ATLANTIS PRESS
Atlantis Studies in Uncertainty Modelling, volume 3
Joint Proceedings of the 19th World Congress of the International Fuzzy Systems Association (IFSA), the 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and the 11th International Summer School on Aggregation Operators (AGOP)

A Fuzzy Order for Graphs Based on the Continuous Entropy of Gaussian Markov Random Fields

Juan Baz[a] and **Irene Díaz**[b] and **Susana Montes**[a] and **Raúl Pérez-Fernández**[a]
[a]Department of Statistics and O.R. and Mathematics Didactics, University of Oviedo, Spain
[b]Department of Computer Science, University of Oviedo, Spain
{bazjuan,sirene,montes,perezfernandez}@uniovi.es

Abstract

Gaussian Markov Random Fields over graphs have been widely used in the context of both theoretical and applied Statistics. In this paper, we study the influence of the graph on the continuous entropy of the distribution. In particular, since the continuous entropy is highly dependant on the correlations between adjacent nodes in the graph, we consider the particular case of Gaussian Markov Random Fields with uniform correlation, i.e., Gaussian Markov Random Fields in which such correlations are equal. We define a partial order relation on the set of graphs that orders the graphs according to their contribution to the continuous entropy. We also present a graded version of this order relation that allows to compare incomparable graphs with respect to the original order relation. Finally, an example for the illustration of the graded order relation is provided.

Keywords: Gaussian Markov Random Fields, Uniform correlation, Continuous Entropy, Fuzzy order of graphs.

1 Introduction

A Markov Random Field (MRF) is a random vector for which the conditional statistical independence structure can be expressed as a graph in which the nodes are the indices of the variables [10]. More precisely, the statistical dependence of two variables is linked to the paths that connect those variables in the graph. If the random vector is multivariate Gaussian, the notion of statistical dependence becomes equivalent to that of linear dependence (see, e.g., [16]) and, thus, the conditional independence of the variables can be studied by identifying null elements of the inverse of the covariance matrix [22]. In this case, the random vector is called a Gaussian Markov Random Field (GMRF).

The modelling of problems by means of a GMRF is, in general, computationally attractive, mainly because the inverse of the covariance matrix is positive definite and typically sparse. For this type of matrices efficient algorithms have been developed, for example, to factorize or solve linear systems of equations [8, 21]. For this reason, many applications of GMRFs have been addressed in different areas, such as image processing [9, 18] and disease control [12, 15]. In recent years, many advances have been made towards learning structures [14], sampling methods [2], simulation [19, 23, 24] and inference [7] over a GMRF model.

The continuous entropy is an extension of the Shannon entropy [25] to continuous random vectors. This quantity measures the average surprisal of the random vector [17]. Finding the distribution that maximizes the continuous entropy given some restrictions is a highly studied field, some examples can be found in [1, 4, 29]. Given the variances and some of the covariances between a collection of variables, the GMRF distribution (over the graph where the nodes are adjacent if and only if the covariance is specified) maximizes the continuous entropy, since the resultant covariance matrix maximizes the determinant [6].

We are interested in studying how the graph structure influences the value of the continuous entropy. Unfortunately, the continuous entropy is highly dependant on the variances and the values of the correlations between adjacent nodes in the graph. To avoid the influence of the variances, we focus the study on the correlation matrix rather than the covariance matrix. Additionally, we consider a particular type of GMRF (GMRFs with uniform correlation) in which the correlation between variables associated with adjacent nodes in the graph is always the same. We thus study the correlation matrix of GMRFs with uniform correlation in order to compare the contribution of different graphs to the continuous entropy of the associated distribution. In particular, we focus our study on the function $\tau(G,\rho_0)$, defined in [27], that equals the determinant of the correlation matrix of the distribution as a function of the graph G and the value ρ_0 of the uniform correlation. We compare the contribution to the continuous entropy of two graphs G_1 and G_2 by comparing $\tau(G_1,\rho_0)$ and $\tau(G_2,\rho_0)$ as functions of ρ_0.

The remainder of the paper is organized as follows. In Section 2, we introduce some basic concepts related to GMRFs and, in particular, to GMRFs with uniform correlation. In Section 3, the relation between the continuous entropy of the distribution and the graph is explained in terms of the function $\tau(G,\rho_0)$. The most important properties of $\tau(G,\rho_0)$ are presented. A preorder relation for ordering graphs in terms of their contribution to the continuous entropy is defined in Section 4. A graded extension of this preorder relation is provided in Section 5, focusing on the set of graphs of order five. Finally, some conclusions and open questions are provided in Section 6.

2 Gaussian Markov Random Fields with uniform correlation

A random vector $(X_1,\ldots,X_n)$ has a Multivariate Gaussian random distribution if any linear combination of its components has a Univariate Gaussian distribution [16]. The joint probability density function of a Multivariate Gaussian random vector has the following expression:

$$f(\vec{x}) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp\left(-\frac{(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})}{2}\right),$$

for any $\vec{x}\in\mathbb{R}^n$, where $\vec{\mu}$ is the mean vector and Σ is the covariance matrix, which is always a positive semidefinite matrix. We denote the set of positive definite matrices of order n by $\mathscr{P}_n$. The diagonal elements of the covariance matrix are equal to the variances of the variables (i.e $\Sigma_{i,i}=\sigma_i^2$). The elements of the correlation matrix P are $P_{i,j}=\rho_{i,j}$, where $\rho_{i,j}$ denotes the Pearson's correlation coefficient between the variables X_i and X_j.

Two continuous random vectors $\vec{X}_A$ and $\vec{X}_B$ of dimensions n_A and n_B are said to be conditionally independent [20] given another continuous random vector $\vec{X}_C$ of dimension n_C if there exist $h:\mathbb{R}^{n_A+n_C}\to[0,\infty]$ and $g:\mathbb{R}^{n_B+n_C}\to[0,\infty]$ such that, for any $\vec{x}_a\in\mathbb{R}^{n_A}$, $\vec{x}_b\in\mathbb{R}^{n_B}$ and $\vec{x}_c\in\mathbb{R}^{n_C}$, it holds that the joint density function can be decomposed as follows: $f(\vec{x}_a,\vec{x}_b,\vec{x}_c)=h(\vec{x}_a,\vec{x}_c)g(\vec{x}_b,\vec{x}_c)$.

Given $\vec{X}_V=\{X_i \mid i\in V=\{1,\ldots,n\}\}$ and $A,B,C\subset V$ three pairwise disjoint subsets of V, we denote by $\vec{X}_A\perp\vec{X}_B\mid\vec{X}_C$ the fact that $\vec{X}_A$ and $\vec{X}_B$ are conditionally independent given $\vec{X}_C$. The shorthand $\vec{X}_{-C}$ is used for referring to $\vec{X}_{\{1,\ldots,n\}\setminus C}$.

Interestingly, for a Multivariate Gaussian distribution, conditionally independent variables given all other variables are characterized by the null elements of the inverse of the covariance matrix.

Theorem 1. *[22] Let $\vec{X}$ be a multivariate Gaussian random vector with mean vector $\vec{\mu}$ and covariance matrix Σ. For any $i\neq j$, it holds that $X_i\perp X_j\mid\vec{X}_{-\{i,j\}}\iff(\Sigma^{-1})_{ij}=0$.*

It is possible to associate the components of a Multivariate Gaussian random vector with the nodes of the graph, representing the conditional dependence of the components by means of the edges of the graph. Before explaining how to associate a node and a random vector, we are going to recall some necessary concepts concerning graphs.

A graph is a pair $G=(V,E)$, where V is any set and E is a set of subsets of V of cardinality equal to 2. In particular, we will consider simple graphs, which are graphs containing no graph loops ($(i,i)\notin E$ for any $i\in V$) or multiple edges (E is not a multiset). Any element in V is called a node and the number of nodes is called the order of the graph. Unless otherwise is specified, we assume that the order of the graph G is n. Any element in E is called an edge.

For any $i_1,i_k\in V$, it is said that they are connected if there exists a sequence of nodes $(i_2,\ldots,i_{k-1})$ such that $(i_j,i_{j+1})\in E$ for any $j\in\{1,\ldots,k-1\}$. The sequence of vertices and edges from i_1 to i_k is called a walk. A path is a walk in which none of the vertices or edges is repeated. The length of a path is the number of edges that it contains. If two vertices are connected by a path of length 1 (i.e., by a single edge), the vertices are called adjacent. The distance between two connected nodes is the minimum length among all the paths between them. If $i_1=i_k$, then the path is called a cycle. A graph is called connected if any two nodes are connected and it is called acyclic if it contains no cycles.

The subgraph induced by $A\subset V$ is the graph $G'=(A,E')$, where $E'=\{(i,j)\in E\mid i,j\in A\}$. Given three pairwise disjoint subsets $A,B,C\subset V$, it is said that C separates A and B if any path between any node of A and any node of B contains a node of C. For more details on Graph Theory we refer to [13].

Formally, given a graph $G=(V,E)$, a Gaussian Markov Random Field over G is a Multivariate Gaussian random vector $\vec{X} = (X_i \mid i \in V)$ such that the following restriction, referred to as the Global Markov Property, is fulfilled: $X_A \perp X_B \mid \vec{X}_C$, for any pairwise disjoint $A,B,C \subset V$ such that C separates A and B. This is equivalent to fulfilling the Pairwise Markov Property [22]: $X_i \perp X_j \mid \vec{X}_{-\{i,j\}}$, for any $i,j \in V$ such that $(i,j) \notin E$ and $i \neq j$.

As a result of Theorem 1, given a GMRF, the Global Markov Property is characterized by the null elements of Σ^{-1}.

Example 1. *Consider the bull graph and the 4-complete graph with an isolated node, $K_4 \cup K_1$ (see Figure 1) and the Multivariate Gaussian random vectors $\vec{X}_A$ and $\vec{X}_B$ with mean vectors $\vec{\mu}_A = \vec{\mu}_B = \vec{0}$ and the following covariance matrices:*

$$\Sigma_A = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.35 & 0.4 \\ 0.5 & 1 & 0.5 & 0.7 & 0.4 \\ 0.5 & 0.5 & 1 & 0.35 & 0.8 \\ 0.35 & 0.7 & 0.35 & 1 & 0.28 \\ 0.4 & 0.4 & 0.8 & 0.28 & 1 \end{pmatrix}$$

$$\Sigma_B = \begin{pmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 1 & -0.1 & 0.2 & 0.3 \\ 0 & -0.1 & 1 & 0.4 & 0.5 \\ 0 & 0.2 & 0.4 & 1 & 0.6 \\ 0 & 0.3 & -0.5 & 0.6 & 1 \end{pmatrix}$$

It holds that $\vec{X}_A$ is a GMRF over the bull graph and $\vec{X}_B$ is a GMRF over $K_4 \cup K_1$. This can be verified by inverting their covariance matrices and checking that the elements associated with non adjacent nodes in the graph are equal to 0.

$$(\Sigma_A)^{-1} \approx \begin{pmatrix} 1.5 & -0.5 & -0.5 & 0 & 0 \\ -0.5 & 2.46 & -0.5 & -1.37 & 0 \\ -0.5 & -0.5 & 3.28 & 0 & -2.22 \\ 0 & -1.37 & 0 & 1.96 & 0 \\ 0 & 0 & -2.22 & 0 & 2.78 \end{pmatrix}$$

$$(\Sigma_B)^{-1} \approx \begin{pmatrix} 0.1 & 0 & 0 & 0 & 0 \\ 0 & 1.22 & -0.42 & -0.10 & -0.51 \\ 0 & -0.42 & 1.51 & -0.25 & -0.73 \\ 0 & -0.10 & -0.25 & 1.60 & -0.81 \\ 0 & -0.51 & -0.73 & -0.81 & 2.00 \end{pmatrix}$$

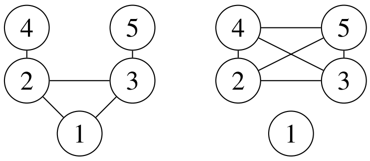

Figure 1: Representation of the bull graph (left-hand side) and $K_4 \cup K_1$ (right-hand side).

We focus on the structure of the graph and just consider GMRFs with uniform correlation, which are GMRFs in which all Pearson's correlation coefficients between adjacent variables are equal, as in [27].

Definition 1. *Let $G = (V,E)$ be a graph. A multivariate Gaussian random vector $\vec{X} = (X_i \mid i \in V)$ is called a GMRF with uniform correlation $\rho_0 \in (-1,1)$ over G if the corresponding correlation matrix P satisfies that:*
- *$P_{i,i} = 1$, for any $i \in V$;*
- *$P_{i,j} = \rho_0$, for any $(i,j) \in E$;*
- *$(P^{-1})_{i,j} = 0$, for any $(i,j) \notin E$ with $i \neq j$.*

Given a graph G and a value ρ_0 for the uniform correlation, we aim at finding the correlation matrix of a GMRF with uniform correlation ρ_0 over G. This problem is a particular case of the GMRF construction problem defined in [26].

Theorem 2. *[26] Let $R,S \in \mathscr{P}_n$ and a graph $G = (V,E)$ of order n. There exists a unique $F \in \mathscr{P}_n$ that satisfies:*
(1) $F_{i,j} = R_{ij}$, for any $(i,j) \in E$ or $i = j$,
(2) $F^{-1}_{i,j} = S_{i,j}$, for any $(i,j) \notin E$ with $i \neq j$.

In particular, setting $S = I_n$, it is concluded that there exists a (unique) correlation matrix of a GMRF with uniform correlation ρ_0 over G as long as there exists a positive definite matrix R satisfying that $R_{i,i} = 1$ for any $i \in V$ and $R_{i,j} = \rho_0$ for any $(i,j) \in E$. From now onward, we focus on the case of positive correlation (i.e., $\rho_0 \geq 0$) since such matrix R always exists. For instance, consider the definite positive matrix R such that $R_{i,i} = 1$ for any $i \in V$ and $R_{i,j} = \rho_0$ for any $i,j \in V$ with $i \neq j$. The uniqueness of the solution implies that any GMRF with positive uniform correlation ρ_0 over a graph G has the same correlation matrix, which will be denoted by $P(G,\rho_0)$. A very important concept to our purposes is the determinant of the correlation matrix $P(G,\rho_0)$, which will be denoted by $\tau(G,\rho_0)$, as in [27]. We extend the function to the case in which $\rho_0 = 1$, as follows. If the graph has at least one edge, the associated correlation matrix is not invertible, so we set $\tau(G,1) = 0$. If G has no edges, the correlation matrix is the identity matrix, so we set $\tau(G,1) = 1$.

3 Entropy of a GMRF with uniform correlation

The continuous entropy, also known as differential entropy, is the extension of the Shannon entropy [25] to continuous random vectors. Given a random vector $\vec{X}$ with probability density function $f(\vec{x})$, the continuous entropy is defined as

$$H(\vec{X}) = \int \log(f(\vec{x}))f(\vec{x})d\vec{x}.$$

Given a covariance matrix, the Multivariate Gaussian random vector maximizes the continuous entropy between those with that covariance matrix [4]. The continuous entropy of a Multivariate Gaussian random vector $\vec{X}$ only depends on the covariance matrix and has the following expression [25]: $H(\vec{X}) = \frac{n}{2}\log(2\pi e) + \frac{1}{2}\log|\Sigma|$.

Example 2. *Consider the random vectors $\vec{X}_A$ and $\vec{X}_B$ from Example 1. The continuous entropy has the following values: $H(\vec{X}_A) = \frac{5}{2}\log(2\pi e) + \frac{1}{2}\log|\Sigma_A| \approx 5.901$ and $H(\vec{X}_B) = \frac{5}{2}\log(2\pi e) + \frac{1}{2}\log|\Sigma_B| \approx 7.771$.*

Considering a GMRF with uniform correlation ρ_0 over a graph G, the relation between the covariance matrix and the correlation matrix leads us to the following expression:

$$H(\vec{X}) = \frac{n}{2}\log(2\pi e) + \frac{1}{2}\log\left(\tau(G,\rho_0)\prod_{i=1}^{n}\sigma_i^2\right).$$

Aside of the variances of the components, the continuous entropy of a GMRF with uniform correlation depends only on the value of $\tau(G,\rho_0)$. In addition, if the value of ρ_0 is fixed, the continuous entropy only depends on the graph.

We are interested in studying the effect of the graph on the continuous entropy, focusing on the study of $\tau(G,\rho_0)$. As a preliminary study, we verify that greater values of ρ_0 imply smaller values of $\tau(G,\rho_0)$ and that adding edges to the graph also makes $\tau(G,\rho_0)$ to take smaller values. The first statement is proven in [27]. In particular, letting G fixed, $\tau_G(x) = \tau(G,x)$ is a continuous, monotone decreasing and logarithmically concave function [27]. The second statement can be proven by applying that the determinant of the matrix F is greater than or equal to the determinant of the matrix R in Theorem 2, see [6].

Proposition 1. *Let $G_1 = (V,E_1)$ and $G_2 = (V,E_2)$ be two graphs with the same set of nodes such that $E_1 \subseteq E_2$. Then, $\tau(G_1,\rho_0) \geq \tau(G_2,\rho_0)$ for any $\rho_0 \in [0,1]$.*

Another interesting property is that two isomorphic graphs G and G' (see [13]) satisfy that $\tau(G',\rho_0) = \tau(G,\rho_0)$ for any $\rho_0 \in [0,1]$. Moreover, it is clear from the properties of the determinants that, adding an isolated node, denoted by K_1, to the graph does not change $\tau(G,\rho_0)$, i.e. $\tau(G\cup K_1,\rho_0) = \tau(G,\rho_0)$ for any graph G and $\rho_0 \in [0,1]$. Thus, the converse of the previous statement is not fulfilled, that is, there are also pairs of non-isomorphic graphs for which the determinants of the correlation matrices are equal. It is possible to find a counterexample even in the case in which the graphs are of the same order. In fact, we will see in Proposition 2 that all trees of a fixed order share the same function $\tau(G,\rho_0)$.

Given a graph G, the function $\tau(G,\rho_0)$ is, in general, not easy to compute. Firstly, we need to determine $P(G,\rho_0)$. Although there are numerical methods to compute the correlation matrix for a specific value of ρ_0 (see [3, 26, 27, 28]), a method to find the explicit expression in terms of ρ_0 has not been developed yet. Here, in order to find $P(G,\rho_0)$, we identify which of the elements of $P(G,\rho_0)$ must have the same value, typically by using the automorphism group of the graph. Subsequently, we compute the inverse of the matrix and apply Theorem 1 for obtaining a non-linear system of equations. This system typically has more than one solution, although only one of these solutions satisfies the conditions of the correlation matrix of a GMRF with uniform correlation. Choosing the adequate solution of the system, if there are more than one, allows us to determine $P(G,\rho_0)$. Afterwards, we only have to compute the determinant of $P(G,\rho_0)$ in order to obtain $\tau(G,\rho_0)$.

Example 3. *Consider the graphs in Figure 1. It can be proven, by inverting the matrices as in Example 1, that the expression of $P(bull,\rho_0)$ and $P(K_4\cup K_1,\rho_0)$ are the following:*

$$P(bull,\rho_0) = \begin{pmatrix} 1 & \rho_0 & \rho_0 & \rho_0^2 & \rho_0^2 \\ \rho_0 & 1 & \rho_0 & \rho_0 & \rho_0^2 \\ \rho_0 & \rho_0 & 1 & \rho_0^2 & \rho_0 \\ \rho_0^2 & \rho_0 & \rho_0^2 & 1 & \rho_0^3 \\ \rho_0^2 & \rho_0^2 & \rho_0 & \rho_0^3 & 1 \end{pmatrix}$$

$$P(K_4\cup K_1,\rho_0) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & \rho_0 & \rho_0 & \rho_0 \\ 0 & \rho_0 & 1 & \rho_0 & \rho_0 \\ 0 & \rho_0 & \rho_0 & 1 & \rho_0 \\ 0 & \rho_0 & \rho_0 & \rho_0 & 1 \end{pmatrix}$$

The computation of the determinant of the matrices above leads to the following expressions of $\tau(bull,\rho_0) = (1+2\rho_0)(1-\rho_0)^2(1-\rho_0^2)^2$ and $\tau(K_4\cup K_1,\rho_0) = \tau(K_4,\rho_0) = (1+2\rho_0)^2(1-\rho_0)^4$.

We end this section by providing the expression of $\tau(G,\rho_0)$ for the families of tree graphs and complete graphs.

A tree, denoted by T_n, is a connected acyclic graph. For obtaining the general expression of $\tau(T_n,\rho_0)$, we need to determine the structure of $P(T_n,\rho_0)$.

Lemma 1. *Let $\vec{X}$ be a GMRF with uniform correlation ρ_0 over a tree graph T_n. Two variables X_i, X_j for which their associated nodes are at a distance d in G have a Pearson's correlation coefficient of $\rho_{i,j} = \rho_0^d$.*

The iterative construction of $P(G,\rho_0)$ allows us to compute the expression of $\tau(G,\rho_0)$ in a simple way.

Proposition 2. *Let T_n be a tree graph of order n. It holds that $\tau(T_n,\rho_0) = (1-\rho_0^2)^{n-1}$.*

A complete graph of order n, denoted by K_n, is a graph in which any two nodes are adjacent. Determining the expression of $P(K_n, \rho_0)$ is straightforward from Definition 1, so we just need to compute the determinant.

Proposition 3. *Let K_n be a complete graph of order n. It holds that $\tau(K_n, \rho_0) = (1 + (n-1)\rho_0)(1-\rho_0)^{n-1}$.*

4 Ordering graphs by using continuous entropy

The aim of this section is to use the relation between $\tau(G, \rho_0)$ and the continuous entropy of Gaussian Markov Random Fields to define an order relation on the set of graphs of certain order n.

Definition 2. *Let A_n be the set of graphs of order n, up to isomorphism. The following binary relation $\preceq$ is defined over A_n:*

$$G \preceq G' \text{ if } \tau(G, \rho_0) \leq \tau(G', \rho_0), \quad \forall \rho_0 \in [0, 1].$$

The fact that $G \preceq G'$ is interpreted, in the context of GMRFs with uniform correlation, as *the contribution of G to the continuous entropy is smaller than or equal to the contribution of G'*. Equivalently, it means that, given the same marginal distributions and the same value of ρ_0, the continuous entropy of a GMRF with uniform correlation ρ_0 over G is smaller than or equal to the continuous entropy of a GMRF with uniform correlation ρ_0 over G'.

The relation defined above is a preorder relation, since it trivially holds that $\tau(G, \rho_0) \leq \tau(G, \rho_0)$ for any $\rho_0 \in [0, 1]$ and, given G_1, G_2, G_3 such that $G_1 \preceq G_2$ and $G_2 \preceq G_3$, it then holds that $G_1 \preceq G_3$ since $\tau(G_1, \rho_0) \leq \tau(G_2, \rho_0) \leq \tau(G_3, \rho_0)$ for any $\rho_0 \in [0, 1]$.

Additionally, if $n \leq 3$, the relation is a total order relation, and, if $n \leq 4$, the relation is a total preorder relation (also called a weak order relation). This can be easily proven by calculating the expressions of $\tau(G, \rho_0)$ for all $G \in A_3$ and $G \in A_4$. The corresponding Hasse diagram for $n = 4$ is shown in Figure 2. For the names of the graphs, we take as a reference the Information System on Graph Classes and their Inclusions, an online encyclopedia of graph classes that can be found at www.graphclasses.org. We also do not include in the notation any isolated node, i.e. we denote $G \cup K_1$ simply by G. We recall that adding an isolated node to a graph does not change the expression of $\tau(G, \rho_0)$.

As an example, the Hasse diagram of $\preceq$ for $n = 5$, shown in Figure 3, reveals that there are some pairs of graphs, located in the middle of the diagram, that cannot be ordered by using $\preceq$. This implies that, for any two incomparable graphs G and G', there exist two uniform correlation values $\rho_1, \rho_2 \in (0, 1)$ such that $\tau(G, \rho_1) < \tau(G', \rho_1)$ and $\tau(G, \rho_2) > \tau(G', \rho_2)$.

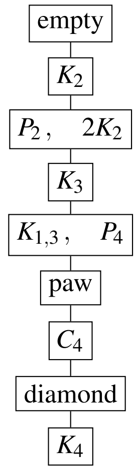

Figure 2: Hasse diagram of $\preceq$ for non isomorphic graphs of order 4.

Example 4. *The representation of the functions $\tau(G, \rho_0)$ on the interval $[0, 1]$ associated with (a) $K_4 \cup K_1$ (or, simply, K_4 when restricted to A_4), and (b) the bull graph, can be found in Figure 4. As can be seen, both functions intersect at a certain point ρ_c, thus both graphs are incomparable with respect to the order $\preceq$. In particular, for $\rho_0 < \rho_c$, it holds that $\tau(K_4 \cup K_1, \rho_0) < \tau(bull, \rho_0)$ and, for $\rho_0 > \rho_c$, it holds that $\tau(K_4 \cup K_1, \rho_0) > \tau(bull, \rho_0)$.*

Interestingly, after calculating and comparing the expressions of $\tau(G, \rho_0)$ for all $G \in A_5$, it is concluded that there do not exist two graphs of order 5 (or less) such that their respective $\tau(G, \rho_0)$ functions intersect at two or more different points (aside of the trivial intersection points $\rho_0 = 0$ and $\rho_0 = 1$).

Remark 1. *It is reasonable to think that if G has a large contribution to the continuous entropy, then $\overline{G}$ should have a small contribution. However, some of the properties that we might expect are not satisfied. For instance, the complementary of $K_4 \in A_5$ is $K_{1,4}$ and both graphs are incomparable (see Figure 3). Therefore, the property $G_1 \preceq \overline{G_1}$ or $\overline{G_1} \preceq G_1$ does not hold in general. Moreover, if we consider the graphs K_3 and claw, in A_5, and pay attention to Figure 3, we can see that $K_3 \preceq$ claw and $\overline{K_3} \preceq \overline{claw}$. This is a counterexample for the property: $G_1 \preceq G_2 \implies \overline{G_2} \preceq \overline{G_1}$.*

5 A fuzzy order defined on the measure of the support

For those cases in which the relation defined in the previous section is not total, we propose a fuzzy order relation by measuring the support in which one $\tau(G, \rho_0)$ function is smaller than or equal to the other function. First of all, we introduce some preliminary concepts

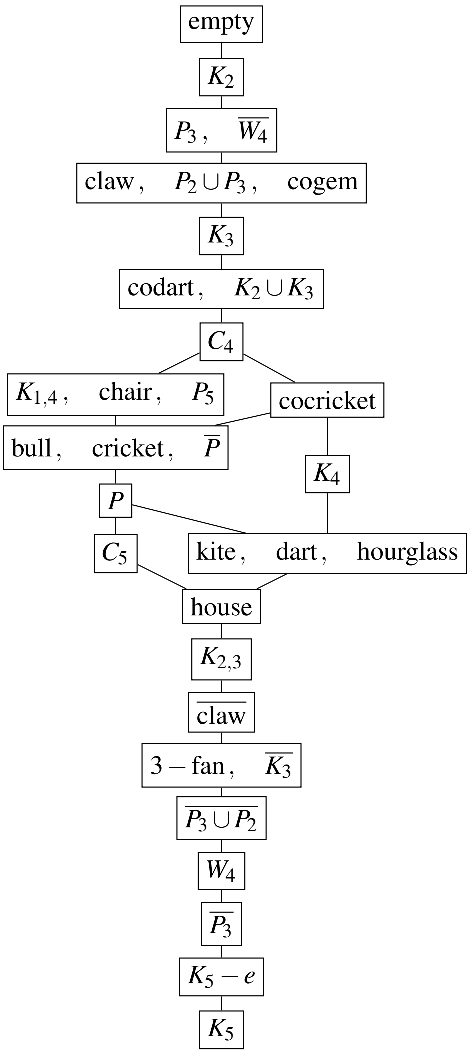

Figure 3: Hasse diagram of $\preceq$ for non isomorphic graphs of order 5.

concerning fuzzy binary relations and fuzzy preorder relations. For more details, we refer to [5].

Definition 3. *Let X be a non-empty set. A mapping $R : X \times X \to [0,1]$ is called a binary fuzzy relation on X.*

Definition 4. *An associative, commutative and increasing binary operation $T : [0,1] \times [0,1] \to [0,1]$ that has 1 as a neutral element is called a t-norm.*

The Łukasiewicz t-norm is one of the most common t-norms [11], and it is defined as $T_L(a,b) = \max\{a+b-1, 0\}$. Other common t-norms are the minimum t-norm $T_{min}(a,b) = \min(a,b)$ and the product t-norm $T_{prod}(a,b) = ab$ [11].

Definition 5. *Let $T : [0,1] \times [0,1] \to [0,1]$ be t-norm. A binary fuzzy relation $R : X \times X \to [0,1]$ is called a T-preorder relation if it satisfies:*

- *Reflexivity: $R(x,x) = 1$ for any $x \in X$.*

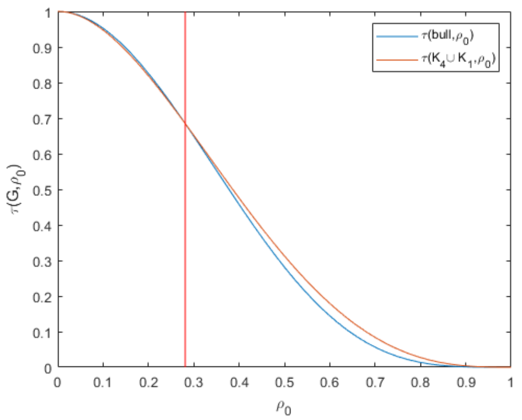

Figure 4: Representation of $\tau(K_4, \rho_0)$ on the interval $[0,1]$, where the red line represents the intersection point.

- *T-transitivity: $T(R(x,y), R(y,z)) \leq R(x,z)$ for any $x,y,z \in X$.*

In particular, we propose a fuzzy preorder relation on the set of graphs that is Łukasiewicz-transitive.

Proposition 4. *Let A_n be the set of graphs of order n, up to isomorphism. The fuzzy binary relation $R : A_n \times A_n \to [0,1]$, defined as $R(G,G') = \mu(\{x \in [0,1] | \tau(G,x) \leq \tau(G',x)\})$, where μ is the Lebesgue measure, is a Łukasiewicz-preorder relation.*

We can interpret the value of $R(G,G')$ as the degree in which the contribution to the continuous entropy of G is smaller than or equal to the contribution of G'.

Example 5. *The intersection point $\rho_c \in (0,1)$ of $\tau(bull, \rho_0)$ and $\tau(K_4, \rho_0)$ is $\rho_c \approx 0.281$. Therefore, since $\tau(K_4, \rho_0) < \tau(bull, \rho_0)$ for any $\rho_0 < \rho_c$ and $\tau(K_4, \rho_0) > \tau(bull, \rho_0)$ for any $\rho_0 > \rho_c$, it holds that $R(K_4, bull) \approx 0.281$ and $R(bull, K_4) \approx 0.719$.*

The behaviour of this fuzzy order relation is coherent with respect to the preorder relation defined in the previous section, i.e., $G \preceq G'$ implies $R(G,G') = 1$. Returning to the Hasse diagram in Figure 3, we provide the values of $R(G,G')$ for the graphs that are incomparable with respect to $\preceq$ in Table 1. All the values have been computed approximately, except for the graphs *kite* and C_5, for which $R(kite, C_5)$ equals 0.5.

Unfortunately, R is not transitive with respect to the product t-norm. For instance, consider the graphs of order 5 denoted previously as cocricket, $K_{1,4}$ and K_4. Computing the corresponding values of R, see Table 1, results in $R(\text{cocricket}, K_{1,4}) \approx 0.39$, $R(K_{1,4}, K_4) \approx 0.48$ and $R(\text{cocricket}, K_4) = 0 < 0.39 \cdot 0.48$. Therefore, R is also not transitive with respect to the minimum t-norm

$R(G_1, G_2)$	$K_{1,4}$	cocr.	bull	P	K_4	C_5	kite
$K_{1,4}$	1	0.61	0	0	0.48	0	0
cocricket	0.39	1	0	0	0	0	0
bull	1	1	1	0	0.72	0	0
P	1	1	1	1	0.81	0	0
K_4	0.52	1	0.28	0.19	1	0.17	0
C_5	1	1	1	1	0.83	1	0.5
kite	1	1	1	1	1	0.5	1

Table 1: Values of $R(G_1, G_2)$ for the graphs of order 5 that are incomparable with respect to $\preceq$, where G_1 is the graph at the header of the row and G_2 is the graph at the header of the column.

since it is pointwise larger (i.e. $T_{min}(a,b) \geq T_{prod}(a,b)$ for any $a, b \in [0,1]$) than the product t-norm.

6 Conclusions and open problems

The relation between a graph and the continuous entropy of a GMRFs with uniform correlation over this graph has been studied. In particular, we use the function $\tau(G, \rho_0)$ for defining a preorder relation on the set of all graphs of the same order in terms of their contribution to the continuous entropy. This preorder relation is proven to be a total preorder relation if $n \leq 4$ and a total order relation if $n \leq 3$. For $n \geq 5$, incomparable graphs appear. For such cases, we have defined a fuzzy relation based on measuring the support in which one $\tau(G, \rho_0)$ function is smaller than or equal to the other one. This fuzzy relation is proven to be a fuzzy preorder relation with respect to the Łukasiewicz t-norm.

There are still some open problems in this work. Firstly, although the Lebesgue measure is the most intuitive choice of measure in Proposition 4, an alternative definition arises if other (probability) measure on $[0,1]$ is considered. We wonder if there exists a measure for which R is transitive with respect to the product or minimum t-norms.

Another problem is to further study the conditions under which two graphs are either comparable or incomparable according to $\preceq$. It was already discussed that $G_2 \subseteq G_1$ implies $G_1 \preceq G_2$. An easy way to find pairs of incomparable graphs is to select a connected but sparse graph and a dense but unconnected graph. However, there are pairs of connected graphs that cannot be compared, for instance, C_5 and kite in Figure 3.

Acknowledgement

This research has been partially supported by the Spanish Ministry of Science and Innovation (TIN-2017-87600-P and PGC2018-098623-B-I00) and FICYT (FC-GRUPIN-IDI/2018/000176).

References

[1] A. H. Bajgiran, M. Mardikoraem, E. S. Soofi, Maximum entropy distributions with quantile information, European Journal of Operational Research 290 (1) (2021) 196–209.

[2] D. A. Brown, C. S. McMahan, S. Watson Self, Sampling strategies for fast updating of Gaussian Markov random fields, The American Statistician (2019) 1–24.

[3] A. P. Dempster, Covariance selection, Biometrics 28 (1) (1972) 157–175.

[4] D. Dowson, A. Wragg, Maximum-entropy distributions having prescribed first and second moments (corresp.), IEEE Transactions on Information Theory 19 (5) (1973) 689–693.

[5] J. C. Fodor, M. Roubens, Fuzzy preference modelling and multicriteria decision support, Vol. 14, Springer Science & Business Media, Luxembourg, 1994.

[6] R. Grone, C. R. Johnson, E. M. Sá, H. Wolkowicz, Positive definite completions of partial hermitian matrices, Linear Algebra and its Applications 58 (1984) 109–124.

[7] Y. Guo, H. Xiong, N. Ruozzi, Marginal inference in continuous Markov random fields using mixtures, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 7834–7841.

[8] P. Hénon, P. Ramet, J. Roman, Pastix: a high-performance parallel direct solver for sparse symmetric positive definite systems, Parallel Computing 28 (2) (2002) 301–321.

[9] S. Hong, W. M. Moon, Application of Gaussian Markov random field model to unsupervised classification in polarimetric sar image, in: Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium, Vol. 2, 2003, pp. 929–931.

[10] R. Kindermann, J. Snell, Markov random fields and their applications, American Mathematical Society, Providence, 1980.

[11] E. P. Klement, R. Mesiar, E. Pap, Triangular norms, Vol. 8, Springer Science & Business Media, Luxembourg, 2013.

[12] L. Knorr-Held, H. Rue, On block updating in Markov random field models for disease mapping, Scandinavian Journal of Statistics 29 (4) (2002) 597–614.

[13] W. Kocay, D. L. Kreher, Graphs, algorithms, and optimization, CRC Press, Boca Raton, 2016.

[14] S. Lee, P. Sobczyk, M. Bogdan, Structure learning of Gaussian Markov random fields with false discovery rate control, Symmetry 11 (10) (2019) 1311.

[15] Y. C. MacNab, On Gaussian Markov random fields and Bayesian disease mapping, Statistical Methods in Medical Research 20 (1) (2011) 49–68.

[16] K. V. Mardia, J. T. Kent, J. M. Bibby, Multivariate Analysis, Academic Press INC, San Diego, 1979.

[17] J. V. Michalowicz, J. M. Nichols, F. Bucholtz, Handbook of differential entropy, Crc Press, Boca Raton, 2013.

[18] R. Molina, A. K. Katsaggelos, J. Mateos, J. Abad, Compound Gauss Markov random fields for astronomical image restoration, Vistas in Astronomy 40 (4) (1996) 539–546.

[19] M. Pereira, N. Desassis, Efficient simulation of Gaussian Markov random fields by chebyshev polynomial approximation, Spatial Statistics 31 (2019) 100359.

[20] V. K. Rohatgi, A. M. E. Saleh, An introduction to probability and statistics, John Wiley & Sons, New York, 2015.

[21] D. Rose, G. Whitten, A. Sherman, R. Tarjan, Algorithms and software for in-core factorization of sparse symmetric positive definite matrices, Computers & Structures 11 (6) (1980) 597–608.

[22] H. Rue, L. Held, Gaussian Markov Random Fields: Theory and Applications, CRC press, Boca Ratón, 2005.

[23] P. Salemi, First-order intrinsic Gaussian Markov random fields for discrete optimisation via simulation, Journal of Simulation 13 (4) (2019) 272–285.

[24] P. Salemi, E. Song, B. L. Nelson, J. Staum, Gaussian Markov random fields for discrete optimization via simulation: Framework and algorithms, Operations Research 67 (1) (2019) 250–266.

[25] C. E. Shannon, A mathematical theory of communication, The Bell System Technical Journal 27 (3) (1948) 379–423.

[26] T. P. Speed, H. T. Kiiveri, Gaussian Markov distributions over finite graphs, The Annals of Statistics (1986) 138–150.

[27] B. Szegedy, On sidorenko's conjecture for determinants and gaussian markov random fields, arXiv preprint arXiv:1701.03632.

[28] N. Wermuth, E. Scheidt, Algorithm as 105: Fitting a covariance selection model to a matrix, Journal of the Royal Statistical Society. Series C (Applied Statistics) 26 (1) (1977) 88–92.

[29] S. Zhang, S. Mallat, Maximum entropy models from phase harmonic covariances, Applied and Computational Harmonic Analysis.