Mutual dependence of random variables

Chung Chan 陳聰

Institute of Network Coding at CUHK
香港中文大學網絡碼研究所

Ph.D. work at Massachusetts Institute of Technology

Nov 3rd, 2010

A secret agreement game

- A group of users vs a wiretapper
- Everyone writes something in private.
- Users can speak in public.
- Users win iff

> everyone writes the same thing except the wiretapper.

A secret agreement game

- A group of users vs a wiretapper
- Everyone writes something in private.
- Users can speak in public.
- Users win iff

> everyone writes the same thing except the wiretapper.

Key questions

1. secret $\stackrel{?}{=}$ mutual dependence
2. meaningful characterization?
3. other operational meanings?

Information Theory

C. E. Shannon

A mathematical theory of communication (1948)

$$H(\mathsf{X}) = \mathsf{E}\ [\ -\log P_{\mathsf{X}}(\mathsf{X})\ ] \qquad \text{Entropy}$$

$$I(\mathsf{X} \wedge \mathsf{Y}) = H(\mathsf{X}) + H(\mathsf{Y}) - H(\mathsf{XY}) \qquad \text{Mutual Information}$$

$$= \mathsf{E}\left[\log \frac{P_{\mathsf{XY}}}{P_{\mathsf{X}} P_{\mathsf{Y}}}\right]$$

$$=: D\left(P_{\mathsf{XY}} \| P_{\mathsf{X}} P_{\mathsf{Y}}\right) \qquad \text{Divergence}$$

Information Theory

C. E. Shannon

A mathematical theory of communication (1948)

$$H(\mathsf{X}) = \mathsf{E}\ [\ -\log P_{\mathsf{X}}(\mathsf{X})\ ] \qquad \text{Entropy}$$

$$I(\mathsf{X} \wedge \mathsf{Y}) = H(\mathsf{X}) + H(\mathsf{Y}) - H(\mathsf{XY}) \qquad \text{Mutual Information}$$

$$= \mathsf{E}\left[\log \frac{P_{\mathsf{XY}}}{P_{\mathsf{X}} P_{\mathsf{Y}}}\right]$$

$$=: D\left(P_{\mathsf{XY}} \| P_{\mathsf{X}} P_{\mathsf{Y}}\right) \qquad \text{Divergence}$$

Information Theory

C. E. Shannon

A mathematical theory of communication (1948)

$$H(\mathsf{X}) = \mathsf{E}\ [\ -\log P_{\mathsf{X}}(\mathsf{X})\ ] \qquad \text{Entropy}$$

$$I(\mathsf{X} \wedge \mathsf{Y}) = H(\mathsf{X}) + H(\mathsf{Y}) - H(\mathsf{XY}) \qquad \text{Mutual Information}$$

$$= \mathsf{E}\left[\log \frac{P_{\mathsf{XY}}}{P_{\mathsf{X}} P_{\mathsf{Y}}}\right]$$

$$=: D\left(P_{\mathsf{XY}} \| P_{\mathsf{X}} P_{\mathsf{Y}}\right) \qquad \text{Divergence}$$

Generalization

Mutual information for ≥ 2 r.v.s

1. natural extension
2. secret key agreement
3. network coding
4. combinatorics

Secret key agreement

M93 U. M. Maurer. Secret key agreement by public discussion from common information. (1993)

AC93 R. Ahlswede and I. Csiszár. Common randomness in information theory and cryptography Part I: Secret Sharing. (1993)

2-user example

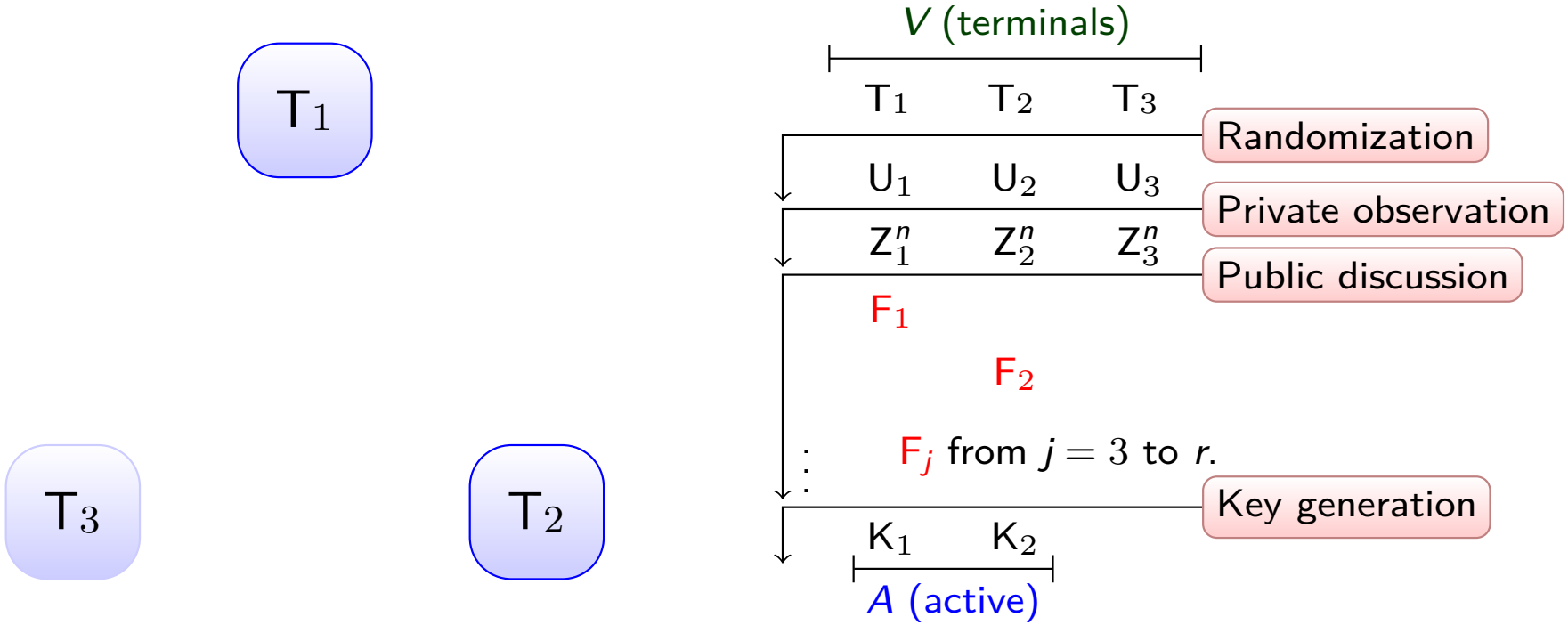
- B_0, B_1, J are independent uniformly random bits.
- Alice observes $Z_1 := (B_0, B_1)$.
- Bob observes $Z_2 := (J, B_J)$.
- Eve observes nothing.

$$\begin{aligned} I(Z_1 \wedge Z_2) &= H(B_0, B_1) + H(B_J, J) - H(B_0, B_1, B_J, J) \\ &= 2 + 2 - 3 \\ &= 1 \end{aligned}$$

Multiterminal secret key agreement

CN04 I. Csiszár and P. Narayan. Secrecy capacities for multiple terminals. (2004)

TGN10 H. Tyagi, P. Gupta and P. Narayan. When is a Function Securely Computable? (2010)

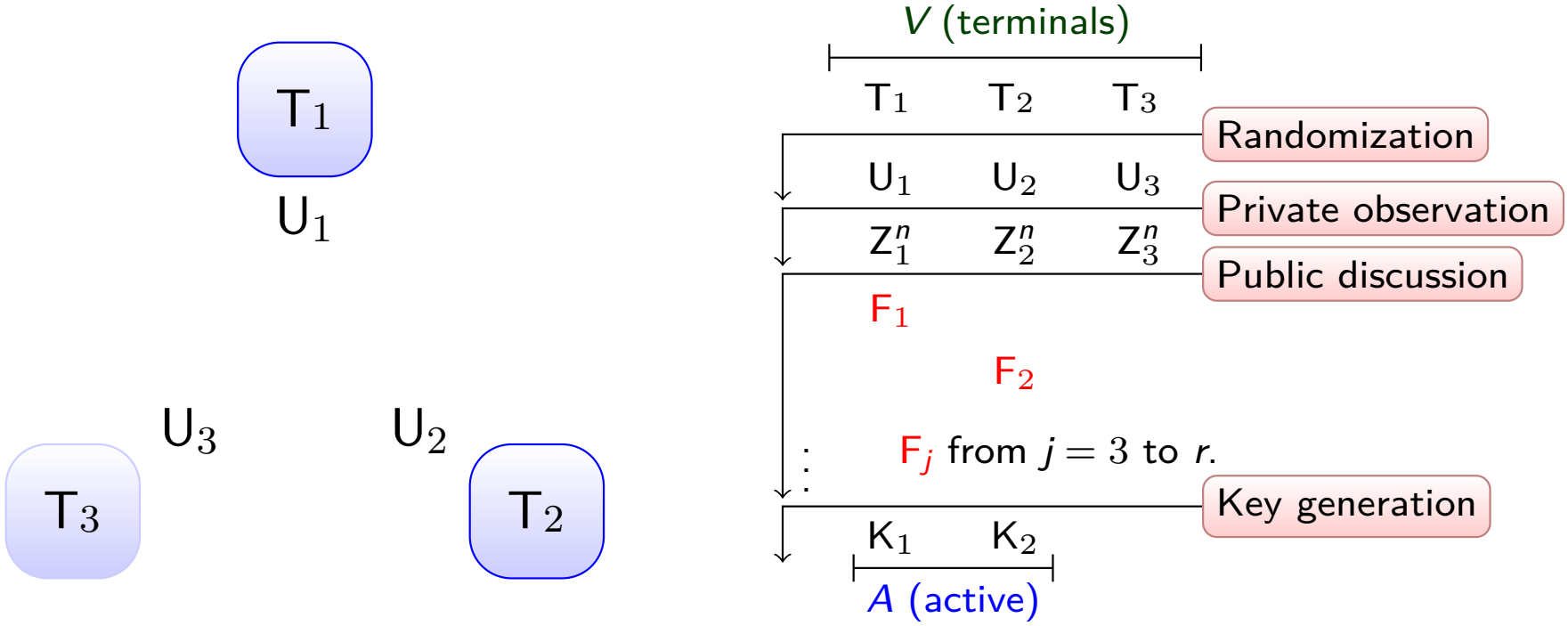


Randomizations: U_i for $i \in V$ with $P_{\mathsf{U}_V} = \prod_{i \in V} P_{\mathsf{U}_i}$

Private sources: Z_i for $i \in V$ with P_{Z_V}

Public messages: $\mathsf{F}_j := F_j(\mathsf{U}_{i_j}, \mathsf{Z}_{i_j}^n, \mathsf{F}_{[j-1]}),\ j \in [r]$

Individual keys: $\mathsf{K}_i := K_i(\mathsf{U}_i, \mathsf{Z}_i^n, \mathsf{F}_{[r]}),\quad i \in A$

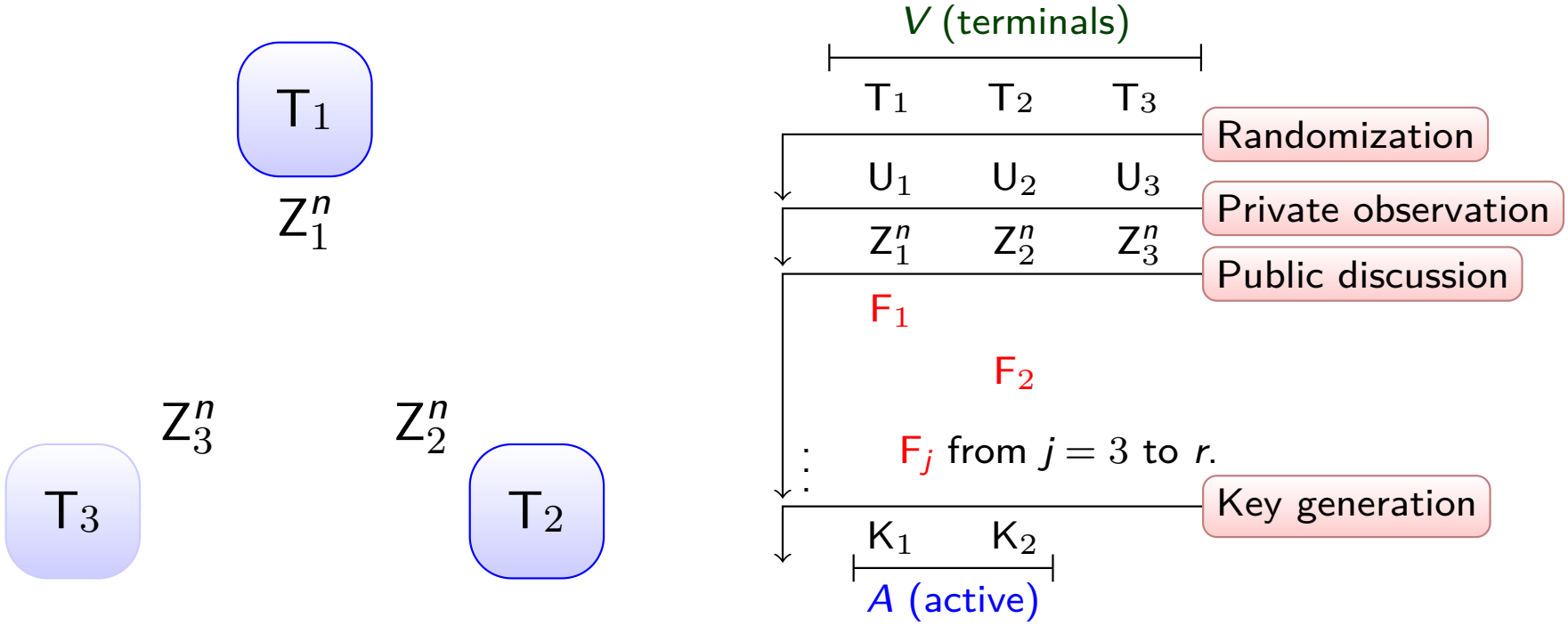


Randomizations: U_i for $i \in V$ with $P_{\mathsf{U}_V} = \prod_{i \in V} P_{\mathsf{U}_i}$

Private sources: Z_i for $i \in V$ with P_{Z_V}

Public messages: $\mathsf{F}_j := F_j(\mathsf{U}_{i_j}, \mathsf{Z}_{i_j}^n, \mathsf{F}_{[j-1]}),\ j \in [r]$

Individual keys: $\mathsf{K}_i := K_i(\mathsf{U}_i, \mathsf{Z}_i^n, \mathsf{F}_{[r]}),\quad i \in A$

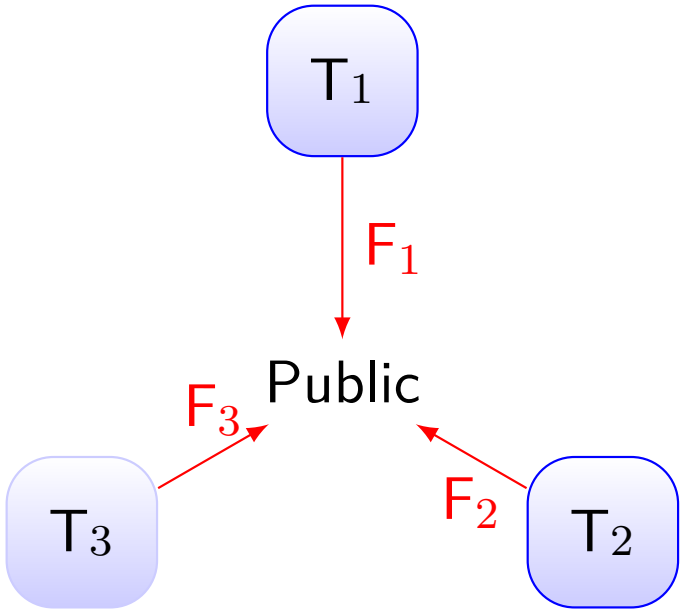


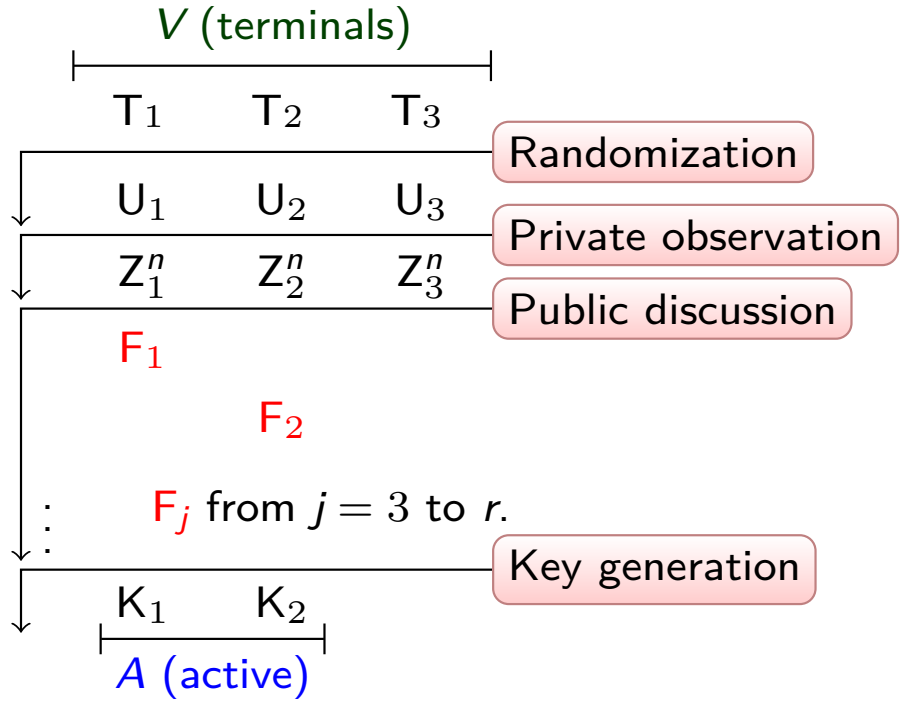
Randomizations: U_i for $i \in V$ with $P_{\mathsf{U}_V} = \prod_{i \in V} P_{\mathsf{U}_i}$

Private sources: Z_i for $i \in V$ with P_{Z_V}

Public messages: $\mathsf{F}_j := F_j(\mathsf{U}_{i_j}, \mathsf{Z}_{i_j}^n, \mathsf{F}_{[j-1]}),\ j \in [r]$

Individual keys: $\mathsf{K}_i := K_i(\mathsf{U}_i, \mathsf{Z}_i^n, \mathsf{F}_{[r]}),\quad i \in A$



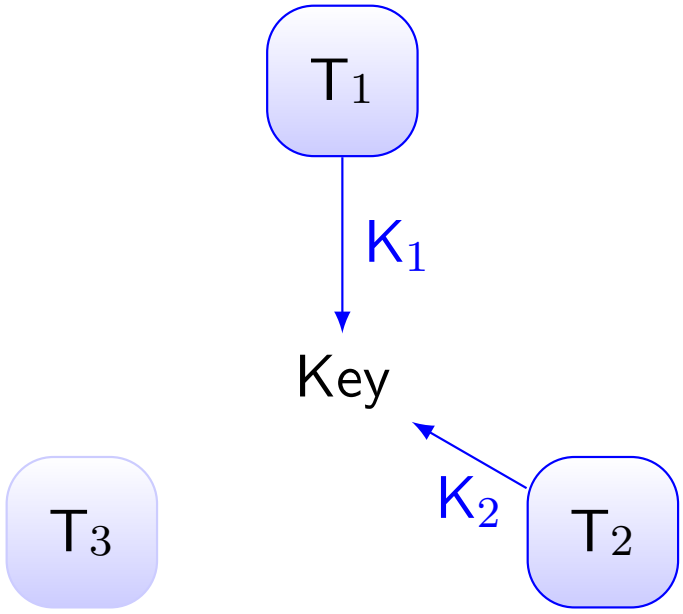


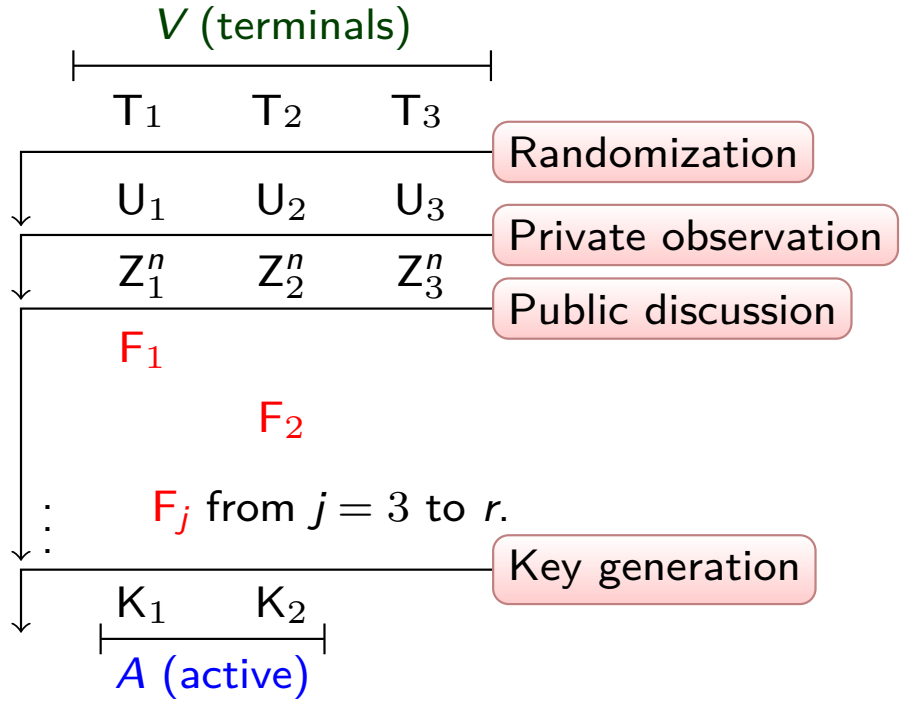
Randomizations: U_i for $i \in V$ with $P_{\mathsf{U}_V} = \prod_{i \in V} P_{\mathsf{U}_i}$

Private sources: Z_i for $i \in V$ with P_{Z_V}

Public messages: $\mathsf{F}_j := F_j(\mathsf{U}_{i_j}, \mathsf{Z}_{i_j}^n, \mathsf{F}_{[j-1]}),\ j \in [r]$

Individual keys: $\mathsf{K}_i := K_i(\mathsf{U}_i, \mathsf{Z}_i^n, \mathsf{F}_{[r]}),\quad i \in A$





Randomizations: U_i for $i \in V$ with $P_{\mathsf{U}_V} = \prod_{i \in V} P_{\mathsf{U}_i}$

Private sources: Z_i for $i \in V$ with P_{Z_V}

Public messages: $\mathsf{F}_j := F_j(\mathsf{U}_{i_j}, \mathsf{Z}_{i_j}^n, \mathsf{F}_{[j-1]}),\ j \in [r]$

Individual keys: $\mathsf{K}_i := K_i(\mathsf{U}_i, \mathsf{Z}_i^n, \mathsf{F}_{[r]}),\quad i \in A$

Secrecy capacity

Recoverability:

$$\Pr\{\exists i \in A,\ \mathsf{K}_i \neq \mathsf{K}\} \to 0$$

Secrecy:

$$\frac{1}{n}\left[\log|K| - H(\mathsf{K}|\mathsf{F})\right] \to 0$$

Recall: $\log|K| \overset{(a)}{\geq} H(\mathsf{K}) \overset{(b)}{\geq} H(\mathsf{K}|\mathsf{F})$

$$C_s := \sup_{\mathsf{U}_i,\mathsf{F}_i,\mathsf{K}_i} \liminf_{n\to\infty} \frac{1}{n} \log|K|$$

LP Characterization

Notation: for $B \subseteq V$, denote $\mathsf{Z}_B := (\mathsf{Z}_i : i \in B)$.

$$C_{\mathsf{s}} = \min_{\lambda} \left[H(\mathsf{Z}_V) - \sum_B \lambda_B \underbrace{H(\mathsf{Z}_B|\mathsf{Z}_{B^c})}_{H(\mathsf{Z}_V)-H(\mathsf{Z}_{B^c})} \right]$$

where

$$\lambda := (\lambda_B \geq 0 : \underbrace{B \not\supseteq A}_{A\text{-co-intersecting}}) \quad \text{with} \quad \underbrace{\sum_{B \ni i} \lambda_B = 1, \forall i \in V}_{\text{fractional partition}}$$

Two-user case:

$$C_s = I(Z_1 \wedge Z_2) = D(P_{Z_1 Z_2} \| P_{Z_1} P_{Z_2})$$

Three-user case:

$$C_s = \min \left\{ I(Z_1 \wedge Z_2 Z_3), I(Z_2 \wedge Z_1 Z_3), I(Z_3 \wedge Z_1 Z_2), \frac{1}{2} D(P_{Z_1 Z_2 Z_3} \| P_{Z_1} P_{Z_2} P_{Z_3}) \right\}$$

Two-user case:

$$C_s = I(Z_1 \wedge Z_2) = D(P_{Z_1 Z_2} \| P_{Z_1} P_{Z_2})$$

Three-user case:

$$C_s = \min \left\{ I(Z_1 \wedge Z_2 Z_3), I(Z_2 \wedge Z_1 Z_3), I(Z_3 \wedge Z_1 Z_2), \frac{1}{2} D(P_{Z_1 Z_2 Z_3} \| P_{Z_1} P_{Z_2} P_{Z_3}) \right\}$$

e.g. $Z_1 = Z_2 = Z_3$

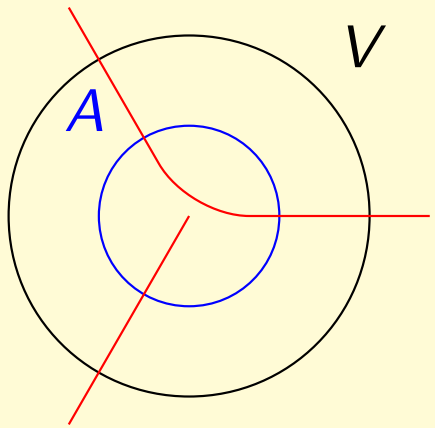
$$\begin{aligned} D(P_{Z_1 Z_2 Z_3} \| P_{Z_1} P_{Z_2} P_{Z_3}) &= H(Z_1) + H(Z_2) + H(Z_3) - H(Z_1 Z_2 Z_3) \\ &= 2H(Z_1) \end{aligned}$$

Mutual dependence

Mutual dependence upper bound [CN04]

$$C_s \leq \min_{\mathcal{P}} \frac{1}{|\mathcal{P}| - 1} D\left(P_{Z_V} \middle\| \prod_{C \in \mathcal{P}} P_{Z_C} \right)$$

$\mathcal{P}$: set partition, A-intersecting



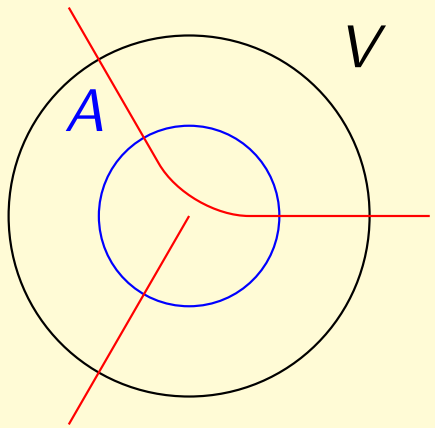
1. When is it tight?
2. Any counter-example?

Mutual dependence

Mutual dependence upper bound [CN04]

$$C_s \leq \min_{\mathcal{P}} \frac{1}{|\mathcal{P}| - 1} D\left(P_{Z_V} \middle\| \prod_{C \in \mathcal{P}} P_{Z_C} \right)$$

$\mathcal{P}$: set partition, A-intersecting



1. When is it tight? $A = V$.
2. Any counter-example? Yes with $V = [6]$, $A = [3]$.

[CZ10] C. Chan and L. Zheng. Mutual dependence for secret key agreement. (2010)

General identity for submodular functions

Submodularity of entropy [Fujishige 78]

$$\forall B_1, B_2 \subseteq V, \quad H(Z_{B_1}) + H(Z_{B_2}) \geq H(Z_{B_1 \cap B_2}) + H(Z_{B_1 \cup B_2})$$

$$\iff I(Z_{B_1} \wedge Z_{B_2} | Z_{B_1 \cap B_2}) \geq 0$$

General identity for submodular functions

Submodularity of entropy [Fujishige 78]

$$\forall B_1, B_2 \subseteq V, \quad \underbrace{H(Z_{B_1})}_{h(B_1)} + \underbrace{H(Z_{B_2})}_{h(B_2)} \geq \underbrace{H(Z_{B_1 \cap B_2})}_{h(B_1 \cap B_2)} + \underbrace{H(Z_{B_1 \cup B_2})}_{h(B_1 \cup B_2)}$$

General identity for submodular functions

Submodularity of entropy [Fujishige 78]

$$\forall B_1, B_2 \subseteq V, \quad \underbrace{H(Z_{B_1})}_{h(B_1)} + \underbrace{H(Z_{B_2})}_{h(B_2)} \geq \underbrace{H(Z_{B_1 \cap B_2})}_{h(B_1 \cap B_2)} + \underbrace{H(Z_{B_1 \cup B_2})}_{h(B_1 \cup B_2)}$$

Theorem ([CZ10])

$$\min_{\lambda} \sum_{B \in \mathcal{F}} \lambda_B h(B) = \min_{\mathcal{P}} \frac{1}{|\mathcal{P}| - 1} \sum_{C \in \mathcal{P}} h(C^c)$$

- $\mathcal{F}$: *co-intersecting*
- λ: *fractional partition*
- $\mathcal{P}$: *set partition*

Other multivariate correlations

- McGill's interaction information (1954)

$$I(Z_i \wedge Z_j | Z_k) - I(Z_i \wedge Z_j)$$

- Watanabe's total correlation (1960)

$$\sum_{i \in V} H(Z_i) - H(Z_V) = D\left(P_{Z_V} \middle\| \prod_{i \in V} P_{Z_i} \right)$$

- Yeung's co-information (1991)

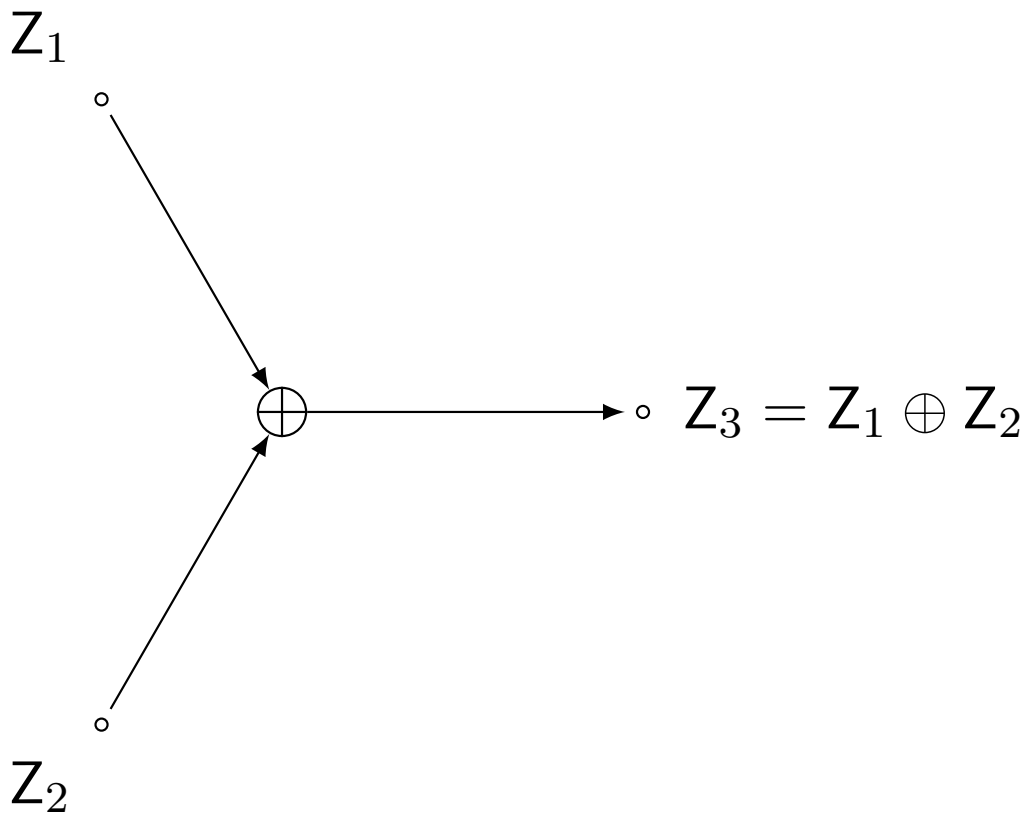
$$-\sum_{B \subseteq V} (-1)^{|B|} H(Z_B)$$

- Jakulin and Bratko's m-way interaction information (2004)

$$-\sum_{B \subseteq V} (-1)^{|V|-|B|} H(Z_B)$$

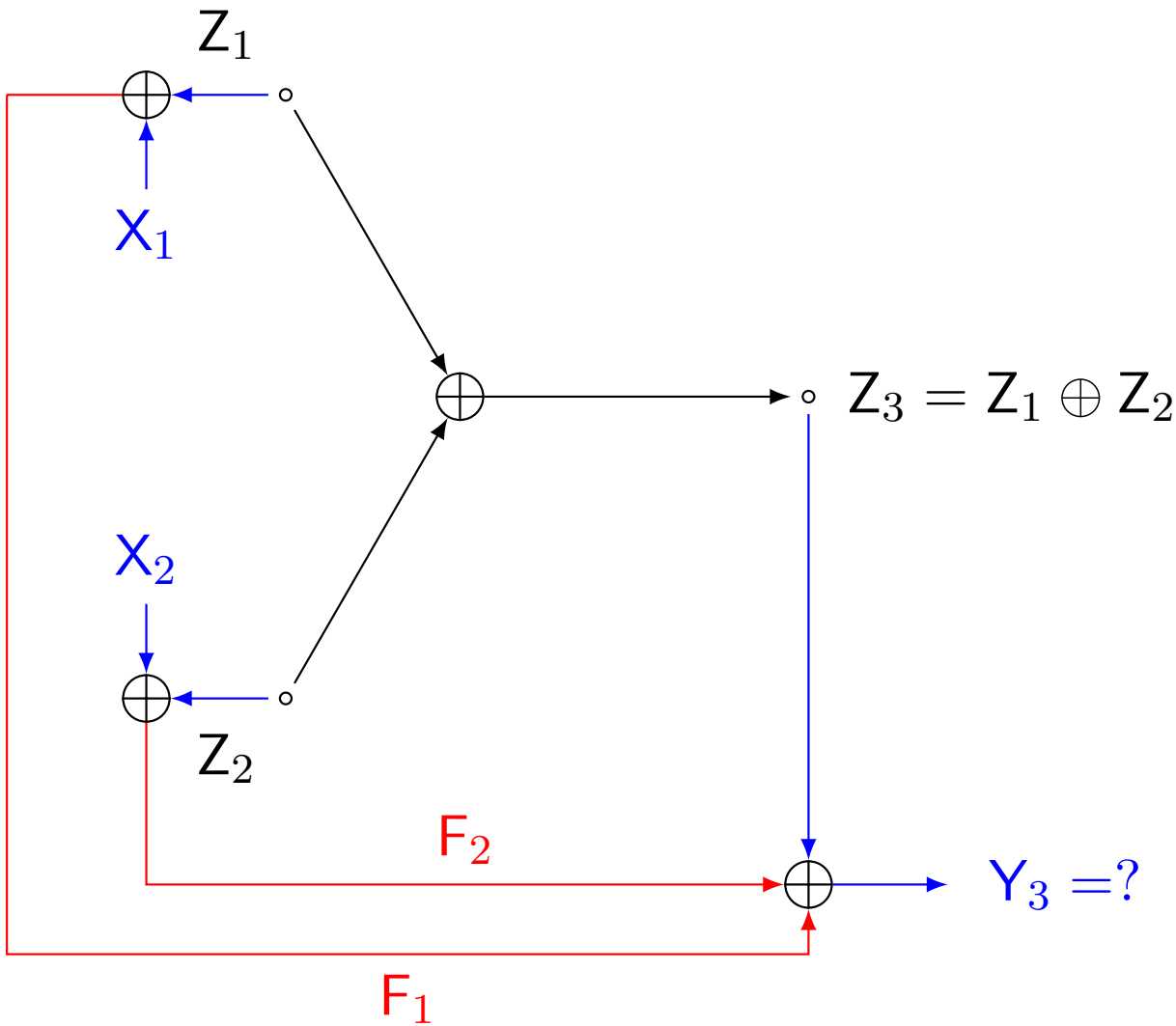
Linear source model

Z_1, Z_2: independent, uniformly random bits

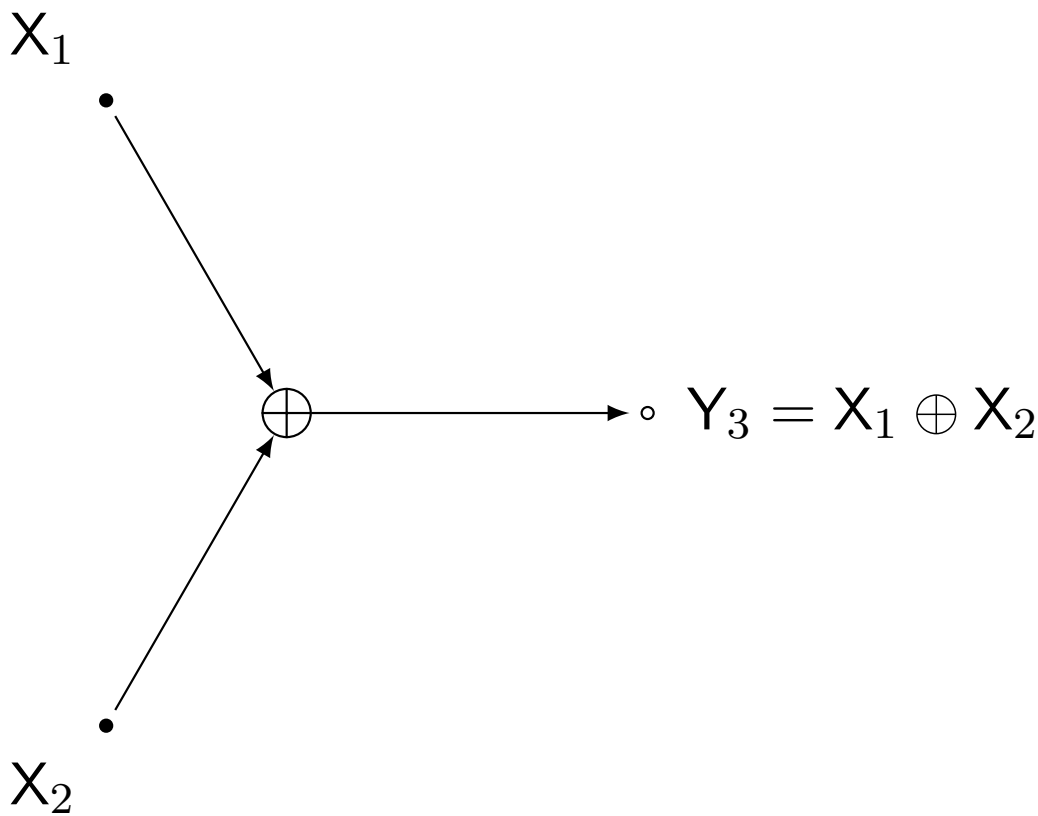


Turn private source into private channel

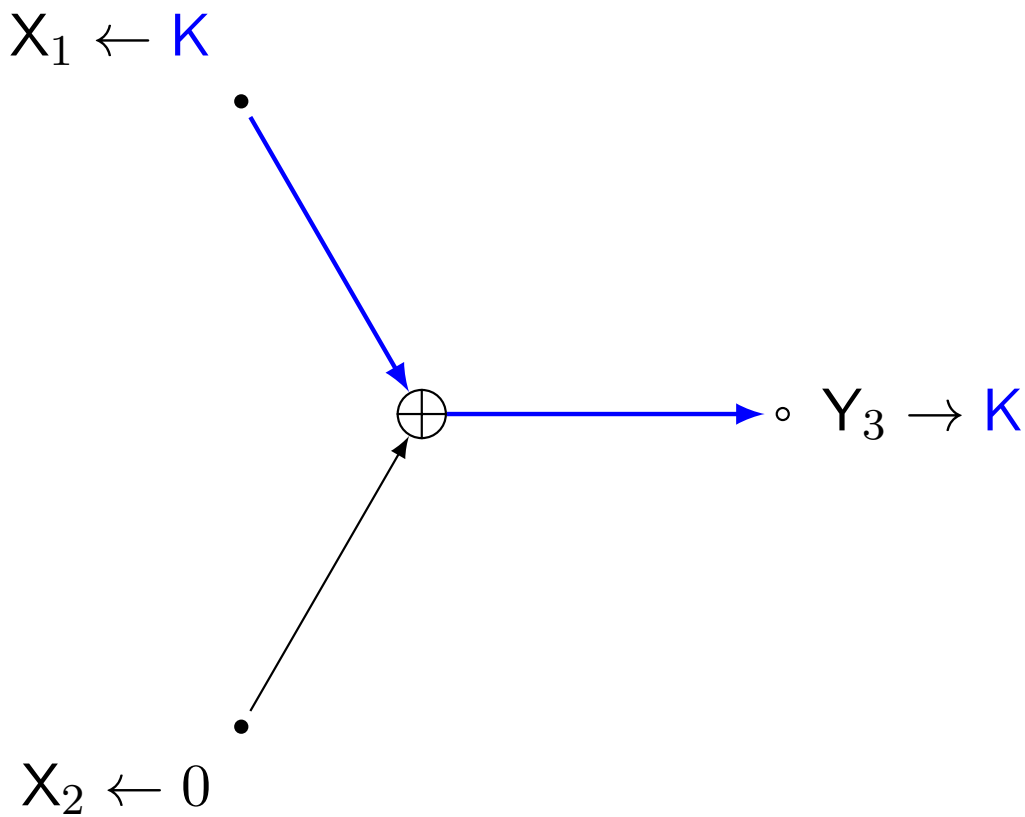
X_1, X_2: input bits Y_3: output bit



Effective private channel

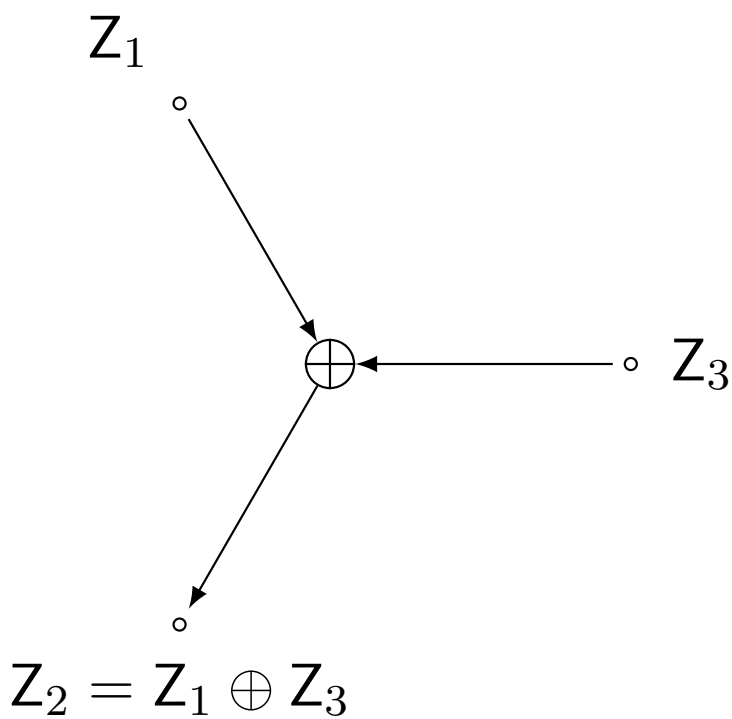


Route a secret bit from T_1 to T_3

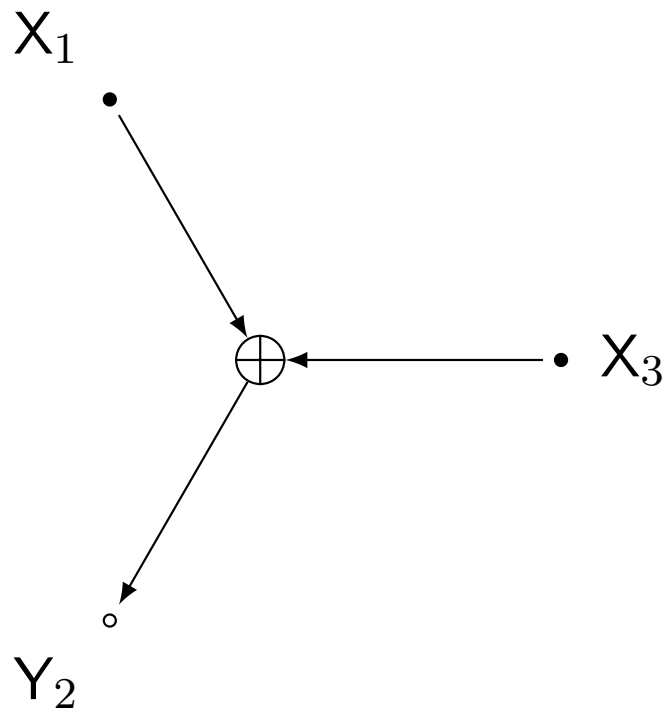


Reorient the inputs

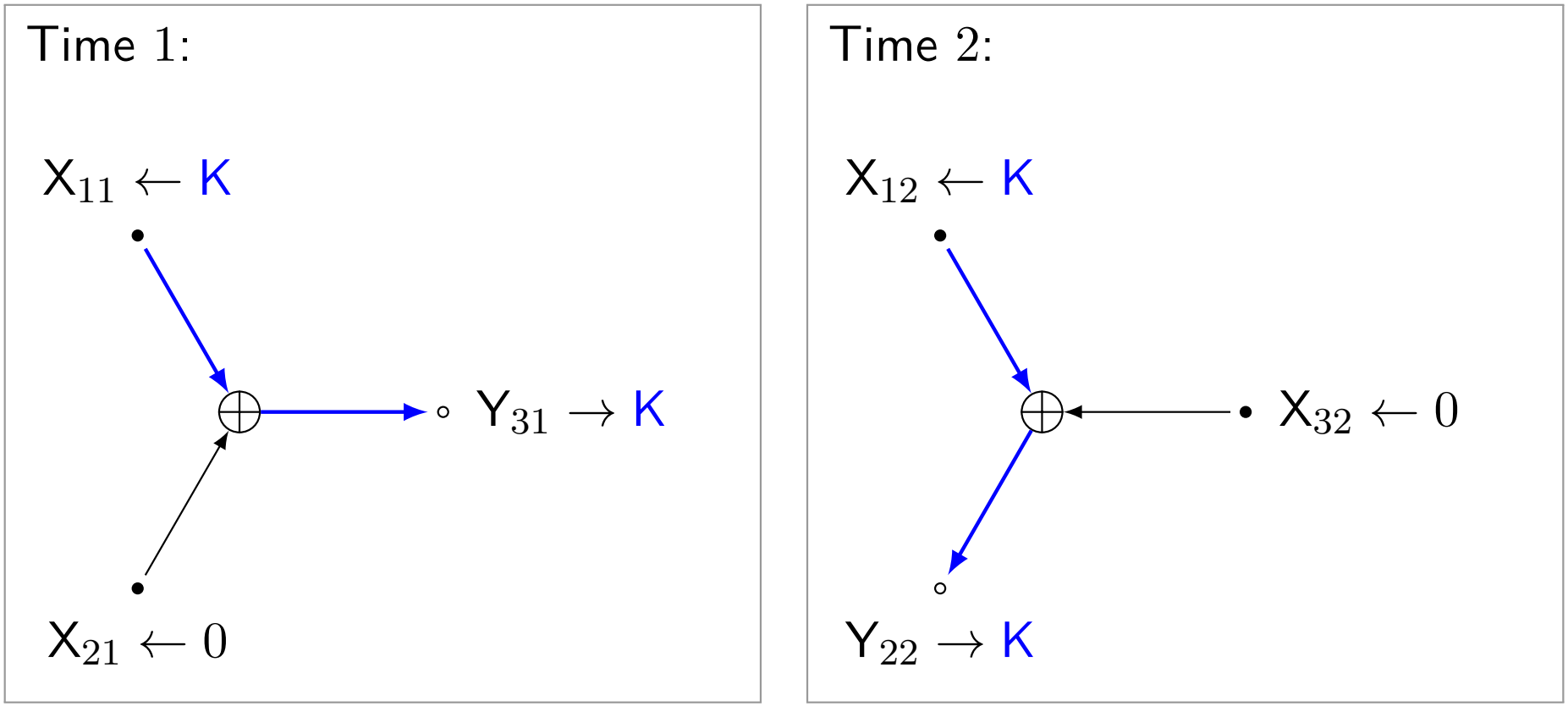
Equivalent source:



Effective channel:



Network coding for secret key agreement



Network throughput $= 0.5$

Is it optimal?

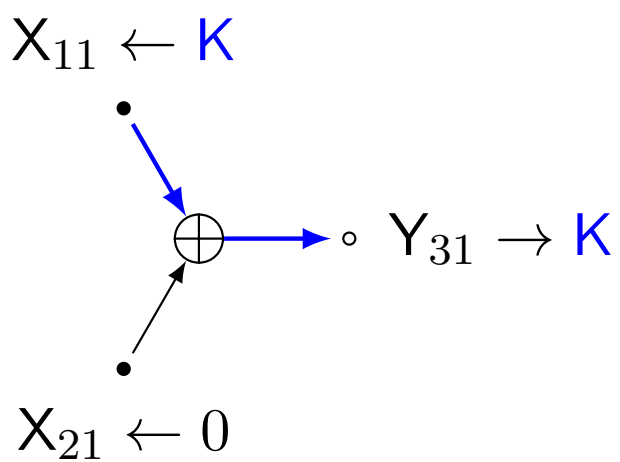
$$
\begin{aligned}
C_s &\leq \frac{1}{2} D(P_{Z_1 Z_2 Z_3} \| P_{Z_1} P_{Z_2} P_{Z_3}) \\
&= \frac{1}{2} \left[H(Z_1) + H(Z_2) + H(Z_3) - H(Z_1 Z_2 Z_3) \right] \\
&= \frac{1}{2} \left[1 + 1 + 1 - 2 \right] \\
&= 0.5
\end{aligned}
$$

Single-source network coding

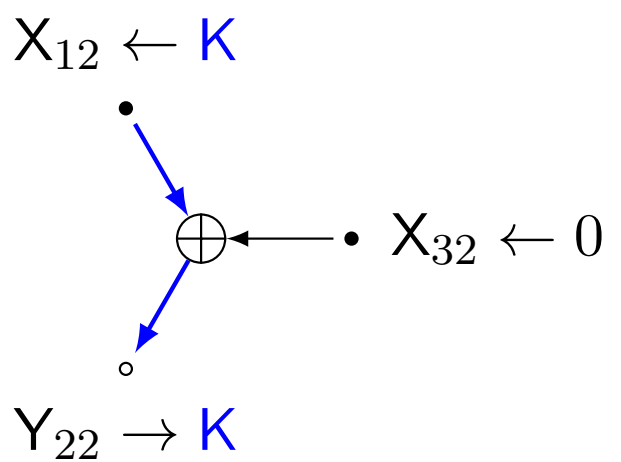
1. Extend the source n times to Z_V^n.
2. Pick a base X_V^n from Z_V^n as inputs.
3. Pick a source node $s \in A$ to generate K and multicast it to A.

- $n = 2$
- $X_{V1} = (Z_{11}, Z_{21}, \emptyset)$
  $X_{V2} = (Z_{12}, \emptyset, Z_{32})$
- $s = 1$

Time 1:



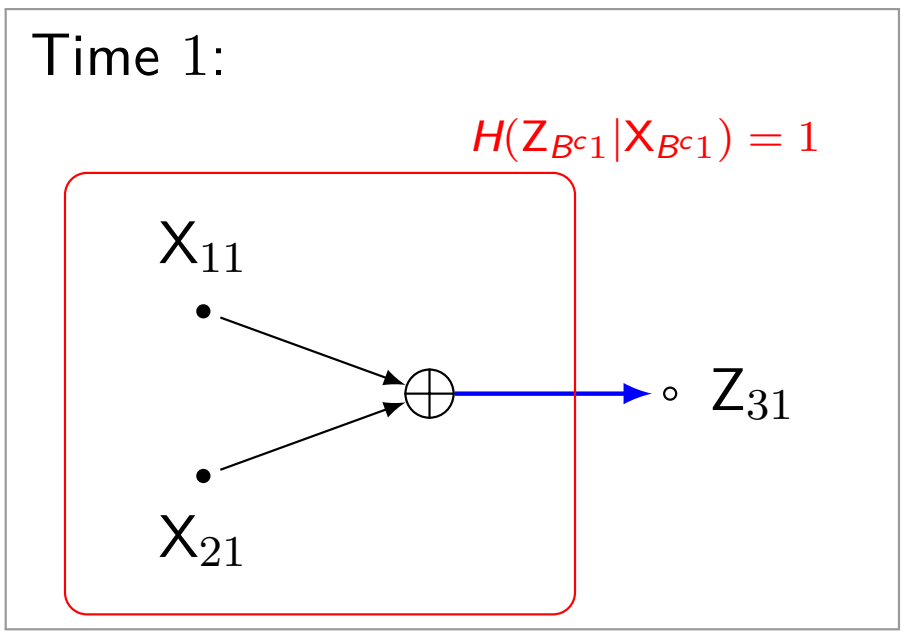
Time 2:

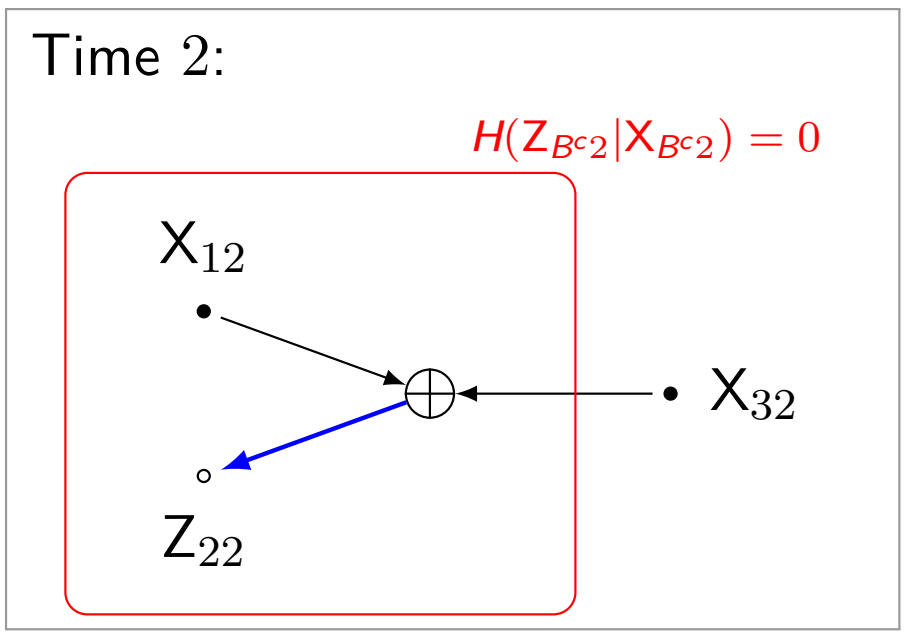


Min-cut characterization

$$\text{Throughput} = \frac{1}{n} \min_{B \subseteq V : s \in B \not\supseteq A} H(Z^n_{B^c} | X^n_{B^c}) \quad \text{given } n,\ X^n_V,\ s$$

Avestimehr, Diggavi, Tse. Wireless network information flow (2007)

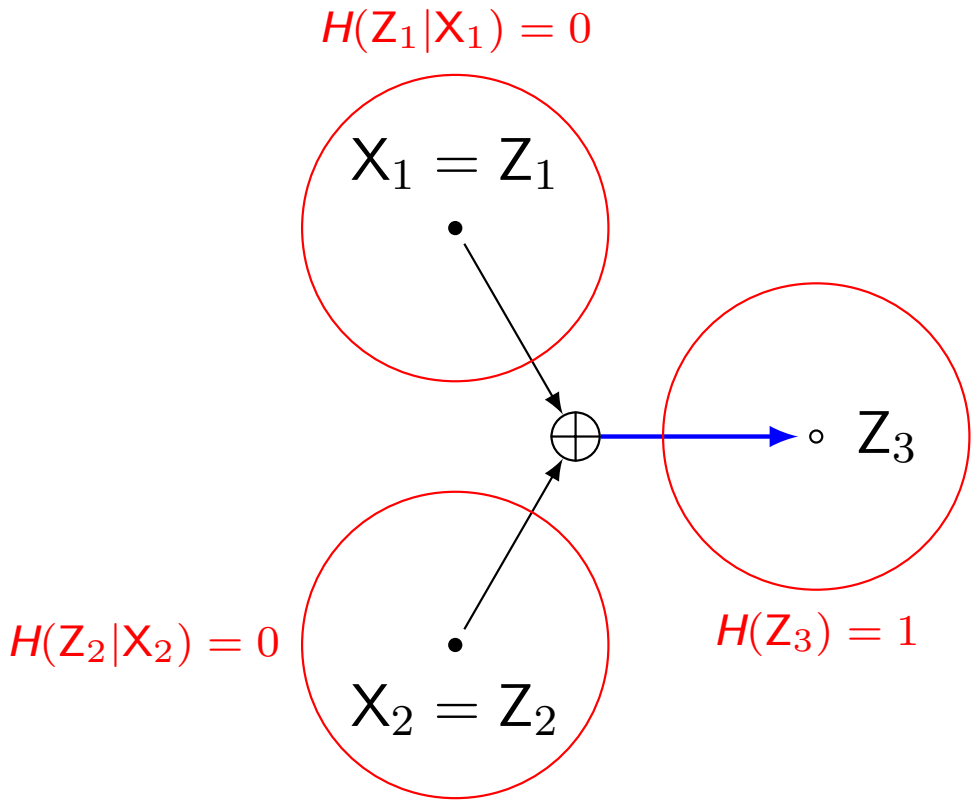




Partition connectivity

When $V = A$ [CZ10]

$$C_s = \min_{\mathcal{P}} \frac{1}{|\mathcal{P}| - 1} \sum_{C \in \mathcal{P}} H(Z_C | X_C) \qquad \text{invar. to } X_V.$$



General identity for matroids

$$A = V \implies \text{max. throughput} = \text{secrecy capacity}$$

For rank function r of a matroid with ground set Z_V,

$$\max_{X_V} \min_{s \in B \not\supseteq V} r(Z_{B^c}|X_{B^c}) = \left\lfloor \max_{X_V} \min_{\mathcal{P}} \frac{1}{|\mathcal{P}| - 1} \sum_{C \in \mathcal{P}} r(Z_C|X_C) \right\rfloor$$

An extension of

- T.K.A. Frank, T. Király and M. Kriesell. On decomposing a hypergraph into k-connected sub-hypergraphs. (2003)
- J. Bang-Jensen and S. Thomassé. Decompositions and orientations of hypergraphs. (2001)

Summary

Key questions

1. secret $\stackrel{?}{=}$ mutual dependence
2. meaningful characterization?
3. other operational meanings?

- secret = minimum normalized divergence
- network throughput
- partition connectivity

Extensions

1. Is single-source network coding optimal when $A \subsetneq V$?
2. Given a private channel instead of a source.
3. What is the minimum public discussion?
4. Other applications?

Thank you!

Email: cchan@inc.cuhk.edu.hk

Reference: `sites.google.com/site/chungcmit/documents/`