}$. given by H. In the special case $\mathscr{K} = \mathscr{M} = \{I, \varnothing\}$ this is the problem of testing the GS model given by G against the one given by H (cf. Section A.7).

5. Independent repetitions and other examples.

5.1. *Independent repetitions*. Let N be a nonempty finite set, and consider the model (2.11). The set

$$\mathscr{K}(N) = \{K \times N | K \in \mathscr{K}\}$$

is a lattice of subsets of $I \times N$ with $\mathscr{J}(\mathscr{K}(N)) = \{K \times N | K \in \mathscr{J}(\mathscr{K})\}$, $\langle K \times N \rangle = \langle K \rangle \times N$ and $[K \times N] = [K] \times N$, $K \in \mathscr{J}(\mathscr{K})$. It then follows that the model

$$\big(\mathrm{N}(\Gamma) | \Gamma \in \mathrm{P}_{G(N), \mathscr{K}(N)}(I \times N)\big) \tag{5.1}$$

is a GS-LCI model given by $G(N)$ and $\mathscr{K}(N)$ (cf. Section A.6). It is easy to see that $\mathrm{P}_{G(N), \mathscr{K}(N)}(I \times N) = \{\Sigma \otimes 1_N | \Sigma \in \mathrm{P}_{G, \mathscr{K}}(I)\}$. Note that $\Sigma \otimes 1_N = \mathrm{diag}(\Sigma | \nu \in N)$, $\Sigma \in \mathrm{P}_{G, \mathscr{K}}(I)$ [cf. (A.5)].

Thus the GS-LCI model (5.1) is, except for the reparametrization

$$\begin{gathered} \mathrm{P}_{G, \mathscr{K}}(I) \leftrightarrow \mathrm{P}_{G(N), \mathscr{K}(N)}(I \times N), \\ \Sigma \mapsto \Sigma \otimes 1_N, \end{gathered} \tag{5.2}$$

the same as N independent repetitions of the model (2.11), that is, the model

$$\big(\mathrm{N}(\Sigma)^{\otimes N} | \Sigma \in \mathrm{P}_{G, \mathscr{K}}(I)\big) \tag{5.3}$$

with observation space $\mathbb{R}^{I \times N}$. The smoothing mapping Ψ corresponding to $G(N)$ is determined in (A.8). The structure constants for the model (5.1) become $(((p_\mu^{[K]} | K \in \mathscr{J}(\mathscr{K})), d_\mu, n_\mu n) | \mu \in \mathfrak{M})$, where $n = |N|$, (cf. Section A.6).

Thus the smoothing function ψ, the partially ordered set $\mathscr{J}(\mathscr{K})$ and the structure constants $(((p_\mu^{[K]} | K \in \mathscr{J}(\mathscr{K})), d_\mu, n_\mu) | \mu \in \mathfrak{M})$ for the model (2.9), determine the smoothing function Ψ, the partially ordered set $\mathscr{J}(\mathscr{K}(N)) \simeq \mathscr{J}(\mathscr{K})$ and the structure constants $(((p_\mu^{[K]} | K \in \mathscr{J}(\mathscr{K})), d_\mu, n_\mu n) | \mu \in \mathfrak{M})$ for the model (5.3). In particular for this model they determine the following.

1. The regularity condition

$$\forall\ K \in \mathscr{J}(\mathscr{K}) \forall\ \mu \in M\colon\ p_\mu^{[K]} > 0 \Rightarrow n n_\mu \geq p_\mu^{K}. \tag{5.4}$$

2. The likelihood equations (3.10) and (3.11) for $\hat{\Sigma} = \hat{\Sigma}(y)$ where xx' is replaced by

$$\overline{S}(y) = \frac{1}{n} \sum (x_\nu x_\nu' | \nu \in N)$$

   and $y = (x_\nu | \nu \in M) \in \mathbb{R}^{I \times N}$ is a partition of y according to (A.5).
3. The moments of the generalized variance $\det(\hat{\Sigma})$ [cf. the reparametrization (5.2)] in the model (5.3).

REMARK 5.1. The inequalities (5.4) determine the minimum number of repetitions for regularity of the model (5.3).

5.2. *Other examples*. Let J and T be finite sets, γ an orthogonal group representation of G on $\mathbb{R}^J$ and $\mathscr{L}$ a lattice of subsets of T. Then ρ: $G \to \mathrm{O}(J \times T)$ defined by $\rho(g) = \gamma(g) \otimes 1_T$ is an orthogonal group representation of G on $\mathbb{R}^{J \times T}$ and $\mathscr{K} = \{J \times L | L \in \mathscr{L}\}$ a lattice of subsets of $J \times T$. The model (2.11) with $I = J \times T$ and ρ and $\mathscr{K}$ defined as above is then a GS-LCI model. This model is called the *tensor product* of the GS model $(\mathrm{N}(\Delta)|\Delta \in \mathrm{P}_G(J))$ and the LCI model $(\mathrm{N}(\Lambda)|\Lambda \in \mathrm{P}_{\mathscr{K}}(T))$. The smoothing function corresponding to ρ is easily obtained in terms of the smoothing function corresponding to γ, the partially ordered set $\mathscr{J}(\mathscr{K})$ is isomorphic to the partially ordered set $\mathscr{J}(\mathscr{L})$, and the structure constants for the tensor product model are given by $(((|[L]|\, p_\mu | L \in \mathscr{J}(\mathscr{L})), d_\mu, n_\mu) \,|\, \mu \in \mathfrak{M})$ where $((p_\mu, d_\mu, n_\mu) \,|\, \mu \in \mathfrak{M})$ are the structure constants for γ. The fundamental quantities (smoothing function, join-irreducible elements and structure constants) for the tensor product model can thus be obtained from the fundamental quantities for each of the models in the tensor product. The model consisting of n independent repetitions of the tensor product model is then regular if and only if $n \geq |L|\, p_\mu$, $L \in \mathscr{J}(\mathscr{L})$, $\mu \in \mathfrak{M}$.

The above tensor product construction allows one to construct numerous specific examples by using the examples of GS models listed in Section A.6 and the examples of LCI models presented in AP (1993). Examples 1.1 and 1.2 in the introduction are tensor products of a trivariate complete symmetry model and the LCI models in Example 2.5 and Example 2.4, respectively, of AP (1993). Example 1.3 is not of this type.

Let $\mathbb{C}$ be the complex numbers, $G = \{\pm 1, \pm i\}$, γ the representation on $\mathbb{C} = \mathbb{R}^J$, $J = \{1, i\}$, given by $\gamma(g)(z) = gz$, $g \in G$, $z \in \mathbb{C}$ and $\mathscr{L}$ a lattice of subsets of T. The tensor product model is then the *complex* LCI-model, that is, the extension of the LCI-model in AP (1993) to complex variables [see also Massam and Neher (1995)]. The structure constants for the GS model are $(1, 2, 1)$. The model consisting of n independent repetitions of the tensor product model is then regular if and only if $n \geq |L|$, $L \in \mathscr{J}(\mathscr{L})$. Note that this condition is the same as (3.3) in AP (1993).

In the same way, one can define the *quaternion* LCI model. The structure constants for the GS model in this tensor product model are $(1, 4, 1)$, and the regularity condition for the independent repetitions model is thus the same, namely, $n \geq |L|$, $L \in \mathscr{J}(\mathscr{L})$.

REMARK 5.2. From the structure theorem in Andersson (1975b), Theorem 4.18, and the invariant formulation of the LCI model in Section 4 of AP (1993), it can be seen that every GS-LCI model is an independent product of GS-LCI models, each being independent repetitions of a real, complex or quaternion LCI model. This structure result for GS-LCI models is a direct extension of the structure result for GS models in Andersson (1975b).

6. Concluding remarks. The GS-LCI models may be generalized in several ways.

1. As in the case of the LCI models [cf. Andersson and Perlman (1994)], the assumption that the expectations of the normal distributions in the model (2.11) are zero can be removed. The *linear GS-LCI model*

$$(\mathrm{N}(\xi, \Sigma) \mid (\xi, \Sigma) \in L \times \mathrm{P}_{G,\mathscr{K}}(I)),$$

where $L \in \mathbb{R}^I$ is an $\mathrm{M}_{G,\mathscr{K}}(I)$-*subspace*, that is, $\mathrm{M}_{G,\mathscr{K}}(I)L = L$, admits an explicit likelihood analysis similar to the GS-LCI models.
2. The conditional independence (CI) restrictions given by the lattice $\mathscr{K}$ may be extended to CI restrictions given by an acyclic directed graph $\mathfrak{R} = (V, E)$, where V represents the vertex set and E the directed edges (arrows) between vertices. Suppose that $I = \dot{\bigcup}(I_v | v \in V)$. Then $\mathbb{R}^I = \times(\mathbb{R}^{I_v} | v \in V)$ and CI restrictions of normal distributions on $\mathbb{R}^I$ w.r.t. (V, E) can be defined as in, for example, Andersson and Perlman (1995b) or Lauritzen (1989, 1996). Let $\mathrm{P}_{\mathfrak{R}}(I) \subseteq \mathrm{P}(I)$ be the corresponding set of nonsingular covariance matrices. Let ρ be an orthogonal group representation of G on $\mathbb{R}^I$ such that $u_{I_v}(\mathbb{R}^{I_v})$ is a G-subspace, $v \in V$ (cf. Lemma 2.1). The model

$$(\mathrm{N}(\Sigma) | \Sigma \in \mathrm{P}_{G,\mathfrak{R}}(I)) \tag{6.1}$$

where $\mathrm{P}_{G,\mathfrak{R}}(I) = \mathrm{P}_G(I) \cap \mathrm{P}_{\mathfrak{R}}(I)$ admits an explicit likelihood analysis similar to the GS-LCI models. If $\mathscr{K}$ is a lattice of subsets of I, then $\mathscr{J}(\mathscr{K})$ is the vertex set for an acyclic directed graph $\mathfrak{R}(\mathscr{K})$, whose directed edges are defined as follows: $K_1 \to K_2$ if and only if $K_1 \subset K_2$, $K_1, K_2 \in \mathscr{J}(\mathscr{K})$. With the definition $I_K = [K]$, $\mathrm{K} \in \mathscr{J}(\mathscr{K})$, it is seen that the GS-LCI model (2.11) is a special case of the model (6.1) with $\mathfrak{R} = \mathfrak{R}(\mathscr{K})$. Note that $\mathfrak{R}(\mathscr{K})$ is a *transitive* directed acyclic graph cf. Andersson, Madigan, Perlman and Triggs (1995a).
3. The condition that all matrices $\rho(g)$, $g \in G$, are $\mathscr{K}$-preserving (cf. Section 2.4) may be weakened as follows: Let $\mathrm{End}(\mathscr{K})$ denote the set of all lattice homomorphisms $f\colon \mathscr{K} \to \mathscr{K}$, that is, $f(L \cap M) = f(L) \cap f(M)$ and $f(L \cup M) = f(L) \cup f(M)$ for all $L, M \in \mathscr{K}$. Now, suppose that G also is represented on $\mathscr{K}$:

$$\begin{aligned} G &\to \mathrm{End}(\mathscr{K}), \\ g &\mapsto (K \mapsto gK), \end{aligned} \tag{6.2}$$

such that

$$\forall\ x \in \mathbb{R}^I,\ \forall\ K \in \mathscr{K},\ \forall\ g \in G\colon\ x_K = 0 \Rightarrow (\rho(g)x)_{gK} = 0. \tag{6.3}$$

In the special case where the representation (6.2) is trivial, that is, $gK = K$, $K \in \mathscr{K}$, (6.3) reduces to the condition that $\rho(g)$ is $\mathscr{K}$-preserving, $g \in G$ [cf. Proposition 2.2 in AP (1993)]. Thus the model (2.11) with the weaker condition (6.3) is a generalization of the GS-LCI model.

The extension given by the combination of (2) and (3) is under investigation and is studied in Madsen (1996). We close this section with a presentation of an example of a model with GS and CI restrictions which is *not* a GS-LCI model.

EXAMPLE 6.1. Let $x = (x_a, x_b, x_c)$ be a trivariate normal distributed variable with mean 0 and covariance matrix $\Sigma = (\sigma_{lm} | l, m = a, b, c)$. For example, x_a, x_b, x_c, could be measurements of some variable on objects a, b, c where b and c are "symmetric." For example, b and c could be two identical subengines of a large (airplane) engine with main engine a. The joint distribution should thus not depend on the (probably irrelevant) labelling of the two subengines; that is, it should remain invariant under the simple linear transformation that corresponds to permutation of b and c. This implies that Σ has the restrictions $\sigma_{ba} = \sigma_{ca}$ and $\sigma_{bb} = \sigma_{cc}$; that is, Σ takes the form

$$\mathrm{H}_{GS}\colon \Sigma = \begin{pmatrix} \sigma & \gamma & \gamma \\ \gamma & \alpha & \beta \\ \gamma & \beta & \alpha \end{pmatrix}.$$

Next suppose that the two subengines b and c are connected to each other only through the main engine a. In this case it would be reasonable to assume that x_b and x_c are conditionally independent given x_a. This implies that Σ has the restriction

$$\mathrm{H}_{\mathrm{LCI}}\colon \sigma_{bc} = \sigma_{ba}\sigma_{aa}^{-1}\sigma_{ac}.$$

The restriction imposed on Σ by H_{GS} and $\mathrm{H}_{\mathrm{LCI}}$ could then be expressed as H_{GS} together with the restriction

$$\mathrm{H}_{\text{GS-LCI}}\colon \beta = \gamma\sigma^{-1}\gamma.$$

Under the model $\mathrm{H}_{\mathrm{LCI}}$, Σ is uniquely determined by the regression parameters $r_b = \sigma_{ba}\sigma_{aa}^{-1}$ and $r_c = \sigma_{ca}\sigma_{aa}^{-1}$, the conditional variances $\lambda_b = \sigma_{b\cdot a}$ and $\lambda_c = \sigma_{c\cdot a}$, and the marginal variance $\lambda_a = \sigma_{aa}$. Thus $\mathrm{H}_{\text{GS-LCI}}$ could equivalently be expressed as $\mathrm{H}_{\mathrm{LCI}}$ together with the restrictions

$$\begin{aligned} r_b &= r_c = r, \\ \lambda_b &= \lambda_c = \lambda. \end{aligned} \tag{6.4}$$

Now consider N i.i.d. observations $x_1, \dots, x_N$ of the trivariate random observation x. Since the model H_{GS} is a special case of multivariate compound symmetry, it is well known from classical multivariate analysis that the ML estimator exists and is unique if and only if $N \geq 2$. In the model $\mathrm{H}_{\mathrm{LCI}}$, the required condition is also $N \geq 2$. Using (6.4), it can easily be seen that in the case of the model $\mathrm{H}_{\text{GS-LCI}}$, the condition is $N \geq 1$. As in the case of the GS-LCI models, the ML estimator can be found using a combination of the techniques from GS models and LCI models. Thus if $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\sigma}$ are the ML-estimates of the unknown parameters under H_{GS}, the ML-estimator under $\mathrm{H}_{\text{GS-LCI}}$ is determined by setting $\hat{r} = \hat{\gamma}\hat{\sigma}^{-1}$, $\hat{\lambda} = \hat{\alpha} - \hat{\gamma}\hat{\sigma}^{-1}\hat{\gamma}$ and $\hat{\lambda}_a = \hat{\sigma}$, respectively.

APPENDIX A

Group symmetry models. The theory of group symmetry (GS) models presented in this Appendix is a small but central part of an unpublished general algebraic theory of normal models developed by Andersson, Brøns, and Jensen in the years 1972–1985 (cf. the Introduction and the list of references). Note that the definition of a GS model [cf. (2.9)], can be extended to a continuous representation ρ: $G \to \mathrm{M}(I)$ of a compact group G on $\mathbb{R}^I$. The definition of the smoothing function ψ [cf. (2.10)], should then be replaced by $\psi(S) = \int \rho(g) S \rho(g)' \beta(ds)$ where β is the unique normalized Haar measure on G.

PROPOSITION A.1. *The action of* $\mathrm{GL}_G(I)$ *on* $\mathrm{P}_G(I)$ *given by the restriction of the action* (2.1) *is well defined, transitive and proper.*

PROOF. It is trivial that the action is well defined. Since $\mathrm{GL}_G(I)$ and $\mathrm{P}_G(I)$ are closed subsets of $\mathrm{GL}(I)$ and $\mathrm{P}(I)$ and the action (2.1) is proper, it follows that the action of $\mathrm{GL}_G(I)$ on $\mathrm{P}_G(I)$ is proper. That the action is transitive follows from Lemma A.1. □

LEMMA A.1. *Let* $\Sigma \in \mathrm{PS}_G(I)$. *Then there exists* $A \in \mathrm{M}_G(I)$ *such that* $\Sigma = AA'$.

PROOF. There exists $A \in \mathrm{M}(I)$ such that $\Sigma = AA'$. Then $AA' \in \mathrm{M}_G(I)$. Since $(AA')^{1/2}$ is a polynomial in AA', it follows that $(AA')^{1/2} \in \mathrm{M}_G(I)$. Thus $\Sigma = (AA')^{1/2}(AA')^{1/2}$. □

A.1. *Maximum likelihood estimation.*

THEOREM A.1. *The maximum likelihood estimator* $\hat{\Sigma}(x)$ *of* $\Sigma \in \mathrm{P}_G(I)$ *in the model* (2.9) *exists if and only if* $\psi(xx')$ *is nonsingular. In this case it is unique and given by* $\hat{\Sigma}(x) = \psi(xx')$.

PROOF. The likelihood function $\mathbf{L}$: $\mathrm{P}_G(I) \times \mathbb{R}^I \to (0, \infty)$ is given by $\mathbf{L}(\Sigma, x) = \det(\Sigma)^{-1/2}\exp(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}xx'))$. Since $\Sigma^{-1} \in \mathrm{P}_G(I)$, it follows that

$$\text{(A.1)} \qquad \mathrm{tr}(\Sigma^{-1}xx') = \mathrm{tr}\big(\Sigma^{-1}\psi(S(x))\big),$$

where $S(x) = xx'$ and ψ is defined in (2.10). The problem is then to maximize

$$\text{(A.2)} \qquad \mathbf{L}(\Sigma, x) = \det(\Sigma)^{-1/2}\exp\big(-\tfrac{1}{2}\mathrm{tr}\big(\Sigma^{-1}\psi(S(x))\big)\big),$$

in $\Sigma \in \mathrm{P}_G(I)$ for fixed $S(x) \in \mathrm{PS}(I)$.

Assume that $\psi(S(x))$ is nonsingular. Then the maximum of $L(\Sigma, x)$ for $\Sigma \in \mathrm{P}(I)$ is well known to be attained at the unique point $\hat{\Sigma}(x) = \psi(S(x))$. Since $\psi(S(x)) \in \mathrm{P}_G(I)$, $\hat{\Sigma}(x)$ is also the solution to the original maximization problem.

Next suppose that $\psi(S(x))$ is singular and set $\Sigma_n = n^{-1}1_I + \psi(S(x))$. Then $\Sigma_n \in \mathrm{P}_G(I)$ and $L(\Sigma_n, x) \to \infty$ for $n \to \infty$. $\square$

REMARK A.1. The maximum of the likelihood function is

$$L\big(\hat{\Sigma}(x), x\big) = \det\big(\hat{\Sigma}(x)\big)^{-1/2}\exp\big(-\tfrac{1}{2}|I|\big).$$

From the likelihood function (A.2) it can easily be verified that the GS model (2.9) is an exponential family with canonical statistic $\psi(xx')$ and with full and open parameter space $\mathrm{P}_G(I)$ [cf. Barndorff-Nielsen (1978)]. In particular the canonical statistic is sufficient and complete.

PROPOSITION A.2. *Consider the GS model* (2.9). *Either the ML estimator* $\hat{\Sigma}(x)$ *exists with probability one w.r.t. all* $\mathrm{N}(\Sigma)$, $\Sigma \in \mathrm{P}_G(I)$, *or else it does not exist for any* $x \in \mathbb{R}^I$.

The proof follows from Theorem A.1, and from Lemmas A.2 and A.4.

REMARK A.2. The distribution of $\hat{\Sigma}$ can be described as a *generalized Wishart distribution* on $\mathrm{P}_G(I)$ cf. the final paragraph of Appendix B. The generalized Wishart distributions $\mathrm{W}_{\Sigma,\chi}$ are parametrized by pairs $(\Sigma, \chi) \in \mathrm{P}_G(I) \times \mathfrak{D}\mathfrak{F}$ where $\mathfrak{D}\mathfrak{F}$ (for generalization of degrees of freedom) is a subset of all multipliers χ on $\mathrm{GL}_G(I)$; that is, all continuous functions $\chi\colon \mathrm{GL}_G(I) \to (0,\infty)$ with the properties $\chi(1_I) = 1$ and $\chi(A_1A_2) = \chi(A_1)\chi(A_2)$, $A_1, A_2 \in \mathrm{GL}_G(I)$. The expectation of $\mathrm{W}_{\Sigma,\chi}$ is Σ. It is beyond the scope of the present paper to discuss these distributions further.

A.2. *Regular elements*. Let $\mathscr{A}(G) \subseteq \mathrm{M}(I)$ be the algebra generated by $\rho(G)$. An element $x \in \mathbb{R}^I$ is called *regular* if $\mathscr{A}(G)x = \mathbb{R}^I$. Let $\Omega \subseteq \mathbb{R}^I$ denote the set of regular elements.

LEMMA A.2. An element $x \in \mathbb{R}^I$ is regular if and only if $\psi(xx')$ is nonsingular.

PROOF. First note that we have the following biimplication: for all $x, z \in \mathbb{R}^I$,

$$\begin{aligned} z'\psi(xx')z = 0 \quad &\Leftrightarrow \quad \sum\big(z'\rho(g)xx'\rho(g)'z \,|\, g \in G\big) = 0 \\ &\Leftrightarrow \quad \forall\, g \in G\colon \big(z'\rho(g)x\big)\big(z'\rho(g)x\big)' = 0 \\ &\Leftrightarrow \quad \forall\, g \in G\colon z'\rho(g)x = 0. \end{aligned}$$

Let $x \in \Omega$ and let $z \in \mathbb{R}^I$. It is enough to show that $z'\psi(xx')z = 0$ implies that $z = 0$. Since $\operatorname{span}\{\rho(g)x | g \in G\} = \mathbb{R}^I$, the above bimultiplication from the left to the right provides this result. Next, assume that $x \in \mathbb{R}^I \setminus \Omega$. Then $\mathscr{A}(G)x \subset \mathbb{R}^I$ and therefore there exists a $z \in \mathbb{R}^I$ with $z \neq 0$ such that $z'(\rho(g)x) = 0$ for all $g \in G$. The biimplication from the right to the left then shows that $\psi(xx')$ is singular. $\square$

LEMMA A.3. *Let $\mathscr{A} \subseteq \mathrm{M}(I)$ be a subalgebra with $1_I \in \mathscr{A}$ and let $\mathscr{A}^* \subseteq \mathscr{A}$ denote the group of nonsingular elements in $\mathscr{A}$. Then $\mathscr{A}^*$ is an open subset of $\mathscr{A}$, and the Lebesgue measure on $\mathscr{A}$ is concentrated on $\mathscr{A}^*$; that is, $\mathscr{A} \setminus \mathscr{A}^*$ has Lebesgue measure zero.*

PROOF. Let the mapping L_a: $\mathscr{A} \to \mathscr{A}$ denote the left multiplication with $a \in \mathscr{A}$. Then $a \in \mathscr{A}^*$ if and only if $\det(L_a) \neq 0$. Since $\det(L_a)$ is a continuous function of $a \in \mathscr{A}$ it follows that $\mathscr{A}^*$ is an open subset of $\mathscr{A}$. Furthermore $\det(L_a)$ is a polynomial in the coordinates of $a \in \mathscr{A}$ with respect to a basis for $\mathscr{A}$. Since $\det(L_{1_I}) \neq 0$, the result follows from Bourbaki (1963), Ch. VII, Section 3, Number 3, Lemma 9. □

LEMMA A.4. *The set Ω is an open subset of $\mathbb{R}^I$. If $\Omega \neq \emptyset$ then the Lebesgue measure on $\mathbb{R}^I$ is concentrated on Ω; that is, $\mathbb{R}^I \setminus \Omega$ has Lebesgue measure zero.*

PROOF. Since $\Omega = \{x \in \mathbb{R}^I | \psi(xx') \text{ is nonsingular}\}$ and $\mathrm{P}_G(I)$ is open in $\mathrm{PS}_G(I)$, it follows that Ω is open in $\mathbb{R}^I$. Let $x \in \Omega$. Then the mapping f: $\mathscr{A}(G) \to \mathbb{R}^I$ given by $A \mapsto Ax$ is linear and surjective. Since the Lebesgue measure on $\mathscr{A}(G)$ is concentrated on the nonsingular elements $\mathscr{A}(G)^*$ of $\mathscr{A}(G)$ (cf. Lemma A.3), it then follows that the Lebesgue measure on $\mathbb{R}^I$ is concentrated on $f(\mathscr{A}(G)^*) \subseteq \Omega$, and thus that $\mathbb{R}^I \setminus \Omega$ has Lebesgue measure zero. □

A GS model with $\Omega \neq \emptyset$, that is, the ML estimator exists with probability 1, is called a *regular* GS model. Note that the concept of regularity only depends on the representation ρ of G on $\mathbb{R}^I$.

A.3. *Structure constants*. Let E and F be finite-dimensional vector spaces over the real numbers. Let $\mathrm{Hom}(E, F)$ denote the vector space of all linear mappings from E to F. The vector space $\mathrm{Hom}(E, E)$ is also an algebra and is denoted by $\mathrm{End}(E)$. The group of all nonsingular elements in $\mathrm{End}(E)$ is denoted by $\mathrm{GL}(E)$ (the *general linear group over* E).

Let G be a finite group and let ρ: $G \to \mathrm{GL}(E)$ be a *group representation of* G *on* E; that is, ρ is a group homomorphism. It is standard to call E a G-*space* and to write gx instead of $\rho(g)(x)$, $g \in G$, $x \in \mathbb{R}^I$. By extension of ρ to the group algebra $\mathbb{R}^{(G)}$, E becomes, in a canonical way, a (left) module over $\mathbb{R}^{(G)}$ [cf., e.g., Bourbaki (1958), Section 13, Number 1, Remarque]. On the other hand, if E is a module over $\mathbb{R}^{(G)}$, then by restriction of $\mathbb{R}^{(G)}$-multiplication to G, we have a group representation of G on E. Since G is finite (compact) the module is semisimple. We can thus use all well-known results from the theory of semisimple modules.

Suppose that we also have a group representation τ: $G \to \mathrm{GL}(F)$ of G on F. A linear mapping f: $E \to F$ that commutes with G, that is, $f(gx) = gf(x)$ for all $x \in E$ and $g \in G$, is called G-linear. Let $\mathrm{Hom}_G(E, F)$ denote the subspace of $\mathrm{Hom}(E, F)$ consisting of all G-linear mappings f: $E \to F$. The

G-spaces E and F are called *disjoint* if $\mathrm{Hom}_G(E, F) = \{0\}$. The G-spaces E and F are called *equivalent* if there exists a bijective mapping in $\mathrm{Hom}_G(E, F)$. A subspace $N \subseteq E$ is called a G-subspace if $GN = N$. Every G-subspace has a complement which is also a G-subspace (this is in fact the definition of semisimple). A representation of G on a vector space $S \neq \{0\}$ is called *irreducible* if there are no nontrivial G-subspaces, that is, if S and $\{0\}$ are the only G-subspaces of S. By Schur's Lemma, $D = \mathrm{End}_G(S)$ is then a division algebra over the real numbers, and is therefore, by Frobenius' Theorem, isomorphic either to the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$ or the quaternion division algebra $\mathbb{H}$. The dimension d of D as a vector space over $\mathbb{R}$ is thus 1, 2 or 4, and S is also a vector space over D. Any two irreducible representations are thus either equivalent or disjoint.

Let $\mathfrak{M}$ be a complete set of disjoint irreducible representations $\rho_\mu: G \to \mathrm{GL}(S_\mu)$ of G on real vector spaces S_μ, $\mu \in \mathfrak{M}$. Since G is finite, $\mathfrak{M}$ is also finite. Let $\mu \in \mathfrak{M}$ and denote by d_μ the dimension of $D_\mu = \mathrm{End}_G(S_\mu)$ as a vector space over the real numbers, and by n_μ the dimension of S_μ as a vector space over D_μ. The dimension of S_μ as a vector space over $\mathbb{R}$ is thus $d_\mu n_\mu$. Note that the integers $d_\mu, n_\mu, \mu \in \mathfrak{M}$, only depend on the group G and not on the representation ρ.

A representation of G on a vector space T is called *isotypic* if all irreducible G-subspaces of T are equivalent. If $\mu \in \mathfrak{M}$ is equivalent with an irreducible G-subspace of T, we say that T is of type μ or of type S_μ. In this case T is a direct sum of equivalent irreducible G-subspaces; that is, T is isomorphic to $S_\mu^{\oplus p}$ where $p \in \mathbb{N}$ is the number of times the irreducible representation ρ_μ is contained in the representation of G on T. The number p is called the *G-dimension of* T. Note that $p = 0$ corresponds to $T = \{0\}$. The decomposition into irreducible components is in general not unique. Two nonzero isotypic representations are disjoint if and only if they are of different type.

Any G-space E has a *unique* decomposition into a direct sum $E = \oplus(T_\mu | \mu \in \mathfrak{M})$ of disjoint isotypic G-subspaces, where T_μ is of type S_μ, $\mu \in \mathfrak{M}$. The unique G-subspaces T_μ are called *isotypic components* of E. The family of *structure constants* for the G-space E are the family $((p_\mu, d_\mu, n_\mu) \mid \mu \in \mathfrak{M})$ of triples of nonnegative integers, where p_μ is the G-dimension of T_μ, $\mu \in \mathfrak{M}$. The family $(p_\mu | \mu \in \mathfrak{M})$ of nonnegative integers is called the *G-dimension* of the G-space E. When $E = \mathbb{R}^I$ and the representation is orthogonal, that is, $\rho(G) \subseteq \mathrm{O}(I)$, the decompositions of $\mathbb{R}^I$ into its isotypic components becomes orthogonal. Note that the dimension of E is $\sum(p_\mu d_\mu n_\mu | \mu \in \mathfrak{M})$.

Let $F = \oplus(U_\mu | \mu \in \mathfrak{M})$ be the unique decomposition of the G-space F into isotypic components. For every $f \in \mathrm{Hom}_G(E, F)$ we have $f(T_\mu) \subseteq U_\mu$ for all $\mu \in \mathfrak{M}$. On the other hand, if $f_\mu \in \mathrm{Hom}_G(T_\mu, U_\mu)$, $\mu \in \mathfrak{M}$, then the direct sum $f = \oplus(f_\mu | \mu \in \mathfrak{M})$: $E \to F$ is G-linear. Thus E and F are disjoint if and only if for all $\mu \in \mathfrak{M}$, either $T_\mu = 0$ or $U_\mu = 0$.

Some of the details of the above can be found in Andersson (1975b). A main reference to the theory of semisimple modules is Bourbaki (1958), Sections 1–4.

A.4. *Regularity and structure constants*. The family of structure constants $((p_\mu, d_\mu, n_\mu) \mid \mu \in \mathfrak{M})$ for the G-space $\mathbb{R}^I$ is also called the *structure constants for the GS-model* (2.9). It follows from the structure theorem for GS models in Andersson (1975b), Theorem 4.18, that the GS model (2.9) is regular if and only if

$$\forall\, \mu \in \mathfrak{M}\colon\ n_\mu \geq p_\mu. \tag{A.3}$$

A.5. *The generalized variance*. From the norming constants in Andersson, Brøns and Jensen [(1983), page 414], one can easily derive the moments of the determinant of a real, complex or quaternion Wishart distributed variable, that is, the moments of the *generalized variance*. These moments are well known for the real (that is, the standard) Wishart distribution and the complex Wishart distribution. Furthermore it then follows from the structure theorem in Andersson [(1975b), Theorem 4.18], that the moments $\mathrm{E}(\det(\hat{\Sigma})^\alpha)$, where $\hat{\Sigma}$ is the ML estimator in the model (2.9) and $\alpha > 0$, are given by

$$\begin{aligned} &\mathrm{E}\Big(\det(\hat{\Sigma})^{\alpha}\Big)\Big/\det(\Sigma)^{\alpha} \\ &\quad = \prod\left(\prod\left(\left(\frac{2}{d_\mu n_\mu}\right)^{d_\mu n_\mu \alpha} \frac{\Gamma\big(d_\mu(n_\mu - j_\mu + 1)/2 + d_\mu n_\mu \alpha\big)}{\Gamma\big(d_\mu(n_\mu - j_\mu + 1)/2\big)} \right| j_\mu = 1, \dots, p_\mu\right) \Bigg| \mu \in \mathfrak{M}\Bigg). \end{aligned} \tag{A.4}$$

In fact $\det(\hat{\Sigma})/\det(\Sigma)$ has the same distribution as a product of independent variables

$$\prod\Big(\prod\Big(X_{j_\mu}^{d_\mu n_\mu} | j_\mu = 1, \dots, p_\mu\Big)\Big|\, \mu \in \mathfrak{M}\Big),$$

where X_{j_μ} follows a χ^2 distribution with $d_\mu(n_\mu - j_\mu + 1)$ degrees of freedom and scale $(d_\mu n_\mu)^{-1}$.

REMARK A.3. Note that (A.3) and (A.4) only depend on the *essential* subfamily of structure constants $((p_\mu, d_\mu, n_\mu) | \mu \in \mathfrak{M}^*)$, where $\mathfrak{M}^* = \{\mu \in \mathfrak{M} | p_\mu > 0\}$. Any subfamily of the structure constants containing the essential subfamily will therefore also be called the structure constants for the model (2.9).

A.6. *Independent repetitions*. Let N be a nonempty finite set and let $\mathscr{S}(N)$ be the symmetric group over N, that is, the group of all permutations of the set N. Let $\rho(\sigma) \in \mathrm{O}(N)$ be the permutation matrix corresponding to $\sigma \in \mathscr{S}(N)$, that is, the matrix with entries $\rho(\sigma)_{\nu\nu'} = 1$ when $\sigma(\nu) = \nu'$ and equal to zero otherwise, $\nu, \nu' \in N$. Let $\mathscr{G}_N$ denote the semidirect product of $\{-1, 1\}^N$ and $\mathscr{S}(N)$, that is, the group with underlying set $\{-1, 1\}^N \times \mathscr{S}(N)$

and composition given by

$$\big((\varepsilon_\nu|\nu\in N),\sigma\big)\circ\big((\varepsilon'_\nu|\nu\in N),\sigma'\big)=\Big(\big(\varepsilon_\nu\varepsilon'_{\sigma^{-1}(\nu)}|\nu\in N\big),\sigma\sigma'\Big),$$

where $(\varepsilon_\nu|\nu\in N),(\varepsilon'_\nu|\nu\in N)\in\{-1,1\}^N$ and $\sigma,\sigma'\in\mathscr{S}(N)$. The group $\mathscr{G}_N$ has a representation $\rho^{\mathscr{G}(N)}$ on $\mathbb{R}^N$ given by

$$\rho^{\mathscr{G}(N)}\big((\varepsilon_\nu|\nu\in N),\sigma\big)\big)=\operatorname{diag}\big(\varepsilon_\nu|\nu\in N\big)\rho(\sigma),$$

where $((\varepsilon_\nu|\nu\in N),\sigma)\in\mathscr{G}_N$. The set of matrices $\rho^{\mathscr{G}(N)}(\mathscr{G}_N)$ is the group of all $N\times N$ permutation matrices with signs $\pm$ on the nonzero entries.

The product group $G(N)=G\times\mathscr{G}_N$ has an orthogonal representation $\rho^{G(N)}$ on $\mathbb{R}^{I\times N}$ given by

$$\rho^{G(N)}((g,\mathfrak{g}))=\rho(g)\otimes\rho^{\mathscr{G}(N)}(\mathfrak{g}),$$

where $(g,\mathfrak{g})\in G(N)$.

It is easy to see that $\mathrm{P}_{G(N)}(I\times N)=\{\Sigma\otimes 1_N|\Sigma\in\mathrm{P}_G(I)\}$. Note that

$$I\times N=\dot{\bigcup}(I|\nu\in N), \tag{A.5}$$

and hence $\Sigma\otimes 1_N=\operatorname{diag}(\Sigma|\nu\in N)$, $\Sigma\in\mathrm{P}_G(I)$.

Thus the GS model $(\mathrm{N}(\Gamma)|\Gamma\in\mathrm{P}_{G(N)}(I\times N))$ is, except for the reparametrization

$$\begin{aligned}\mathrm{P}_G(I)&\leftrightarrow\mathrm{P}_{G(N)}(I\times N),\\ \Sigma&\mapsto\Sigma\otimes 1_N,\end{aligned} \tag{A.6}$$

the same as N independent repetitions of the model (2.9), that is, the model

$$\big(\mathrm{N}(\Sigma)^{\otimes N}|\Sigma\in\mathrm{P}_G(I)\big) \tag{A.7}$$

with observation space $\mathbb{R}^{I\times N}$. The smoothing mapping Ψ corresponding to $G(N)$ is determined by

$$\begin{aligned}\Psi(S)&=\left(\frac{1}{n}\sum\big(\psi(S_{\nu\nu})|\nu\in N\big)\right)\otimes 1_N\\ &=\psi\left(\frac{1}{n}\sum\big(S_{\nu\nu}|\nu\in N\big)\right)\otimes 1_N,\end{aligned} \tag{A.8}$$

where $S=(S_{\nu\nu'}\mid(\nu,\nu')\in N\times N)\in\mathrm{PS}(I\times N)$ is the partition of S according to (A.5) and $n=|N|$.

Finally the structure constants for the model (A.7) become $((p_\mu,d_\mu,n_\mu n)\mid\mu\in\mathfrak{M})$. This follows since the representation $\rho_\mu\otimes\rho^{\mathscr{G}(N)}$ of $G(N)$ on $S_\mu\otimes\mathbb{R}^N$ is irreducible and its commutator is isomorphic to D_μ, $\mu\in M$.

Thus the smoothing mapping ψ and the structure constants $((p_\mu,d_\mu,n_\mu)\mid\mu\in\mathfrak{M})$ for the model (2.9) determine the smoothing function Ψ and the structure constants $((p_\mu,d_\mu,n_\mu n)|\mu\in\mathfrak{M})$ for the model (A.7). In particular, they determine the regularity condition

$$\forall\,\mu\in M:\ nn_\mu\geq p_\mu, \tag{A.9}$$

the estimator $\hat{\Sigma}$ for $\Sigma\in\mathrm{P}_G(I)$:

$$\hat{\Sigma}(y)=\psi\big(\overline{S}(y)\big),$$

where $\overline{S}(y) = (1/n)\Sigma(x_\nu x_\nu' | \nu \in N)$ and $y = (x_\nu | \nu \in N) \in \mathbb{R}^{I \times N}$ is a partition of y according to (A.5), and the moments of the generalized variance $\det(\hat{\Sigma})$ [cf. the reparametrization (A.6)] in the model (A.7).

REMARK A.4. The inequalities (A.9) determine the minimum number of repetitions for regularity of the model (A.7).

The literature contains several specific examples of GS models and testing problems.

1. Complete symmetry, Wilks (1946).
2. Compound symmetry of types I and II, Votaw (1948).
3. Circular symmetry, Olkin and Press (1969).
4. Independent multivariate observations, Anderson (1984).
5. Independent multivariate observations with the same covariance matrix, Anderson (1984).
6. Complex normal distributions, Goodman (1963), Khatri (1965a, b), Andersson (1975b), Andersson, Brøns and Jensen (1983), Andersson and Perlman (1984).
7. Quaternion normal distributions, Andersson (1975b), Andersson, Brøns and Jensen (1983).
8. Multivariate complete symmetry, Arnold (1973). This model is, except for a simple rearrangement, the same as compound symmetry of type II, (cf. (2) above).
9. Circular symmetry in blocks, Olkin (1973). This model is not the multivariate version of circular symmetry. It is in fact the model of *dihedral* symmetry, that is, circular symmetry and invariance under reverse ordering [cf. Perlman (1987)].

For all of these models, the smoothing function and the structure constants can easily be obtained.

A.7. *Testing additional symmetry*. Let $G \subseteq G_0$ be a subgroup of the finite group G_0 and let $\rho_0\colon G_0 \to \mathrm{O}(I)$ be an orthogonal group representation extending $\rho\colon G \to \mathrm{O}(I)$. The GS model

$$\left(\mathrm{N}(\Sigma) | \Sigma \in \mathrm{P}_{G_0}(I)\right) \tag{A.10}$$

is a submodel of the GS model (2.9) since $\mathrm{P}_{G_0}(I) \subseteq \mathrm{P}_G(I)$. Since $\mathscr{A}(G)x \subseteq \mathscr{A}(G_0)x$ for all $x \in \mathbb{R}^I$, it follows that if $x \in \mathbb{R}^I$ is regular for the model (2.9), then it is also regular for the model (A.10). In particular regularity of the model (2.9) implies regularity of the submodel (A.10). We shall consider the statistical problem of testing

$$\mathrm{H}_0\colon \Sigma \in \mathrm{P}_{G_0}(I)$$

against

$$\mathrm{H}\colon \Sigma \in \mathrm{P}_G(I).$$

Let ψ_0 denote the smoothing mapping corresponding to the representation ρ_0 [cf. (2.10)]. Assume that (2.9), and therefore also the submodel (A.10), is regular. Then the likelihood ratio statistic Q exists with probability 1, and from Remark A.1 it follows that Q is given by

$$Q = \frac{\det(\hat{\Sigma})^{1/2}}{\det(\hat{\Sigma}_0)^{1/2}},$$

where $\hat{\Sigma}_0(x) = \psi_0(xx') = \psi_0(S(x))$ is the ML estimator for $\Sigma \in \mathrm{P}_{G_0}(I)$ in the model (A.10). Since the distribution of $\det(\hat{\Sigma})$ and $\det(\hat{\Sigma}_0)$ both have scale $\det(\Sigma_0)$ under H_0, it follows that Q is an ancillary statistic under H_0. Since $\hat{\Sigma}_0$ is complete and sufficient, it follows from Basu's lemma that $\hat{\Sigma}_0$ and Q are independent under H_0. [This also can be obtained using Lemma 3 in Andersson, Brøns and Jensen (1983) with $X = \Omega$, $Y = \mathrm{P}_{G_0}(I)$ and $G = \mathrm{GL}_{G_0}(I)$, which implies that $\hat{\Sigma}_0$ is independent of the maximal invariant statistic of the action of $G = \mathrm{GL}_{G_0}(I)$ on $\mathrm{P}_G(I)$. Since Q is invariant under this action, the above result follows.]

Thus $\det(\hat{\Sigma}_0)$ is independent of Q and $\det(\hat{\Sigma})^{1/2} = \det(\hat{\Sigma}_0)^{1/2}Q$, hence it follows that the αth moment of Q, $\alpha \in [0, \infty)$, is given by

$$\mathrm{E}(Q^\alpha) = \frac{\mathrm{E}\Big(\det(\hat{\Sigma})^{\alpha/2}\Big)}{\mathrm{E}\Big(\det(\hat{\Sigma}_0)^{\alpha/2}\Big)}$$

so

$$\begin{aligned}\mathrm{E}(Q^\alpha) &= \Bigg[\prod\Bigg(\prod\Bigg(\left(\frac{2}{d_\mu n_\mu}\right)^{d_\mu n_\mu \alpha/2} \frac{\Gamma\big(d_\mu(n_\mu - j_\mu + 1)/2 + d_\mu n_\mu \alpha/2\big)}{\Gamma\big(d_\mu(n_\mu - j_\mu + 1)/2\big)} \Bigg| \\ &\qquad\qquad j_\mu = 1, \dots, p_\mu\Bigg)\Bigg|\mu \in \mathfrak{M}\Bigg)\Bigg] \\ &\quad\times\Bigg[\prod\Bigg(\prod\Bigg(\left(\frac{2}{d_\nu n_\nu}\right)^{d_\nu n_\nu \alpha/2} \frac{\Gamma\big(d_\nu(n_\nu - j_\nu + 1)/2 + d_\nu n_\nu \alpha/2\big)}{\Gamma\big(d_\nu(n_\nu - j_\nu + 1)/2\big)} \Bigg| \\ &\qquad\qquad j_\nu = 1, \dots, p_\nu\Bigg)\Bigg|\nu \in \mathfrak{N}\Bigg)\Bigg]^{-1},\end{aligned}$$

where $((p_\nu, d_\nu, n_\nu)|\nu \in \mathfrak{N})$ are the structure constants for the representation ρ_0 of G_0. Note that the moments satisfy the classical Box conditions since $\Sigma(p_\mu d_\mu n_\nu | \mu \in \mathfrak{M}) = |I| = \Sigma(p_\nu d_\nu n_\nu | \nu \in \mathscr{N})$. One can thus use the approximation to the distribution of $-2\log(Q)$ given for example in Anderson (1984), pages 311–316. Another, and probably better, method is the saddle point approximation given in Jensen (1991).

A.8. *Linear GS models.*

LEMMA A.5. *Let $L \subseteq \mathbb{R}^I$ be an $\mathrm{M}_G(I)$-subspace, that is, $\mathrm{M}_G(I)L = L$, and let $D \in \mathrm{P}_G(I)$. Then $L^\perp = \{y \in \mathbb{R}^I | \forall\ x \in L\colon\ y'Dx = 0\}$ does not depend on $D \in \mathrm{P}_G(I)$, that is, $L^\perp = \{y \in \mathbb{R}^I \mid \forall\ x \in L\colon\ y'x = 0\}$, and it is an $\mathrm{M}_G(I)$-subspace.*

PROOF. First note that $A \in \mathrm{M}_G(I)$ if and only if $A' \in \mathrm{M}_G(I)$. It follows from Lemma A.1 that there exists $A \in \mathrm{GL}_G(I)$ such that $D = A'A$. Then

$$\begin{aligned} L^\perp &= \{y \in \mathbb{R}^I \mid \forall\ x \in L\colon\ y'A'Ax = 0\} \\ &= \{y \in \mathbb{R}^I \mid \forall\ x \in L\colon\ y'(A'Ax) = 0\} \\ &= \{y \in \mathbb{R}^I \mid \forall\ x \in L\colon\ y'x = 0\}, \end{aligned}$$

since $(A'A)L = L$. Next let $y \in L^\perp$ and $A \in \mathrm{M}_G(I)$. For every $x \in L$ we have that $(Ay)'x = y'A'x = 0$ since $A'x \in L$. □

Let $L \subseteq \mathbb{R}^I$ be an $\mathrm{M}_G(I)$-subspace and let P_L denote the orthogonal projection matrix determined by L, that is, $(x - P_L x)'z = 0$ for all $z \in L$. The statistical model

$$(\mathrm{N}(\xi, \Sigma) \mid (\xi, \Sigma) \in L \times \mathrm{P}_G(I)) \tag{A.11}$$

is called the linear group symmetry (LGS) model determined by L and G.

THEOREM A.2. *The maximum likelihood estimator $(\hat{\xi}(x), \hat{\Sigma}(x))$ of $(\xi, \Sigma) \in L \times \mathrm{P}_G(I)$ in the model (A.11) exists if and only if $\psi((x - P_L x)(x - P_L x)')$ is nonsingular. In this case, it is unique and given by $\hat{\xi}(x) = P_L(x)$ and $\hat{\Sigma}(x) = \psi((x - P_L x)(x - P_L x)')$.*

PROOF. The likelihood function $\mathbf{L}\colon (L \times \mathrm{P}_G(I)) \times \mathbb{R}^I \to (0, \infty)$ is given by

$$\mathbf{L}((\xi, \Sigma), x) = \det(\Sigma)^{-1/2}\exp\left(-\tfrac{1}{2}\mathrm{tr}\left(\Sigma^{-1}(x - \xi)(x - \xi)'\right)\right).$$

Since $\Sigma^{-1} \in \mathrm{P}_G(I)$ and $L^\perp$ is the orthogonal complement to L w.r.t. all $\Sigma^{-1} \in \mathrm{P}_G(I)$, it follows that

$$\begin{aligned} &\mathrm{tr}\left(\Sigma^{-1}(x - \xi)(x - \xi)'\right) \\ &\qquad = \mathrm{tr}\left(\Sigma^{-1}\psi\left((x - P_L x)(x - P_L x)'\right)\right) \\ &\qquad\quad + \mathrm{tr}\left(\Sigma^{-1}(P_L x - \xi)(P_L x - \xi)'\right). \end{aligned} \tag{A.12}$$

The problem is then to maximize

$$\begin{aligned} \mathbf{L}((\xi, \Sigma), x) = \det(\Sigma)^{-1/2}\exp\Big(&-\tfrac{1}{2}\mathrm{tr}\left(\Sigma^{-1}\psi\left((x - P_L x)(x - P_L x)'\right)\right) \\ &-\tfrac{1}{2}\mathrm{tr}\left(\Sigma^{-1}(P_L x - \xi)(P_L x - \xi)'\right)\Big), \end{aligned}$$

where $(\xi, \Sigma) \in L \times \mathrm{P}_G(I)$ and $x \in \mathbb{R}^I$ is fixed. Now it is seen that the ML estimator for $\xi \in L$ always exists and is uniquely given by $\hat{\xi}(x) = P_L(x)$. The theorem now follows from Theorem A.1 by replacing $\psi(xx')$ by $\psi((x - P_L x)(x - P_L x)')$. $\square$

REMARK A.5. The maximum of the likelihood function is

$$\mathbf{L}\big(\hat{\Sigma}(x), x\big) = \det\big(\hat{\Sigma}(x)\big)^{-1/2} \exp\big(-\tfrac{1}{2}|I|\big).$$

An element $x \in \mathbb{R}^I$ is called *regular* w.r.t. L and G if $x - P_L x \in \Omega$. The set of regular elements w.r.t. L then has the form $\Omega_L \times L \subseteq \mathbb{R}^I$, where $\Omega_L \subseteq L^\perp$. An element $x \in \mathbb{R}^I$ is then regular w.r.t. L and G if and only if $\psi((x - P_L x)(x - P_L x)')$ is nonsingular.

LEMMA A.6. *The set $\Omega_L \times L$ is an open subset of $\mathbb{R}^I$. If $\Omega_L \times L \neq \emptyset$ then the Lebesgue measure on $\mathbb{R}^I$ is concentrated on $\Omega_L \times L$,* that is, $\mathbb{R}^I \setminus (\Omega_L \times L)$ *has Lebesgue measure zero.*

PROOF. Since $\Omega_L = \{z \in L^\perp \,|\, \psi(zz') \text{ is nonsingular}\}$ and $\mathrm{P}_G(I)$ is open in $\mathrm{PS}_G(I)$, it follows that Ω_L is open in $L^\perp$ and thus that $\Omega_L \times L$ is open in $\mathbb{R}^I$. Let $x \in \Omega_L \times L$. Then the mapping $f\colon \mathscr{A}(G) \to \mathbb{R}^I$ given by $A \mapsto A(x - P_L x)$ is linear and surjective. Since the Lebesgue measure on $\mathscr{A}(G)$ is concentrated on the nonsingular elements $\mathscr{A}(G)^*$ of $\mathscr{A}(G)$ (cf. Lemma A.3) and f is linear and surjective, it follows that the Lebesgue measure on $\mathbb{R}^I$ is concentrated on $f(\mathscr{A}(G)^*) \subseteq \Omega_L \times L$, and thus the $\mathbb{R}^I \setminus (\Omega_L \times L)$ has Lebesgue measure zero. $\square$

PROPOSITION A.3. *Consider the LGS model* (A.11). *Either the ML estimator $(\hat{\xi}(x), \hat{\Sigma}(x))$ exists with probability one w.r.t. all* $\mathrm{N}(\xi, \Sigma)$, $(\xi, \Sigma) \in L \times \mathrm{P}_G(I)$, *or else it does not exist for any $x \in \mathbb{R}^I$.*

The proof follows from Theorem A.2 and Lemma A.6.

A LGS model with $\Omega_L \neq \emptyset$, that is, such that the ML estimator exists with probability 1, is called a *regular* LGS model. Note that the concept of regularity only depends on the representation of G on $\mathbb{R}^I$ and L.

REMARK A.6. It follows from the theory of semisimple modules, for example as explained in Bourbaki (1958) or Andersson (1975b), that any $\mathrm{M}_G(I)$-module L corresponds to D_μ-subspaces $L_\mu \subseteq S_\mu$, $\mu \in \mathfrak{M}$, where we define $L_\mu = \{0\}$ when $p_\mu = 0$. Let l_μ be the D_μ-dimension of L_μ, $\mu \in \mathfrak{M}$. We shall call the family $(l_\mu | \mu \in \mathfrak{M})$, the $\mathrm{M}_G(I)$-*dimension* of L. In particular, an $\mathrm{M}_G(I)$-subspace L of $\mathbb{R}^I$ is an $\mathrm{M}_G(I)$-module.

The group representation ρ on $\mathbb{R}^I$ and the $\mathrm{M}_G(I)$-subspace L thus define a family $((p_\mu, d_\mu, n_\mu, l_\mu) | \mu \in \mathfrak{M})$ of four-tuples of nonnegative integers (cf. Section A.3). This family is called the family of *structure constants* for the model (A.11). Note that the dimension of L is $\Sigma(p_\mu d_\mu l_\mu | \mu \in \mathfrak{M})$.

REMARK A.7. It follows from Andersson [(1975b), Theorem 4.18], that the model (A.11) is regular if and only if

$$\forall\ \mu \in \mathfrak{M}\colon\ n_\mu \geq p_\mu + l_\mu .$$

Also the distribution of $\det(\hat{\Sigma})/\det(\Sigma)$ can be obtained from Andersson (1975b) as a product of independent variables

$$\prod\Big(\prod\Big(X_{j_\mu}^{d_\mu n_\mu} | j_\mu = 1,\dots,p_\mu\Big)\Big|\ \mu \in \mathfrak{M}\Big),$$

where X_{j_μ} follows a χ^2 distribution with $d_\mu(f_\mu - j_\mu + 1)$ degrees of freedom and scale $(d_\mu n_\mu)^{-1}$ and $f_\mu = n_\mu - l_\mu$. Thus

$$\begin{aligned} &\mathrm{E}\Big(\det(\hat{\Sigma})^{\alpha}\Big)\Big/\det(\Sigma)^{\alpha} \\ \text{(A.13)}\qquad &= \prod\left(\prod\left(\left(\frac{2}{d_\mu n_\mu}\right)^{d_\mu n_\mu \alpha} \frac{\Gamma\big(d_\mu(f_\mu - j_\mu + 1)/2 + d_\mu n_\mu \alpha\big)}{\Gamma\big(d_\mu(f_\mu - j_\mu + 1)/2\big)}\right| \right. \\ &\qquad\qquad \left.\left. j_\mu = 1,\dots,p_\mu\right)\right|\mu \in \mathfrak{M}\Big). \end{aligned}$$

The theory of LGS models can be developed in a similar fashion to the theory of GS models presented above.

APPENDIX B

Proof of Theorem 4.1. The proof follows the same ideas as the proof of Theorem 3.1 in AP (1995a). Simply follow the Appendix in AP (1995a) with the following changes and improvements:

Replace $\mathrm{M}(I \times N)$ by $\mathbb{R}^I$ and replace the definition of Ω in (A.1) of AP (1995a) with the definition (3.18). Furthermore replace $\mathbf{GL}(\mathscr{K})$, $\mathbf{P}(\mathscr{K})$ and $\mathbf{P}_{[K]}$, $K \in \mathscr{I}(\mathscr{K})$, by $\mathrm{GL}_{G,\mathscr{K}}(I)$, $\mathrm{P}_{G,\mathscr{K}}(I)$ and $\mathrm{P}_G([K])$, $K \in \mathscr{I}(\mathscr{K})$, respectively. Note that $\mathrm{GL}_{G,\mathscr{K}}(I)$ acts on Ω since for $A \in \mathrm{GL}_{G,\mathscr{K}}(I)$ and $x \in \Omega$, $\mathscr{A}_K(G) A_K x_K = A_K \mathscr{A}_K(G) x_K \supseteq A_K E_K = E_K$ for all $K \in \mathscr{I}(\mathscr{K})$.

The next step is to transform the normal distributions in the model (4.1) by $\pi_T\colon \Omega \to \Omega/\mathscr{T}$. The hypothesis $H_{\mathscr{K}}$ in AP (1995a) is replaced by $\mathrm{H}_{G,\mathscr{K}}$.) We notice that it is not necessary to represent the orbit projection $\pi_{\mathscr{T}}$ explicitly. Thus as we shall see, Lemma A.1 in AP (1995a) is unnecessary in the present proof and in the proof of Theorem 3.1 in AP (1995a). The calculations from (A.7) to (A.8) in AP (1995a) are now replaced by the following.

First note that $\beta = \otimes(\lambda_K | K \in \mathscr{I}(\mathscr{K}))$ is a Haar measure on $\mathscr{T}$, where λ_K is a Lebesgue measure on $\mathrm{M}_G([K] \times \langle K \rangle)$, $K \in \mathscr{I}(\mathscr{K})$. Let $x \in \Omega$. We shall first use that $(Tx)_{[K]} = x_{[K]} + T_{[K\rangle} x_{\langle K \rangle}$ and $(Tx)_{\langle K \rangle} = T_{\langle K \rangle} x_{\langle K \rangle}$ together with (2.6) and the translation invariance of λ_K, $K \in \mathscr{I}(\mathscr{K})$, to obtain

$$\begin{aligned} & q(\pi_{\mathscr{T}}(x)) \\ &\quad = \det(\Sigma)^{-1/2} \int \exp\left(-\tfrac{1}{2}\operatorname{tr}(\Sigma^{-1}(Tx)(Tx)')\right)\beta(dT) \\ &\quad = \prod \Bigg(\det(\Sigma_{[K]\cdot})^{-1/2} \int \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}(x_{[K]} + T_{[K\rangle} x_{\langle K\rangle})(\cdots)'\right)\right) \\ &\qquad\qquad \times \lambda_K(dT_{[K\rangle}) \Big| K \in \mathscr{I}(\mathscr{K})\Bigg), \end{aligned} \tag{B.1}$$

where the order of integration should be determined by a never-increasing listing of the elements in $\mathscr{I}(\mathscr{K})$. Let $K \in \mathscr{I}(\mathscr{K})$. It follows from (3.5), (3.7) and (A.1) that

$$\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot} x_{[K]\circ} T_{[K\rangle} x_{\langle K\rangle}\right) = 0, \tag{B.2}$$

and that

$$\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot} x_{[K]\circ} x'_{[K]\circ}\right) = \operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot} \psi(x)_{[K]\circ}\right), \tag{B.3}$$

where

$$x_{[K]\circ} = x_{[K]} - \hat{R}_{[K\rangle} x_{\langle K\rangle}, \tag{B.4}$$

and $\hat{R}_{[K\rangle}$ is the unique solution in $\mathrm{M}_G([K] \times \langle K\rangle)$ to (3.10). The factor corresponding to $K \in \mathscr{I}(\mathscr{K})$ in (B.1) can be rewritten as follows:

$$\begin{aligned} & \int \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}(x_{[K]} + T_{[K\rangle} x_{\langle K\rangle})(\cdots)'\right)\right)\lambda_K(dT_{[K\rangle}) \\ &\quad = \int \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}\left(x_{[K]\circ} + \left(\hat{R}_{[K\rangle} + T_{[K\rangle}\right)x_{\langle K\rangle}\right)(\cdots)'\right)\right)\lambda_K(dT_{[K\rangle}) \\ &\quad = \int \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}(x_{[K]\circ} + T_{[K\rangle} x_{\langle K\rangle})(\cdots)'\right)\right)\lambda_K(dT_{[K\rangle}) \\ &\quad = \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}\psi(xx')_{[K]\circ}\right)\right) \\ &\qquad \times \int \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}(T_{[K\rangle} x_{\langle K\rangle})(\cdots)'\right)\right)\lambda_K(dT_{[K\rangle}), \end{aligned}$$

where the first equality follows from (B.4), the second from the translation invariance of λ_K and the third from (B.2) and (B.3). The density q of $\pi_{\mathscr{T}}(P)$ w.r.t. the quotient measure λ/β on $\Omega/\mathscr{T}$ is thus given by

$$\begin{aligned} & q(\pi_{\mathscr{T}}(x)) \\ &\quad = \prod \Bigg(\det(\Sigma_{[K]\cdot})^{-1/2} \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}\psi(xx')_{[K]\circ}\right)\right) \\ &\qquad\qquad \times \int \exp\left(-\tfrac{1}{2}\operatorname{tr}\left(\Sigma^{-1}_{[K]\cdot}(T_{[K\rangle} x_{\langle K\rangle})(\cdots)'\right)\right)\lambda_K(dT_{[K\rangle}) \Big| \\ &\qquad\qquad\qquad K \in \mathscr{I}(\mathscr{K})\Bigg). \end{aligned} \tag{B.5}$$

The next step is to represent the transformed measure $\pi_{\mathscr{T}}(P) = q \cdot (\lambda/\beta)$ as $q_1 \cdot \nu$, where ν is an invariant measure under the induced action of $\mathscr{A}$ on $\Omega/\mathscr{T}$. The first part of the second paragraph on page 35 in AP (1995a) is unchanged until the formula for mod φ_A. This formula is in general not valid for the present paper. Next define

$$n(\Lambda_{[K]}) = \int \exp\Big(-\tfrac{1}{2}\mathrm{tr}\big(\Lambda_{[K]}^{-1} T_{[K\rangle} T'_{[K\rangle}\big)\Big)\lambda_K(dT_{[K\rangle})$$

and

$$\begin{aligned} &c(\Lambda_{[K]}, x_{\langle K\rangle}) \\ &\quad = n(\Lambda_{[K]})^{-1}\int \exp\Big(-\tfrac{1}{2}\mathrm{tr}\big(\Lambda_{[K]}^{-1}(T_{[K\rangle}x_{\langle K\rangle})(T_{[K\rangle}x_{\langle K\rangle})'\big)\Big)\lambda_K(dT_{[K\rangle}), \end{aligned}$$

where $\Lambda_{[K]} \in \mathrm{P}_G([K])$. Since $c(A_{[K]}\Lambda_{[K]}A'_{[K]}, x_{\langle K\rangle}) = c(\Lambda_{[K]}, x_{\langle K\rangle})$, $A_{[K]} \in \mathrm{GL}_G([K])$, it follows from Proposition A.2 that $c(\Lambda_{[K]}, x_{\langle K\rangle})$ does not depend on $\Lambda_{[K]} \in \mathrm{P}_G([K])$. Furthermore, replace m in AP (1995a) by

$$\begin{aligned} &m(\pi_{\mathscr{T}}(x)) \\ &\quad = \prod\Bigg(\det\big(\hat{\Sigma}(x)_{[K]\cdot}\big)^{1/2}\bigg(\int\exp\Big(-\tfrac{1}{2}\mathrm{tr}\big(\hat{\Sigma}(x)_{[K]\cdot}^{-1}(T_{[K\rangle}x_{\langle K\rangle})(\cdots)'\big)\Big) \\ &\qquad\qquad\qquad\qquad \times\lambda_K(dT_{[K\rangle})\bigg)^{-1}\Bigg| K \in \mathscr{I}(\mathscr{K})\Bigg). \end{aligned}$$

The density q_1 is therefore given by

$$q_1(\pi_{\mathscr{T}}(x)) = \frac{n(\Sigma_{[K]\cdot})}{n\big(\hat{\Sigma}(x)_{[K]\cdot}\big)}\frac{\det\big(\hat{\Sigma}(x)_{[K]\cdot}\big)^{1/2}}{\det\big(\Sigma_{[K]\cdot}\big)^{1/2}}\exp\bigg(-\frac{1}{2}\mathrm{tr}\Big(\Sigma_{[K]\cdot}^{-1}\hat{\Sigma}(x)_{[K]\cdot}\Big)\bigg).$$

Now apply Lemma A.2 in AP (1995a) with the identifications $G = G' = \mathscr{A}$, $X = \Omega/\mathscr{T}$, $X' = \times(\mathrm{P}_G([K]) | K \in \mathscr{I}(\mathscr{K}))$, $\varphi = \mathrm{id}_{\mathscr{A}}$, and with ψ given by $\psi(\pi_{\mathscr{T}}(x)) = (\hat{\Sigma}(x)_{[K]\cdot} | K \in \mathscr{I}(\mathscr{K}))$, to conclude that the induced action of $\mathscr{A}$ on $\Omega/\mathscr{T}$ is proper.

Next replace the right-hand side of (A.11) in AP (1995a) by

$$\text{(B.6)} \qquad \frac{n(\Sigma_{[K]\cdot})}{n(\Lambda_{[K]})}\frac{\det\big(\hat{\Sigma}_{[K]\cdot}\big)^{1/2}}{\det\big(\Lambda_{[K]}\big)^{1/2}}\exp\bigg(-\frac{1}{2}\mathrm{tr}\Big(\Sigma_{[K]\cdot}^{-1}\Lambda_{[K]}\Big)\bigg).$$

Note that ν_K is an invariant measure on $\mathrm{P}_G([K])$ under the action of $\mathrm{GL}_G([K])$ and that the distribution of $\hat{\Sigma}_{[K]\cdot}$ has the density (B.6) w.r.t. $\nu_K(d\Lambda_{[K]})$, $K \in \mathscr{I}(\mathscr{K})$. These distributions are all generalized Wishart distributions (cf. Remark A.2). This completes the proof of Theorem 4.1. □

Acknowledgments. We thank Michael D. Perlman and G. Gerard Wojnar for careful reading of the manuscript and Hans Brøns for many insightful suggestions.

REFERENCES

ANDERSEN, H. H., HØJBJERRE, M., SØRENSEN, D. and ERIKSEN, P. S. (1995). *Linear and Graphical Models for the Multivariate Complex Normal Distribution*. *Lecture Notes in Statist*. **101**. Springer, New York.

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

ANDERSSON, S. A. (1975a). Forelæningnoter i flerdimensional statistisk analyse. Lecture notes, Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

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

ANDERSSON, S. A. (1976). Forelæningsnoter i flerdimensional statistisk analyse. Lecture notes, Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

ANDERSSON, S. A. (1978). On the mathematical foundation of multivariate statistical analysis. Inst. Mathematical Statistics, Univ. Copenhagen. (Lecture notes presented at the meeting in Lunteren, Holland, Nov. 1978.)

ANDERSSON, S. A. (1990). The lattice structure of orthogonal linear models and orthogonal variance component models. *Scand. J. Statist*. **17** 287–319.

ANDERSSON, S. A. (1992). Normal statistical models given by group symmetry. Lecture notes, DMW–Seminar, Gunzburg, Germany.

ANDERSSON, S. A., BRØNS, H. and JENSEN, S. T. (1975). En algebraisk teori for normale statistiske modeller. Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

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

ANDERSSON, S. A., MADIGAN, D., PERLMAN, M. D. and TRIGGS, C. M. (1995a). On the relation between conditional independence models determined by finite distributive lattices and directed acyclic graphs. *J. Statist. Plann. Inference* **48** 25–46.

ANDERSSON, S. A., MADIGAN, D., PERLMAN, M. D. and TRIGGS, C. M. (1995b). A graphical characterization of lattice conditional independence models. *Ann. Math. Artificial Intelligence* **21** 27–50.

ANDERSSON, S. A. and PERLMAN, M. D. (1984). Two testing problems relating the real and complex multivariate normal distributions. *J. Multivariate Anal*. **14** 21–51.

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. (1994). Normal linear models with lattice conditional independence restrictions. In *Multivariate Analysis and Its Applications*. **24** 97–110. IMS, Hayward, CA.

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). Normal linear regression models with recursive graphical Markov structure. *J. Multivariate Anal*. To appear.

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

BARNDORFF-NIELSEN, O. E. (1978). *Information and Exponential Families in Statistical Theory*. Wiley, New York.

BERTELSEN, A. (1989). On non-null distributions connected with testing reality of a complex normal distribution. *Ann. Statist*. **17** 929–936.

BOURBAKI, N. (1958). *Éléments de Mathématique*. Algébre, Ch. 8. Hermann, Paris.

BOURBAKI, N. (1963). *Éléments de Mathématique*. Integration, Ch. 7 à 8. Hermann, Paris.

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

COX, D. R. and WERMUTH, N. (1993). Linear dependencies represented by chain graphs. *Statist. Sci*. **8** 204–283.

FRYDENBERG, M. (1990). The chain graph Markov property. *Scand. J. Statist*. **17** 333–353.

GOODMAN, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Statist*. **34** 152–177.

HYLLEBERG, B., JENSEN, M. and ØRNBØL, E. (1993). Graphical symmetry models. M.S. thesis, Aalborg Univ.

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

JENSEN, S. T. (1973). Forelæsninger i matematisk statistik. Lecture notes, Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

JENSEN, S. T. (1974). Forelæsningsnoter i flerdimensional statistisk analyse. Lecture notes, Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

JENSEN, S. T. (1977). Flerdimensional statistisk analyse. Lecture notes, Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

JENSEN, S. T. (1983). Symmetrimodeller. Lecture notes, Inst. Mathematical Statistics, Univ. Copenhagen. (In Danish.)

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

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

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

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

LAURITZEN, S. L. and WERMUTH, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *Ann. Statist.* **17** 31–57.

MADSEN, J. (1996). Invariant normal models with recursive graphical Markov structure. Unpublished manuscript.

MASSAM, H. and NEHER, E. (1995). Lattice conditional independence models with covariance matrices on symmetric cones. Unpublished manuscript.

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

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

PERLMAN, M. D. (1987). Group symmetry covariance models. *Statist. Sci.* **2** 421–425.

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

WHITTAKER, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley, Chichester.

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

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405
E-MAIL: standers@ucs.indiana.edu
jmadsen@math.ku.dk