Mutual Dependence for Secret Key Agreement
Chung Chan and Lizhong Zheng

Abstract—A mutual dependence expression is established for the secret key agreement problem when all users are active. In certain source networks, the expression can be interpreted as certain notions of connectivity and network information flow. In particular, the secrecy problem can be mapped to a new class of network coding problems with selectable links and undirected broadcast links. For such networks, the secrecy capacities serve as upper bounds on the maximum network throughputs, while the network coding solutions can be used for secret key agreement.

Index Terms—Secret key agreement, mutual dependence, network coding, partition connectivity, supermodularity

I. MUTUAL DEPENDENCE

We first consider a measure of correlation among a set of random variables, and establishes its operational meaning with the problems of secret key agreement and communication for omniscience in [2].

In information theory, the dependence between any two random variables is captured by the mutual information

\[ I(Z_1, Z_2) := D(P_{Z_1Z_2} \| P_{Z_1}P_{Z_2}) = H(Z_1) - H(Z_1 | Z_2) \]  

(1.1)

where \( P_{Z_1Z_2} \) denotes the distribution of \( Z_1 \) and \( Z_2 \), \( D(\cdot \| \cdot) \) is the information divergence, and \( H(\cdot) \) is the entropy measure.[3] It has various operational meanings spanning over the channel and source coding theories. A heuristically appealing extension[2] to the multivariate case with more than two random variables is the following mutual dependence expression.

Definition 1.1 (Mutual Dependence) For any finitely-valued random vector \( Z_V := (Z_i : i \in V) \) with \( |V| \geq 2 \), the mutual dependence of \( Z_V \) is defined as,

\[ I(Z_V) := \min_{\mathcal{P} \in \Pi} \frac{1}{|\mathcal{P}|-1} D \left( P_{Z_V} \middle\| \prod_{C \in \mathcal{P}} P_{Z_C} \right) \]  

(1.2)

where \( \Pi \) is the collection of set-partitions \( \mathcal{P} \) of \( V \) into at least 2 non-empty sets.

Example 1.1 Mutual dependence (1.2) reduces to the usual mutual information when \( |V| = 2 \), i.e., \( I(Z_{\{1,2\}}) = I(Z_1 \cap Z_2) \).

With \( V := \{1,2,3\} \), we have \( I(Z_{\{1,2\}}) = \min \{ I(Z_1 \cap Z_2), I(Z_2 \cap Z_3), I(Z_3 \cap Z_1), I(Z_1 \cap Z_3), I(Z_3 \cap Z_1Z_2), \} \) and \( \frac{1}{2} |\sum_{i \in \{3\}} H(Z_i) - H(Z_{\{3\}})| \).

(1.1) and (1.2) are similar in the sense that both can be expressed in terms of the information divergence from the joint distribution to the product distribution of certain marginals. We will establish the following operational meaning for (1.2) in the problem of secret key agreement (SK) and communication for omniscience (CO) considered in [2].

Operational meaning of mutual dependence:
The mutual dependence expression \( I(Z_V) \) equals

- **SK**: the maximum rate (secrecy capacity) of secret key that can be generated from the discrete multiple memoryless source (DMMS) \( Z_V \) with unlimited authenticated public discussion, and

- **CO**: the maximum savings in rate, below \( H(Z_V) \), of the public discussion required to attain omniscience of the DMMS \( Z_V \), i.e. \( H(Z_V) \) subtracted by the smallest rate of communication for omniscience.

We have considered here the special case when all terminals are active in the sense that they all want to share the secret in the SK problem, and attain omniscience in the CO problem.

Theorem 1.1 Given a finite ground set \( V : |V| \geq 2 \), the mutual dependence in (1.2) satisfies,

\[ I(Z_V) = H(Z_V) - \max_{\lambda \in \Lambda} \sum_{B \in \mathcal{F}} \lambda_B H(Z_B | Z_{B^c}) \]  

(1.3)

where \( \mathcal{F} := 2^V \setminus \{V\}, \Lambda \) is defined as the collection of fractional partitions \( \lambda := (\lambda_B : B \in \mathcal{F}) \) of \( V \), i.e. \( \lambda_B \geq 0 \) for all \( B \in \mathcal{F} \) and \( \sum_{B \in \mathcal{F}} \lambda_B = 1 \) for all \( i \in V \).

This establishes the result since the expression on the R.H.S. of (1.3) bears the desired operation meaning by [2].

Proof (Theorem 1.1) Define \( h : \mathcal{F} \rightarrow \mathbb{R} \) as

\[ h(B) := H(Z_B | Z_{B^c}) \quad \forall B \in \mathcal{F} \]  

(1.4)

with the convention \( h(\emptyset) = 0 \). We will use the following well-known[4] supermodularity property of \( h \) to prove the theorem.

Subclaim 1.1A \( h \) is supermodular, i.e.

\[ h(B_1) + h(B_2) \leq h(B_1 \cap B_2) + h(B_1 \cup B_2) \]  

(1.5)

for all \( B_1, B_2 \in \mathcal{F} : B_1 \cap B_2, B_1 \cup B_2 \in \mathcal{F} \).

Proof (Subclaim 1.1A) Consider proving the non-trivial case where \( B_1 \) and \( B_2 \) are non-empty. We have the positivity of mutual information that,

\[ I(Z_{B_1} \cap Z_{B_2} | Z_{B_1 \cap B_2}) \geq 0 \implies H(Z_{B_1}) + H(Z_{B_2}) \geq H(Z_{B_1 \cup B_2}) + H(Z_{B_1 \cap B_2}) \]

1 More generally, there is a duality that \( C_{CH}(R) = C_S(R) + R \) where \( C_{CH}(R) \) and \( C_S(R) \) are the common randomness and secrecy capacities respectively under the same rate constraint \( R \) on public discussion and the same source model. (See [1] for details)

This work is suppprted in part by Project #MMT-p2-09 of the Shun Hing Institute of Advanced Engineering, The Chinese University of Hong Kong.

Manuscript written for CISS 2010. Related work at [1].

Chung Chan (chungc@mit.edu) is with Shun Hing Institute of Advanced Engineering, The Chinese University of Hong Kong, and Research Laboratory of Electronics at MIT, Massachusetts Institute of Technology. Lizhong Zheng is with Research Laboratory of Electronics at MIT, Massachusetts Institute of Technology.
(1.5) follows since \( h(B) = H(Z_V) - H(Z_{B^c}) \).

By the Strong Duality Theorem[5], the maximization in (1.3) is equal to its linear programming dual,

\[
\begin{align}
\text{minimize} & \quad \sum_{i \in V} r_i \\
\text{subject to} & \quad \sum_{i \in B} r_i \geq h(B) \quad \forall B \in \mathcal{F} \quad (1.6a)
\end{align}
\]

The supermodularity property of \( h \) translates to the following property on the tight relations of the dual problem.

**Subclaim 1.1B** For any feasible solution \( r \) to the dual linear program (1.6), and \( B_1, B_2 \in \mathcal{F} : B_1 \cap B_2, B_1 \cup B_2 \notin \mathcal{F} \), if \( B_1 \) and \( B_2 \) are tight constraints, i.e.

\[
\sum_{i \in B_j} r_i = h(B_j) \quad \text{for } j = 1, 2 \quad (1.7a)
\]

then \( B_1 \cup B_2 \) is also a tight constraint,

\[
\sum_{i \in B_1 \cup B_2} r_i = h(B_1 \cup B_2) \quad (1.7b)
\]

n.b. \( B_1 \cap B_2 \) is also tight but we do not need it for the proof of Theorem 1.1.

**Proof (Subclaim 1.1B)** Since \( B_1 \cup B_2 \in \mathcal{F} \), we immediately have \( \sum_{i \in B_1 \cup B_2} r_i \geq h(B_1 \cup B_2) \) by (1.6b). The reverse inequality can be proved as follows,

\[
\sum_{i \in B_1 \cup B_2} r_i = \sum_{i \in B_1} r_i + \sum_{i \in B_2} r_i - \sum_{i \in B_1 \cap B_2} r_i \\
\leq h(B_1) + h(B_2) - h(B_1 \cap B_2) \quad (a) \\
\leq h(B_1 \cup B_2) \quad (b)
\]

where (a) is by (1.7a) and (1.6b) on \( B_1 \cap B_2 \notin \mathcal{F} \), and (b) is by Subclaim 1.1A. With a similar argument, we also have \( \sum_{i \in B_1 \cap B_2} r_i = h(B_1 \cap B_2) \).

Let \( \lambda^* \) be an optimal solution to the maximization in (1.3). Define its support set as,

\[
\mathcal{B} := \{ B \in \mathcal{F} : \lambda_B^* > 0 \} \quad (1.8)
\]

and the corresponding partition of \( V \) as,

\[
\mathcal{P}^* := \left\{ \left( \bigcup \{ B \in \mathcal{B} : B \ni i \} \right)^c : i \in V \right\} \quad (1.9)
\]

**Subclaim 1.1C** \( \mathcal{P}^* \) in (1.9) belongs to \( \Pi \) in Definition 1.1.

**Proof (Subclaim 1.1C)** Define the relation \( \sim_R \) on \( V \) as,

\[
i \sim_R j \iff i \in C_j \quad \text{for } i, j \in V
\]

where \( C_i := (\bigcup \{ B \in \mathcal{B} : B \ni i \})^c \). By definition (1.9), \( \mathcal{P}^* = \{ C_i : i \in V \} \). To show that \( \mathcal{P}^* \) is a partition of \( V \), it suffices to show that \( \sim_R \) is an equivalence relation on \( V \) as follows.

\[
i \sim_R j \iff \{ B \in \mathcal{B} : B \ni i \} \supseteq \{ B \in \mathcal{B} : B \ni j \} \\
\iff \{ B \in \mathcal{B} : i \in B \} \subseteq \{ B \in \mathcal{B} : j \in B \}
\]

\( \lambda^* \) exists or equivalently \( \Lambda \) is non-empty. For example, \( \lambda_{\{i\}} = 1 \) for \( i \in V \) is a fractional partition in \( \Lambda \). For the more general case considered in Theorem 3.1, \( \Lambda \) may be empty.

i.e. any set in \( \mathcal{B} \) that contains \( i \) also contains \( j \). Using this simplification, it is easy to see that \( \sim_R \) satisfies the defining properties of an equivalence relation:

- **Reflexivity:** \( R \) is reflexive since \( i \in C_i \) trivially for \( i \in V \).
- **Transitivity:** Suppose \( i \sim_R j \) and \( j \sim_R k \) for some \( i, j, k \in V \). Then,

\[
\{ B \in \mathcal{B} : i \in B \} \subseteq \{ B \in \mathcal{B} : j \in B \} \subseteq \{ B \in \mathcal{B} : k \in B \}
\]

which implies \( i \sim_R k \) as desired.

- **Symmetry:** suppose to the contrary that \( i \sim_R j \) but \( j \not\sim_R i \) for some \( i, j \in V \). Then,

\[
\{ B \in \mathcal{B} : i \in B \} \subseteq \{ B \in \mathcal{B} : j \in B \}
\]

This implies, by definition (1.8) of \( \mathcal{B} \) that

\[
\sum_{B \ni i} \lambda_B^* < \sum_{B \ni j} \lambda_B^*
\]

which is the desired contradiction since both sides equal 1 by the definition of \( \Lambda \) in Theorem (1.1).

Finally, to argue that \( |\mathcal{P}^*| \geq 2 \), note that \( \mathcal{B} \neq \emptyset \) since \( \sum_{B \notin \mathcal{F}} \lambda_B^* > 0 \). Since any \( B \in \mathcal{F} \) satisfies \( B \neq \emptyset \), we have \( C_i \neq V \) for all \( i \in V \) as desired.

The supermodularity of \( h \) implies the following property on every part of \( \mathcal{P}^* \).

**Subclaim 1.1D** For any optimal \( r^* \) to the dual problem (1.6),

\[
\sum_{i \in C} r_i^* = h(C^c) \quad \forall C \in \mathcal{P}^* \quad (1.10)
\]

**Proof (Subclaim 1.1D)** By the Complementary Slackness Theorem[5, Theorem 5.4], \( \sum_{i \in B} r_i^* = h(B) \) for all \( B \in \mathcal{B} \).

By Subclaim 1.1B, for all \( i \in V \), we have

\[
\sum_{i \in \bigcup \{ B \in \mathcal{B} : B \ni i \}} r_i^* = h\left( \bigcup \{ B \in \mathcal{B} : B \ni i \} \right)
\]

which gives the desired equality (1.10) under (1.9).

This completes the proof since the Primal/Dual Optimality Criteria[5, Theorem 5.5] implies \( \{ B^c \in \mathcal{P}^*: |\mathcal{P}^*| - 1 : B \in \mathcal{F} \} \) is an optimal solution in \( \Lambda \). More precisely, for all feasible \( r \) to the dual (1.6), and \( \mathcal{P} \in \Pi \),

\[
H(Z_V) - \max_{\lambda \in \Lambda} \sum_{B \in \mathcal{F}} \lambda_B H(Z_B | Z_{B^c})
\]

\[
\leq H(Z_V) - \sum_{i \in V} r_i = H(Z_V) - \frac{1}{|\mathcal{P}| - 1} \sum_{C \in \mathcal{P}} \sum_{i \in C} r_i
\]

\[
\leq H(Z_V) - \frac{1}{|\mathcal{P}| - 1} \left( \sum_{C \in \mathcal{P}} H(Z_C) - H(Z_V) \right)
\]

When we set \( r \) to an optimal solution \( r^* \), (a) is satisfied with equality by the Strong Duality Theorem. When we also set \( \mathcal{P} \) to \( \mathcal{P}^* \), which is valid by Subclaim 1.1C, (b) is also satisfied with equality by Subclaim 1.1D. This gives the desired equality (1.3) and completes the proof of Theorem 1.1.
II. INTERPRETATION VIA EMULATED SOURCE NETWORK

In this section, we will show that under certain classes of source networks, we can interpret mutual dependence (1.2) as certain notions of connectivity, and more concretely as the amount of information flow in certain types of networks. To make the notion of “flow” explicit, we consider the following class of source networks.

**Definition 2.1 (Emulated source network)** Terminal \( i \in V \) observes \( X_i = (X_i, Y_i) \) such that,

\[
P_{X_i|Y_i} = \left( \prod_{i \in V} P_{X_i|Y_i|X_i} \right) \tag{2.1}
\]

This can be viewed as a source emulated by having terminal \( i \in V \) send \( X_i \) independently over a channel that returns \( Y_i = f_i(X_i, N_i) \) to terminal \( i \), where \( f_i \) is deterministic and \( N_i \)'s are independent channel noises that satisfy \( P_{N_i|X_i} = \prod_{i \in V} P_{N_i} \).

**Proposition 2.1** The mutual dependence (1.2) of the emulated source network in Definition 2.1 is,

\[
I(Z_V) = \min_{\mathcal{P} \subseteq \Pi} \frac{1}{|\mathcal{P}| - 1} \sum_{C \in \mathcal{P}} I(X_C \wedge Y_C|X_C) \tag{2.2}
\]

which is also the secrecy capacity when all terminals are active by Theorem 1.1. □

**Proof** For all \( \mathcal{P} \subseteq \Pi \),

\[
\sum_{C \in \mathcal{P}} H(X_C|Y_C) - H(X_V|Y_V)
\]

\[
= \sum_{C \in \mathcal{P}} [H(X_C) + H(Y_C|X_C)] - H(X_V) - H(Y_V|X_V)
\]

\[
= \sum_{C \in \mathcal{P}} [H(Y_C|X_C) - H(Y_V|X_V)]
\]

where the last equality is by (2.1) that \( \sum_{C \in \mathcal{P}} H(X_C) = H(X_V) \) and \( \sum_{C \in \mathcal{P}} H(Y_C|X_C) = H(Y_V|X_V) \).

\[I(X_C \wedge Y_C|X_C)\] is intuitively the flow of information from \( C \) to \( C \). For a more concrete interpretation, we will consider some specific classes of networks for which the dependency among the observations can be abstracted by a hypergraph.

**Hypergraph:** A hypergraph \( H := (V, E, \phi) \) is defined by the mapping \( \phi : E \rightarrow 2^V \setminus \{\emptyset\} \) from an edge \( e \) to a non-empty subset \( \phi(e) \) of the vertices. The star hypergraph\( [6] \)

\( H^* := (V, E, \phi, \rho) \) of \( H \) has an additional mapping \( \rho : E \rightarrow V \) from an edge \( e \) to a root node \( \rho(e) \in \phi(e) \) for the edge.

A. Interference network

**Definition 2.2 (Interference network)** Given a hypergraph \( H := (V, E, \phi) \) and a finite additive group \( \mathbb{G} \) of order \( q \), terminal \( i \in V \) observes,

\[
Z_i := \{Z_i^e : e \in E, i \in \phi(e)\} \tag{2.3}
\]

where \( Z_i^e \) for \( i \in V \) and \( e \in E \) are random variables taking values from \( \mathbb{G} \) such that,

1) \( Z_i^e := (Z_i^e : i \in \phi(e)) \) for \( e \in E \) are independent, and
2) for all \( e \in E \)

\[
\sum_{i \in \phi(e)} Z_i^e = 0 \tag{2.4}
\]

and for all \( j \in \phi(e) \), \( Z_i^e|_{\phi(e) \setminus \{j\}} \) is uniformly distributed over \( \mathbb{G}^{\phi(e) \setminus \{j\}} \). □

**Proposition 2.2** The mutual dependence of the interference network in Definition 2.2 is \( (\log q) p^+(H) \) where \( p^+(H) \) is the strength of \( H \) defined as,

\[
p^+(H) := \min_{\mathcal{P} \subseteq \Pi} \frac{\sum_{C \in \mathcal{P}} |\delta_H^+(C)|}{|\mathcal{P}|-1} \tag{2.5a}
\]

\[
= \min_{\mathcal{P} \subseteq \Pi} |\delta_H^+(\mathcal{P})| \tag{2.5b}
\]

where \( H^* := (V, E, \phi, \rho) \) is any star hypergraph of \( H \),

\[
\delta_H^+(C) := \{e \in E : \rho(e) \in C \supseteq \phi(e)\} \tag{2.6}
\]

\[
\delta_H^+(\mathcal{P}) := \{e \in E : \forall C \in \mathcal{P}, C \supseteq \phi(e)\} \tag{2.7}
\]

which are the set of outgoing edges of \( C \) and crossing edges of \( \mathcal{P} \) respectively. □

**Proof** The interference network is a special case of the emulated source network in Definition 2.1 with,

\[
Y_i := (Z_i^e : e \in E, i = \rho(e))
\]

\[
X_i := (Z_i^e : e \in E, i \in \phi(e) \setminus \{\rho(e)\})
\]

where \( \rho \) is an arbitrarily chosen orientation of \( H \). We have,

\[
I(X_C \wedge Y_C|X_C) = H(Y_C|X_C) = |\delta_H^+(C)|
\]

\[
\sum_{C \in \mathcal{P}} |\delta_H^+(C)| = \sum_{e \in E} \sum_{C \supseteq \phi(e)} I(C \supseteq \phi(e)) = |\delta_H^+(\mathcal{P})|
\]

which completes the proof with Proposition 2.1. □

The strength \( p^+(H) \) of \( H \) has the following immediate interpretation of partition connectivity.

Partition connectivity\( [7][8] \):

\( p^+(H) \) is the maximum rational number \( x \in \mathbb{Q} \) such that \( H \) is \( x \)-partition-connected, i.e. \( [x(k-1)] \) edges need to be removed from \( H \) to yield \( k \) or more disconnected components for any \( k \in |V| \).

Each edge in \( H \) corresponds to a link of secret information flow with the following linear public discussion scheme.

\[3\]This pure source emulation approach to SK when terminals are given a channel instead of a source can sometimes be optimal. For instance, uniform input is optimal for finite linear channel \( Y_i = M_i(X_i, N_i) \) for \( i \in V \) where \( M_i \) is a homomorphism between finite abelian groups and \( N_i \)'s are arbitrarily correlated noise. This covers the interference and broadcast networks to be defined later. (See [1] for details)
Perfectly recover

I secret, i.e.

M

Thus, the secret key agreement problem turns into a broadcast and receiver, the terminals can agree on a common secret

e

is used at each time as follows: for each edge

with selectable links defined by the hypergraph

H

in independent link from

j

φ

∈

G

H

extension[8][6] to hypergraphs summarized below.

Given an edge

j

i

∈

min

extended hypergraph with

P

˜

1)

(123

Z

M

φ

∈

G

H

1)

follows from [6, Lemma 3.1 and Theorem 4.2, 5.1].

2) Up to

p+(H) = np+(H) by (2.5), and

1) ˜

H

can be represented by a graph

G

:= (V, ˜E, θ) in the following sense: p+(G) = p+(H) and for all

e

θ(e) ⊆ φ(e) : 1 ≤ |θ(e)| ≤ 2

2) Up to

p+(G) edge-disjoint spanning trees can be packed in

G

i.e. there exists spanning trees

Tj := (V, ˜Ej, θ) for

j

[p+(G)] with disjoint edge sets

Ej ⊆ ˜E

3) For all optimal

P*

∈ P that attains the minimum in

(2.5b) and any excess edge

˜

e

* ∈ ˜E \ \bigcup \ [e \in p+(G)] \ ˜Ej not used in a maximal tree packing. ∴C ∈ P*, C ⊇ φ(˜e*) □

PROOF

1) follows from [6, Lemma 3.1 and Theorem 4.2, 5.1].

2) follows from the Disjoint Tree Theorem[7, Corollary 51.1a] of Tutte and Nash-Williams.

3) Suppose to the contrary that for some optimal

P*

and excess edge

˜

e

*, we have

C ̸⊆ φ(˜e*) for all

C ∈ P*.

|δ̂H(P*)| = {e ∈ ˜E : ∀C ∈ P*, C ̸⊆ φ(˜e*)}

(a) ≥

|{e ∈ ˜E : ∀C ∈ P*, C ̸⊆ φ(˜e*)} |

(b) = (|P*| − 1)p+(G)

where (a) is by excluding

e*

and (b) is because each spanning tree

Tj contributes \((|P*| − 1)\) distinct crossing edges. Substituting

p+(G) = np+(H) and

δ̂H(P*) = n[δH(P)] into the last inequality, we have the desired contradiction to the optimality of

P*.

Since every spanning tree packed in

H

supports one unit of secret information flow from any designated root terminal to all other terminals, the tree packing result implies that

p+(H)

units of secret key can be broadcast to all terminals in total, achieving the secrecy capacity.

We now have the desired interpretation of mutual dependence as the secret information flow in a broadcast session of a network with selectable links. The optimal partition

P*

to (2.5b) also has the intuitive meaning as classes of well-connected terminals: two terminals must be in the same class in any optimal

P* if there is an excess private link between them. This connects and extends the results of [6][8][9][10].

Theorem 2.1

Given hypergraph

H

with strength

p+(H) defined in (2.5), let

C_{N,sl} be the maximum throughput of a broadcast session of the delay-free network with selectable links defined in Definition 2.3, and

C_{S,if} be the secrecy capacity of the interference network defined in Definition 2.2, then

C_{N,sl} = C_{S,if} = (log q)p+(H)

Furthermore, the maximum throughput and secrecy capacity can be attained non-asymptotically with delay at most

|V| − 1.

The maximum throughput, in particular, can be achieved by routing.

□

Example 2.1

For the interference network defined in Definition 2.2, let

G

be the binary field

F2, and

H := (V, E, φ)

be the hypergraph on

V = [3] with edge

E := \{123\}

and

φ(123) = \{1, 2, 3\}.

From (2.3),

Z3 = Z1 + Z2.

Since

p+(H) = 1/2, H can be extended as described in Proposition 2.3 to

H := (V, ˜E, φ)

with

n = 2, and represented by the graph

G := (V, ˜E, θ) that can be maximally packed with

p+(G) = 1 spanning tree

T1 := (V, ˜E1, θ) where

= \{(123, 1), (123, 2)\}

θ(123, 1) := \{1, 2\} and θ(123, 2) := \{1, 3\}

Thus, one secret key bit

K

can be propagated from terminal 1 through the tree network

T1 to terminal 2 and 3. In particular, we can have

K

set to

Z1 = \{(123, 1), \{(123, 2), \{Z2, \{(123, 1), \{Z3, \{Z3, \{(123, 1) + Z2, \{(123, 2) + (Z1 + Z2)\}\}\}\}\}\}\}\]\n
revealed by terminal 1, 2 and 3 respectively. Terminal 2 and 3 can recover

K

by the linear operations

Z1 + Z2

and

Z3 + Z2

respectively. K is perfectly secret because

Z1

is independent of the public messages.

□
B. Broadcast Network

We now show another class of emulated source networks for which the mutual dependence has a different interpretation of connectivity, and can be mapped to a network coding problem with undirected broadcast links.

Definition 2.4 (Broadcast Network) Given \( H := (V, E, \phi) \) and a finite field \( \mathbb{F}_q \) of order \( q \), terminal \( i \in V \) observes,

\[
Z_i := \{ Z^e : e \in E, i \in \phi(e) \}
\]

where \( \{ Z^e : e \in E \} \) is uniformly distributed over \( \mathbb{F}_q^{\exists E} \). \( \Box \)

Proposition 2.4 The mutual dependence of the broadcast network in Definition 2.4 is \((\log q) p^-(H)\) where,

\[
p^-(H) := \min_{P \in \Pi} \frac{\sum_{C \in P} |\delta_H^*(C)|}{|P| - 1}\tag{2.10a}
\]

\[
= \min_{P \in \Pi} \frac{\sum_{e \in E} (|\pi_P(\phi(e))| - 1)}{|P| - 1}\tag{2.10b}
\]

\( H^* := (V, E, \phi, \rho) \) is any star hypergraph of \( H \),

\[
\delta_H^*(C) := \{ e \in E : \rho(e) \in C^c \not\supseteq \phi(e) \}
\]

\[
\pi_P(\phi(e)) := \{ C \cap \phi(e) : C \in P \} \setminus \{\emptyset\}
\]

are the set of in-cut of \( C \) and the partition of \( e \) respectively. \( \Box \)

Proof The broadcast network is a special emulated source network in Definition 2.1 with,

\[
Y_i := \{ Z^e : e \in E, i \in \phi(e) \} \setminus \{ \rho(e) \}
\]

\[
X_i := \{ Z^e : e \in E, i = \rho(e) \}
\]

for any arbitrary orientation \( \rho \) of \( H \). With the equalities

\[
I(\chi_{C^c}^e \wedge \chi_{C^c}^e) = I(Y_i^e | X_i^e) = |\delta_H^*(C)|
\]

\[
= \sum_{e \in E} \sum_{C \in P} 1_{\{ C^c \not\supseteq \phi(e) \}} = \sum_{e \in E} (|\pi_P(\phi(e))| - 1)
\]

the desired result follows from Proposition 2.1. \( \Box \)

\( p^-(H) \) expresses an alternative notion of connectivity: the maximum \( x \in Q \) such that any partitioning of the vertices \( V \) into \( k \) parts split the edges into a total of at least \( \lceil x(k - 1) \rceil \) additional parts for any \( k \in \lceil |V| \rceil \).

Each edge in \( H \) corresponds to a broadcast link of secret information flow as follows.

Hyperedge as undirected broadcast link:

Given an edge \( e \), select a sender \( i \in \phi(e) \) to encrypt an independent secret \( M \in \mathbb{F}_q \) into the public message \( M + Z^e \). The remaining terminals in \( \phi(e) \) can perfectly recover \( M \) knowing \( Z^e \). Since \( Z^e \) is uniformly distributed, the encryption is perfectly secret. We effectively have a private broadcast link from \( i \) to \( \phi(e) \setminus \{ i \} \) with unit capacity.

Viewing each edge as a broadcast link, the terminals can agree on a common secret key by broadcasting it through the network. In particular, at least one unit of secret information flow to all terminals is supported by an edge-connected spanning subhypergraph defined as follows.

Edge-connected spanning subhypergraphs:

An edge-connected spanning subhypergraph \( H' := (V, E', \phi) \) of a hypergraph \( (V, E, \phi) \) satisfies \( E' \subseteq E \), \( \bigcup_{e \in E'} \phi(e) = V \) and \( |\delta_H^*(C, C')| \geq 1 \) for every \( C \subseteq V \).

Definition 2.5 (Network with undirected broadcast links) A network with undirected broadcast links defined by the hypergraph \( H = (V, E, \phi) \) is used at each time as follows: for all \( e \in E \), a sender \( i \in \phi(e) \) can be selected to send a unit \((\log q)\) bits of data noiselessly to all receivers \( j \in \phi(e) \). \( \Box \)

Although there is no analogous packing result to Proposition 2.3, this network coding approach to secret key agreement is also optimal by the following min-cut characterization of \( p^-(H) \) in [6, Theorem 5.2].

Min-cut characterization of \( p^-(H) \):

For any \( s \in V \), there is a star hypergraph \( H^* \) of the extension \( H \) of \( H \) (defined in (2.8) with \( p^* \) replaced by \( p^- \)) such that \( \delta^+_{H^*}(C) \geq p^-(H) \) for any \( C \subseteq V : s \in C \).

Theorem 2.2 Given hypergraph \( H \) with \( p^-(H) \) defined in (2.5), let \( C_{S,ub} \) be the maximum throughput of a broadcast session of the delay-free network with undirected broadcast links in Defintion 2.5, and \( C_{S,bc} \) be the secrecy capacity of the broadcast network defined in Definition 2.4, then

\[
C_{S,ub} = C_{S,bc} = (\log q) p^-(H)
\]

Furthermore, the maximum throughput and secrecy capacity can be attained asymptotically with finite delay at most \( |V|^3 |E| p^-(H)^2 \log_q |V|^2 |E| p^-(H) q \). The throughput, in particular, can be achieved by the convolutional code in [11]. \( \Box \)

Proof (Sketch from [11]) This result follows from an extension of the algebraic argument in [11] using the extension[6, Theorem 4.1] of the Menger’s Theorem for star hypergraphs. The delay is the product \( nk(\mu + 1) \) of the extension \( n \) to turn \( H \) to \( \tilde{H} \), the extension \( k \) to turn the field \( \mathbb{F}_q \) to \( \mathbb{F}_{q^k} \) for existence of the desired network code, and the delay \( \mu \) required to avoid cyclic dependency in the information flow of the network. \( \Box \)

In the following, we give a simple example for which one cannot decompose any extension \( \tilde{H} \) of \( H \) into \( p^-(\tilde{H}) \) spanning edge-connected subhypergraphs. This implies that coding may be necessary to attain the maximum throughput of the network with undirected broadcast links in a broadcast session.

Example 2.2 For the broadcast network defined in Definition 2.4, let \( q = 2 \) and \( H := (V, E, \phi) \) be the hypergraph on \( V = [4] \) with edges \( E := \{ 123, 134, 124 \} \) and \( \phi(ijk) := \{ i, j, k \} \) for \( i, j, k \in V \). Then, \( p^-(H) = 2 \) but it is not possible to pack two spanning edge-connected subhypergraphs, for that requires four edges. Indeed, we can maximally pack three spanning edge-connected subhypergraphs \( H_i := (V, E_i, \phi) \) for \( i \in [3] \) after extending \( H \) with \( n = 2 \), where \( E_1 := \{ (123, 1), (134, 1) \} \), \( E_2 := \{ (124, 1), (123, 2) \} \) and \( E_3 := \{ (134, 2), (124, 2) \} \). Thus, a pure routing solution only achieves a key rate of 1.5 bits per use of the broadcast network in Definition 2.4.
Let \( H^* := (V, E, \phi, \rho) \) be the star hypergraph of \( H \) where \( \rho(e) = 1 \) for all \( e \in E \), n.b. it satisfies \( |\delta_{H^*}^+(C)| \geq 2 \) for all \( C \subseteq V : s \in C \). We can propagate two secret key bits \( K_1, K_2 \in \mathbb{F}_2 \) from \( s := 1 \) using the following linear network code: send \( K_1 \) through the broadcast link 123, \( K_2 \) through 134, and \( K_1 + K_2 \) through 124. Since every terminal has access to at least two links, they can recover the key bits perfectly.

### III. RELATED WORK

(See [1] for details.) Consider the more general secret key agreement problem in [2] where only a subset \( A \subseteq V \) of the terminals are active. As shown in [2], inequality \( \leq \) for (1.3) holds with \( \mathcal{F} \) replaced by

\[ \mathcal{F}(A) := \{ B \subseteq V : \emptyset \neq B \subseteq A \} \]

Theorem 1.1 asserts that equality holds for \( A = V \). However, equality may not hold when \( A \subset V \) as shown by the counterexample below.\(^4\) This resolves an open question in [2].

**Example 3.1** Given uniformly random bits \( X_1, X_2 \) and \( X_3 \) in \( \mathbb{F}_2 \), define the source network \( Z'_V \) for \( V := \{ 6 \} \) as follows:

\[ Z_1 := X_1 + X_2, \quad Z_2 := X_1 + X_3, \quad Z_3 := X_2 + X_3, \quad Z_4 := X_3, \quad Z_5 := X_2 \text{ and } Z_6 := X_1. \]

With \( A := [3] \), the mutual dependence in (1.2) (with \( \mathcal{F} \) replaced by \( \mathcal{F}(A) \)) and secrecy capacity on the R.H.S. of (1.3) are 1 bit and 0.75 bits respectively.\(^5\)

Indeed, Theorem 1.1 can be extended in a slightly different direction to a general identity for supermodular function optimizations using the following notion of partitions.

**Definition 3.1** Given a finite ground set \( V : |V| \geq 2 \), define \( \Phi(A) \) for \( A \subseteq V : |A| \geq 2 \) as the collection of all families \( \mathcal{F} \subseteq 2^V \setminus \{ V \} \) that satisfy for all \( B, B' \in \mathcal{F} \) that \( B \not\supseteq A \) and

\[ B \cup B' \not\supseteq A \implies B \cap B' \in \mathcal{F} \]

It follows that \( \Phi(A) \subseteq \Phi(A') \) for all \( A \subseteq A' \). In particular, \( \mathcal{F}(A') \setminus \Phi(A) \) and \( \mathcal{F} := 2^V \setminus \{ V \} = \mathcal{F}(V) \setminus \Phi(V) \).

Denote \( \mathcal{F} := \{ B \in \mathcal{F} : (B \in \mathcal{F}) \} \). Define \( \Pi(\mathcal{F}, U) \) for \( \mathcal{F} \subseteq \Phi(V) \) and \( U \subseteq V \) as the collection of all families \( \mathcal{P} \) such that \( (C \cup U : C \in \mathcal{P}) \) is a set-partition of \( U \) into at least 2 non-empty disjoint sets in \( \mathcal{F} \), i.e.

\[ \mathcal{P} \subseteq \mathcal{F} : |\mathcal{P}| \geq 2, \bigcup \mathcal{P} \supseteq U \text{ and } \forall i \in U, \exists C \in \mathcal{P} : i \in C \]

It follows that \( \Pi(\mathcal{F}, U) \supseteq \Pi(\mathcal{F}, U') \) for all \( U \subseteq U' \).

Define \( \Lambda(\mathcal{F}, U) := \lambda(B : B \in \mathcal{F}) \) satisfying

\[ \forall B \in \mathcal{F}, \lambda_B \geq 0 \text{ and } \forall i \in U, \sum_{B \in \mathcal{F}, i \in B} \lambda_B = 1 \]

It follows that \( \Lambda(\mathcal{F}, U) \supseteq \Lambda(\mathcal{F}, U') \) for all \( U \subseteq U' \).

**Theorem 3.1** Given a finite ground set \( V : |V| \geq 2 \), we have for all \( A \subseteq V : |A| \geq 2 \), \( \mathcal{F} \in \Phi(A) \), and supermodular function \( h : \mathcal{F} \mapsto \mathbb{R} \) that,

\[ \max_{\lambda \in \Lambda(\mathcal{F}, A)} \sum_{B \in \mathcal{F}} \lambda_B h(B) = \max_{\mathcal{P} \in \Pi(\mathcal{F}, A)} \frac{1}{|\mathcal{P}| - 1} \sum_{C \in \mathcal{P}} h(C) \]

with the convention that \( \max \) over an empty set is \(-\infty\).\(^3\)

**REFERENCES**