

## Research Article

# Gaussian Covariance Faithful Markov Trees

Dhafer Malouche<sup>1,2</sup> and Bala Rajaratnam<sup>2</sup>

<sup>1</sup> *Unité de Recherche Signaux et Systèmes (U2S), Ecole Supérieure de la Statistique et de l'Analyse de l'Information (ESSAI), Ecole Nationale d'Ingénieurs de Tunis (ENIT), 6 Rue des Métiers, Charguia II 2035, Tunis Carthage, Ariana, Tunis 1002, Tunisia*

<sup>2</sup> *Department of Statistics, Department of Environmental Earth System Science, Woods Institute for the Environment, Stanford University, Stanford, CA 94305, USA*

Correspondence should be addressed to Bala Rajaratnam, brajarat@stanford.edu

Received 30 May 2011; Accepted 9 August 2011

Academic Editor: Junbin B. Gao

Copyright © 2011 D. Malouche and B. Rajaratnam. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical models are useful for characterizing conditional and marginal independence structures in high-dimensional distributions. An important class of graphical models is covariance graph models, where the nodes of a graph represent different components of a random vector, and the absence of an edge between any pair of variables implies marginal independence. Covariance graph models also represent more complex conditional independence relationships between subsets of variables. When the covariance graph captures or reflects all the conditional independence statements present in the probability distribution, the latter is said to be *faithful* to its covariance graph—though in general this is not guaranteed. Faithfulness however is crucial, for instance, in model selection procedures that proceed by testing conditional independences. Hence, an analysis of the faithfulness assumption is important in understanding the ability of the graph, a discrete object, to fully capture the salient features of the probability distribution it aims to describe. In this paper, we demonstrate that multivariate Gaussian distributions that have trees as covariance graphs are necessarily faithful.

## 1. Introduction

Markov random fields and graphical models are widely used to represent conditional independences in a given multivariate probability distribution (see [1–5], to name just a few). Many different types of graphical models have been studied in the literature. Concentration graphs encode conditional independence between pairs of variables given the remaining ones. Formally, let us consider a random vector  $\mathbf{X} = (X_v, v \in V)'$  with a probability distribution  $P$  where  $V$  is a finite set representing the random variables in  $\mathbf{X}$ . An undirected graph  $G_0 = (V, E_0)$  is called the *covariance graph* (see [1, 6–11]) associated with the probability

distribution  $P$  if the set of edges  $E_0$  is constructed as follows:

$$(u, v) \notin E_0 \iff X_u \perp\!\!\!\perp X_v. \quad (1.1)$$

Note that  $(u, v) \notin E$  means that the vertices  $u$  and  $v$  are not adjacent in  $G$ .

The *concentration graph* associated with  $P$  is an undirected graph  $G = (V, E)$ , where  $V$  is the set of vertices and each vertex represents one variable in  $\mathbf{X}$ . The set  $E$  is the set of edges (between the vertices in  $V$ ) constructed using the pairwise rule: for pair  $(u, v) \in V \times V$ ,  $u \neq v$ ,

$$(u, v) \notin E \iff X_u \perp\!\!\!\perp X_v \mid \mathbf{X}_{V \setminus \{u, v\}}, \quad (1.2)$$

where  $\mathbf{X}_{V \setminus \{u, v\}} := (X_w, w \neq u \text{ and } w \neq v)'$ .

Note that the subscript zero is invoked for covariance graphs (i.e.,  $G_0$  versus  $G$ ) as the definition of covariance graphs does not involve conditional independences.

Both concentration and covariance graphs not only are used to encode pairwise relationships between pairs of variables in the random vector  $\mathbf{X}$ , but as we will see below, these graphs can also be used to encode conditional independences that exist between subsets of variables of  $\mathbf{X}$ . First, we introduce some definitions.

The multivariate distribution  $P$  is said to satisfy the ‘‘intersection property’’ if for, any subsets  $A, B, C$ , and  $D$  of  $V$  which are pairwise disjoint,

$$\mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_{C \cup D} \text{ and } \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_C \mid \mathbf{X}_{B \cup D} \text{ then } \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_{B \cup C} \mid \mathbf{X}_D. \quad (1.3)$$

We will call the *intersection* property (see [2]) in (1.3) above the *concentration intersection* property in this paper in order to differentiate it from another property that is satisfied by  $P$  when studying covariance graph models. Though this property can be further relaxed, we will retain the terminology used in [2].

We first define the concept of separation on graphs. Let  $A, B$ , and  $S$  denote a pairwise disjoint set of vertices. We say that a set  $S$  *separates*  $A$  and  $B$  if all paths connecting  $A$  and  $B$  in  $G$  intersect  $S$ , that is,  $A \perp_G B \mid S$ . (This is not to be confused with stochastic independence which is denoted by  $\perp\!\!\!\perp$  as compared to  $\perp_G$ .) Now, let  $P$  satisfy the concentration intersection property. Then, for any triplet  $(A, B, S)$  of subsets of  $V$  pairwise disjoint, if  $S$  separates  $A$  and  $B$  in the concentration graph  $G$  associated with  $P$ , then the random vector  $\mathbf{X}_A = (X_v, v \in A)'$  is independent of  $\mathbf{X}_B = (X_v, v \in B)'$  given  $\mathbf{X}_S = (X_v, v \in S)'$ . This latter property is called *concentration global Markov* property and is formally defined as

$$A \perp_G B \mid S \implies \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_S. \quad (1.4)$$

Kauermann [6] shows that if  $P$  satisfies the following property: for any triplet  $(A, B, S)$  of subsets of  $V$  pairwise disjoint,

$$\text{if } \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \text{ and } \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_C \text{ then } \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_{B \cup C}, \quad (1.5)$$

then, for any triplet  $(A, B, S)$  of subsets of  $V$  pairwise disjoint, if  $S$  separates  $A$  and  $B$  in the covariance graph  $G_0$  associated with  $P$ , then  $\mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}$ . This latter property is called the *covariance global Markov property* and can be written formally as follows:

$$A \perp_{G_0} B \mid S \implies \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}. \quad (1.6)$$

In parallel to the concentration graph case, property (1.5) will be called the *covariance intersection* property and is sometimes also referred to as the *composition* property. Even if  $P$  satisfies both intersection properties, the covariance and concentration graphs may not be able to capture or reflect all the conditional independences present in the distribution; that is, there may exist one or more conditional independences present in the probability distribution that does not correspond to any separation statement in either  $G$  or  $G_0$ . Equivalently, a lack of a separation statement in either  $G$  or  $G_0$  does not necessarily imply a conditional independence. On the contrary case when no other conditional independence exists in  $P$  except the ones encoded by the graph, we classify  $P$  as a *faithful* probability distribution to its graphical model (see [12]). More precisely, we say that  $P$  is concentration faithful to its concentration graph if, for any triplet  $(A, B, S)$  of subsets of  $V$  pairwise disjoint, the following statement holds:

$$A \perp_G B \mid S \iff \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_S. \quad (1.7)$$

Similarly,  $P$  is said to be *covariance faithful* to its covariance graph  $G_0$  if, for any triplet  $(A, B, S)$  of subsets of  $V$  pairwise disjoint, the following statement holds:

$$A \perp_{G_0} B \mid S \iff \mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}. \quad (1.8)$$

A natural question of both theoretical and applied interest in probability theory is to understand the implications of the *faithfulness* assumption. This assumption is fundamental since it yields a bijection between the probability distribution  $P$  and the graph  $G$  in terms of the independences that are present in the distribution. In this paper, we show that when  $P$  is a multivariate Gaussian distribution, whose covariance graph is a tree, it is necessarily covariance faithful, that is, such probability distributions satisfy property (1.8). Equivalently, the associated covariance graph  $G$  is fully able to capture all the conditional independences present in the multivariate distribution  $P$ . This result can be considered as a dual of a previous probabilistic result proved by Becker et al. [13] for concentration graphs that demonstrates that Gaussian distributions having concentration trees (i.e., the concentration graph is a tree) are necessarily *concentration faithful* to its concentration graph (implying that property (1.7) is satisfied). This result was proved by showing that Gaussian distributions satisfy two types of conditional independence properties: the intersection property and the decomposable transitivity property. The approach in the proof of the main result of this paper is vastly different from the one used for concentration graphs (see [13]). Indeed, a naïve or unsuspecting reader could mistakenly think that the result for covariance trees follows simply by replacing the covariance matrix with its inverse in the result in Becker et al. [13]. This is of course incorrect and, in some sense, equivalent to saying that a matrix and its inverse are the same. The covariance matrix encodes marginal independences whereas the inverse covariance matrix encodes conditional independences. These are very different

models. Moreover, the former is a curved exponential family model whereas the latter is a natural exponential family model.

The outline of this paper is as follows. Section 2 presents graph theory preliminaries. Section 2.2 gives a brief overview of covariance and concentration graphs associated with multivariate Gaussian distributions. The proof of the main result of this paper is given in Section 3. Section 4 concludes by summarizing the results in the paper and the implications thereof.

## 2. Preliminaries

### 2.1. Graph Theoretic Concepts

This section introduces notation and terminology that is required in subsequent sections. An undirected graph  $G = (V, E)$  consists of two sets  $V$  and  $E$ , with  $V$  representing the set of vertices, and  $E \subseteq (V \times V) \setminus \{(u, u), u \in V\}$  the set of edges satisfying, for all  $(u, v) \in E \iff (v, u) \in E$ . For  $u, v \in V$ , we write  $u \sim_G v$  when  $(u, v) \in E$  and we say that  $u$  and  $v$  are *adjacent* in  $G$ . A *path* connecting two distinct vertices  $u$  and  $v$  in  $G$  is a sequence of distinct vertices  $(u_0, u_1, \dots, u_n)$ , where  $u_0 = u$  and  $u_n = v$ , where, for every  $i = 0, \dots, n - 1$ ,  $u_i \sim_G u_{i+1}$ . Such a path will be denoted  $p = p(u, v, G)$  and we say that  $p(u, v, G)$  connects  $u$  and  $v$  or alternatively  $u$  and  $v$  are *connected* by  $p(u, v, G)$ . We also denote by  $\mathcal{P}(u, v, G)$  the set of paths between  $u$  and  $v$ . We now proceed to define the subclass of graphs known as trees. Let  $G = (V, E)$  be an undirected graph. The graph  $G$  is called a *tree* if any pair of vertices  $(u, v)$  in  $G$  is connected by exactly one path; that is,  $|\mathcal{P}(u, v, G)| = 1$  for all  $u, v \in V$ . A subgraph of  $G$  *induced* by a subset  $U \subseteq V$  is denoted by  $G_U = (U, E_U)$ ,  $U \subseteq V$  and  $E_U = E \cap (U \times U)$ . A *connected component* of a graph  $G$  is the largest subgraph  $G_U = (U, E_U)$  of  $G$  such that each pair of vertices can be connected by at least one path in  $G_U$ . We now state a lemma, without proof, that is needed in the proof of the main result of this paper.

**Lemma 2.1.** *Let  $G = (V, E)$  be an undirected graph. If  $G$  is a tree, then any subgraph of  $G$  induced by a subset of  $V$  is a union of connected components, each of which are trees (or what we will refer to as a “union of tree connected components”).*

For a connected graph, a *separator* is a subset  $S$  of  $V$  such that there exists a pair of nonadjacent vertices  $u$  and  $v$  such that  $u, v \notin S$  and

$$\forall p \in \mathcal{P}(u, v, G), \quad p \cap S \neq \emptyset. \quad (2.1)$$

If  $S$  is a separator, then it is easily verified that every  $S' \supseteq S$  such that  $S' \subseteq V \setminus \{u, v\}$  is also a separator.

### 2.2. Gaussian Concentration and Covariance Graphs

In this section, we present a brief overview of concentration and covariance graphs in the case when the probability distribution  $P$  is multivariate Gaussian. Consider a random variable  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , where  $\boldsymbol{\mu} \in \mathbb{R}^{|V|}$  and  $\Sigma = (\sigma_{uv}) \in \mathbb{P}^+$ , where  $\mathbb{P}^+$  denotes the cone of positive definite matrices. Without loss of generality, we will assume that  $\boldsymbol{\mu} = \mathbf{0}$ . Gaussian distributions can also be parameterized by the inverse of the covariance matrix  $\Sigma$  denoted by  $K = \Sigma^{-1} = (k_{uv})$ . The matrix  $K$  is called the *precision* or *concentration* matrix. It is well known

(see [2]) that for any pair of variables  $(X_u, X_v)$ , where  $u \neq v$ ,  $X_u \perp\!\!\!\perp X_v \mid \mathbf{X}_{V \setminus \{u,v\}} \Leftrightarrow k_{uv} = 0$ . Hence, the concentration graph  $G = (V, E)$  can be constructed simply using the precision matrix  $K$  and the following rule:  $(u, v) \notin E \Leftrightarrow k_{uv} = 0$ . Furthermore, it can be easily deduced from a classical result (see [2]) that for any Gaussian concentration graph model the pairwise Markov property in (1.2) is equivalent to the concentration global Markov property in (1.4).

As seen earlier in (1.1) covariance graphs on the other hand are constructed using pairwise marginal independence relationships. It is also well known that, for multivariate Gaussian distributions,  $X_u \perp\!\!\!\perp X_v \Leftrightarrow \sigma_{uv} = 0$ . Hence, in the Gaussian case, the covariance graph  $G_0 = (V, E_0)$  can be constructed using the following rule:  $(u, v) \notin E_0 \Leftrightarrow \sigma_{uv} = 0$ . It is also easily seen that Gaussian distributions satisfy the covariance intersection property defined in (1.5). Hence, Gaussian covariance graphs can also encode conditional independences according to the following rule: for any triplet  $(A, B, S)$  of subsets of  $V$  pairwise disjoint, if  $S$  separates  $A$  and  $B$  in the covariance graph  $G_0$ , then  $\mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}$ .

### 3. Gaussian Covariance Faithful Trees

We now proceed to study the faithfulness assumption in the context of multivariate Gaussian distributions and when the associated covariance graphs are trees. The main result of this paper, presented in Theorem 3.1, proves that multivariate Gaussian probability distributions having tree covariance graphs are necessarily faithful to their covariance graphs; that is, all of the independence and dependences in  $P$  can be read by using graph separation. We now formally state Theorem 3.1. The proof follows shortly after a series of lemmas/theorem(s) and an illustrative example.

**Theorem 3.1.** *Let  $\mathbf{X}_V = (X_v, v \in V)'$  be a random vector with Gaussian distribution  $P = \mathcal{N}_{|V|}(\mu, \Sigma)$ . Let  $G_0 = (V, E_0)$  be the covariance graph associated with  $P$ . If  $G_0$  is a disjoint union of trees, then  $P$  is covariance faithful to  $G_0$ .*

The proof of Theorem 3.1 requires, among others, a result that gives a method to compute the covariance matrix  $\Sigma$  from the precision matrix  $K$  using the paths in the concentration graph  $G$ . The result can also be easily extended to show that the precision matrix  $K$  can be computed from the covariance matrix  $\Sigma$  using the paths in the covariance graph  $G_0$ . We now formally state this result.

**Lemma 3.2.** *Let  $\mathbf{X}_V = (X_v, v \in V)'$  be a random vector with Gaussian distribution  $P = \mathcal{N}_{|V|}(\mu, \Sigma)$ , where  $\Sigma$  and  $K = \Sigma^{-1}$  are positive definite matrices. Let  $G = (V, E)$  and  $G_0 = (V, E_0)$  denote, respectively, the concentration and covariance graph associated with the probability distribution of  $\mathbf{X}_V$ . For all  $(u, v)$  in  $V \times V$ ,*

$$k_{uv} = \sum_{p \in \mathcal{P}(u,v,G_0)} (-1)^{|p|+1} |\sigma|_p \frac{|\Sigma \setminus p|}{|\Sigma|}, \quad \sigma_{uv} = \sum_{p \in \mathcal{P}(u,v,G)} (-1)^{|p|+1} |k|_p \frac{|K \setminus p|}{|K|}, \quad (3.1)$$

where, if  $p = (u_0, \dots, u_n)$ ,  $|\sigma|_p = \sigma_{u_0 u_1} \sigma_{u_1 u_2} \dots \sigma_{u_{n-1} u_n}$ ,  $|k|_p = k_{u_0 u_1} k_{u_1 u_2} \dots k_{u_{n-1} u_n}$ ,  $K \setminus p = (k_{uv}, (u, v) \in (V \setminus p) \times (V \setminus p))$  and  $\Sigma \setminus p = (\sigma_{uv}, (u, v) \in (V \setminus p) \times (V \setminus p))$  denote, respectively,  $K$  and  $\Sigma$  with rows and columns corresponding to variables in path  $p$  omitted. The determinant of a zero-dimensional matrix is defined to be 1.

The lemma above follows immediately from a basic result in linear algebra which gives the cofactor expression for the inverse of a square matrix. In particular, for an invertible matrix  $A$ , its inverse  $A^{-1}$  can be expressed as follows:

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A), \quad \text{where } \text{adj}(A) \text{ denotes the adjoint of } A. \quad (3.2)$$

A simple proof can be found in Brualdi and Cvetkovic [14]. The result has been rediscovered in other contexts (see [15]), but, as noted above, it follows immediately from the expression for the inverse of a matrix.

The proof of our main theorem (Theorem 3.1) also requires the results proved in the lemma below.

**Lemma 3.3.** *Let  $\mathbf{X}_V = (X_v, v \in V)'$  be a random vector with Gaussian distribution  $P = \mathcal{N}_{|V|}(\mu, K = \Sigma^{-1})$ . Let  $G_0 = (V, E_0)$  and  $G = (V, E)$  denote, respectively, the covariance and concentration graphs associated with  $P$ , then*

- (i)  $G$  and  $G_0$  have the same connected components,
- (ii) if a given connected component in  $G_0$  is a tree, then the corresponding connected component in  $G$  is complete and vice versa.

*Proof.* Proof of (i): the fact that  $G_0$  and  $G$  have the same connected components can be deduced from the matrix structure of the covariance and the precision matrix. The connected components of  $G_0$  correspond to block diagonal matrices in  $\Sigma$ . Since  $K = \Sigma^{-1}$ , then, by properties of inverting partitioned matrices,  $K$  also has the same block diagonal matrices as  $\Sigma$  in terms of the variables that constitute these matrices. These blocks correspond to distinct components in  $G$  and  $G_0$ . Hence, both matrices have the same connected components.

Proof of (ii): let us assume now that the covariance graph  $G_0$  is a tree, hence it is a connected graph with only one connected component. We will prove that the concentration graph  $G$  is complete by using Lemma 3.2 and computing any coefficient  $k_{uv}$  ( $u \neq v$ ). Since  $G_0$  is a tree, there exists exactly one path between any two vertices  $u$  and  $v$ . We will denote this path as  $p = (u_0 = u, \dots, u_n = v)$ . Then, by Lemma 3.2,

$$k_{uv} = (-1)^{n+1} \sigma_{u_0 u_1} \dots \sigma_{u_{n-1} u_n} \frac{|\Sigma \setminus p|}{|\Sigma|}. \quad (3.3)$$

First, note that the determinants of the matrices in (3.3) are all positive since principal minors of positive definite matrices are positive. Second, since we are considering a path in  $G_0$ ,  $\sigma_{u_{i-1} u_i} \neq 0$ , for all  $i = 1, \dots, n$ . Using these two facts, we deduce from (3.3) that  $k_{uv} \neq 0$  for all  $(u, v) \in E$ . Hence,  $u$  and  $v$  are adjacent in  $G$  for all  $(u, v) \in E$ . The concentration graph  $G$  is therefore complete. The proof that when  $G$  is assumed to be a tree implying that  $G_0$  is complete follows similarly.  $\square$

We now give an example illustrating the main result in this paper (Theorem 3.1).

*Example 3.4.* Consider a Gaussian random vector  $\mathbf{X} = (X_1, \dots, X_8)'$  with covariance matrix  $\Sigma$  and its associated covariance graph (which is a tree) as given in Figure 1(a).

Consider the sets  $A = \{1, 2\}$ ,  $B = \{5\}$ , and  $S = \{4, 6\}$ . Note that  $S$  does not separate  $A$  and  $B$  in  $G_0$  as any path from  $A$  and  $B$  does not intersect  $S$ . Hence, we cannot

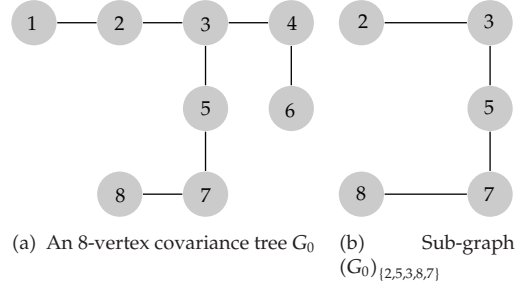


Figure 1: Covariance graph in Example 3.4.

use the covariance global Markov property to claim that  $\mathbf{X}_A$  is not independent of  $\mathbf{X}_B$  given  $\mathbf{X}_{V \setminus (A \cup B \cup S)}$ . This is because the covariance global Markov property allows us to read conditional independences present in a distribution if a separation is present in the graph. It is not an “if and only if” property in the sense that the lack of a separation in the graph does not necessarily imply the lack of the corresponding conditional independence. We will show however that in this example  $\mathbf{X}_A$  is indeed not independent of  $\mathbf{X}_B$  given  $\mathbf{X}_{V \setminus (A \cup B \cup S)}$ . In other words, we will show that the graph has the ability to capture this conditional dependence present in the probability distribution  $P$ .

Let us now examine the relationship between  $X_2$  and  $X_5$  given  $\mathbf{X}_{\{3,7,8\}}$ . Note that in this example  $V \setminus (A \cup B \cup S) = \{3, 8, 7\}$ ,  $2 \in A$ , and  $5 \in B$ . Note that the covariance graph associated with the probability distribution of the random vector  $(X_2, X_5, \mathbf{X}_{\{3,8,7\}})'$  is the subgraph represented in Figure 1(b) and can be obtained directly as a subgraph of  $G_0$  induced by the subset  $\{2, 5, 3, 7, 8\}$ .

Since 2 and 5 are connected by exactly one path in  $(G_0)_{\{2,5,3,7,8\}}$ , that is,  $p = (2, 3, 5)$ , then the coefficient  $k_{25|387}$ , that is, the coefficient between 2 and 5 in inverse of the covariance matrix of  $(X_2, X_5, \mathbf{X}_{\{3,8,7\}})'$ , can be computed using Lemma 3.2 as follows:

$$k_{25|387} = (-1)^{2+1} \sigma_{23} \sigma_{35} \frac{|\Sigma(\{8, 7\})|}{|\Sigma(\{2, 5, 3, 8, 7\})|} \tag{3.4}$$

where  $\Sigma(\{7, 8\})$  and  $\Sigma(\{2, 5, 3, 8, 7\})$  are, respectively, the covariance matrices of the Gaussian random vectors  $(X_7, X_8)'$  and  $(X_2, X_5, \mathbf{X}_{\{3,8,7\}})'$ . Hence,  $k_{25|387} \neq 0$  since the right hand side of the equation in (3.4) is different from zero. Hence,  $X_2 \not\perp\!\!\!\perp X_5 \mid \mathbf{X}_{\{3,8,7\}}$ .

Now, recall that, for any Gaussian random vector  $\mathbf{X}_V = (X_u, u \in V)'$ ,

$$\mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B \mid \mathbf{X}_S \quad \text{iff } \forall (u, v) \in A \times B, \quad X_u \perp\!\!\!\perp X_v \mid \mathbf{X}_S, \tag{3.5}$$

where  $A, B$ , and  $S$  are pairwise disjoint subsets of  $V$ . The contrapositive of (3.5) yields

$$X_2 \not\perp\!\!\!\perp X_5 \mid \mathbf{X}_{\{3,7,8\}} \implies X_{\{1,2\}} \not\perp\!\!\!\perp X_5 \mid \mathbf{X}_{\{3,7,8\}}. \tag{3.6}$$

Hence, we conclude that since  $\{4, 6\}$  does not separate  $\{1, 2\}$  and  $\{5\}$ ,  $\mathbf{X}_{\{1,2\}}$  is not independent of  $X_5$  given  $\mathbf{X}_{\{3,7,8\}}$ . Thus, we obtain the desired result:

$$\{1, 2\} \not\perp_{G_0} \{5\} \mid \{4, 6\} \implies \mathbf{X}_{\{1,2\}} \not\perp X_5 \mid \mathbf{X}_{\{3,7,8\}}. \quad (3.7)$$

We now proceed to the proof of Theorem 3.1.

*Proof of Theorem 3.1.* Without loss of generality, we assume that  $G_0$  is a connected tree. Let us assume to the contrary that  $P$  is not covariance faithful to  $G_0$ , then there exists a triplet  $(A, B, S)$  of pairwise disjoint subsets of  $V$ , such that  $\mathbf{X}_A \perp \mathbf{X}_B \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}$ , but  $S$  does not separate  $A$  and  $B$  in  $G_0$ , that is,

$$\mathbf{X}_A \perp \mathbf{X}_B \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}, \quad A \not\perp_{G_0} B \mid S. \quad (3.8)$$

As  $S$  does not separate  $A$  and  $B$  and since  $G_0$  is a connected tree, then there exists a pair of vertices  $(u, v) \in A \times B$  such that the single path  $p$  connecting  $u$  and  $v$  in  $G_0$  does not intersect  $S$ ; that is,  $S \cap p = \emptyset$ . Hence,  $p \subseteq V \setminus S = (A \cup B) \cup (V \setminus (A \cup B \cup S))$ . Thus, two cases are possible with regard to where the path  $p$  can lie: either  $p \subseteq A \cup B$  or  $p \cap (V \setminus (A \cup B \cup S)) \neq \emptyset$ . Let us examine both cases separately.

*Case 1* ( $p \subseteq A \cup B$ ). In this case, the entire path between  $u$  and  $v$  lies in  $A \cup B$  and hence we can find a pair of vertices  $(u', v')$  belonging to  $p$  and  $(u', v') \in A \times B$  such that  $u' \sim_{G_0} v'$ . (As an illustration of this point, consider the graph presented in Figure 1(a). Let  $A = \{1, 2\}$ ,  $B = \{3, 5\}$ , and  $S = \{4, 6\}$ . We note that the path  $p = (1, 2, 3, 5)$  lies entirely in  $A \cup B$  and hence we can find two vertices, namely,  $2 \in A$  and  $3 \in B$ , belonging to path  $p$  that are adjacent in  $G_0$ ).

Recall that since  $G_0$  is a tree, any induced graph of  $G_0$  by a subset of  $V$  is a union of tree connected components (see Lemma 2.1). Hence, the subgraph  $(G_0)_W$  of  $G_0$  induced by  $W = \{u', v'\} \cup V \setminus (A \cup B \cup S)$  is a union of tree connected components. As  $u'$  and  $v'$  are adjacent in  $G_0$ , they are also adjacent in  $(G_0)_W$  and belong to the same connected component of  $(G_0)_W$ . (In our example in Figure 1(a) with  $W = \{2, 3, 8, 7\}$ ,  $(G_0)_W$  consists of a union of two connected components with its respective vertices being  $\{2, 3\}$  and  $\{8, 7\}$ .) Hence, the only path between  $u'$  and  $v'$  is precisely the edge  $(u', v')$ . Using Lemma 3.2 to compute the coefficient  $k_{u'v' \mid V \setminus (A \cup B \cup S)}$ , that is,  $(u', v')$ th coefficient in the inverse of the covariance matrix of the random vector  $\mathbf{X}_W = (X_w, w \in W)' = (X_{u'}, X_{v'}, \mathbf{X}_{V \setminus (A \cup B \cup S)})'$ , we obtain,

$$k_{u'v' \mid V \setminus (A \cup B \cup S)} = (-1)^{1+1} \sigma_{u'v'} \frac{|\Sigma(W \setminus \{u', v'\})|}{|\Sigma(W)|}, \quad (3.9)$$

where  $\Sigma(W)$  denotes the covariance matrix of  $\mathbf{X}_W$  and  $\Sigma(W \setminus \{u', v'\})$  denotes the matrix  $\Sigma(W)$  with the rows and the columns corresponding to variables  $X_{u'}$  and  $X_{v'}$  omitted. We can therefore deduce from (3.9) that  $k_{u'v' \mid V \setminus (A \cup B \cup S)} \neq 0$ . Hence,  $X_{u'} \not\perp X_{v'} \mid \mathbf{X}_{V \setminus (A \cup B \cup S)}$ . Now, since  $P$  is Gaussian,  $u' \in A$ , and  $v' \in B$ , we can apply (3.5) to arrive at a contradiction to our initial assumption in (3.8). Note in the case that  $V \setminus (A \cup B \cup S)$  is empty the path  $p$  has to lie entirely in  $A \cup B$ . This is because by assumption  $p$  does not intersect  $S$ . The case when  $p$  lies in  $A \cup B$  is covered in Case 1 and hence it is assumed that  $V \setminus (A \cup B \cup S) \neq \emptyset$ . (As an illustration of this



point, consider once more the graph presented in Figure 1(a). Consider  $A = \{1, 2\}$ ,  $B = \{7, 8\}$ , and  $S = \{4, 6\}$ . Here,  $V \setminus (A \cup B \cup S) = \{3, 5\}$  and the path  $p = (1, 2, 3, 5, 7, 8)$  connecting  $A$  and  $B$  intersects  $V \setminus (A \cup B \cup S)$ .

*Case 2* ( $p \cap (V \setminus (A \cup B \cup S)) \neq \emptyset$  and  $V \setminus (A \cup B \cup S) \neq \emptyset$ ). In this case, there exists a pair of vertices  $(u', v') \in A \times B$  with  $u', v' \in p$ , such that the vertices  $u'$  and  $v'$  are connected by exactly one path  $p' \subseteq p$  in the induced graph  $(G_0)_W$  of  $G_0$  by  $W = \{u', v'\} \cup V \setminus (A \cup B \cup S)$  (see Lemma 2.1). (In our example in Figure 1 with  $A = \{1, 2\}$ ,  $B = \{7, 8\}$ , and  $S = \{4, 6\}$ , the vertices  $u'$  and  $v'$  will correspond to vertices 2 and 7, respectively, and  $p' = (2, 3, 5, 7)$ , which is a path entirely contained in  $\{u', v'\} \cup (V \setminus (A \cup B \cup S))$ ).

Let us now use Lemma 3.2 to compute the coefficient  $k_{u'v'|V \setminus (A \cup B \cup S)}$ , that is, the  $(u', v')$ -coefficient in the inverse of the covariance matrix of the random vector  $\mathbf{X}_W = (X_w, w \in W)' = (X_{u'}, X_{v'}, \mathbf{X}_{V \setminus (A \cup B \cup S)})'$ . We obtain that

$$k_{u'v'|V \setminus (A \cup B \cup S)} = (-1)^{|p'|+1} |\sigma_{p'}| \frac{|\Sigma(W \setminus p')|}{|\Sigma(W)|}, \quad (3.10)$$

where  $\Sigma(W)$  denotes the covariance matrix of  $\mathbf{X}_W$  and  $\Sigma(W \setminus p')$  denotes  $\Sigma(W)$  with the rows and the columns corresponding to variables in path  $p'$  omitted. One can therefore easily deduce from (3.10) that  $k_{u'v'|V \setminus (A \cup B \cup S)} \neq 0$ . Thus,  $X_{u'}$  is not independent of  $X_{v'}$  given  $\mathbf{X}_{V \setminus (A \cup B \cup S)}$ . Hence, once more we obtain a contradiction to (3.5) since  $u' \in A$  and  $v' \in B$ .  $\square$

*Remark 3.5.* The dual result of the theorem above for the case of concentration trees was proved by Becker et al. [13]. We note however that the argument used in the proof of Theorem 3.1 cannot also be used to prove faithfulness of Gaussian distributions that have trees as concentration graphs. The reason for this is as follows. In our proof, we employed the fact that the subgraph  $(G_0)_{\{u,v\} \cup S}$  of  $G_0$  induced by a subset  $\{u, v\} \cup S \subseteq V$  is also the covariance graph associated with the Gaussian subrandom vector of  $\mathbf{X}_V$  as denoted by  $\mathbf{X}_{\{u,v\} \cup S} = (X_w, w \in \{u, v\} \cup S)'$ . Hence, it was possible to compute the coefficient  $k_{uv|S}$  which quantifies the conditional (in)dependence between  $u$  and  $v$  given  $S$ , in terms of the paths in  $(G_0)_{\{u,v\} \cup S}$  and the coefficients of the covariance matrix of  $\mathbf{X}_{\{u,v\} \cup S} = (X_w, w \in \{u, v\} \cup S)'$ . On the contrary, in the case of concentration graphs the sub-graph  $G_{\{u,v\} \cup S}$  of the concentration graph  $G$  induced by  $\{u, v\} \cup S$  is not in general the concentration graph of the random vector  $\mathbf{X}_{\{u,v\} \cup S} = (X_w, w \in \{u, v\} \cup S)'$ . Hence our approach is not directly applicable in the concentration graph setting.

## 4. Conclusion

In this note we looked at the class of multivariate Gaussian distributions that are Markov with respect to covariance graphs and prove that Gaussian distributions which have trees as their covariance graphs are necessarily faithful. The method of proof used in the paper is also vastly different in nature from the proof of the analogous result for concentration graph models. Hence, the approach that is used could potentially have further implications. Future research in this area will explore if the analysis presented in this paper can be extended to other classes of graphs or distributions.

## Acknowledgments

D. Malouche was supported in part by a Fullbright Fellowship Grant 68434144. B. Rajaratnam was supported in part by NSF grants DMS0906392, DMS(CMG)1025465, AGS1003823, NSA H98230-11-1-0194, and SUFSC10-SUSHSTF09SMSCVISG0906.

## References

- [1] D. R. Cox and N. Wermuth, *Multivariate Dependencies*, vol. 67 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, London, UK, 1996.
- [2] S. L. Lauritzen, *Graphical Models*, vol. 17 of *Oxford Statistical Science Series*, The Clarendon Press Oxford University Press, New York, NY, USA, 1996.
- [3] J. Whittaker, *Graphical Models in Applied Multivariate Statistics*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Ltd., Chichester, UK, 1990.
- [4] D. Edwards, *Introduction to Graphical Modelling*, Springer Texts in Statistics, Springer, New York, NY, USA, 2nd edition, 2000.
- [5] B. Rajaratnam, H. Massam, and C. M. Carvalho, "Flexible covariance estimation in graphical Gaussian models," *The Annals of Statistics*, vol. 36, no. 6, pp. 2818–2849, 2008.
- [6] G. Kauermann, "On a dualization of graphical Gaussian models," *Scandinavian Journal of Statistics*, vol. 23, no. 1, pp. 105–116, 1996.
- [7] M. Banerjee and T. Richardson, "On a dualization of graphical Gaussian models: a correction note," *Scandinavian Journal of Statistics*, vol. 30, no. 4, pp. 817–820, 2003.
- [8] N. Wermuth, D. R. Cox, and G. M. Marchetti, "Covariance chains," *Bernoulli*, vol. 12, no. 5, pp. 841–862, 2006.
- [9] D. Malouche, "Determining full conditional independence by low-order conditioning," *Bernoulli*, vol. 15, no. 4, pp. 1179–1189, 2009.
- [10] K. Khare and B. Rajaratnam, "Covariance trees and Wishart distributions on cones," in *Algebraic Methods in Statistics and Probability II*, vol. 516 of *Contemporary Mathematics*, pp. 215–223, American Mathematical Society, Providence, RI, USA, 2010.
- [11] K. Khare and B. Rajaratnam, "Wishart distributions for decomposable covariance graph models," *The Annals of Statistics*, vol. 39, no. 1, pp. 514–555, 2011.
- [12] M. Studený, *Probabilistic Conditional Independence Structures*, Springer, New York, NY, USA, 2004.
- [13] A. Becker, D. Geiger, and C. Meek, "Perfect tree-like Markovian distributions," *Probability and Mathematical Statistics*, vol. 25, no. 2, pp. 231–239, 2005.
- [14] R. A. Brualdi and D. Cvetkovic, *A Combinatorial Approach to Matrix Theory and Its Applications*, Chapman & Hall/CRC, New York, NY, USA, 2008.
- [15] B. Jones and M. West, "Covariance decomposition in undirected Gaussian graphical models," *Biometrika*, vol. 92, no. 4, pp. 779–786, 2005.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

