Terence Chan; Qi Chen; Raymond Yeung

Characterisation of conditional independence structures for polymatroids using vanishing sets

Kybernetika, Vol. 56 (2020), No. 6, 1022–1044

Persistent URL: http://dml.cz/dmlcz/148497

KYBERNETIKA — VOLUME 56 (2020), NUMBER 6, PAGES 1022–1044

CHARACTERISATION OF CONDITIONAL INDEPENDENCE STRUCTURES FOR POLYMATROIDS USING VANISHING SETS

Terence H. Chan, Qi Chen, Raymond W. Yeung

In this paper, we characterise and classify a list of full conditional independences via the structure of the induced set of vanishing atoms. Construction of Markov random subfield and minimal characterisation of polymatroids satisfying a MRF will also be given.

Keywords: full conditional independence, markov random field, polymatroids

Classification: 62B10, 62-09

1. INTRODUCTION

In many applications, random variables involved in a system are often not arbitrary but satisfy some constraints, among which conditional independence is one major class. Therefore, characterising the conditional independence structure for random variables is a very important problem.

There is a variety of "conditional independence" in the literature. Perhaps the most well known conditional independence concept is in the context of random variables, where conditional independence refers to that the joint probability distribution can be factorised in a specific way (as determined by the conditional independence). For example, that X and Y are conditionally independent given Z means that

$$\Pr(x, y, z) = \frac{\Pr(x, z)\Pr(y, z)}{\Pr(z)}.$$

Alternatively, in the context of undirected graphs, one can define "conditional independence" using graph separation. For example, two nodes X and Y are called "conditionally independent"[1] given Z if the two nodes are disconnected after removing node Z (i. e., all paths connecting $X$and Y, if exist, must pass through Z). Conditional independence can also be defined in the context of database relation [6].

Clearly, different "classes" of conditional independence concepts are not necessarily the same (and in fact they are not). However, they do share many similarities – many

DOI: 10.14736/kyb-2020-6-1022

[1] Here, we borrow the convention from Markov random field (which is a graphical model for random variables satisfying Markov property specified by an undirected graph).

conditional independence implication rules hold in both contexts. This is in fact one reason why Markov random field (MRF) can be used as a graphical model for random variables (which satisfy the conditional independence relations specified by the graph).

The concept of Markov random field first appeared in a work in statistical physics by Ernst Ising, who used MRF to model some physical properties (spins) of ferromagnetic materials. In his model, the underlying graph is a rectangular lattice. The idea behind this is a belief or a model that atomic spin of a particle will be affected (primarily) only by its neighbours. In other words, conditionally on the states of its neighbours, the spin is conditionally independent of the other particles. Nowadays, MRFs are used in many different areas including image modeling and processing [1], wireless and ad hoc networking [4], modeling for social networks [7,8], genomics [3] etc.

Instead of defining conditional independence relation as a property of a mathematical object (whether it can be statistical independence among a set of random variables or graph separation in an undirected graph), one may also take an "axiomatic approach" by viewing each conditional independence as a predicate and then define a dependency model as a set of conditional independence relations satisfying some conditional independence implication rules or axioms. For example, a set of conditional independence relations is called a *semi-graphoid* if it satisfies the following axioms [5]:

- (Symmetry) $\mathcal{X} \perp \mathcal{Y}|\mathcal{Z} \Leftrightarrow \mathcal{Y} \perp \mathcal{X}|\mathcal{Z}$
- (Decomposition) $\mathcal{X} \perp \mathcal{Y} \cup \mathcal{W}|\mathcal{Z} \Rightarrow \mathcal{X} \perp \mathcal{Y}|\mathcal{Z}$ and $\mathcal{X} \perp \mathcal{W}|\mathcal{Z}$
- (Weak Union) $\mathcal{X} \perp \mathcal{Y} \cup \mathcal{W}|\mathcal{Z} \Rightarrow \mathcal{X} \perp \mathcal{Y}|\mathcal{Z} \cup \mathcal{W}$
- (Contraction) $\mathcal{X} \perp \mathcal{Y}|\mathcal{Z}$ and $\mathcal{X} \perp \mathcal{W}|\mathcal{Z} \cup \mathcal{Y} \Rightarrow \mathcal{X} \perp \mathcal{Y} \cup \mathcal{W}|\mathcal{Z}$

and a *graphoid* if it additionally satisfies

$$\text{(Intersection)} \qquad \mathcal{X} \perp \mathcal{Y}|\mathcal{Z}, \mathcal{W} \text{ and } \mathcal{X} \perp \mathcal{W}|\mathcal{Z} \cup \mathcal{Y} \Rightarrow \mathcal{X} \perp \mathcal{Y} \cup \mathcal{W}|\mathcal{Z}.$$

Here, we use the notation $\mathcal{X} \perp \mathcal{Y}|\mathcal{Z}$ as a shorthand to that "$\mathcal{X}$ and $\mathcal{Y}$ are conditionally independent given $\mathcal{Z}$". Also, it is well known that the set of conditional independence for random variables is a semi-graphoid while the conditional independence for undirected graph is actually a graphoid.

In some senses, the above axioms can be regarded as conditional independence implication/inference rules. Since the conditional independence structures for random variables and for undirected graphs are not exactly the same, a valid conditional independence implication rule for undirected graph could be invalid for random variables. As an example, the intersection axiom (or inference rule) does not hold in the context of conditional independence structure for random variables. Another example is

$$\text{(Strong Union)} \qquad \mathcal{X} \perp \mathcal{Y}|\mathcal{Z} \Rightarrow \mathcal{X} \perp \mathcal{Y}|\mathcal{Z}\mathcal{W}$$

which is satisfied by the conditional independence structure for undirected graph, but not necessarily for random variables.

This paper alternatively uses polymatroids (and the vanishing atoms defined using which) as a tool to help capture conditional independence structure for random variables.

It is well known that for any set of discrete random variables, it is associated with a Shannon entropy function. Any conditional independence among random variables can be equivalently written as an equality related to Shannon entropies. Hence, if one can characterise all Shannon entropy functions, then the conditional independence structure for random variables can also be determined in theory. However, in practice, it is well known that obtaining such a characterisation is nearly impossible. Therefore, we instead use polymatroids which include entropy functions as special cases.

Since all entropy functions are polymatroids, concepts of conditional independence can be naturally extended in the context of polymatroids. Also, any conditional independence implication rules or structure for polymatroids will remain valid for random variables. It turns out that each full conditional independence (FCI) can be equivalently represented by a collection of vanishing atomic constraint, each of which is a polymatroidal equality (and is indexed by a subset of polymatroidal variables). In other words, FCIs are merely special cases of vanishing atomic constraints. The main objective of this paper is to identify the conditional independence structures for polymatroids using vanishing atomic constraints. Specifically, we aim to answer the following question:

> *What are the properties of the set of vanishing atomic constraints that correspond to a set of FCIs (or a set of conditional independence induced by an undirected graph)*?

The organisation of our paper is as follows: In Section 2, we introduce the background on important concepts such as polymatroids and full conditional independence. Section 3 contains the main results, where we offer a characterisation and classification of various FCIs based on the structure of vanishing atom sets. In Section 4, we extend our work to subsystems where we re-discover (using a different approach) the graph construction of Markov random subfield. We also derive the minimal set of inequalities to characterise polymatroids subject to a MRF constraint.

2. BACKGROUND AND PRELIMINARIES

Consider a nonempty set $\mathcal{N} = \{1, \dots, n\}$ whose elements will be referred to as variables or ground set elements. Sets will often be written with calligraphic typeface (e. g., $\mathcal{A}, \mathcal{B}$). Singletons $\{A\}$ will be written without enclosing braces. Set union will interchangeably be written $\mathcal{A}, \mathcal{B} = \mathcal{A} \cup \mathcal{B}$. Set complement of $\mathcal{A}$ (i. e., $\mathcal{N} \setminus \mathcal{A}$) will be denoted by $\bar{\mathcal{A}}$.

Definition 1. Let $h : 2^{\mathcal{N}} \mapsto \mathbb{R}$ be a real-valued function defined on the non-empty subsets of $\mathcal{N}$. The function h is called a *polymatroid* if it satisfies the following conditions:

$$h(\emptyset) = 0 \tag{1}$$

$$h(\mathcal{N}) - h(\mathcal{N} \setminus A) \geq 0, \quad \forall A \in \mathcal{N} \tag{2}$$

$$h(A, \mathcal{C}) + h(B, \mathcal{C}) - h(A, B, \mathcal{C}) - h(\mathcal{C}) \geq 0, \quad \forall \mathcal{C} \subseteq \mathcal{N} \text{ and } A, B \notin \mathcal{C}. \tag{3}$$

The set of polymatroids will be denoted by $\Gamma(\mathcal{N})$.

Examples of polymatroids include entropy functions where $h(\mathcal{A})$ is defined as the joint entropy of a set of discrete random variables $(X_i, i \in \mathcal{A})$. For any $h \in \Gamma(\mathcal{N})$, one

can uniquely define a function $\mathcal{A} \mapsto I_h(\wedge\bar{\mathcal{A}}\,|\mathcal{A})$ for any $\mathcal{A} \neq \mathcal{N}$ such that

$$h(\mathcal{B}) = \sum_{\mathcal{A}:\mathcal{B}\setminus\mathcal{A}\neq\emptyset} I_h(\wedge\bar{\mathcal{A}}\,|\mathcal{A}). \tag{4}$$

For example, let $\mathcal{N} = \{1, 2, 3\}$. Then, as illustrated in Figure 1,

$$h(1) = I_h(1|23) + I_h(\wedge 12|3) + I_h(\wedge 13|2) + I_h(\wedge 123|\emptyset).$$

Remark 1. The term $I_h(\wedge\bar{\mathcal{A}}\,|\mathcal{A})$ is called atomic because for any $\mathcal{B} \subseteq \mathcal{N}$, $h(\mathcal{B})$ can be written as a sum of these atomic terms.

In addition, we will also use the following notations. For any disjoint subsets $\mathcal{A}, \mathcal{B}, \mathcal{C} \subseteq \mathcal{N}$, we denote

$$I_h(\wedge\mathcal{A}|\mathcal{C}) \triangleq \sum_{\mathcal{D}:\,\mathcal{C}\subseteq\mathcal{D} \text{ and } \mathcal{A}\subseteq\bar{\mathcal{D}}} I_h(\wedge\bar{\mathcal{D}}\,|\mathcal{D}) \tag{5}$$

$$I_h(\mathcal{A}|\mathcal{C}) \triangleq \sum_{\mathcal{D}:\,\mathcal{C}\subseteq\mathcal{D} \text{ and } \mathcal{A}\setminus\mathcal{D}\neq\emptyset} I_h(\wedge\bar{\mathcal{D}}\,|\mathcal{D}) \tag{6}$$

$$I_h(\mathcal{A}\wedge\mathcal{B}|\mathcal{C}) \triangleq \sum_{\mathcal{D}:\,\mathcal{C}\subseteq\mathcal{D} \text{ and } \mathcal{A}\setminus\mathcal{D}\neq\emptyset \text{ and } \mathcal{B}\setminus\mathcal{D}\neq\emptyset} I_h(\wedge\bar{\mathcal{D}}\,|\mathcal{D}). \tag{7}$$

Example 1. Using our convention, we have

$$\begin{aligned} I_h(1|2) &= I_h(\wedge 13|2) + I_h(1|23) \\ h(1) = I_h(1|\emptyset) &= I_h(1|23) + I_h(\wedge 12|3) + I_h(\wedge 13|2) + I_h(\wedge 123|\emptyset) \\ I_h(\wedge 12|\emptyset) &= I_h(\wedge 123|\emptyset) + I_h(\wedge 12|3) \\ I_h(12|\emptyset) &= I_h(1|23) + I_h(\wedge 12|3) + I_h(2|13) + I_h(\wedge 123|\emptyset). \end{aligned}$$

We can use set notations and Venn diagram to illustrate the meaning of above notations. Roughly speaking, one can imagine that there exist N sets $\{S_1, \ldots, S_N\}$ such that

- $h(\mathcal{A}) = I_h(\mathcal{A}|\emptyset)$ *corresponds to the set measure for the set* $\left(\bigcup_{i\in\mathcal{A}} S_i\right)$
- $I_h(\wedge\bar{\mathcal{D}}|\mathcal{D})$ *corresponds to the set measure for the set* $\left(\bigcap_{i\in\bar{\mathcal{D}}} S_i\right) \setminus \left(\bigcup_{i\in\mathcal{D}} S_i\right)$.
- $I_h(\wedge\mathcal{A}|\mathcal{C})$ *corresponds to the set measure for the set* $\left(\bigcap_{i\in\mathcal{A}} S_i\right) \setminus \left(\bigcup_{i\in\mathcal{C}} S_i\right)$.
- $I_h(\mathcal{A}|\mathcal{C})$ *corresponds to the set measure for the set* $\left(\bigcup_{i\in\mathcal{A}} S_i\right) \setminus \left(\bigcup_{i\in\mathcal{C}} S_i\right)$.
- $I_h(\mathcal{A}\wedge\mathcal{B}|\mathcal{C})$ *corresponds to the set measure for the set* $\left(\bigcup_{i\in\mathcal{A}} S_i\right) \cap \left(\bigcup_{i\in\mathcal{B}} S_i\right) \setminus \left(\bigcup_{i\in\mathcal{C}} S_i\right)$.

Remark 2. If h is well understood from the context, then we may drop h and simplify the notations as

$$I(\wedge\mathcal{A}|\mathcal{C}), I(\mathcal{A}|\mathcal{C}), \text{ and } I(\mathcal{A}\wedge\mathcal{B}|\mathcal{C}).$$

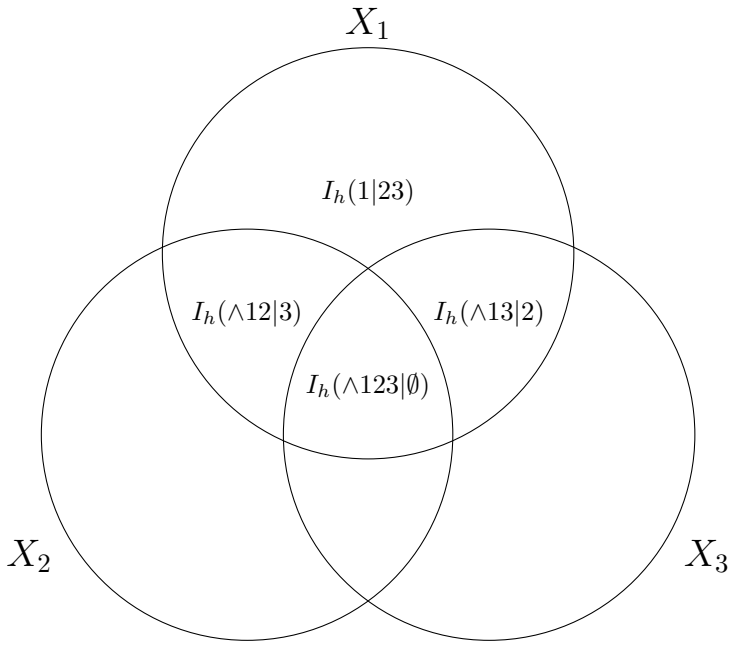

Fig. 1. Illustration for
$h(1) = I_h(1|23) + I_h(\wedge 12|3) + I_h(\wedge 13|2) + I_h(\wedge 123|\emptyset)$.

Example 2. Suppose h is the entropy function for random variables $\{X_i, i \in \mathcal{N}\}$. Hence, for any $\mathcal{A} \subseteq \mathcal{N}$, $h(\mathcal{A})$ is the joint entropy of random variables $(X_i, i \in \mathcal{A})$. Let $\mathcal{A} = \{1, 2\}$ and $\mathcal{C} = \{3\}$. Then

- $I_h(\wedge\mathcal{A}|\mathcal{C})$ (or $I_h(1\wedge 2|3)$) will be the conditional mutual information of X_1 and X_2 given X_3, and
- $I_h(\mathcal{A}|\mathcal{C})$ (or $I_h(1, 2|3)$) will be the joint entropy of X_1, X_2 given X_3.

Definition 2. (Full conditional independence) A *full conditional independence* (FCI) is denoted by a tuple

$$(\mathcal{B}_1 \perp \ldots \perp \mathcal{B}_r|\mathcal{A})$$

where:

1. $\mathcal{A}$ is a (possibly empty) subset of $\mathcal{N}$,
2. $\mathcal{B}_1, \ldots, \mathcal{B}_r, \mathcal{A}$ are all mutually disjoint
3. $\mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r \cup \mathcal{A} = \mathcal{N}$.

Definition 3. A polymatroid h is said to satisfy the FCI $\psi = (\mathcal{B}_1 \perp \ldots \perp \mathcal{B}_r|\mathcal{A})$ if and only if

$$h(\mathcal{N}) = h(\mathcal{A}) + \sum_{i=1}^{r} \left(h(\mathcal{B}_i \cup \mathcal{A}) - h(\mathcal{A})\right). \tag{8}$$

In fact, if h is the entropy function for random variables $\{X_i, i \in \mathcal{N}\}$, then h satisfies the FCI if and only if conditionally on $(X_i, i \in \mathcal{A})$, the groups of random variables $(X_i, i \in \mathcal{B}_1), \ldots, (X_i, i \in \mathcal{B}_r)$ are mutually independent.

In this paper, we will consider a list of full conditional independences

$$\Psi = \{\psi_m, m = 1, \ldots, M\} \tag{9}$$

where

$$\psi_m \triangleq (\mathcal{B}_1^m \perp \ldots \perp \mathcal{B}_{r_m}^m |\mathcal{A}_m). \tag{10}$$

Consider the set of polymatroids in $\Gamma(\mathcal{N})$ satisfying the list of FCIs in Ψ. Denote the set as $\Lambda(\Psi, \mathcal{N})$ or simply $\Lambda(\Psi)$.

Lemma 1. (Yeung et al. [9], [10, Chapter 12], Yeung [12]) A polymatroid $h \in \Gamma(\mathcal{N})$ satisfies the set of full conditional independences Ψ in (9) (or we can also say that Ψ represents h) if and only if $I_h(\wedge\bar{\mathcal{D}}\,|\mathcal{D}) = 0$ for all $\mathcal{D}$ such that there exists m for which

1. $\mathcal{A}_m \subseteq \mathcal{D}$ and
2. there exist two distinct $\mathcal{B}_i^m$ and $\mathcal{B}_j^m$ such that both $\mathcal{B}_i^m \backslash \mathcal{D}$ and $\mathcal{B}_j^m \backslash \mathcal{D}$ are nonempty.

Motivated by the above lemma, by abusing our notation, we define the vanishing atomic constraint as follows:

Definition 4. (Vanishing atomic constraint) Let V be a subset of $\mathcal{N}$. Then, by abusing our notation, we define the set $\Lambda(V, \mathcal{N})$ as the set of all functions h where $I_h(\wedge\bar{\mathcal{D}}\,|\mathcal{D}) = 0$ for all $\mathcal{D} \in V$. For simplicity, we will call elements $\mathcal{D} \in V$ (or the corresponding term $I(\wedge\bar{\mathcal{D}}\,|\mathcal{D})$[2]) the vanishing atoms.

Definition 5. (Vanishing atoms for FCIs) Consider the set of full conditional independences Ψ in (9). Then Ψ induces a set of vanishing atoms V such that $\mathcal{D} \in V$ if and only if there exists m for which

1. $\mathcal{A}_m \subseteq \mathcal{D}$ and
2. there exist two distinct $\mathcal{B}_i^m$ and $\mathcal{B}_j^m$ such that both $\mathcal{B}_i^m \backslash \mathcal{D}$ and $\mathcal{B}_j^m \backslash \mathcal{D}$ are nonempty.

We denote the vanishing atom set as $\mathcal{I}m(\Psi)$. By direct verification, Lemma 1 can be restated as that

$$\Lambda(\mathcal{I}m(\Psi), \mathcal{N}) = \Lambda(\Psi, \mathcal{N}).$$

[2]We drop the function h in the notation as it is only a generic function.

With respect to a vanishing atom set V, we will use

$$I_V(\wedge\bar{\mathcal{D}}\,|\mathcal{D}) \equiv 0 \tag{11}$$

to denote that $\mathcal{D} \in V$. Now, recalling (5) – (7), the terms such as $I(\wedge A|\mathcal{C}), I(A|\mathcal{C})$ and $I(A\wedge B|\mathcal{C})$ can always be written as a sum of atomic terms of the form $I(\wedge\bar{\mathcal{D}}\,|\mathcal{D})$ for some specific choices of $\mathcal{D}$. Therefore, for simplicity, we will also call these terms *vanishing* if all the atomic terms involved in the summation are also vanishing. We will also use

$$I_V(\wedge A|\mathcal{C}) \equiv 0 \tag{12}$$

to denote that the term $I_V(\wedge A|\mathcal{C})$ is vanishing (with respect to V). Also, if $I_V(\wedge A|\mathcal{C})$ is not vanishing, then we will instead use

$$I_V(\wedge A|\mathcal{C}) \not\equiv 0 \tag{13}$$

to denote the relation. Similar notation applies to other terms as well.

Remark 3. Note that $I_V(\wedge A|\mathcal{C}) \equiv 0$ is defined with respect to the vanishing atom set V and is equivalent to that

$$I_h(\wedge A|\mathcal{C}) = 0$$

for all polymatroids h where $I_h(\wedge\mathcal{D}|\mathcal{D}) = 0$ for all $\mathcal{D} \in V$.

For any vanishing atom set V, a list of FCIs Ψ is said to represent V if

$$\mathcal{I}m(\Psi) \subseteq V.$$

In this case, it is obvious that $\Lambda(V) \subseteq \Lambda(\mathcal{I}m(\Psi))$. In addition, if $\mathcal{I}m(\Psi) = V$, then we call the representation *faithful*.

3. CHARACTERISATION AND CLASSIFICATION OF FCI

In this section, we show that one can characterise and classify different sets of FCIs using vanishing atom sets.

3.1. Properties of FCIs

Proposition 1. Consider a set of FCIs Ψ and its induced set of vanishing atoms $V = \mathcal{I}m(\Psi)$. For any $\mathcal{C}$ and $A \notin \mathcal{C}$, let

$$\mathcal{K} \triangleq \{B \notin \mathcal{C} :\ I_V(A\wedge B|\mathcal{C}) \not\equiv 0\}. \tag{14}$$

Then $I_V(\wedge\mathcal{K}|\mathcal{C}) \not\equiv 0$.

Proof. See [2]. □

Definition 6. (Perfect Vanishing Set) A subset of atoms V is called "*perfect*" if for any proper subset $\mathcal{C} \subset \mathcal{N}$, one can partition $\mathcal{N} \setminus \mathcal{C}$ into disjoint components $\mathcal{A}_1, \ldots \mathcal{A}_k$ such that

1. For $i = 1, \ldots, k$, $I_V(\wedge \mathcal{A}_i | \mathcal{C}) \not\equiv 0$.
2. For any distinct $i, j = 1, \ldots, k$, $X \in \mathcal{A}_i$ and $Y \in \mathcal{A}_j$, $I_V(X \wedge Y | \mathcal{C}) \equiv 0$. Or equivalently, $I_V(\mathcal{A}_i \wedge \mathcal{A}_j | \mathcal{C}) \equiv 0$ as all atomic terms involved in $I_V(X \wedge Y | \mathcal{C})$ are vanishing.

Lemma 2. Consider two partitions of $\mathcal{N} \setminus \mathcal{C}$, denoted by

$$\{\mathcal{A}_1, \ldots, \mathcal{A}_k\} \text{ and } \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$$

such that

1. For $r = 1, \ldots, k$, $I_V(\wedge \mathcal{A}_r | \mathcal{C}) \not\equiv 0$.
2. For $s = 1, \ldots, m$, $I_V(\wedge \mathcal{B}_s | \mathcal{C}) \not\equiv 0$.
3. For any distinct $i, j = 1, \ldots, k$, $X \in \mathcal{A}_i$ and $Y \in \mathcal{A}_j$, $I_V(X \wedge Y | \mathcal{C}) \equiv 0$.
4. For any distinct $i, j = 1, \ldots, m$, $X \in \mathcal{B}_i$ and $Y \in \mathcal{B}_j$, $I_V(X \wedge Y | \mathcal{C}) \equiv 0$.

Then the two partitions are equivalent. In other words, for any component $\mathcal{A}_r$, there exists $\mathcal{B}_s$ such that $\mathcal{A}_r = \mathcal{B}_s$. As a corollary, $k = m$ as well.

P r o o f. Consider any component $\mathcal{A}_r$. We want to show that there must exist some s such that $\mathcal{A}_r = \mathcal{B}_s$. First, it is obvious that there exists at least one $s = 1, \ldots, m$ such that $\mathcal{A}_r \cap \mathcal{B}_s \neq \emptyset$. Next, we will show that $\mathcal{A}_r = \mathcal{B}_s$.

Let $X \in \mathcal{A}_r \cap \mathcal{B}_s$. Then as $I_V(\wedge \mathcal{A}_r | \mathcal{C}) \not\equiv 0$, we have $I_V(A \wedge X | \mathcal{C}) \neq 0$ for all $A \in \mathcal{A}_r$. Similarly, $I_V(B \wedge X | \mathcal{C}) \neq 0$ for all $B \in \mathcal{B}_s$. On the other hand, for any $i \neq r$ and $Y \in \mathcal{A}_i$, $I_V(Y \wedge X | \mathcal{C}) \equiv 0$. Hence, we prove that $\mathcal{B}_s$ must be a subset of $\mathcal{A}_r$. By symmetry, $\mathcal{A}_r \subseteq \mathcal{B}_s$. Thus, we prove that $\mathcal{A}_r = \mathcal{B}_s$ and the lemma is proved. □

Remark 4. Due to Lemma 2, the variable k (i. e., the number of components partitioning $\mathcal{N} \setminus \mathcal{C}$) in Definition 6 is in fact uniquely defined. Hence, we call k the component order of $\mathcal{C}$ (with respect to V) and it will be denoted as $\omega(\mathcal{C}; V)$ or $\omega(\mathcal{C})$ directly if V is understood from the context. Also, for notation simplicity, the components will also be referred to as

$$\Omega_1(\mathcal{C}), \ldots, \Omega_k(\mathcal{C}).$$

Theorem 1. (Vanishing atoms of FCIs are perfect) For any list of FCIs Ψ, its induced set of vanishing atoms (i. e., $\mathcal{I}m(\Psi)$) is perfect.

P r o o f. Let $V = \mathcal{I}m(\Psi)$. According to Proposition 1, for any proper subset $\mathcal{C}$ of $\mathcal{N}$, one can partition the set $\mathcal{N} \setminus \mathcal{C}$ into equivalent classes such that $X, Y \in \mathcal{N} \setminus \mathcal{C}$ is in the same classes (or component) if and only if $I_V(X \wedge Y | \mathcal{C}) \not\equiv 0$. The theorem then directly follows. □

Proposition 2. Let V be a perfect set of vanishing atoms. Then the atomic term $I_V(\wedge \bar{\mathcal{C}} \, | \mathcal{C}) \equiv 0$ (i. e., is vanishing) if and only if $\omega(\mathcal{C}) \geq 2$.

P r o o f . Clearly, if $I_V(\wedge\bar{\mathcal{C}}\,|\mathcal{C}) \not\equiv 0$, then for any $X, Y \notin \mathcal{C}$, $I_V(X\wedge Y|\mathcal{C}) \not\equiv 0$ as well. This implies that $\omega(\mathcal{C}) = 1$.

On the other hand, if $\omega(\mathcal{C}) = 1$, then by definition, $I_V(X\wedge Y|\mathcal{C}) \not\equiv 0$ for all $X, Y \notin \mathcal{C}$. Proposition 1 thus implies that $I_V(\wedge\bar{\mathcal{C}}\,|\mathcal{C}) \not\equiv 0$. The proposition thus follows. □

Remark 5. Proposition 2 can be viewed as an alternative definition or characterisation for when a subset $\mathcal{C}$ or the atomic term $I_V(\wedge\bar{\mathcal{C}}\,|\mathcal{C})$ is vanishing when V is perfect.

The following theorem is the converse of Theorem 1 where we will prove that if the set of vanishing atoms V is perfect, one can construct a list of FCIs Ψ that faithfully represents V (i. e., $\mathcal{I}m(\Psi) = V$).

Theorem 2. Let V be a perfect set of vanishing atoms. Define Ψ as the list of the following FCIs

$$\psi_{\mathcal{C}} \triangleq (\Omega_1(\mathcal{C}), \ldots \Omega_{\omega(\mathcal{C})}(\mathcal{C})\,|\mathcal{C})$$

for any $\mathcal{C}$ with component order $\omega(\mathcal{C}) \geq 2$. Then

$$V = \mathcal{I}m(\Psi).$$

P r o o f . Suppose $\mathcal{C} \in V$ (i. e., $I_V(\wedge\bar{\mathcal{C}} : |\mathcal{C}) \equiv 0$). By Proposition 2, its component order $\omega(\mathcal{C})$ is at least 2. Thus, $\psi_{\mathcal{C}}$ is a FCI in the list Ψ, proving that $\mathcal{C} \in \mathcal{I}m(\Psi)$.

On the other hand, suppose $\mathcal{C} \notin V$, (i. e., $I_V(\wedge\bar{\mathcal{C}} : |\mathcal{C}) \not\equiv 0$). By Proposition 1, $\mathcal{C}$ has component order 1. We now want to prove that $\mathcal{C} \notin \mathcal{I}m(\Psi)$. Suppose to the contrary that $\mathcal{C} \in \mathcal{I}m(\Psi)$. By Definition 5, there exists $\mathcal{W}$ and the corresponding FCI $\psi_{\mathcal{W}}$ such that

1. $\mathcal{W} \subseteq \mathcal{C}$
2. there exist two distinct and disjoint subsets $\mathcal{B}_1$ and $\mathcal{B}_2$ of $\mathcal{N} \setminus \mathcal{C}$ such that both $\mathcal{B}_1 \setminus \mathcal{W}$ and $\mathcal{B}_2 \setminus \mathcal{W}$ are nonempty.

Suppose $X \in \mathcal{B}_1 \setminus \mathcal{W}$ and $Y \in \mathcal{B}_2 \setminus \mathcal{W}$. By the construction of $\psi_{\mathcal{W}}$, we know that $I_V(X\wedge Y\,|\mathcal{C}) \equiv 0$ and further implies that $I_V(\wedge\bar{\mathcal{C}}|\mathcal{C}) \equiv 0$, contradicting to the assumption that $I_V(\wedge\bar{\mathcal{C}} : |\mathcal{C}) \not\equiv 0$. The theorem is thus proved. □

Theorems 1 and 2 proved that the "perfectness" property is a necessary and sufficient condition for when a set of vanishing atoms V can be faithfully represented by a set of FCIs. In the following example, we illustrate that the condition perfectness is indeed critical.

Example 3. Let $V = \{\emptyset\}$. Note that, $I_V(X\wedge Y\wedge Z|\emptyset) \equiv 0$. The information diagram is shown in Figure 2, where the vanishing atom will be denoted by the symbol "*". It can be verified directly that V is not perfect and that there does not exist any list of FCIs Ψ that faithfully represent V, i. e., $\mathcal{I}m(\Psi) = V$.

Before we end this subsection, we should reiterate again that the conditional independence structure is defined with respect to polymatroids. In this case, a conditional independence implication rules is valid, if and only if it is valid for all polymatroids.

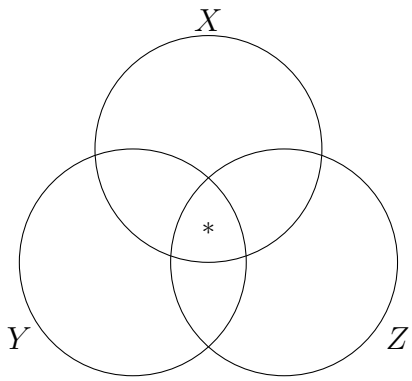

Fig. 2. Information Diagram showing the vanishing atoms in Example 3. Here, V is not perfect.

3.2. Markov random fields

Let $\mathcal{G}$ be an undirected graph with vertex set $\mathcal{N}$. Here, we assume that $\mathcal{G}$ does not have any edge joining a vertex to itself. In the graph, each node represents a variable. For any proper subset $\mathcal{C}$ of $\mathcal{N}$, we call $\mathcal{C}$ a *cut set* if one can partition $\mathcal{N}$ into partitions $\mathcal{C}, \mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_k$ for $k \geq 2$ such that

1. $\mathcal{W}_i$ is connected in the subgraph $\mathcal{G} \setminus \mathcal{C}$, which is obtained by removing all the vertices in $\mathcal{C}$ and also all corresponding incident edges.
2. for any distinct i, j, vertices $X_i \in \mathcal{W}_i$ and $X_j \in \mathcal{W}_j$ are disconnected in $\mathcal{G} \setminus \mathcal{C}$. We call $\mathcal{W}_i$ a connected component in $\mathcal{G} \setminus \mathcal{C}$.

Definition 7. (FCIs induced by MRF) Consider a graph $\mathcal{G}$ and any cut set $\mathcal{C}$ of $\mathcal{G}$, it induces the following FCI

$$\psi_{\mathcal{C}} \triangleq (\mathcal{W}_1, \ldots \mathcal{W}_k \,|\mathcal{C})$$

where $\mathcal{W}_1, \ldots, \mathcal{W}_k$ are the connected components in $\mathcal{G} \setminus \mathcal{C}$.

We refer such a collection of FCIs as the MRF induced by $\mathcal{G}$ and will denote it by $\mathbb{L}(\mathcal{G})$.

Remark 6. Some may refer to a Markov Random Field represented by $\mathcal{G}$ as a set of random variables satisfying the FCIs in $\mathbb{L}(\mathcal{G})$. In this paper, our focus is on the properties of FCIs in $\mathbb{L}(\mathcal{G})$. Therefore, we will instead define a MRF as the set of FCIs that are induced by a graph directly.

As a MRF is a set of FCIs, we can consider the set of vanishing atoms that it induces (i. e., the set $\mathcal{I}m(\mathbb{L}(\mathcal{G}))$).

Proposition 3. (Graphical interpretation) Consider a MRF $\mathcal{G}$. Then $\mathcal{C} \in \mathcal{I}m(\mathbb{L}(\mathcal{G}))$ if and only if the subgraph $\mathcal{G} \setminus \mathcal{C}$ has only 1 connected component.

P r o o f . Direct verification. □

By Theorem 1, we know that $\mathcal{I}m(\mathbb{L}(\mathcal{G}))$ must be perfect. However, the following example will show that the converse is not true – there exists a perfect set of vanishing atoms V such that $V \neq \mathcal{I}m(\mathbb{L}(\mathcal{G}))$ for any possible $\mathcal{G}$.

Example 4. (A list of FCIs that are not MRF) Consider the following list of FCIs Ψ:

$$\psi_1 = (X, Y|Z)$$
$$\psi_2 = (X, Z|Y)$$
$$\psi_3 = (Y, Z|X).$$

The set of induced vanishing atoms V is illustrated in the information diagram in Figure 3. It can be checked by brute-force that there is no graph $\mathcal{G}$ that can faithfully represent V. In other words, for any $\mathcal{G}$,

$$\mathcal{I}m(\Psi) \neq \mathcal{I}m(\mathbb{L}(\mathcal{G})).$$

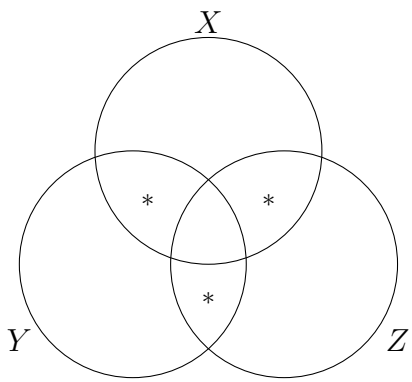

Fig. 3. Information Diagram of FCIs in Example 4.

Motivated by Example 4, the natural question is: *What is a necessary and sufficient condition for a perfect set of vanishing atoms to be faithfully represented by a graph?* In the following, we will answer this question. Before that, we will need the following definition.

Definition 8. (Intersection Property) Let V be a set of vanishing atoms. The set V is said to satisfy the "intersection property" if for any proper subset $\mathcal{C}$ of $\mathcal{N}$ such that

$$I_V(\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}, \mathcal{C}) \equiv 0$$
$$I_V(\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}, \mathcal{C}) \equiv 0,$$

then

$$I_V(\,\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}|\mathcal{C}) \equiv 0.$$

Proposition 4. (Intersection property of MRF) Let $\mathcal{G}$ be a MRF and $V = \mathcal{I}m(\mathbb{L}(\mathcal{G}))$ be the associated vanishing atom set. Then V satisfies the intersection property.

Proof. Note that $I_V(\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y},\mathcal{C}) \equiv 0$ is equivalent to that any nodes in $\mathcal{X}$ and $\mathcal{Z}$ are disconnected in the subgraph $\mathcal{G} \setminus (\mathcal{C},\mathcal{Y})$. Similarly, $I_V(\mathcal{Y}\wedge\mathcal{Z}|\mathcal{Z},\mathcal{C}) \equiv 0$ is equivalent to that any nodes in $\mathcal{Y}$ and $\mathcal{Z}$ are disconnected in the subgraph $\mathcal{G}\setminus(\mathcal{C},\mathcal{X})$. According to the intersection property in MRF, we have that both nodes in $\mathcal{X}$ and $\mathcal{Y}$ will be disconnected from the subgraph $\mathcal{G}\setminus\mathcal{C}$. This in turns implies that $I_V(\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}|\mathcal{C}) \equiv 0$, proving that the set V satisfies the intersection property. □

Remark 7. As shown in the proof, Proposition 4 is analogous and based on the well known result that the set of conditional independence defined by an undirected graph satisfies the intersection property. The difference is only that the context is in conditional independence for polymatroids.

Corollary 1. Let $\mathcal{G}$ be a MRF and $V = \mathcal{I}m(\mathbb{L}(\mathcal{G}))$ be the associated vanishing atom set. Suppose $\mathcal{C}$ has a component order 1 with respect to V, and $X, Y, Z \notin \mathcal{C}$. Then $I_V(X\wedge Z|Y,\mathcal{C}) \equiv 0$ implies that $I_V(Y\wedge Z|X,\mathcal{C}) \not\equiv 0$.

Proof. Suppose to the contrary that $I_V(Y\wedge Z|X,\mathcal{C}) \equiv 0$. According to the intersection property in Proposition 4, $I_V(X, Y\wedge Z|\mathcal{C}) \equiv 0$. This further implies that $I_V(\wedge\bar{\mathcal{C}}|\mathcal{C}) \equiv 0$ and also the component order of $\mathcal{C}$ is at least 2 by Proposition 2. A contradiction is thus established and the corollary is proved. □

Theorem 1 and Proposition 4 proved that the set of vanishing atoms induced by a MRF must be perfect and satisfy the intersection property. In the following, we will show that the converse is also true. To achieve this goal, we first describe a method to construct a graph from a set of vanishing atoms.

Definition 9. (Minimal graph construction) Let V be a set of vanishing atoms. We will construct a graph $\mathcal{G}$ as follows:

- The set of vertices is $\mathcal{N}$
- For any $X, Y \in \mathcal{N}$, there is an edge (X, Y) in the graph $\mathcal{G}$ if and only if

$$I_V(X\wedge Y|\mathcal{N} \setminus X, Y) \not\equiv 0.$$

The graph we constructed using the above method will be denoted as $\mathcal{G}_V$ to highlight the dependency on V in the construction.

Remark 8. The graph construction was also proposed in [11]. However, there are some minor differences in the context where the graph is constructed with respect to a set of random variables. Having said that, the spirit is essentially the same.

Following the same argument used in [11], we can easily prove that if $\mathcal{G}'$ is another graph that represents (not necessarily faithfully) V, then $\mathcal{G}_V$ must be a subgraph of $\mathcal{G}'$. For this reason, we call the above construction the minimal graph construction. However, as we shall illustrate, $\mathcal{G}_V$ does not necessarily represent V.

Theorem 3. (Minimal graph construction for perfect vanishing atom set) Let V be a perfect set of vanishing atoms. Then

$$V \subseteq \mathcal{I}m(\mathbb{L}(\mathcal{G}_V)).$$

As a corollary, if $\mathcal{G}_V$ represents V, then it also faithfully represents V.

Proof. Let $\mathcal{C} \in V$ be the vanishing atom and hence its component order must be at least 2. Consider any two distinct components say $\Omega_i(\mathcal{C})$ and $\Omega_j(\mathcal{C})$. Let $X \in \Omega_i(\mathcal{C})$ and $Y \in \Omega_j(\mathcal{C})$. Then $I_V(X \wedge Y|\mathcal{C}) \equiv 0$ and thus $I_V(X \wedge Y|\mathcal{N} \setminus X, Y) \equiv 0$. It means that there are no edges connecting X and Y in $\mathcal{G}_V$. Using the same argument, one can conclude that in the graph $\mathcal{G}_V \setminus \mathcal{C}$, $\Omega_i(\mathcal{C})$ and $\Omega_j(\mathcal{C})$ are disconnected. This implies that $\mathcal{C}$ is also vanishing with respect to $\mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$. The proof thus follows. □

Example 5. Consider the following list of FCIs Ψ:

$$\psi_1 = (X, Y|Z)$$
$$\psi_2 = (X, Z|Y).$$

Its corresponding set of vanishing atoms V is illustrated in the information diagram in Figure 4. By Theorem 1, V is perfect. With respect to V, $\mathcal{G}_V$ has only 1 edge connecting Y and Z. See Figure 5. The set of vanishing atoms is shown in Figure 4 and 5, which clearly indicate that

$$V \subseteq \mathcal{I}m(\mathbb{L}(\mathcal{G}_V)).$$

Example 6. Consider the set of vanishing atoms V displayed in Figure 6. Note that $I_V(X \wedge Y) \not\equiv 0$ and $I_V(X \wedge Z) \not\equiv 0$ but $I_V(X \wedge Y \wedge Z) \equiv 0$. Therefore, V is not perfect. On the other hand, the graph $\mathcal{G}_V$ is shown in Figure 7. Also, it is clear that $\mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$ is an empty set. This illustrates that the "perfectness" condition V in Theorem 3 is important in order for the theorem to hold.

Proposition 5. Let V be a set of perfect vanishing atom set and satisfy the intersection property. Let $\mathcal{X}$ and $\mathcal{Y}$ be two disjoint nonempty subsets such that

$$I_V(X \wedge Y|\mathcal{N} \setminus XY) \equiv 0$$

for all $X \in \mathcal{X}, Y \in \mathcal{Y}$. Then

$$I_V(\mathcal{X} \wedge \mathcal{Y}|\mathcal{N} \setminus \mathcal{X}\mathcal{Y}) \equiv 0 \tag{15}$$

Proof. We will prove the proposition by recursion. Assume without loss of generality that

$$\mathcal{X} = (X_1, \ldots, X_k) \text{ and } \mathcal{Y} = (Y_1, \ldots, Y_m).$$

For any $1 \leq i \leq k$, we will use $\mathcal{X}_{i^c}$ to denote $\mathcal{X} \setminus X_i$ and $\mathcal{X}_{[i,j]}$ to denote $X_i, X_{i+1}, \ldots X_j$. Let $\mathcal{C} = \mathcal{N} \setminus \mathcal{X}\mathcal{Y}$.

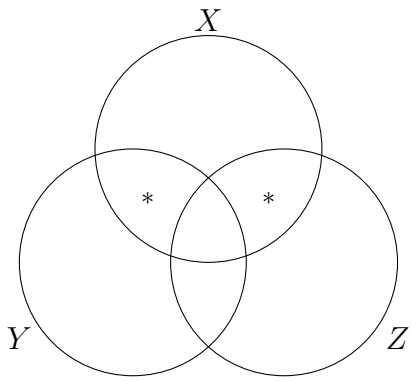

Fig. 4. Example 5: Information Diagram and vanishing atom set V.

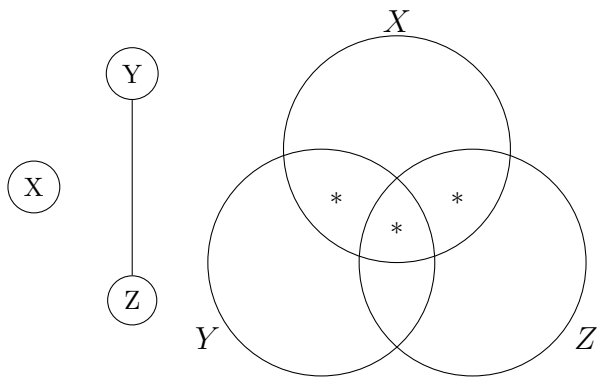

Fig. 5. Example 5: MRF $\mathcal{G}_V$ and the corresponding vanishing atom set $\mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$.

As our first step, we aim to prove that

$$I_V(X_1 \wedge \mathcal{Y} | \mathcal{C}, \mathcal{X}_{1^c}) \equiv 0. \tag{16}$$

First, by construction,

$$I_V(X_1 \wedge Y_1 | \mathcal{C}, \mathcal{X}_{1^c}, \mathcal{Y}_{1^c}) \equiv 0$$
$$I_V(X_1 \wedge Y_2 | \mathcal{C}, \mathcal{X}_{1^c}, \mathcal{Y}_{2^c}) \equiv 0.$$

Using the intersection property, we have

$$I_V(X_1 \wedge Y_{[1,2]} | \mathcal{C}, \mathcal{X}_{1^c}, \mathcal{Y}_{[3,m]}) \equiv 0.$$

Similarly,

$$I_V(X_1 \wedge Y_3 | \mathcal{C}, \mathcal{X}_{1^c}, \mathcal{Y}_{3^c}) \equiv 0.$$

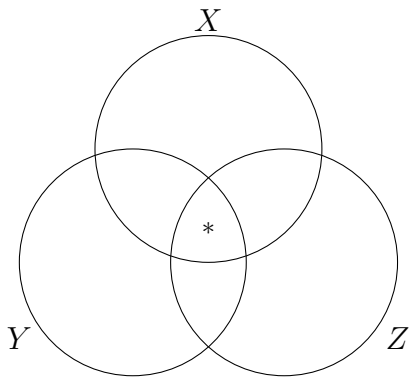

Fig. 6. Example 6: Vanishing atom set V.

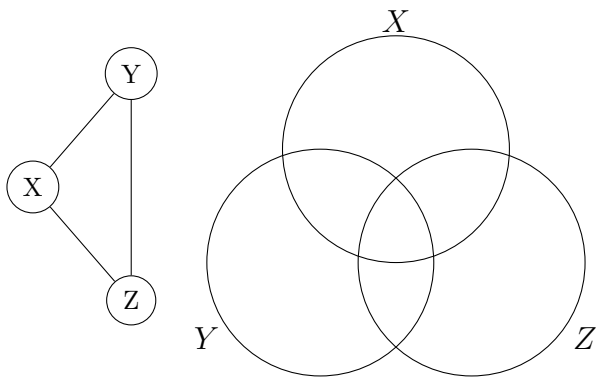

Fig. 7. Example 6: Graph $\mathcal{G}_V$ and its associated information diagram.

Invoking the intersection property, we can prove that

$$I_V(X_1 \wedge Y_{[1,3]} | \mathcal{C}, \mathcal{X}_{1^c}, \mathcal{Y}_{[4,k]}) \equiv 0.$$

Repeating the argument recursively, we can then prove that

$$I_V(X_1 \wedge \mathcal{Y} | \mathcal{C}, \mathcal{X}_{1^c}) \equiv 0$$

Similarly, we can also prove that

$$I_V(X_2 \wedge \mathcal{Y} | \mathcal{C}, \mathcal{X}_{2^c}) \equiv 0$$

Then using the intersection property,

$$I(X_1 X_2 \wedge \mathcal{Y} | \mathcal{C}, X_{[3,k]}) = 0.$$

Together with

$$I_V(X_3 \wedge \mathcal{Y} | \mathcal{C}, \mathcal{X}_{3^c}) \equiv 0,$$

the intersection property implies that

$$I_V(\mathcal{X}_{[1,3]} \wedge \mathcal{Y} | \mathcal{C}, \mathcal{X}_{[4,k]}) \equiv 0.$$

Again, repeat the procedure, we can prove that

$$I_V(\mathcal{X} \wedge \mathcal{Y} | \mathcal{C}) \equiv 0.$$

The result thus follows. □

Theorem 4. Let V be a perfect set of vanishing atoms that also satisfies also the intersection property. Then

$$V = \mathcal{I}m(\mathbb{L}(\mathcal{G}_V)).$$

P r o o f . As V is perfect, by Theorem 3,

$$V \subseteq \mathcal{I}m(\mathbb{L}(\mathcal{G}_V)).$$

It remains to prove that

$$\mathcal{I}m(\mathbb{L}(\mathcal{G}_V)) \subseteq V.$$

Let $\mathcal{C} \in \mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$. Then $\mathcal{C}$ is the cut set in the graph $\mathcal{G}_V$. By definition, there are at least two disjoint subsets $\mathcal{X}$ and $\mathcal{Y}$ such that $\mathcal{X}$ and $\mathcal{Y}$ are not connected in the subgraph $\mathcal{G}_V \setminus \mathcal{C}$. In other words, for any $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$,

$$I_V(X_i \wedge Y_j | \mathcal{N} \setminus XY) \equiv 0.$$

Then by Proposition 5, we have that

$$I_V(\mathcal{X} \wedge \mathcal{Y} | \mathcal{C}) \equiv 0$$

and hence $\mathcal{C}$ is also vanishing. The theorem is thus proved. □

3.3. Summary

In [11], FCIs induced by a graph are studied. This paper extends the results and also offers a different perspective to the same problem. First, we will restate some of the results in [11] using our terminology.

Proposition 6. (Yeung et al. [11]) Consider any set of vanishing atoms V. For any graph $\mathcal{G}$, if

$$\mathcal{I}m(\mathbb{L}(\mathcal{G})) \subseteq V,$$

then $\mathcal{G}$ contains $\mathcal{G}_V$ as a subgraph where $\mathcal{G}_V$ is constructed as in Definition 9. In other words, if $\mathcal{G}$ is a representation for V, then $\mathcal{G}$ contains $\mathcal{G}_V$ as a subgraph.

Proposition 7. (Yeung et al. [11]) If there exists the smallest subgraph $\mathcal{G}$ representing V, then $\mathcal{G} = \mathcal{G}_V$. Also, if $\mathcal{G}_V$ represents V, then $\mathcal{G}_V$ is the smallest subgraph representing V.

Remark 9. We need to point out that Proposition 7 does not imply that

$$\mathcal{I}m(\mathbb{L}(\mathcal{G})) = V,$$

as illustrated by the following example.

Example 7. Consider the same set of vanishing atom set V as in Example 6. $\mathcal{G}_V$ is given in Figure 7. It is obvious that $\mathcal{G}_V$ is the smallest. However,

$$\mathcal{I}m(\mathbb{L}(\mathcal{G}_V)) \neq V.$$

We will summarise our results in the following theorem.

Theorem 5. (Classification of vanishing atom sets) Let V be a set of vanishing atoms. Then there are three distinct cases:

Case 1: V is perfect and satisfies the intersection property. In this case,

1. $V = \mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$.
2. $\mathcal{G}_V$ is the smallest graph that represents V.

Case 2: V is perfect but not does not satisfy the intersection property. In this case,

1. V is a proper subset of $\mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$.
2. $\mathcal{G}_V$ does not represent V and the smallest graph representing V does not exist.

Case 3: V is not perfect. In this case,

1. the smallest graph representing V may or may not exist. However, if it does exists, it is equal to $\mathcal{G}_V$.
2. If it does not exist, then $\mathcal{G}_V$ is not a representation. In other words, $\mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$ is not a subset of V.

Proof. We first prove Case 1. Suppose V is perfect and satisfies the intersection property. Then Theorem 4 showed that

$$V = \mathcal{I}m(\mathbb{L}(\mathcal{G}_V)).$$

Also, according to Proposition 7, $\mathcal{G}_V$ is also the smallest graph representing V.

In Case 2 where V is perfect but does not satisfy the intersection property, Theorem 3 proved that

$$V \subseteq \mathcal{I}m(\mathbb{L}(\mathcal{G}_V)).$$

By Theorem 4, the assumption that V does not satisfy the intersection property means that $V \neq \mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$. Now, since V is a proper subset of $\mathcal{I}m(\mathbb{L}(\mathcal{G}_V))$, $\mathcal{G}_V$ does not represent V. By Proposition 7, the smallest subgraph representing V does not exist.

Finally, we will consider Case 3. When V is not perfect, then there is no guarantee that the smallest graph representing V exists or not. However, Proposition 7 guarantees that $\mathcal{G}_V$ is either the smallest graph representing V or that the smallest graph does not exist. □

4. EXTENSIONS

4.1. Markov random subfield

In the previous section, we consider constraints defined by FCIs assuming that there are N variables indexed by $\mathcal{N} = \{1, \ldots, N\}$. Consider a set of vanishing atomic constraints V. These constraints induce the set

$$\Lambda(V, \Gamma(\mathcal{N}))$$

which is the set of all polymatroids h in $\Gamma(\mathcal{N})$ such that

$$I_h(\wedge \bar{\mathcal{C}} \,|\mathcal{C}) = 0$$

for all $\mathcal{C} \in \mathcal{V}$.

The fundamental question here is to determine the "projection" of this set to $\Gamma(\mathcal{M})$. To be precise, let $\mathcal{H}$ be a subset of $\Gamma(\mathcal{N})$. We define

$$\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\mathcal{H})$$

as the set of polymatroids $h_1 \in \Gamma(\mathcal{M})$ such that there exists $h_2 \in \mathcal{H}$ where

$$h_1(\mathcal{A}) = h_2(\mathcal{A})$$

for all nonempty subsets $\mathcal{A}$ of $\mathcal{M}$.

Definition 10. (Projecting vanishing atom set) Let V be a vanishing atom set for $\mathcal{N}$. Let W be defined as follows: For any $\mathcal{D} \subseteq \mathcal{M}$, $\mathcal{D} \in W$ if and only if

$$I_V(\wedge(\mathcal{M} \setminus \mathcal{D})|\mathcal{D}) \equiv 0.$$

We call W the projection of V from $\mathcal{N}$ to $\mathcal{M}$.

Theorem 6. Let V be a vanishing atom set for $\mathcal{N}$ and W be the projection of V from $\mathcal{N}$ to $\mathcal{M}$. Then

$$\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N}))) \subseteq \Lambda(W, \Gamma(\mathcal{M})). \tag{17}$$

P r o o f. Let $h_2 \in \Lambda(V, \Gamma(\mathcal{N}))$. Then, by definition, $h_2 \in \Gamma(\mathcal{N})$ and $I_{h_2}(\wedge(\mathcal{M}\setminus\mathcal{D})|\mathcal{D}) = 0$ for all $\mathcal{D} \in W$. Let $h_1(\mathcal{A}) = h_2(\mathcal{A})$ for all $\mathcal{A} \subseteq \mathcal{M}$. It is clear that h_1 is also a polymatroid and hence belongs to $\Gamma(\mathcal{M})$. Also, $I_{h_1}(\wedge(\mathcal{M} \setminus \mathcal{D})|\mathcal{D}) = I_{h_2}(\wedge(\mathcal{M} \setminus \mathcal{D})|\mathcal{D}) = 0$ for all $\mathcal{D} \in W$. Hence, $h_1 \in \Lambda(W, \Gamma(\mathcal{M}))$. □

The following is an example showing that subset relation in (17) can be proper.

Example 8. Let $\mathcal{N} = \{1, 2, 3, 4\}$. Define V as the set of vanishing atoms where

$$V = \{\{1\}, \{2\}, \{3\}, \{4\}, \{1, 3\}, \{2, 3\}, \{1, 2\}\}.$$

By definition, if $h \in \Lambda(V, \Gamma(\mathcal{N}))$, then

$$I_h(\wedge \bar{\mathcal{C}} \,|\mathcal{C}) = 0, \quad \forall \mathcal{C} \in V.$$

Let $\mathcal{M} = \{1, 2, 3\}$. Then it is straightforward to prove that $I_V(\wedge(\mathcal{M} \setminus \mathcal{D}) \,|\mathcal{D}) \not\equiv 0$ for all $\mathcal{D} \subseteq \mathcal{M}$. Thus,

$$\Lambda(W, \Gamma(\mathcal{M})) = \Gamma(\mathcal{M}).$$

Let $h_1 \in \Gamma(\mathcal{M})$ such that

$$I_{h_1}(\wedge \mathcal{M} \setminus \mathcal{D} \,|\mathcal{D}) = \begin{cases} 1 & \text{if } \mathcal{D} = \{1\}, \{2\}, \{3\} \\ -1 & \text{if } \mathcal{D} = \emptyset \\ 0 & \text{otherwise .} \end{cases}$$

Now, we will show that there does not exist any $h_2 \in \Lambda(V, \Gamma(\mathcal{N}))$ such that

$$h_1(\mathcal{A}) = h_2(\mathcal{A}), \ \forall \mathcal{A} \in \mathcal{M}.$$

Hence, we prove that $\text{Proj}_{\mathcal{N} \to \mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N})))$ is indeed a proper subset of $\Lambda(W, \Gamma(\mathcal{M}))$.

Suppose to the contrary that there exists such an $h_2 \in \Lambda(V, \Gamma(\mathcal{N}))$. Let $\mathcal{D}$ be a proper subset of $\mathcal{M}$. Then

$$I_{h_1}(\wedge(\mathcal{M} \setminus \mathcal{D})|\mathcal{D}) = I_{h_2}(\wedge(\mathcal{M} \setminus \mathcal{D})|\mathcal{D}, 4) + I_{h_2}(\wedge(4 \cup \mathcal{M} \setminus \mathcal{D}) : \,|\mathcal{D}). \tag{18}$$

By the definition of the vanishing atom set V, if $\mathcal{D} = \emptyset$, then $I_{h_2}(\wedge(\mathcal{M} \setminus \mathcal{D}) : |\mathcal{D}, 4) = 0$. Otherwise, $I_{h_2}(\wedge(4 \cup \mathcal{M} \setminus \mathcal{D}) : \,|\mathcal{D}) = 0$. Together with (18), for any proper subset D of $\mathcal{N}$, we have

$$I_{h_2}(\wedge \bar{\mathcal{D}} \,|\mathcal{D}) = \begin{cases} 1 & \text{if } \mathcal{D} = \{1, 4\} \text{ or } \{2, 4\} \text{ or } \{3, 4\} \\ -1 & \text{if } \mathcal{D} = \emptyset \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

It is now straightforward to see that

$$\begin{aligned} I_{h_2}(1 \wedge 4) &= I_{h_2}(1 \wedge 4|2, 3) + I_{h_2}(1 \wedge 4 \wedge 2|3) + I_{h_2}(1 \wedge 4 \wedge 3|2) + I_{h_2}(1 \wedge 2 \wedge 3 \wedge 4) \\ &= -1. \end{aligned}$$

This implies that h_2 cannot be a polymatroid, contradicting the assumption that $h_2 \in \Lambda(V, \Gamma(\mathcal{N}))$. Hence, the theorem is proved.

Theorem 7. Let V be a vanishing atom set and W be the projection of V from $\mathcal{N}$ to $\mathcal{M}$. If V is perfect, then W is also perfect. If V satisfies the intersection property, then so does W.

P r o o f . Direct verification from definitions. □

Corollary 2. Let V be a vanishing atom set and W be the projection of V from $\mathcal{N}$ to $\mathcal{M}$. If V is perfect and satisfies the intersection property, then $\mathcal{G}_W$ is the smallest MRF representing $\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N})))$.

P r o o f . According to Theorems 6 and 7, W is perfect, satisfies the intersection property and

$$\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N}))) \subseteq \Lambda(W, \Gamma(\mathcal{M})). \tag{20}$$

Now, we will show that W is the largest vanishing atom set for (20) to hold.

Suppose $\mathcal{C} \notin W$. By construction of W,

$$I_V(\wedge(\mathcal{M} \setminus \mathcal{C}) \,|\mathcal{C}) \not\equiv 0.$$

In this case, there exists $\tilde{\mathcal{C}} \subseteq \mathcal{N}$ such that

1. $\mathcal{C} \subseteq \tilde{\mathcal{C}}$
2. $\mathcal{M} \setminus \tilde{\mathcal{C}} = \mathcal{M} \setminus \mathcal{C}$
3. $I_V(\wedge\bar{\tilde{\mathcal{C}}} \,|\tilde{\mathcal{C}}) \not\equiv 0.$

Define the following polymatroid $h_2 \in \Gamma(\mathcal{N})$ such that

$$I_{h_2}(\wedge\bar{\mathcal{D}}|\mathcal{D}) = \begin{cases} 1 & \text{if } \mathcal{D} = \tilde{\mathcal{C}} \\ 0 & \text{otherwise.} \end{cases}$$

It is obvious that $h_2 \in \mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N})))$ and

$$I_{h_1}(\wedge(\mathcal{M} \setminus \mathcal{C}) \,|\mathcal{C}) = 1$$

where $h_1 = \mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(h_2)$. Hence, if W^* is some other vanishing atom set containing $\mathcal{C}$, then h_1 will not be contained in $\Lambda(W^*, \Gamma(\mathcal{M}))$ and hence

$$\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N}))) \not\subseteq \Lambda(W^*, \Gamma(\mathcal{M})).$$

This proves that W is the largest vanishing atom set such that

$$\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N}))) \subseteq \Lambda(W, \Gamma(\mathcal{M})).$$

Recall that W is perfect and satisfies the intersection property. Then Theorem 5 implies that $\mathcal{G}_W$ is the smallest graph representing W.

Now, consider another graph $\mathcal{G}$ (defined with respect to $\mathcal{M}$) which satisfies the following criteria:

1. $W^* = \mathcal{I}m(\mathbb{L}(W^*))$
2. $\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N}))) \subseteq \Lambda(W^*, \Gamma(\mathcal{M}))$

By previous argument, we know that $W^* \subset W$. Hence, it is obvious that $\mathcal{G}_{W^*}$ must contain $\mathcal{G}_W$ as a subgraph. The corollary is thus proved. □

In the following, we will consider the following scenario. Let V be a perfect vanishing atom set satisfying the intersection property, and $\mathcal{G}_V$ be the corresponding MRF which faithfully represents V. Now consider the subsystem/subfield on $\mathcal{M}$. By the above corollary, we know that the subfield $\mathcal{G}_W$ is the smallest MRF representing $\mathrm{Proj}_{\mathcal{N}\to\mathcal{M}}(\Lambda(V, \Gamma(\mathcal{N})))$. The following theorem illustrates how to construct the MRF $\mathcal{G}_W$ directly from the MRF $\mathcal{G}_V$.

Theorem 8. (Constructing MRF subfield) The MRF subfield $\mathcal{G}_W$ can be constructed from $\mathcal{G}_V$ via the following procedure:

> **Markov Random Subfield Construction Procedure:** For any distinct $X, Y \in \mathcal{N}$, the edge (X, Y) is in $\mathcal{G}_W$ if and only if X and Y are connected in $\mathcal{G}_V \setminus (\mathcal{M} \setminus XY)$ (or in other words, X and Y are connected in the graph $\mathcal{G}_V$ after removing nodes from the set $\mathcal{M} \setminus XY$).

Proof. By definition, the graph construction for $\mathcal{G}_W$ is as follows: For any distinct $X, Y \in \mathcal{M}$, the edge (X, Y) is in $\mathcal{G}_W$ if and only if

$$I_W(X \wedge Y | \mathcal{M} \setminus XY) \not\equiv 0$$

or equivalently,

$$I_V(X \wedge Y | \mathcal{M} \setminus XY) \not\equiv 0.$$

By Proposition 3, $I_V(X \wedge Y | \mathcal{M} \setminus XY) \not\equiv 0$ if and only if X and Y are connected in $\mathcal{G}_V \setminus (\mathcal{M} \setminus XY)$. The theorem thus follows. □

Remark 10. The above construction of $\mathcal{G}_W$ from $\mathcal{G}_V$ was the same as the one proposed in [11]. However, the proof for the validity of the construction is different. This paper provides an alternative angle to the same construction.

4.2. Characterisation of minimal inequalities

Definition 11. Let $\mathcal{C} \subseteq \mathcal{N}$ and $A, B \notin \mathcal{C}$. The pair (A, B) is called $\mathcal{C}$-minimal if

1. $\mathcal{C}$ is nonvanishing, i. e., $I_V(\wedge \bar{\mathcal{C}} \,|\mathcal{C}) \not\equiv 0$
2. $I_V(X \wedge B | \mathcal{C}, A) \equiv 0$ implies $I_V(A \wedge B | \mathcal{C}, X) \equiv 0$ for all $X \notin \mathcal{C}$.
3. $I_V(Y \wedge A | \mathcal{C}, B) \equiv 0$ implies $I_V(B \wedge A | \mathcal{C}, Y) \equiv 0$ for all $Y \notin \mathcal{C}$.

Theorem 9. (Minimal characterisation [2]) The set of polymatroids satisfying full conditional independence Ψ is explicitly characterised by the following set of linear equality and inequality constraints

1. Vanishing atomic equality constraint:

$$I_h(\wedge \bar{\mathcal{C}} \,|\mathcal{C}) = 0, \quad \forall \mathcal{C} \in \mathcal{I}m(\Psi).$$

2. Minimal submodular inequalities:

$$I(A \wedge B|\mathcal{C}) \geq 0 \tag{21}$$

where (A, B) is $\mathcal{C}$-minimal.

In the following, we show that when Ψ can be faithfully represented by a graph $\mathcal{G}$, then there is a "graphical" way to identify the $\mathcal{C}$-minimal (A, B) tuples.

Theorem 10. (Graphical characterisation for $\mathcal{C}$-minimal tuple (A, B)) Let V be perfect and satisfies the intersection property. Then (A, B) is called $\mathcal{C}$-minimal if and only if

1. $\mathcal{C}$ is nonvanishing
2. $\mathcal{C}, A$ is nonvanishing
3. $\mathcal{C}, B$ is nonvanishing.

Equivalently, the graph $\mathcal{G}_V \setminus \mathcal{C}$ is connected, and remains so after further removing either A or B from it.

P r o o f. Clearly, if 1) – 3) are true, then (A, B) is $\mathcal{C}$-minimal. The converse is also true. To see this, suppose that (A, B) is $\mathcal{C}$-minimal. Then by definition, $\mathcal{C}$ is nonvanishing.

Suppose to the contrary that $\mathcal{C}, A$ is vanishing. Then there exists X such that $I_V(X \wedge B|\mathcal{C}, A) \equiv 0$. By definition of that (A, B) is $\mathcal{C}$-minimal, $I_V(A \wedge B|\mathcal{C}, X) \equiv 0$. The intersection property will further imply that $I_V(AX \wedge B|\mathcal{C}) \equiv 0$. This contradicts that $\mathcal{C}$ is minimal. So we proved that $\mathcal{C}, A$ is nonvanishing. Similarly, it is proved for that $\mathcal{C}, B$ is nonvanishing. Hence, the theorem is proved. □

5. CONCLUSION

In this paper, we have characterised various kinds of FCIs constraints. We showed that the set of vanishing atoms induced by FCIs are often perfect. In addition, if the set also satisfies the intersection property, then the set of FCIs indeed corresponds to a MRF. We have also identified how to construct the MRF, which happens to be the smallest graph representing the FCIs. We also extend our work to Markov random subfields. We rediscover (and provide an alternative proof for) the Markov random subfield construction in [11]. Our validity proof is much simpler, relying only on properties of vanishing atom set. Finally, we derive the minimal characterisation of inequalities for polymatroids represented by a MRF.

ACKNOWLEDGEMENT

This work is supported by the Australian Research Council under Discovery Projects DP190102896 and National Natural Science Foundation of China (NSFC) under NSFC 61971321.

(Received February 27, 2019)

REFERENCES

[1] A. Blake, P. Kohli, and C. Rother: Markov random fields for vision and image processing. MIT Press, 2011. DOI:10.7551/mitpress/8579.001.0001

[2] T. H. Chan, S. Thakor, and A. Grant: Minimal characterisation of Shannon-type inequalities under functional dependence and full conditional independence structures. IEEE Trans. Inform. Theory *65* (2019), 7, 4041–4051. DOI:10.1109/tit.2019.2900669

[3] M. Chen, J. Cho, and H. Zhao: Incorporating biological pathways via a Markov random field model in genome-wide association studies. PLOS Genetics *7* (2011), 1–13. DOI:10.1371/journal.pgen.1001353

[4] L. E. Doyle, A. C. Kokaram, S. J. Doyle, and T. K. Forde: Ad hoc networking, Markov random fields, and decision making. IEEE Signal Process. Mag. *23* (2006), 63–73. DOI:10.1109/msp.2006.1708413

[5] D. Geiger and J. Pearl: Logical and algorithmic properties of conditional independence and graphical models. Ann. Statist. *21* (1993), 4, 2001–2021. DOI:10.1214/aos/1176349407

[6] F. Malvestuto: A unique formal system for binary decompositions of database relations, probability distributions, and graphs. Inform. Sci. *59* (1992), 1, 21–52. DOI:10.1016/0020-0255(92)90042-7

[7] T. A. Snijders: Statistical models for social networks. Ann. Rev. Sociology *37* (2011), 1, 131–153. DOI:10.1146/annurev.soc.012809.102709

[8] T. Wang, H. Krim, and Y. Viniotis: A generalized Markov graph model: Application to social network analysis. IEEE J. Selected Topics Signal Process. *7* (2013), 318–332. DOI:10.1109/jstsp.2013.2246767

[9] R. W. Yeung: A new outlook on Shannon's information measures. IEEE Trans. Inform. Theory *37* (1991), 466–474. DOI:10.1109/18.79902

[10] R. W. Yeung: Information Theory and Network Coding. Springer, 2008.

[11] R. W. Yeung, A. Al-Bashabsheh, C. Chen, Q. Chen, and P. Moulin: On information-theoretic characterizations of Markov random fields and subfields. IEEE Trans. Inform. Theory *65* (2019), 1493–1511. DOI:10.1109/tit.2018.2866564

[12] R. W. Yeung, T. T. Lee, and Z. Ye: Information-theoretic characterizations of conditional mutual independence and markov random fields. IEEE Trans. Inform. Theory *48* (2002), 1996–2011. DOI:10.1109/tit.2002.1013139

Terence Chan, University of South Australia.
e-mail: terence.chan@unisa.edu.au

Qi Chen, Xidian University, China.
e-mail: qichen@xidian.edu.cn

Raymond Yeung, Department of Information Engineering, The Chinese University of Hong Kong.
e-mail: whyeung@ie.cuhk.edu.hk