Shared Information for the Cliqueylon Graph

Sagnik Bhattacharya and Prakash Narayan†

2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10206503

***Abstract*—Shared information is a measure of mutual dependence among $m \geq 2$ jointly distributed discrete random variables. A new undirected probabilistic graphical model, a cliqueylon graph, is introduced, with potential applications in leader-follower swarms and neuron clusters with correlations of varying strength. Shared information is characterized explicitly for the cliqueylon, relying on structural properties of an underlying optimization. Implications for the data compression problem of omniscience are highlighted.**

***Index Terms*—Shared information, omniscience, cliqueylon graph, undirected probabilistic graphical model.**

I. Introduction

Suppose that $m \geq 2$ parties, possessing varying extents of partial knowledge such as different parts of a large file, all wish to acquire full knowledge of the entire file, thereby achieving *omniscience*. To this end, the parties communicate among themselves with minimal interactive exchange of bits on a public and noiseless broadcast channel. What is the smallest possible rate of communication using which such omniscience can be attained by all the parties? A description of omniscience is provided in Section VII below. Representing the initial information of party i by a random variable (rv) X_i, $1 \leq i \leq m$, it was shown in [10] that the *shared information* (SI) of $X_1, \ldots, X_m$ subtracted from the joint entropy $\mathrm{H}(X_1, \ldots, X_m)$ is a lower bound for this smallest communication rate. Further, an intrinsic connection was observed between achieving omniscience and generating a common "secret key" by means of such interactive communication, with SI serving as an upper bound for the largest rate of shared common randomness that the m terminals could generate. Simultaneous tightness of both bounds was observed for $m = 2$ and 3 in [10], and established for arbitrary $m \geq 2$ in [2], [4], [8]. It was noted already in [10] that shared information, for $m = 2$, particularized to mutual information between X_1 and X_2. This led to the suggestion of SI as a measure of mutual dependence among multiple rvs [10], [13]. In a significant advance, the role of SI was enlarged by properties developed in [5], where it is termed "multivariate mutual information." Appealing properties of shared information derived in [5] include a data processing inequality.

Shared information also appears in a central role in fundamental bounds in contexts other than omniscience and secrecy capacity. These include maximal packing of edge-disjoint spanning trees in a multigraph ([16], [15], and also variant models in [3], [9], [5]); optimum querying exponent for resolving common randomness [17]; strong converse for multiterminal secret key capacity [17], [18]; and also undirected network coding [4], and data clustering [7].

The expression for shared information involves an optimization over all possible nontrivial partitions of $X_1, \ldots, X_m$. In general, a computation of SI can be prohibitive for all but the smallest values of m. An efficient algorithm for computing SI, when the underlying pmf $P_{X_1 \cdots X_m}$ of $X_1, \ldots, X_m$ is known, has been proposed in [5]. Our approach is based on a different viewpoint: In specific settings of $P_{X_1 \cdots X_m}$, can structural insights be drawn concerning the mentioned optimization or the form of SI itself? Such special structure could facilitate, for instance, the design of communication protocols for *successive* omniscience by achieving it first for a subset of suitable parties before extending it to all parties [6]. When $P_{X_1 \cdots X_m}$ is unknown, such structure could point to a means for estimating SI.

Special models of interest that enable such insights into explicit characterizations of SI have been considered, for instance, in [10], [15]. Motivating our present line of work, a simple formula for SI for tree-structured graphical models, viz. Markov chain and Markov chain on a tree, was introduced in [10]. Distinctively separate proofs that relied on special structural properties were obtained in [1] (see also [7]); also, these properties led to an algorithm for estimating SI when $P_{X_1 \cdots X_m}$ was not known. *A similar analysis for characterizing SI for nontree graphical models was not known.*

Main contributions

We introduce a new class of nontree undirected probabilistic graphical models in which the underlying graph is a *cliqueylon*, consisting of a central clique, each vertex of which is also the root of a tree. Such a graphical model can be used to describe settings featuring a fully connected "command center" consisting of leader agents that communicate directly among themselves, with each such leader overseeing a secondary group of agents. The probabilistic behavior of each secondary group and its leader is modeled as a tree-structured undirected graphical model.

A potential application of the model above is in a swarm of mobile robots with a small group of freely communicating leader robots, each of which controls a distinct set of follower robots with limited communication capabilities. Omniscience in this setting would provide state information about each robot to every other robot, enabling, for example, better routing decisions based on the state of the entire swarm. In a neuroscience context, a central clique can be considered to represent a particular lobe of the brain with many strongly interconnected neurons; several smaller sparsely connected

†S. Bhattacharya and P. Narayan are with the Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. E-mail: {sagnikb, prakash}@umd.edu. This work was supported by the U.S. National Science Foundation under Grant CCF1910497.

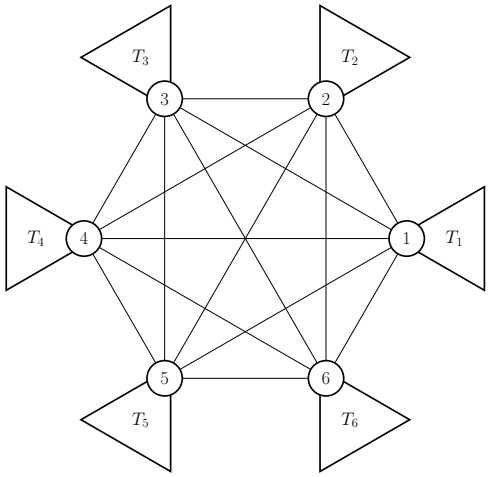

Figure 1. Cliqueylon with 6 vertices in the central clique

neurons connected to the lobe are depicted probabilistically as trees.

In the setting of a cliqueylon graph with known graphical structure and joint pmf, we obtain a structural result concerning the optimization inherent in the definition of SI that leads to a simple and explicit formula. The notion of graph connectivity plays a material role in our analysis. We first show the sufficiency of optimizing over a restricted class of partitions in which atoms and complements of atoms of the partition are all connected. We then exploit this restriction and special graph structure of the cliqueylon to simplify further the class of feasible partitions. The restricted set of partitions underlies our simple formula for SI of such a graph as the minimum of SI of the clique and SI of the trees. Our approach affords a simple method for computing SI when the clique is small; even if the clique is large, the computational savings can be significant.

Section II presents the preliminaries, and statements of a key technical lemma and main results for the shared information of a cliqueylon graph. Complete proofs of all our results are presented in Sections III to VI. Section VII 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, all partitions of $\mathcal{M}$, including those with special properties below, will allude to nontrivial partitions.

Definition 1 (Shared information [10], [14])**.** *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 $\mathrm{SI}(X_{\mathcal{M}}) = \min_{\pi \in \Pi(\mathcal{M})} \mathcal{I}_\pi(X_{\mathcal{M}})$.

In general, there can be multiple partitions in $\Pi(\mathcal{M})$ that attain $\mathrm{SI}(X_{\mathcal{M}})$. By [5, Theorem 5.2], there exists a unique partition $\pi^* \in \Pi(\mathcal{M})$, termed *fundamental partition*, with $\mathcal{I}_{\pi^*}(X_{\mathcal{M}}) = \mathrm{SI}(X_{\mathcal{M}})$ such that every $\mathrm{SI}(X_{\mathcal{M}})$-attaining partition is a coarsening of π^*.

A complete characterization of the properties of π^* is in [5]. For our purposes, the following property is pertinent.

Fact [5, Proposition 5.3]: If $A \subseteq \mathcal{M}$ is such that $\mathrm{SI}(X_A) > \mathrm{SI}(X_{\mathcal{M}})$, then A is either an atom of π^* or a subset of an atom of π^*.

The rvs $X_1, \ldots, X_m$ will be associated with a suitable underlying graph and endowed with Markov properties based on the structure of the graph. The notion of *separation* will be pertinent. 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 $a \in B$, $b \in B$, any path that connects A to B has at least one vertex $s = s(a, b)$ in S.

Definition 2 (Global Markov property [12])**.** *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.

Remark 1. *If $\mathcal{G} = (\mathcal{M}, \mathcal{E})$ is a clique, i.e., with an edge connecting every pair of vertices, no triple A, B and S as above exist.*

Upon augmenting a clique with appropriate trees, a rich class of graphs emerge that are compatible with the Markov property.

Definition 3 (Cliqueylon graph)**.** *Let $\mathcal{G} = (\mathcal{M}, \mathcal{E})$ be a graph consisting of a (central) clique $\mathcal{C}$, with each vertex $i \in \mathcal{C}$ being the root of a (hanging) tree T_i. The vertex set $\mathcal{M}$ is then given by*

$$\mathcal{M} = \mathcal{C} \cup \left(\bigcup_{i \in \mathcal{C}} T_i \setminus \{i\} \right),$$

and the edge set $\mathcal{E}$ is made up of edges between every pair of vertices in $\mathcal{C}$ and the edges in each tree T_i, $i \in \mathcal{C}$.

Hereafter, we shall sometimes denote a subgraph and also its vertices by the same symbol. This abuse of notation should cause no confusion as the distinction will be clear from the context.

We recall that given a graph $\mathcal{G} = (\mathcal{M}, \mathcal{E})$, a subset $A \subseteq \mathcal{M}$ is connected if the subgraph of $\mathcal{G}$ induced by A is connected. Let $\Pi_a(\mathcal{M}) \subseteq \Pi(\mathcal{M})$ be the set of partitions of $\mathcal{M}$ such that every atom of a partition is connected. Further, let $\Pi_{ac}(\mathcal{M}) \subseteq \Pi_a(\mathcal{M})$ be the set of partitions of $\mathcal{M}$ such that every atom of the partition is connected *and* the complement of every atom is also connected. It is clear that $\Pi_{ac}(\mathcal{M}) \neq \phi$ for a cliqueylon.

The notion of *maximal* connectivity will also be useful. Given a graph $\mathcal{G} = (\mathcal{M}, \mathcal{E})$, let $A \subsetneq A' \subsetneq \mathcal{M}$. Then, A is maximally connected in A' if A is connected and the addition to A of any vertex u in $A' \setminus A$ renders $A \cup \{u\}$ disconnected. Any disconnected subset of $\mathcal{M}$ can be written as a disjoint union of maximally connected subsets in it. *For a disconnected atom (of a partition), the corresponding maximally connected constituent subsets hereafter will be called subatoms.*

A cliqueylon has the following key property that we exploit repeatedly in our proofs.

Lemma 1. *Consider a cliqueylon $\mathcal{G} = (\mathcal{M}, \mathcal{E})$. Let A be a disconnected atom of some partition $\pi \in \Pi(\mathcal{M})$. Then there exists at least one subatom $A_i \subsetneq A$ such that cutting a single boundary edge (one that connects a vertex in A_i to a vertex in $A \setminus A_i$) renders A into two connected components, with A_i lying in one component and every other subatom of A lying in the other.*

Our two main results are the following.

Theorem 2. *Let $P_{X_\mathcal{M}}$ satisfy the Markov property with respect to the cliqueylon $\mathcal{G} = (\mathcal{M}, \mathcal{E})$. Then,*

$$\mathrm{SI}(X_\mathcal{M}) = \min_{\pi \in \Pi_a(\mathcal{M})} \mathcal{I}_\pi(X_\mathcal{M}) = \min_{\pi \in \Pi_{ac}(\mathcal{M})} \mathcal{I}_\pi(X_\mathcal{M}). \quad (2)$$

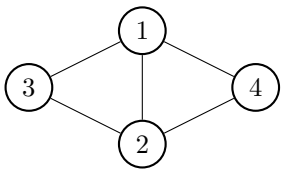

Figure 2. Graph in Remark 2

Remark 2. *Theorem 2 does not hold for general nontree graphs. For example, let X_1, X_2 be Bernoulli rvs with joint pmf given by $P_{X_1X_2}(0,0) = 0.01$, $P_{X_1X_2}(0,1) = 0.4$, $P_{X_1X_2}(1,0) = 0.25$ and $P_{X_1X_2}(1,1) = 0.34$. Let $N_1 = \mathrm{Ber}(0.01)$ and $N_2 = \mathrm{Ber}(0.99)$ be independent Bernoulli rvs, and let $X_3 = X_1 \oplus X_2 \oplus N_1$ and $X_4 = X_1 \cdot X_2 \oplus N_2$ ($\oplus$ denotes $\bmod 2$ addition). The rvs X_1, X_2, X_3, X_4 satisfy the Markov property with respect to the graph in Figure 2. Explicit calculations show that the unique minimizing partition in* (1) *is $\{\{3,4\},\{1\},\{2\}\}$, and clearly $\{3,4\}$ is not a connected subset.*

Remark 3. *Theorem 2 asserts that the minimization over partitions in $\Pi(\mathcal{M})$ can be restricted to partitions in $\Pi_a(\mathcal{M})$ and even further to those in $\Pi_{ac}(\mathcal{M})$.*

Theorem 3. *Let $P_{X_\mathcal{M}}$ satisfy the Markov property with respect to the cliqueylon $\mathcal{G} = (\mathcal{M}, \mathcal{E})$ in Definition 3. Let $\mathcal{E}_\mathcal{C}$ be the set of edges in the clique $\mathcal{C}$. Then,*

$$\mathrm{SI}(X_\mathcal{M}) = \min\left\{\mathrm{SI}(X_\mathcal{C}), \min_{(i,j) \in \mathcal{E} \setminus \mathcal{E}_\mathcal{C}} \mathrm{I}(X_i \wedge X_j)\right\}.$$

Remark 4. *Consider a (new) graph $\mathcal{G}' = (\mathcal{M}, \mathcal{E}')$ derived from the cliqueylon $\mathcal{G} = (\mathcal{M}, \mathcal{E})$ by removing edges from $\mathcal{C}$ so as to turn $\mathcal{C}$ into a (connected) tree. If $P_{X_\mathcal{M}}$ satisfies the Markov property with respect to $\mathcal{G}' = (\mathcal{M}, \mathcal{E}')$, then $\mathcal{G}'$ is termed a* Markov Chain on a Tree *[11], [10], [1]. For this particularization, by Theorem 3,*

$$\mathrm{SI}(X_\mathcal{M}) = \min_{(i,j) \in \mathcal{E}'} \mathrm{I}(X_i \wedge X_j)$$

thereby recovering [1, Theorem 3] as a special case.

Remark 5. *We note that by [1] (and also the previous remark), $\min_{(i,j) \in \mathcal{E} \setminus \mathcal{E}_\mathcal{C}} \mathrm{I}(X_i \wedge X_j)$ is the same as $\min_{i \in \mathcal{C}} \mathrm{SI}(X_{T_i})$.*

III. Proof of Lemma 1

Proof. Let A be a disconnected atom of some partition $\pi \in \Pi(\mathcal{M})$, and let $A_1, \ldots, A_t$ be subatoms. There must exist at least one A_l, $1 \leq l \leq t$, such that $A_l \cap \mathcal{C} = \phi$. Otherwise, every $A_l \cap \mathcal{C} \neq \phi$, and since $\mathcal{C}$ is a (fully connected) clique and each A_l is connected, A must also be connected, a contradiction.

Next, let A_l be a subatom with $A_l \cap \mathcal{C} = \phi$. Being connected and not intersecting $\mathcal{C}$, it must lie completely in one of the trees rooted at the vertices in $\mathcal{C}$. Let this tree be T_l. Note that any subatom that lies inside a tree must be a connected subtree of that tree, owing to connectivity.

Since A_l is a connected subset fully contained in T_l, it is easy to see that cutting any boundary edge of A_l separates $\mathcal{G}$ into two connected components, one of which is a subtree of T_l.

Fix $l \in \mathcal{C}$ as the root in T_l, creating a directed tree T_l^*. Because T_l is a tree, exactly one of the boundary edges of A_l is an incoming edge, and in general it may have any number of outgoing edges.

Case 1: If A_l has no outgoing edges, or if none of the subtrees obtained by cutting an outgoing edge of A_l contains any other subatom of A, then clearly cutting the edge between the root of the subatom and its parent in T_l creates two components, one of which contains A_l and the other contains every other subatom.

Case 2: If one of the subtrees obtained by cutting an outgoing edge of A_l does contain a subatom of A, then recursively pick such a subatom and repeat the argument; since we pick an outgoing edge every time, the depth (in T_l^*) of the root of the subatom under consideration increases with every step, and since each T_i is finite, this process has to terminate after a finite number of steps, yielding a subatom $A_{l'}$, say. Cutting the edge connecting the root of $A_{l'}$ to its parent separates the graph into two connected components T (a subtree of T_l) and $\mathcal{M} \setminus T$. Since the process terminated, none of the subtrees obtained by cutting the outgoing edges from $A_{l'}$ can contain a subatom of A, and therefore the only subatom of A that is

contained in T is $A_{l'}$; the rest of the subatoms of A must be in $\mathcal{M} \setminus T$. □

IV. Proof of Theorem 2: Connected atoms

This is a proof of the first claim in (2). Considering first the case $k = 2$, let $\pi = (\pi_1, \pi_2)$ be a 2-partition with at least one disconnected atom. Suppose that π_1 is disconnected; let A_1 be the subatom of π_1 with the property in Lemma 1, and let A_1 be a subtree of T_1. Let the vertex $1 \in \mathcal{C}$ be the root of T_1, which turns T_1 into a directed tree T_1^*. We have two cases depending on whether A_1 has outgoing boundary edges.

Case 1: If A_1 has no outgoing boundary edges, let j be the parent (in T_1^*) of the root of A_1. Owing to maximal connectivity of A_1, $j \in \pi_2$, and j separates A_1 from $\mathcal{M} \setminus A_1$. Using the Markov property, $\mathrm{I}(X_{\pi_1} \wedge X_{\pi_2}) \geq \mathrm{I}(X_{A_1} \wedge X_j) = \mathrm{I}(X_{A_1} \wedge X_{\mathcal{M}\setminus A_1})$, and therefore the partition $\{A_1, \mathcal{M} \setminus A_1\} \in \Pi_a(\mathcal{M})$ *and* is at least as good as π, i.e., with a lower $\mathcal{I}$-value.

Case 2: If A_1 has outgoing boundary edges, pick one. The subtree T obtained by cutting that edge must be a subset of π_2, and A_1 separates that subtree from $\mathcal{M} \setminus (T \cup A_1)$. Again, by the Markov property, $\mathrm{I}(X_{\pi_1} \wedge X_{\pi_2}) \geq \mathrm{I}(X_T \wedge X_{A_1}) = \mathrm{I}(X_T \wedge X_{\mathcal{M}\setminus T})$, and therefore the partition $\{T, \mathcal{M} \setminus T\} \in \Pi_a(\mathcal{M})$ *and* is at least as good as π.

Now, let $k \geq 3$ and suppose that $\pi = (\pi_1, \ldots, \pi_k)$ is a partition with π_1 being disconnected. Let A_1 be the subatom of π_1 with the property in Lemma 1; let A_1 be a subtree of T_1. As above, let the vertex $1 \in \mathcal{C}$ be the root of T_1, which turns T_1 into a directed tree T_1^*, and let j be the parent of the root of A_1. Because A is maximally connected, j cannot be in π_1; let $j \in \pi_u$ for some $2 \leq u \leq k$. Back in the *undirected* tree, this π_u separates A from every other subatom of π_1; which is guaranteed by Lemma 1. Using the Markov property,

$$A \multimap \pi_u \multimap \pi_1 \setminus A,$$

we get

$$\mathrm{I}(X_A \wedge X_{\pi_1\setminus A}) \leq \mathrm{I}(X_{\pi_u} \wedge X_{\pi_1\setminus A}) \leq \mathrm{I}(X_{\pi_u} \wedge X_{\pi_1}). \quad (3)$$

Next, consider the $(k-1)$-partition π' and the $(k+1)$-partition π'' of $\mathcal{M}$, defined by

$$\pi' = \left(\pi_1 \cup \pi_u, \{\pi_v\}_{v\neq 1, v\neq u}\right), \quad (4)$$

$$\pi'' = \left(\pi_1 \setminus A, A, \pi_u, \{\pi_v\}_{v\neq 1, v\neq u}\right). \quad (5)$$

Then

$$\mathcal{I}(\pi) \geq \min\{\mathcal{I}(\pi'), \mathcal{I}(\pi'')\} \quad (6)$$

which can be seen in a manner similar to [1, Theorem 3].

Referring to (4) and (5), we infer from (6) that for a given k-partition π with a disconnected atom π_1 as above, merging a disconnected atom with another atom (as in (4)) or breaking it to create a connected atom (as in (5)), lead to partitions π' or π'', of which at least one has a lower $\mathcal{I}$-value than π. This argument is repeated until a final partition with connected atoms is reached which has the following form. Consider the set of all maximally connected components of the atoms of $\pi = (\pi_1, \ldots, \pi_k)$; a connected π_i already constitutes such a component. The final partition will consist of *connected* unions of such components. □

V. Proof of Theorem 2: Connected complements

This proof addresses the second claim in (2). If a partition $\pi = (\pi_1, \ldots, \pi_k)$ has an atom π_i such that π_i^c is disconnected, then removing π_i must break $\mathcal{M}$ into multiple disconnected components. By the first assertion in (2), it is sufficient to minimize over partitions in $\Pi_a(\mathcal{M})$. For a partition $\pi = (\pi_1, \ldots, \pi_k) \in \Pi_a(\mathcal{M})$ with a (connected) atom π_i such that π_i^c is disconnected, every other atom π_j, $j \neq i$ must be entirely contained in one of the disconnected components obtained by removing π_i from $\mathcal{M}$. In particular, if one such component is the union of atoms $\pi_1, \ldots, \pi_m$, $m < k-1$, then π_i separates the union of all other atoms from this component.

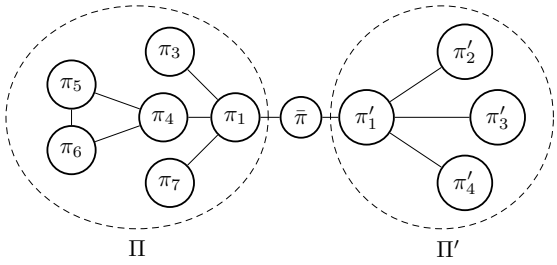

Figure 3. Illustrative graph for the second claim of Theorem 2

With $\bar{\pi}$ in the role of π_i above, let

$$\pi = \{\pi_1, \ldots, \pi_l, \bar{\pi}, \pi_1', \ldots, \pi_{l'}'\}$$

be a partition of $\mathcal{M}$ such that $\bar{\pi}$ separates $\Pi = \cup_{j=1}^{l} \pi_j$ and $\Pi' = \cup_{j=1}^{l'} \pi_j'$ in $\mathcal{M}$. Let two other partitions be given by $\pi' = \{\bar{\pi} \cup \Pi, \pi_1', \ldots, \pi_{l'}'\}$ and $\pi'' = \{\pi_1, \ldots, \pi_l, \bar{\pi} \cup \Pi'\}$. We shall show that either $\mathcal{I}(\pi') \leq \mathcal{I}(\pi)$ or $\mathcal{I}(\pi'') \leq \mathcal{I}(\pi)$.

Assume in contradiction that $\mathcal{I}(\pi') > \mathcal{I}(\pi)$ and $\mathcal{I}(\pi'') > \mathcal{I}(\pi)$. Straightforward manipulation implies that

$$\begin{aligned} &l'(l+l')(\mathcal{I}(\pi) - \mathcal{I}(\pi')) \\ &= l' \sum_{i=1}^{l} \mathrm{H}(X_{\pi_i}) - l \sum_{j=1}^{l'} \mathrm{H}(X_{\pi_j'}) + l' \,\mathrm{H}(X_{\bar{\pi}}) \\ &\qquad - (l+l')\,\mathrm{H}(X_{\bar{\pi}}, X_\Pi) + l\,\mathrm{H}(X_{\mathcal{M}}) < 0 \quad (7) \end{aligned}$$

since $\mathcal{I}(\pi') > \mathcal{I}(\pi)$. Similarly, we also get

$$\begin{aligned} &l(l+l')(\mathcal{I}(\pi) - \mathcal{I}(\pi'')) \\ &= l \sum_{j=1}^{l'} \mathrm{H}(X_{\pi_j'}) - l' \sum_{i=1}^{l} \mathrm{H}(X_{\pi_i}) + l\,\mathrm{H}(X_{\bar{\pi}}) \\ &\qquad - (l+l')\,\mathrm{H}(X_{\bar{\pi}}, X_{\Pi'}) + l'\,\mathrm{H}(X_{\mathcal{M}}) < 0. \quad (8) \end{aligned}$$

because $\mathcal{I}(\pi'') > \mathcal{I}(\pi)$. Adding (7) and (8) and using $\mathcal{M} = \Pi \cup \bar{\pi} \cup \Pi'$, we get that

$$\begin{aligned} &\mathrm{H}(X_{\bar{\pi}}) - \mathrm{H}(X_{\bar{\pi}}, X_\Pi) - \mathrm{H}(X_{\bar{\pi}}, X_{\Pi'}) \\ &\qquad + \mathrm{H}(X_\Pi, X_{\bar{\pi}}, X_{\Pi'}) < 0. \quad (9) \end{aligned}$$

Since $X_\Pi - \circ - X_{\bar{\pi}} - \circ - X_{\Pi'}$, we get $\mathrm{H}(X_\Pi, X_{\bar{\pi}}, X_{\Pi'}) = \mathrm{H}(X_{\bar{\pi}}, X_{\Pi'}) + \mathrm{H}(X_\Pi \,|\, X_{\bar{\pi}})$. As $\mathrm{H}(X_{\bar{\pi}}, X_\Pi) = \mathrm{H}(X_{\bar{\pi}}) + \mathrm{H}(X_\Pi \,|\, X_{\bar{\pi}})$, we see that the left-side of (9) equals zero, which is a contradiction. Therefore, at least one of π' and π'' must be as good as π.

Repeating this process eventually leads to a partition such that the subgraph induced by the complement of each atom of the partition is connected. □

VI. Proof of Theorem 3

By Theorem 2, it suffices to minimize in (1) over partitions in $\Pi_{ac}(\mathcal{M})$ made up of atoms that are each connected and have connected complements. We first claim that $\Pi_{ac}(\mathcal{M})$ consists of (only) two types of partitions: (a) 2-partitions obtained upon cutting an edge that is in some T_i but not in $\mathcal{C}$; and (b) partitions whose atoms are composed of unions of T_is.

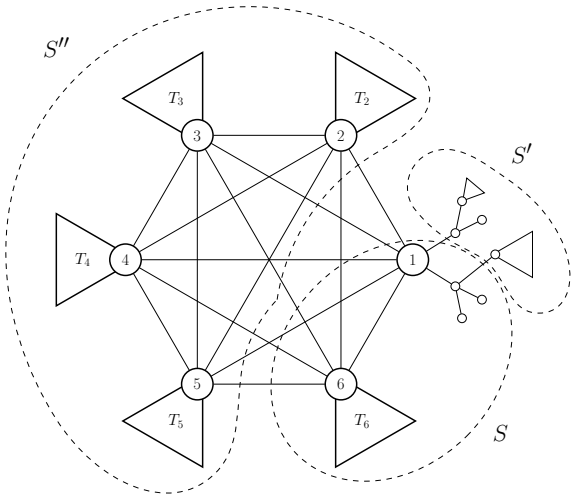

Figure 4. Illustrative graph for the proof of Theorem 3

Consider $S \subsetneq \mathcal{M}$ that contains vertex $1 \in \mathcal{C}$ (also, $1 \in T_1$), say, but does not include all the vertices in T_1. Let $S' \neq \phi$ be the set of vertices in T_1 that are not in S, and let $\mathcal{S}'' = \mathcal{M} \backslash (S \cup S') \neq \phi$. By construction, S separates S' and S'' in $\mathcal{G}$, in which case S cannot be an atom of a partition in $\Pi_{ac}(\mathcal{M})$. Therefore, for S to be an atom of a partition in $\Pi_{ac}(\mathcal{M})$, at least one of S' and S'' must be empty. In case S' is empty, S includes the entirety of T_1. If S'' is empty, then $\{S, S'\}$ partition $\mathcal{M}$. Further, S' must be obtained by cutting a single edge in T_1 and must also be an atom of the partition. This can be seen in two steps: if S' were not obtained by cutting a single edge (i, j) with $i \in S$ and $j \in S'$, S would separate the atoms constituting S', and if S were not an atom of the partition, the atom of the partition containing j (which is necessarily a subset of S' owing to connectivity) would separate the other atoms constituting S' from S. Since vertex $1 \in \mathcal{C}$ was chosen arbitrarily, this argument shows that $S \subsetneq \mathcal{M}$ which includes vertex $i \in \mathcal{C}$ can constitute an atom of a partition in $\Pi_{ac}(\mathcal{M})$ iff $\{S, T_i \setminus S\}$ is a partition of $\mathcal{M}$, or S is a union of T_i and possibly other T_js. This establishes the claim above.

Next, suppose that $\mathrm{SI}(X_\mathcal{C})$ is achieved by a partition $\pi_\mathcal{C} = (\pi_1, \ldots, \pi_l)$ of $\mathcal{C}$. For each atom $\pi_i \in \pi_\mathcal{C}$, let $T_{\pi_i} = \cup_{u \in \pi_i} T_u$ be the collection of the vertices of all trees rooted in π_i, $1 \leq i \leq l$. Then, $\pi_\mathcal{M} = \{T_{\pi_i}, 1 \leq i \leq l\}$ is in $\Pi_{ac}(\mathcal{M})$. We have

$$\mathrm{SI}(X_\mathcal{C}) = \frac{1}{l-1}\left[\sum_{i=1}^{l} \mathrm{H}(X_{\pi_i}) - \mathrm{H}(X_\mathcal{C})\right],$$

and, using the Markov property,

$$\begin{aligned}\mathcal{I}_{\pi_\mathcal{M}}(X_\mathcal{M}) &= \frac{1}{l-1}\left[\sum_{i=1}^{l} \mathrm{H}(X_{T_{\pi_i}}) - \mathrm{H}(X_\mathcal{M})\right] \\ &= \mathrm{SI}(X_\mathcal{C}). \end{aligned} \tag{10}$$

We note that (10) implies $\mathrm{SI}(X_\mathcal{M}) \leq \mathrm{SI}(X_\mathcal{C})$. If $\mathrm{SI}(X_\mathcal{M}) = \mathrm{SI}(X_\mathcal{C})$, we get that $\mathrm{SI}(X_\mathcal{M})$, too, is achieved by the partition $\pi_\mathcal{M}$. On the other hand, if $\mathrm{SI}(X_\mathcal{C}) > \mathrm{SI}(X_\mathcal{M})$, $\mathcal{C}$ must be an atom or a subset of an atom, of the $\mathrm{SI}(X_\mathcal{M})$-achieving fundamental partition (by the fact following Definition 1). The only partitions in $\Pi_{ac}(\mathcal{M})$ that are coarser than a fundamental partition with $\mathcal{C}$ as a feasible subset of an atom are the 2-partitions of type (a) above; and the optimal such partition is obtained by cutting an edge in some $T_i \setminus \mathcal{C}$ with the smallest mutual information across it. □

VII. Closing remarks

Consider $n \geq 2$ independent and identically distributed repetitions $X_\mathcal{M}^n \triangleq (X_1^n, \ldots, X_m^n)$ of $X_\mathcal{M} = (X_1, \ldots, X_m)$, with party i observing the component X_i^n, $i \in \mathcal{M}$. The problem of omniscience [10] entails each party $i \in \mathcal{M}$ reconstructing all of $X_\mathcal{M}^n$ from X_i^n and interactive communication among the parties over a public and noiseless broadcast channel. By [10], [2], the minimal achievable asymptotic (in n) rate of communication is $\mathrm{H}(X_\mathcal{M}) - \mathrm{SI}(X_\mathcal{M})$. A two-stage method for attaining omniscience, introduced in [6] as successive omniscience, proceeds by the atoms of a partition of $\mathcal{M}$, termed local groups, first attaining local omniscience, followed by all the parties in $\mathcal{M}$ acquiring global omniscience. Successive omniscience, in general, is suboptimal in that the minimum aggregate rate of communication in the two stages can exceed $\mathrm{H}(X_\mathcal{M}) - \mathrm{SI}(X_\mathcal{M})$ [6]. However, no penalty in aggregate rate is suffered if the local groups correspond to the atoms of an $\mathrm{SI}(X_\mathcal{M})$-attaining partition of $\mathcal{M}$. Theorem 3 above implies that upon choosing the local groups as connected subsets that are localized in the cliqueylon, it is always possible to attain omniscience optimally by means of successive omniscience.

Finally, Theorem 3 provides recourse to a relatively efficient estimation of $\mathrm{SI}(X_\mathcal{M})$ for a cliqueylon when $P_{X_\mathcal{M}}$ is unknown. A naive procedure would consist of estimating $P_{X_\mathcal{M}}$ from $X_\mathcal{M}^n$ (requiring sample size n to grow exponentially in $|\mathcal{M}|$), followed by computation of $\mathrm{SI}(X_\mathcal{M})$ from $P_{X_\mathcal{M}}$ (for which an efficient algorithm is known [5]). For a cliqueylon, Theorem 3 leads to significant gains in both steps, particularly when $|\mathcal{C}| \ll |\mathcal{M}|$. Specifically, it suffices to estimate only $P_{X_\mathcal{C}}$ and the minimum mutual information among all edges that are not contained in $\mathcal{C}$.

REFERENCES

[1] S. Bhattacharya and P. Narayan, "Shared information for a Markov chain on a tree," in *2022 IEEE International Symposium on Information Theory*, 2022.

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

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

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

[5] 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.

[6] 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.

[7] 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*, 2016.

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

[9] T. A. Courtade and T. R. Halford, "Coded cooperative data exchange for a secret key," *IEEE Transactions on Information Theory*, vol. 62, 2016.

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

[11] H.-O. Georgii, *Gibbs Measures and Phase Transitions*. De Gruyter, 2011.

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

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

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

[15] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy, and Steiner tree packing," *IEEE Transactions on Information Theory*, vol. 56, 2010.

[16] 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.

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

[18] H. Tyagi and S. Watanabe, "Converses for secret key agreement and secure computing," *IEEE Transactions on Information Theory*, vol. 61, 2015.