Shared Information under Simple Markov Independencies

Madhura Pathegama and Sagnik Bhattacharya†

***Abstract*—Shared information is a measure of mutual dependence among $m \geq 2$ jointly distributed discrete random variables. We show that the shared information of a Markov random field in which the underlying graph has at least one cut vertex, is the same as the minimum shared information of its blocks (also called biconnected components). This generalizes prior results on shared information of Markov random fields to a much wider class of nontree graphs.**

***Index Terms*—Shared information, cut vertex, biconnected component.**

I. Introduction

Let $X_1, \ldots, X_m$, $m \geq 2$, be random variables (rvs) with finite alphabets. Their shared information $\mathrm{SI}(X_1, \ldots, X_m)$ is a measure of their mutual dependence. Shared information characterizes the largest rate of shared secret key that can be generated by a set of m terminals, with terminal i only having access to independent and identically distributed repetitions X_i^n of the rv X_i, but each terminal being able to interactively communicate with the others over a public, noiseless broadcast channel [11], [6]. Shared information subtracted from the joint entropy $\mathrm{H}(X_1, \ldots, X_m)$ is also the minimum communication rate necessary for each of the m terminals to achieve omniscience, that is, to learn the information $(X_1^n, \ldots, X_m^n)$ jointly available to all the terminals, with the same noiseless interactive broadcast communication as before [11], [6].

Shared information (SI) along with its expression in terms of divergence were introduced as an upper bound for the largest rate of shared secret key in [11]. The connection to omniscience was also elucidated therein. Tightness of the bound was shown for the case $m = 2$ and 3 in [11], as was the particularization to Shannon's mutual information for $m = 2$. The latter motivated the suggestion of SI as a measure of mutual dependence among multiple rvs (see also [15]).

In a significant advance, tightness for arbitrary m was later established in [3], [5], [9]. A comprehensive study of the properties of shared information, including an axiomatic approach to information measures for multiple rvs and a data processing inequality for SI, can be found in [6].

Myriad other operational interpretations of shared information in disparate areas of information theory include: maximal packing of edge-disjoint spanning trees in a mutigraph ([18], [17], see also [4], [10], [6]); optimum querying exponent for resolving common randomness [19]; strong converse for multiterminal secret key capacity [19], [20]; and also undirected network coding [5], and data clustering [8].

Using submodularity of entropy, an efficient algorithm to compute shared information when the underlying pmf is fully known was proposed in [6], with improvements in [12]. However, when the underlying pmf is only partially known or unknown, computing shared information becomes difficult since the definition involves an optimization over all partitions of a set of size m. Therefore, structural properties of shared information in specific contexts that can simplify the optimization problem are of interest. Such properties enable efficient estimation of shared information when the underlying pmf is unknown (see, for example, [2]). Even when the underlying pmf is known, they also yield avenues for efficiently achieving successive omniscience [7] wherein some subset of terminals achieves omniscience first before extending to all terminals.

Special models that allow explicit characterizations of shared information include the PIN model [17], [18]. Closer to our current work, explicit and simple formulae for SI for a Markov chain and a Markov chain on a tree were found in [11]. Materially different proofs that shed light on the structure of the SI-achieving partitions were obtained in [2], and yet another proof approach was introduced in [8]. The first similar characterization for nontree graphical models, using the approach of [2], was for the cliqueylon graph in [1].

Main contributions

Our main result provides structural insight into the shared information of a general graphical model, that holds whenever the graphical model has at least a single cut vertex. Since all of the models for which such explicit characterizations are known (Markov chain, Markov chain on a tree and the cliqueylon graph) have cutsets consisting of a single vertex, it generalizes all previously known results regarding SI in graphical models.

Our main result is that the shared information of a graph is the minimum of the shared information of its blocks (also known as biconnected components), which are its maximal 2-connected subgraphs. It is easy to see that cut vertices separate the graph into blocks. When the blocks are small, our characterization leads to significant efficiency gains in computing shared information. We also show that an SI-achieving partition of the graph can be easily obtained from the SI-achieving partition of the component achieving the minimum.

Graphs with cut vertices or blocks can be useful in statistically modeling joint distributions in multiple contexts. For

†M. Pathegama and S. Bhattacharya are with the Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. Email: {pankajap,sagnikb}@umd.edu. M. Pathegama is supported by the U.S. National Science Foundation under Grant CCF21004489. S. Bhattacharya is supported by the U.S. National Science Foundation under Grant CCF2310203.

979-8-3503-8284-6/24/$31.00 ©2024 IEEE

example, in mobile networks, when a single antenna services a geographical area, that antenna serves as a cut vertex between mobile phones in that area and all nodes in the rest of the network. When secure communication is necessary, each node in the network may wish to share a secret key, and shared information characterizes the maximum rate of such a secret key. Similarly, in a swarm of robots, a group of robots may have limited communication capabilities and interact with other robots in the swarm via a special robot. A graphical model with the special robot forming a cut vertex would be a good probabilistic model of the swarm. The robots may wish to share a map of their environment based on their collective sensing capabilities, which becomes a problem of omniscience. The minimum amount of communication necessary for such omniscience is then given by the shared information of the knowledge of each robot subtracted from the joint entropy.

Other examples of graphical models in which cut vertices are useful include biological neural networks in the brain, in which a small group of neurons form the connection between two much larger groups of neurons that form a lobe, and power grids, with a substation separating the downstream network from the upstream one.

Section II presents the preliminaries, the statements of the main theorem and a key technical proposition. A complete proof of the proposition is in Section III. Section IV contains closing remarks.

II. Preliminaries and Main Results

Let $X_1, \ldots, X_m$, $m \geq 2$, be rvs with finite alphabets $\mathcal{X}_1, \ldots, \mathcal{X}_m$, respectively, and joint pmf $P_{X_1 \cdots X_m}$. For $A \subseteq \mathcal{M} = \{1, \ldots, m\}$, let $X_A \triangleq (X_i, i \in A)$. Let $\pi = (\pi_1, \ldots, \pi_k)$ denote a k-partition of $\mathcal{M}$, $2 \leq k \leq m$, with atoms π_i, $1 \leq i \leq k$. Let $\Pi(\mathcal{M})$ be the set of all nontrivial partitions of $\mathcal{M}$, i.e., with $k \geq 2$ atoms. Hereafter we will consider only nontrivial partitions of $\mathcal{M}$.

Definition 1 (Shared information [16])**.** *The* shared information *of $X_1, \ldots, X_m$ is defined as*

$$\mathrm{SI}(X_{\mathcal{M}}) = \min_{\pi \in \Pi(\mathcal{M})} \frac{1}{|\pi| - 1} D(P_{X_{\mathcal{M}}} \parallel \prod_{u=1}^{|\pi|} P_{X_{\pi_u}}). \quad (1)$$

Given a partition $\pi \in \Pi(\mathcal{M})$, we denote

$$\mathcal{I}_\pi(X_{\mathcal{M}}) = \frac{1}{|\pi| - 1} D(P_{X_{\mathcal{M}}} \parallel \prod_{u=1}^{|\pi|} P_{X_{\pi_u}}),$$

so that $\mathrm{SI}(X_{\mathcal{M}}) = \min_{\pi \in \Pi(\mathcal{M})} \mathcal{I}_\pi(X_{\mathcal{M}})$.

Remark 1. A useful way to write $\mathcal{I}_\pi(X_{\mathcal{M}})$ in terms of mutual information is

$$\mathcal{I}_\pi(X_{\mathcal{M}}) = \frac{1}{|\pi| - 1} \sum_{i=2}^{k} \mathrm{I}(X_{\pi_i} \wedge X_{\pi_1}, \ldots, X_{\pi_{i-1}}).$$

In the setting of a Markov random field (MRF), the rvs $X_1, \ldots, X_m$ are associated with the vertices of a graph and exhibit Markov properties based on the structure of the graph. The Markov properties rely on the notion of *separation*. Given a graph $\mathcal{G} = (\mathcal{M}, \mathcal{E})$ with vertex set $\mathcal{M} = \{1, \ldots, m\}$ and edge set $\mathcal{E}$, let A, B and S be (pairwise) disjoint, nonempty subsets of $\mathcal{M}$. Then S *separates* A and B if for every $a \in A$, $b \in B$, any path that connects A to B has at least one vertex $s = s(a, b)$ in S.

Definition 2 (Global Markov property [14])**.** *Given a graph $\mathcal{G} = (\mathcal{M}, \mathcal{E})$, assign rv X_i to vertex i, $i \in \mathcal{M}$. The pmf $P_{X_{\mathcal{M}}} = P_{X_1 \cdots X_m}$ satisfies the global Markov property with respect $\mathcal{G}$ if for every triple of disjoint, nonempty subsets A, B, S of $\mathcal{M}$ such that S separates A and B, the following Markov condition holds:*

$$X_A \multimap X_S \multimap X_B.$$

Hereafter, the global Markov property will be termed simply the Markov property.

In this paper we only consider *connected* graphs; by convention, disconnected graphs correspond to independent sets of rvs which have $\mathrm{SI} = 0$.

To precisely state our main result, we need the notion of blocks in graphs. For a general introduction to graph theory, see [13].

Definition 3 (Induced subgraph, cut vertex, block [13])**.** *Given a subset $A \subsetneq \mathcal{M}$ of vertices, the subgraph of $\mathcal{G}$ induced by A, denoted by $\mathcal{G}_A$, is the graph with vertex set A and edge set $\mathcal{E}_A = \{(i, j) \in \mathcal{E} : i, j \in A\}$.*

A cut vertex *is any vertex $v \in \mathcal{M}$ such that there exist nonempty subsets $A, B \subsetneq \mathcal{M}$ with $A \cap B = \{v\}$ and $A \cup B = \mathcal{M}$ such that v separates $A \setminus \{v\}$ from $B \setminus \{v\}$. Equivalently, $v \in \mathcal{M}$ is a cut vertex if the subgraph $\mathcal{G}_{\mathcal{M} \setminus \{v\}}$ is a disconnected graph.*

A block*, or a* biconnected component*, is a maximal subgraph $\mathcal{G}' \subseteq \mathcal{G}$ that cannot be separated into nontrivial disconnected components by erasing any single vertex in $\mathcal{G}'$.*

Remark 2. A block of size 2 consists of two vertices connected by an edge. Any pair of vertices in a block of size ≥ 3 must have two vertex-disjoint paths between them.

Remark 3. Every vertex $v \in \mathcal{M}$ belongs to some block. Further, blocks and cut vertices are intimately connected. In particular, the intersection of two blocks is either empty or a single vertex that is a cut vertex of $\mathcal{G}$. This is easy to see since if the intersection had more than one vertex, for every pair of vertices with one vertex from each block, there would be at least two vertex-disjoint paths between them, contradicting maximality. See Figure 1.

Example 1. A complete graph is itself a block. If $\mathcal{G}$ is a tree, then every adjacent pair of vertices form a block. The cliqueylon graph (introduced in [1]) contains a central clique (complete graph) with each vertex in the clique being the root of a tree. The central clique is a block, and every adjacent pair of vertices in the trees is a block. □

Let $\mathcal{B}(\mathcal{G})$ be the set of blocks in $\mathcal{G}$. The following is our main result.

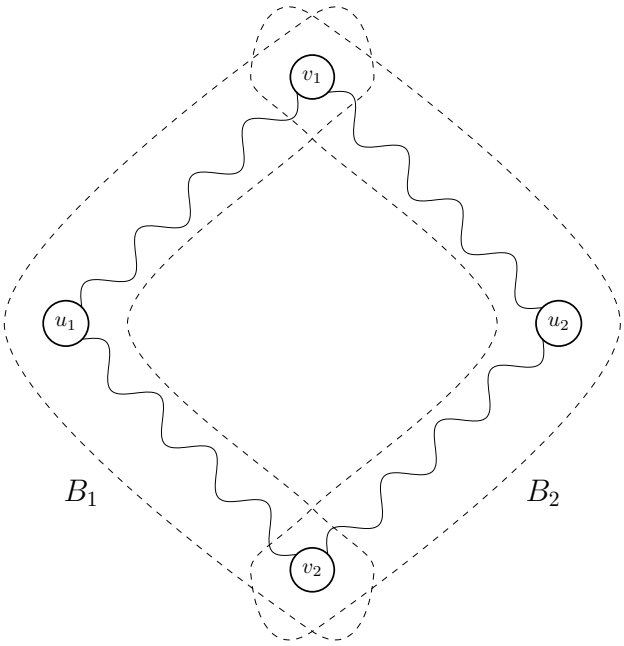

Figure 1. Let B_1 and B_2 be blocks such that they intersect at two vertices v_1 and v_2. For any $u_1 \in B_1$ and $u_2 \in B_2$, there are two vertex-disjoint paths between them.

Theorem 1. *The SI of $X_{\mathcal{M}}$ is equal to the minimum SI among all blocks in the graph.*

$$\mathrm{SI}(X_{\mathcal{M}}) = \min_{B \in \mathcal{B}(\mathcal{G})} \mathrm{SI}(X_B). \tag{2}$$

Example 2. Consider a graph consisting of two cliques with exactly one vertex in common. See Figure 2. Theorem 1 shows that the SI of all the rvs equals the minimum of the SIs of the rvs in the two cliques. □

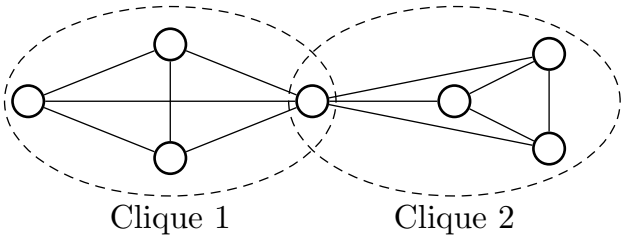

Figure 2. Example with two cliques sharing a common vertex

Remark 4. If the underlying graph is a path graph or a tree, the resulting MRF is called a Markov chain or a Markov chain on a tree. Both are graphs with no cycles, and the SI in both cases is given by $\mathrm{SI}(X_{\mathcal{M}}) = \min_{(i,j)\in\mathcal{E}} \mathrm{I}(X_i, X_j)$ [2], [11]. Among nontree graphical models, the SI for the cliqueylon graph was shown in [1] to be the minimum of the SI of the clique rvs and the minimum mutual information across any tree-edge. In light of Example 1, Theorem 1 generalizes all of these results to any nontree graphical model that is not itself a block. Notice that the graph in Example 2 is neither a tree nor a cliqueylon.

The main ingredient of the proof of Theorem 1 is the following technical result.

Proposition 2. *Let $X_{\mathcal{M}} = X_A \cup \{Z\} \cup X_B$ such that $X_A - \circ - Z - \circ - X_B$. Then,*

$$\mathrm{SI}(X_{\mathcal{M}}) = \min\left\{\mathrm{SI}(X_A, Z), \mathrm{SI}(X_B, Z)\right\}. \tag{3}$$

Further, if $\mathrm{SI}(X_{\mathcal{M}}) = \mathrm{SI}(X_A, Z)$ and $\pi^* = (\pi_1, \ldots, \pi_k)$ achieves SI for (X_A, Z), then

$$\pi'_k = \begin{cases} \pi_k \cup B, & \text{if } Z \in \pi_k \\ \pi_k, & \text{otherwise} \end{cases}$$

is an SI-achieving partition for $\mathcal{M}$.

Remark 5. In Proposition 2, $Z = X_v$ for some $v \in \mathcal{M}$.

The proof of Proposition 2 is given in the next section. Theorem 1 follows from Proposition 2.

Proof of Theorem 1. We prove Theorem 1 by applying induction on the size of $\mathcal{B}(\mathcal{G})$.

If $|\mathcal{B}(\mathcal{G})| = 2$, there are 2 blocks in $\mathcal{G}$ which we call P_1 and P_2. Since $\mathcal{G}$ is connected, this implies that the intersection of these blocks has to be a cut vertex which we call p. Since p is a cut vertex, the Markov property implies $X_{P_1\setminus\{p\}} - \circ - X_p - \circ - X_{P_2\setminus\{p\}}$ and applying Proposition 2 with $A = P_1 \setminus \{p\}$, $B = P_2 \setminus \{p\}$ and $Z = X_p$, we prove the base case.

Now assume that the statement is true for all values of $|\mathcal{B}(\mathcal{G})|$ up to some integer m. If $|\mathcal{B}(\mathcal{G})| = m+1$, because $\mathcal{G}$ is connected, each block contains at least one cut vertex. Choose one such cut vertex p. Let P_1 and P_2 be the sets of vertices separated by p. Again applying Proposition 2,

$$\mathrm{SI}(X_{\mathcal{M}}) = \min\left\{\mathrm{SI}(X_{P_1\cup\{p\}}), \mathrm{SI}(X_{P_2\cup\{p\}})\right\}.$$

Since $|\mathcal{B}(\mathcal{G}_{P_1\cup\{p\}})| \le m$ and $|\mathcal{B}(\mathcal{G}_{P_2\cup\{p\}})| \le m$, by the induction hypothesis,

$$\begin{aligned}
&\mathrm{SI}(X_{\mathcal{M}}) \\
&= \min\left\{\mathrm{SI}(X_{P_1\cup\{p\}}), \mathrm{SI}(X_{P_2\cup\{p\}})\right\} \\
&= \min\left\{\min_{B_1 \in \mathcal{B}(\mathcal{G}_{P_1\cup\{p\}})} \mathrm{SI}(X_{B_1}), \min_{B_2 \in \mathcal{B}(\mathcal{G}_{P_2\cup\{p\}})} \mathrm{SI}(X_{B_2})\right\} \\
&= \min_{B \in \mathcal{B}(\mathcal{G})} \mathrm{SI}(X_B),
\end{aligned}$$

concluding the proof. □

A direct implication of the second part of Proposition 2 is the following corollary which constructs an SI-achieving partition for $X_{\mathcal{M}}$ using an SI-achieving partition of the optimal block.

Corollary 3. *Let $B^* \in \mathcal{B}(\mathcal{G})$ such that $\mathrm{SI}(X_{\mathcal{M}}) = \mathrm{SI}(X_{B^*})$. Let $\pi^*(B^*) = (\pi_1, \ldots, \pi_k)$ be an SI-achieving partition of X_{B^*}. Define the sets $\rho_i \subseteq \mathcal{M} \setminus B^*$, $i = 1, \ldots, k$, such that*

$$\rho_i = \{v \in \mathcal{M} \setminus B^* : \exists \text{ a path } P \text{ between } v \text{ and } x \in \pi_i \text{ such that } P \cap \pi_j = \emptyset \text{ for } i \neq j\}.$$

Then the partition $\pi^*(\mathcal{M}) = (\pi'_1, \ldots, \pi'_k)$ of $X_{\mathcal{M}}$ constructed as follows is an SI-achieving partition of $X_{\mathcal{M}}$:

$$\pi'_i = \pi_i \cup \rho_i.$$

Corollary 3 shows that an SI-achieving partition for $X_{\mathcal{M}}$ can be obtained by picking an SI-achieving partition of the block B^* that achieves the minimum on the right side of (2) and for each cut-vertex in that block, adding to the atom containing the cut-vertex the vertices of $\mathcal{M} \setminus B^*$ that are connected to B^* via the cut-vertex. See Figure 3.

Remark 6. The intersections of the sets ρ_i and B^* partition the set of cut vertices of $\mathcal{G}$ that lie in B^*.

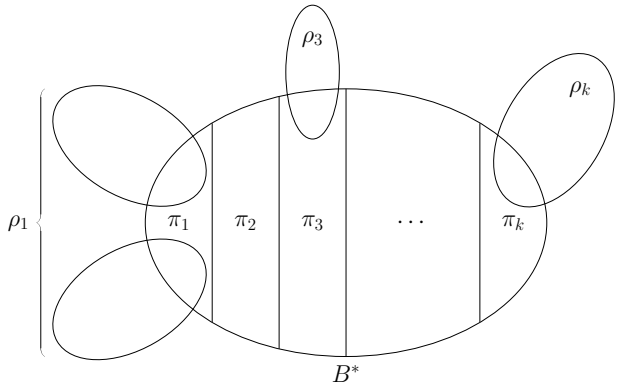

Figure 3. Constructing an optimal partition for $X_{\mathcal{M}}$ using the optimal partition for the block with minimum SI.

III. Proof of Proposition 2

The first part of Proposition 2 is a consequence of the following two lemmas, proved separately.

Lemma 4. *Let $X_{\mathcal{M}} = X_A \cup \{Z\} \cup X_B$. Then,*

$$\mathrm{SI}(X_{\mathcal{M}}) \geq \min\left\{\mathrm{SI}(X_A, Z), \mathrm{SI}(X_B, Z)\right\}. \tag{4}$$

A more general result for *nonsingleton* Z was shown in [6, Corollary 5.1] using submodular optimization. Our proof for singleton Z is elementary and provides some structural interpretation via an auxiliary rv.

Lemma 5. *Let $X_{\mathcal{M}} = X_A \cup \{Z\} \cup X_B$ such that $X_A - \circ - Z - \circ - X_B$. Then,*

$$\mathrm{SI}(X_{\mathcal{M}}) \leq \min\left\{\mathrm{SI}(X_A, Z), \mathrm{SI}(X_B, Z)\right\}. \tag{5}$$

Remark 7. Note that Lemma 4 does *not* assume any structure on the $X_{\mathcal{M}}$.

Proof of Lemma 4. Let $Z' = Z$ be an auxiliary rv. We first show that given an SI-achieving partition $\pi^* = (\pi_1, \ldots, \pi_k)$ of $X_{\mathcal{M}} = (X_A, Z, X_B)$, we can construct a partition π' of $X'_{\mathcal{M}} = (X_A, Z, Z', X_B)$ such that

$$\mathrm{SI}(X_{\mathcal{M}}) = \mathrm{I}_{\pi'}(X'_{\mathcal{M}}). \tag{6}$$

Without loss of generality, let $Z \in \pi_1$. It is easily seen that the partition $\pi' = (\pi'_1, \ldots, \pi'_k) = (\pi_1 \cup \{Z'\}, \pi_2, \ldots, \pi_k)$ has the required property, since $Z' = Z$ implies $\mathrm{H}(Z' \,|\, X_{\pi_1}) = \mathrm{H}(Z' \,|\, X_{\mathcal{M}}) = 0$ and therefore

$$\begin{aligned}
&\mathrm{I}_{\pi'}(X'_{\mathcal{M}}) \\
&\quad = \frac{1}{k-1}\left[\mathrm{H}(X_{\pi_1}, Z') + \sum_{i=2}^{k} \mathrm{H}(X_{\pi_i}) - \mathrm{H}(X_{\mathcal{M}}, Z')\right] \\
&\quad = \frac{1}{k-1}\left[\sum_{i=1}^{k} \mathrm{H}(X_{\pi_i}) - \mathrm{H}(X_{\mathcal{M}})\right] \\
&\quad = \mathrm{SI}(X_{\mathcal{M}}).
\end{aligned}$$

We now claim that

$$\mathrm{I}_{\pi'}(X'_{\mathcal{M}}) \geq \min\left\{\mathrm{SI}(X_A, Z), \mathrm{SI}(X_B, Z')\right\}, \tag{7}$$

which proves Lemma 4 in light of (6).

To prove the claim, let $Y_A = (X_A, Z)$ and $Y_B = (X_B, Z')$. We group the atoms of π' into three sets T_A, T_B and T_{AB} such that T_A (resp. T_B) consists of atoms of π' that only contain rvs from Y_A (resp. Y_B) while the atoms in T_{AB} consists of elements from both Y_A and Y_B.

$$\begin{aligned}
T_A &= \{\pi'_i : \pi'_i \cap Y_B = \emptyset, i = 1, \ldots, k\}, \\
T_B &= \{\pi'_i : \pi'_i \cap Y_A = \emptyset, i = 1, \ldots, k\}, \\
T_{AB} &= \pi' \setminus (T_A \cup T_B).
\end{aligned}$$

Let $|T_{AB}| = p$, $|T_A| = q$ and $|T_B| = r$. Note that $p+q+r = k$ and also that $p \geq 1$, since π'_1 contains both Z and Z' and therefore intersects both Y_A and Y_B.

In what follows, we require $p + q \geq 2$. If $p = 1$ and $q = 0$, we must have $r \geq 1$ since $k \geq 2$. In this case, we henceforth interchange the roles of A and B.

Assume without loss of generality that

$$\begin{aligned}
T_{AB} &= \{\pi'_i, i = 1, \ldots, p\}, \\
T_A &= \{\pi'_i, i = p+1, \ldots, p+q\}, \\
T_B &= \{\pi'_i, i = p+q+1, \ldots, k\}.
\end{aligned}$$

Further, for each $i = 1, \ldots, p$, let $\pi'_i = \pi'_{i,A} \sqcup (\pi'_i \setminus \pi'_{i,A})$ where $\pi'_{i,A} = \pi'_i \cap Y_A$. Letting $\pi'_{i,A} = \pi'_i$ for $i = p+1, \ldots, p+q$, we get the nontrivial partition of Y_A given by $\pi'_A = \{\pi'_{i,A}, i = 1, \ldots, p+q\}$.

Case 1 ($r > 0$): Let $\pi'_{0,B} = \cup_{i=1}^{p}(\pi'_i \setminus \pi'_{i,A})$ and $\pi'_{i,B} = \pi'_{i+p+q}$ for $i = 1, \ldots, r$. The sets $\pi'_B = \{\pi'_{i,B}, i = 0, \ldots, r\}$

are a (nontrivial) partition of Y_B. Then, (see Remark 1)

$$
\begin{aligned}
&\mathrm{I}_{\pi'}(X'_{\mathcal{M}}) \\
&= \frac{1}{k-1}\sum_{i=2}^{k} \mathrm{I}(X_{\pi'_i} \wedge X_{\pi'_1}, \ldots, X_{\pi'_{i-1}}) \\
&= \frac{1}{k-1}\sum_{i=2}^{p+q} \mathrm{I}(X_{\pi'_i} \wedge X_{\pi'_1}, \ldots, X_{\pi'_{i-1}}) \\
&\qquad\qquad + \frac{1}{k-1}\sum_{i=p+q+1}^{k} \mathrm{I}(X_{\pi'_i} \wedge X_{\pi'_1}, \ldots, X_{\pi'_{i-1}}) \\
&\geq \frac{1}{k-1}\sum_{i=2}^{p+q} \mathrm{I}(X_{\pi'_{i,A}} \wedge X_{\pi'_{1,A}}, \ldots, X_{\pi'_{i-1,A}}) \\
&\qquad\qquad + \frac{1}{k-1}\sum_{i=1}^{r} \mathrm{I}(X_{\pi'_{i,B}} \wedge X_{\pi'_{0,B}}, \ldots, X_{\pi'_{i-1,B}}) \\
&= \frac{p+q-1}{k-1}\,\mathrm{I}_{\pi'_A}(Y_A) + \frac{r}{k-1}\,\mathrm{I}_{\pi'_B}(Y_B) \\
&\geq \frac{p+q-1}{k-1}\,\mathrm{SI}(Y_A) + \left(1 - \frac{p+q-1}{k-1}\right)\mathrm{SI}(Y_B) \\
&\geq \min\left\{\mathrm{SI}(Y_A), \mathrm{SI}(Y_B)\right\}.
\end{aligned}
$$

where the first inequality follows because $\pi'_{i,A} \subseteq \pi'_i$, $i = 1, \ldots, p+q$, and $\pi'_{0,B} \subseteq \cup_{i=1}^{p+q}\pi'_{i-1}$.

Case 2 ($r = 0$): In this case, $p + q = k$ and the calculation above yields $\mathrm{I}_{\pi'}(X'_{\mathcal{M}}) \geq \mathrm{SI}(Y_A)$, which implies (7). □

Proof of Lemma 5. Without loss of generality, it is sufficient to prove that $\mathrm{SI}(X_{\mathcal{M}}) \leq \mathrm{SI}(X_A, Z)$. Let $\pi^* = (\pi_1, \ldots, \pi_k)$ be an SI-achieving partition for (X_A, Z), that is, a partition that achieves the minimum on the right side of (1). Assume that $Z \in \pi_1$. Then, $\pi = (\pi_1 \cup X_B, \pi_2, \ldots, \pi_k)$ is a nontrivial partition of $X_{\mathcal{M}}$ and we get

$$
\begin{aligned}
&\mathrm{SI}(X_{\mathcal{M}}) \\
&\leq \mathcal{I}_{\pi}(X_{\mathcal{M}}) \\
&= \frac{1}{k-1}\left[\sum_{i=1}^{k}\mathrm{H}(X_{\pi_i}) - \mathrm{H}(X_{\mathcal{M}})\right] \\
&= \frac{1}{k-1}\left[\mathrm{H}(X_{\pi_1}, X_B) + \sum_{i=2}^{k}\mathrm{H}(X_{\pi_i}) - \mathrm{H}(X_A, Z, X_B)\right] \\
&= \frac{1}{k-1}\left[\mathrm{H}(X_{\pi_1}) + \mathrm{H}(X_B \,|\, Z)\right. \\
&\qquad\qquad \left. + \sum_{i=2}^{k}\mathrm{H}(X_{\pi_i}) - \mathrm{H}(X_A, Z) - \mathrm{H}(X_B \,|\, Z)\right] \quad (8) \\
&= \mathrm{SI}(X_A, Z),
\end{aligned}
$$

where (8) is a consequence of the Markov chains $X_A \multimap Z \circ\!\!-\!\!\circ X_B$ (notice that $\pi_1 \subseteq A$). □

The second part of Proposition 2 also follows from the proof of Lemma 5 above.

IV. Closing Remarks

A pertinent follow-up to Theorem 1 would be to identify general classes of graphs in which blocks are small on average, the setting in which our result leads to the most significant savings.

A generalization of Theorem 1 to graphical models without a cut vertex remains an open problem. Our proofs rely on there being a single cut vertex which then belongs to a single atom of the relevant partition. Particularly, in the proof of Lemma 5, we needed X_A and X_B to be conditionally independent given a single Z, and we could then find an atom of the SI-achieving partition that contained Z. This is no longer the case for a cutset with more than one vertex, which may then be spread out among multiple atoms of an SI-achieving partition.

Acknowledgements

The authors thank Prakash Narayan for many helpful discussions and his comments on this manuscript, and to the reviewers for their valuable comments.

References

[1] S. Bhattacharya and P. Narayan, "Shared information for the cliqueylon graph," in *2023 IEEE International Symposium on Information Theory (ISIT)*, IEEE, Jun. 2023.

[2] S. Bhattacharya and P. Narayan, "Shared information for a Markov chain on a tree," *IEEE Transactions on Information Theory*, 2024.

[3] C. Chan, "On tightness of mutual dependence upper-bound for secret-key capacity of multiple terminals," *ArXiv*, vol. abs/0805.3200, 2008.

[4] C. Chan, "Linear perfect secret key agreement," in *2011 IEEE Information Theory Workshop*, 2011.

[5] C. Chan, "The hidden flow of information," *2011 IEEE International Symposium on Information Theory Proceedings*, 2011.

[6] C. Chan, A. Al-Bashabsheh, J. B. Ebrahimi, T. Kaced, and T. Liu, "Multivariate mutual information inspired by secret-key agreement," *Proceedings of the IEEE*, vol. 103, no. 10, 2015.

[7] C. Chan, A. Al-Bashabsheh, Q. Zhou, N. Ding, T. Liu, and A. Sprintson, "Successive omniscience," *IEEE Transactions on Information Theory*, vol. 62, no. 6, 2016.

[8] C. Chan, A. Al-Bashabsheh, Q. Zhou, T. Kaced, and T. Liu, "Info-clustering: A mathematical theory for data clustering," *IEEE Transactions on Molecular, Biological, and Multi-Scale Communications*, vol. 2, no. 1, 2016.

[9] C. Chan and L. Zheng, "Mutual dependence for secret key agreement," in *2010 44th Annual Conference on Information Sciences and Systems (CISS)*, 2010.

[10] T. A. Courtade and T. R. Halford, "Coded cooperative data exchange for a secret key," *IEEE Transactions on Information Theory*, vol. 62, no. 7, 2016.

[11] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, Dec. 2004.

[12] N. Ding, P. Sadeghi, and T. Rakotoarivelo, "Improving computational efficiency of communication for omniscience and successive omniscience," *IEEE Transactions on Information Theory*, vol. 67, no. 7, 2021.

[13] F. Harary, *Graph Theory*. Reading, MA: Addison-Wesley, 1969.

[14] S. L. Lauritzen, *Graphical Models*. Oxford University Press, 1996.

[15] P. Narayan, "Omniscience and secrecy," Plenary Talk, *IEEE International Symposium on Information Theory*, Cambridge, MA, 2012.

[16] P. Narayan and H. Tyagi, "Multiterminal secrecy by public discussion," *Foundations and Trends in Communications and Information Theory*, vol. 13, no. 2-3, 2016.

[17] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy, and Steiner tree packing," *IEEE Transactions on Information Theory*, vol. 56, no. 12, 2010.

[18] S. Nitinawarat, C. Ye, A. Barg, P. Narayan, and A. Reznik, "Secret key generation for a pairwise independent network model," *IEEE Transactions on Information Theory*, vol. 56, no. 12, Dec. 2010.

[19] H. Tyagi and P. Narayan, "How many queries will resolve common randomness?" *IEEE Transactions on Information Theory*, vol. 59, no. 9, 2013.

[20] H. Tyagi and S. Watanabe, "Converses for secret key agreement and secure computing," *IEEE Transactions on Information Theory*, vol. 61, no. 9, 2015.