Multiterminal Data Compression and Secret Key Generation

Prakash Narayan

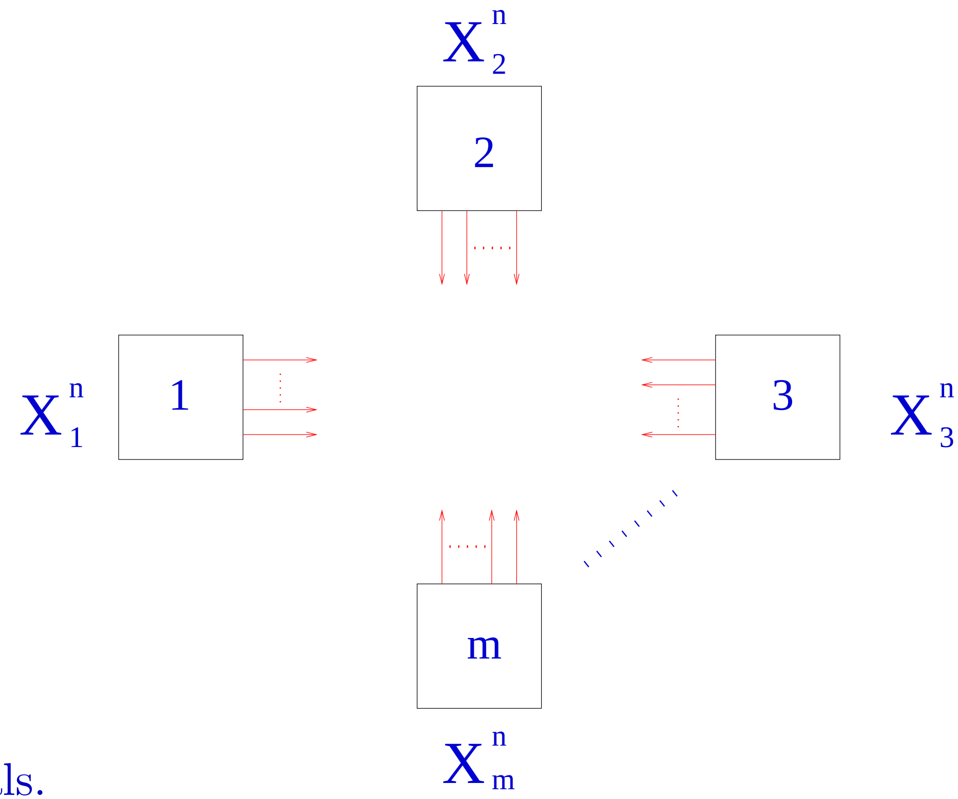
Joint work with Imre Csiszár and Chunxuan Ye

$\boxed{\textbf{Multiterminal Data Compression}}$

Multiterminal Data Compression

- Multiple terminals observe separate but correlated signals, e.g., different noisy versions of a common broadcast signal or measurements of a parameter of the environment.
- The terminals seek to attain omniscience, i.e., to learn *all* the signals.
- To this end, the terminals then transmit to each other.
- Such transmissions occur in a rate-efficient manner, and exploit the correlated nature of the observed signals.
- This problem does not involve any secrecy constraints.

The Model



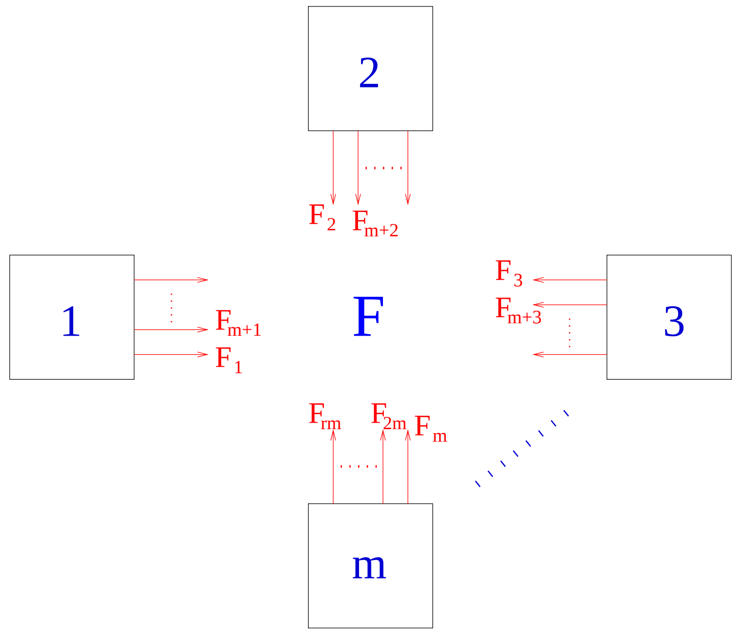
- $m \geq 2$ terminals.

- $X_1, \ldots, X_m$, are finite-valued random variables (rvs) with (known) joint distribution $P_{X_1,\ldots,X_m}$.

- Each terminal i, $i = 1, \ldots, m$, observes a signal comprising n independent and identically distributed versions (say, in time) of the rv X_i, namely the sequence $X_i^n = (X_{i1}, \ldots, X_{in})$.

- The signal components observed by the different terminals at each time are identically distributed according to $P_{X_1,\ldots,X_m}$.

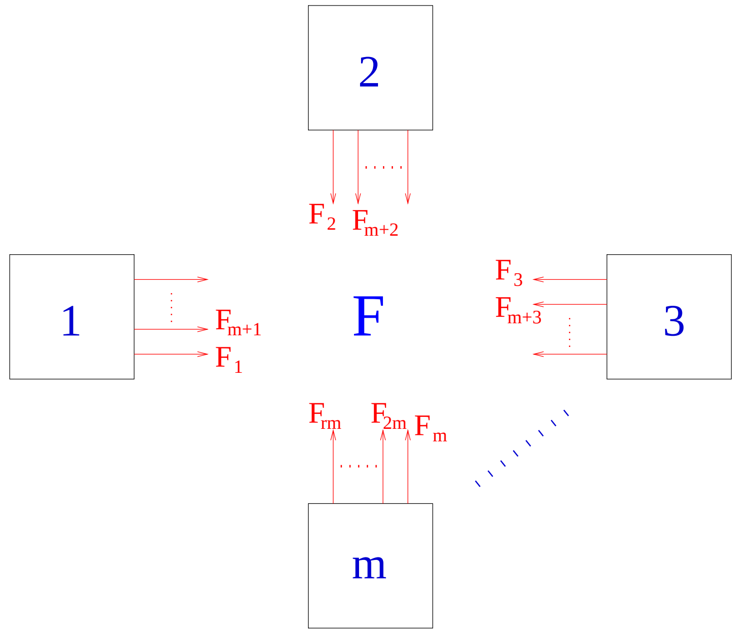
The Model



Objective: Each terminal wishes to become "omniscient," i.e., to reconstruct $(X_1^n, \dots, X_m^n)$ with probability $\cong 1$.

- The terminals are allowed to communicate over a *noiseless* channel, possibly interactively in several rounds.
- The transmissions from any terminal are observed by all the other terminals.
- A transmission from a terminal is allowed to be any function of its own observed signal, and of all previous transmissions.
- No (explicit) rate constraints are imposed on the transmissions.
- Let **F** denote collectively all the transmissions.

Communication for Omniscience



Objective: Each terminal wishes to become "omniscient," i.e., to reconstruct $(X_1^n, \ldots, X_m^n)$ with probability $\cong 1$, using communication $\mathbf{F} = \mathbf{F}(n)$.

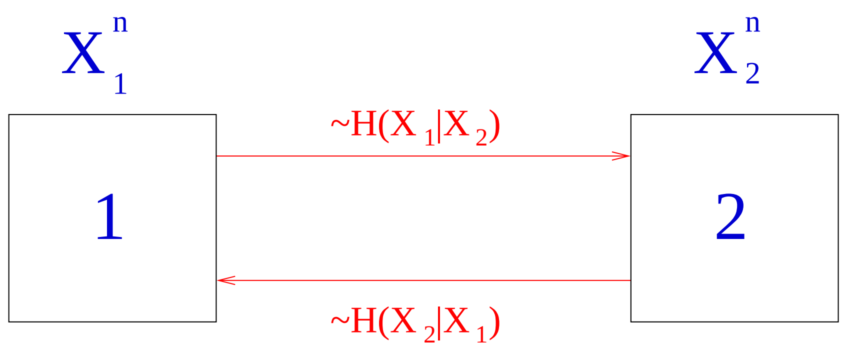
- What is the minimum number of bits of overall communication $\mathbf{F} = \mathbf{F}(n)$ needed for all the terminals to achieve omniscience?

- The smallest achievable rate of *communication for omniscience* (CO-rate):

$$R_{min} \triangleq \inf_n \lim \frac{1}{n} \log_2 \left(\text{cardinality of range of } \mathbf{F}\right).$$

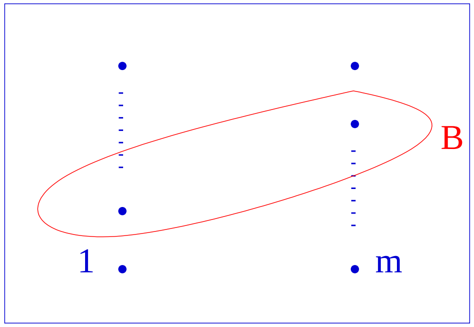
A Special Case: Two Terminals

Slepian-Wolf Data Compression (1973)



$$R_{min} = H(X_1|X_2) + H(X_2|X_1).$$

Minimum Communication for Omniscience



Proposition [I. Csiszár - P. N., '04]: The smallest achievable CO-rate R_{min} is

$$R_{min} = \min_{(R_1,\ldots,R_m)\in\mathcal{R}_{SW}} \sum_{i=1}^{m} R_i,$$

where $\mathcal{R}_{SW} = \left\{(R_1,\cdots,R_m) : \sum_{i\in B} R_i \geq H(X_B|X_{B^c}),\ \forall B\subset\{1,\ldots,m\}\right\},$

and can be achieved with noninteractive communication.

$\boxed{\textbf{Secret Key Generation}}$

Secret Key Generation

- Multiple terminals observe separate but correlated signals, e.g., different noisy versions of a common broadcast signal or measurements of a parameter of the environment.
- The terminals then transmit over a noiseless public channel in order to generate a secret key, i.e.,
  - random variables (rvs) generated at each terminal which agree with probability $\cong 1$; and
  - the rvs are effectively concealed from an eavesdropper with access to the public transmissions.
- The key generation procedure exploits the correlated nature of the observed signals.
- The secret key thereby generated can be used for secure encrypted communication between the terminals.

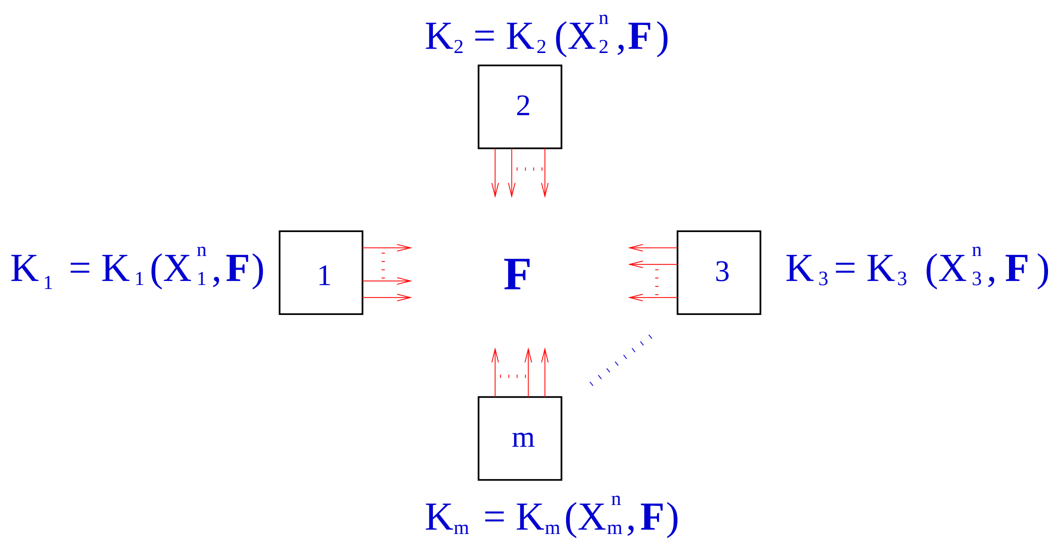
Some Related Work

- Maurer 1990, 1991, 1993, 1994, $\cdots$
- Ahlswede - Csiszár 1993, 1994, 1998, $\cdots$
- Bennett, Brassard, Crépeau, Maurer 1995.
- Csiszár 1996.
- Maurer - Wolf 1997, 2003, $\cdots$
- Venkatesan - Anantharam 1995, 1997, 1998, 2000, $\cdots$
- Csiszár - Narayan 2000, 2004, 2005.
- Renner - Wolf 2003.
- Muramatsu 2004, 2005.
- Ye - Narayan 2004, 2005.

$\vdots$

$\vdots$

What is a Secret Key?

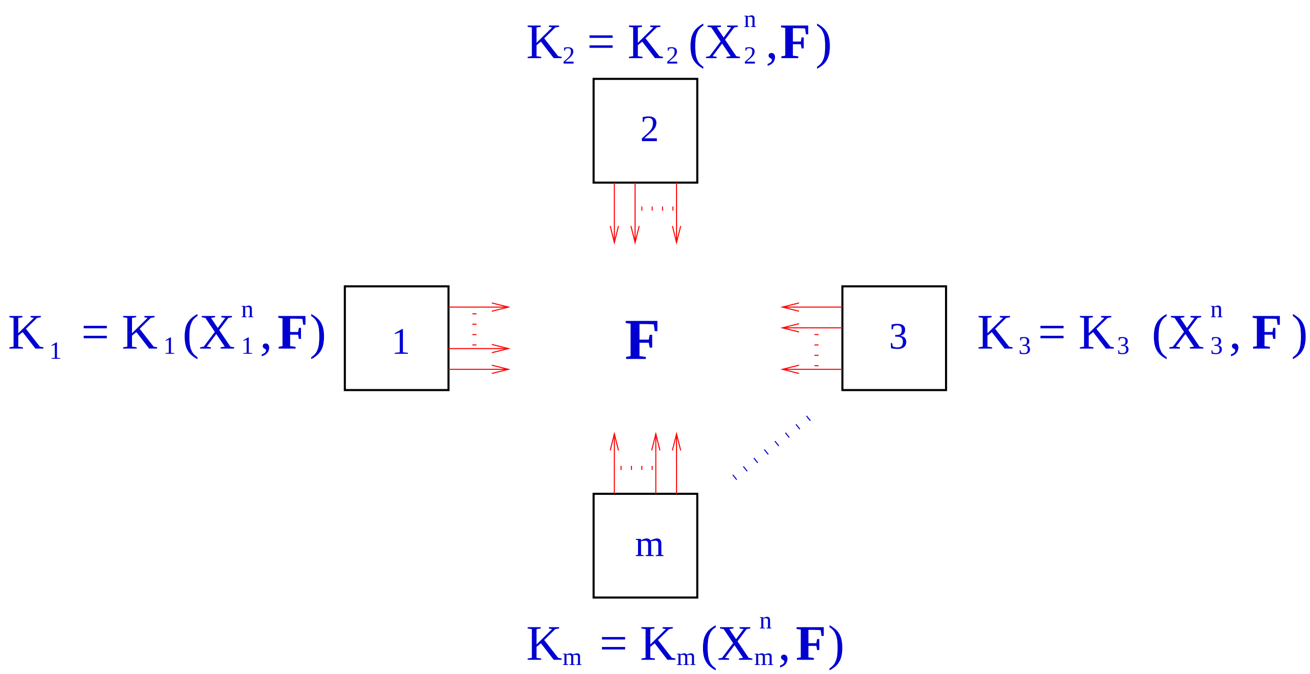


Secret Key (SK): A function K of $(X_1^n, \cdots, X_m^n)$ is a SK, achievable with communication $\mathbf{F}$, if

- $Pr\{K = K_1 = \cdots = K_m\} \cong 1$ ("common randomness")
- $I(K \wedge \mathbf{F}) \cong 0$ ("secrecy")
- $H(K) \cong \log$ (cardinality of key space) . ("uniformity")

Thus, a secret key is effectively concealed from an eavesdropper with access to $\mathbf{F}$, and is nearly uniformly distributed.

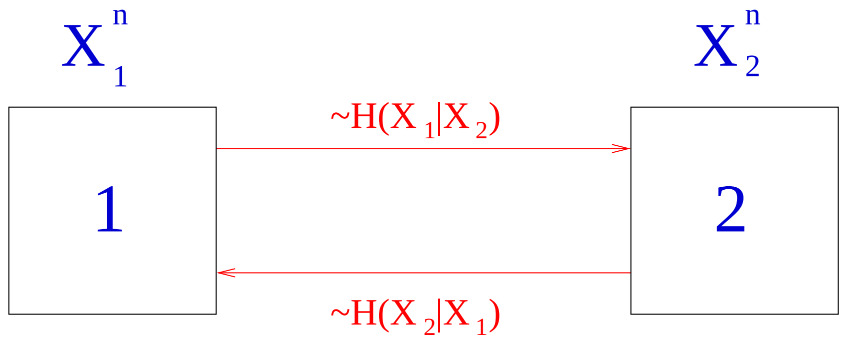
Secret Key Capacity



Objective: Determine the *largest entropy rate of such a* SK which can be achieved with suitable communication: SK-capacity C_{SK}.

$\boxed{\textbf{The Connection}}$

Special Case: Two Terminals



- SK-capacity [Maurer '93, Ahlswede-Csiszár '93]:

$$C_{SK} = I(X_1 \wedge X_2).$$

- **An interpretation:**

$$\begin{aligned} C_{SK} &= I(X_1 \wedge X_2) \\ &= H(X_1, X_2) - [H(X_1|X_2) + H(X_2|X_1)] \\ &= \text{Entropy rate of omniscience} - \text{Smallest achievable CO-rate } R_{min}. \end{aligned}$$

Secret Key Capacity

Theorem [I. Csiszár - P. N., '04]: The SK-capacity C_{SK} for the terminals $1, \ldots, m$ equals

$$C_{SK} = H(X_1, \ldots, X_m) - \text{Smallest achievable CO-rate, } R_{min}\text{, i.e., smallest aggregate rate of communication which enables all the terminals to become omniscient}$$

and can be achieved with noninteractive communication.

- A (single-letter) characterization of R_{min}, thus, leads to the same for C_{SK}.
- The SK-capacity is not increased by randomization at the terminals.

Note: R_{min} is obtained as a solution to a multiterminal data compression problem *not involving any secrecy constraints.*

$\boxed{\textbf{Main Idea}}$

Lemma [I. Csiszár - P. N., '04]: *If L represents "common randomness" for all the terminals, achievable with communication $\mathbf{F}$ for some (signal) observation length n, then $\frac{1}{n}H(L|\mathbf{F})$ is an achievable SK-rate.*

In particular, with $L \cong$ omniscience $= (X_1^n, \ldots, X_m^n)$, we get

$$\frac{1}{n}H(L|\mathbf{F}) \;\cong\; \frac{1}{n}H(X_1^n, \ldots, X_m^n|\mathbf{F}) \;=\; H(X_1, \ldots, X_m) - \frac{1}{n}H(\mathbf{F}).$$

Elementary Constructive Schemes for Secret Key Generation

How is a Secret Key Obtained?

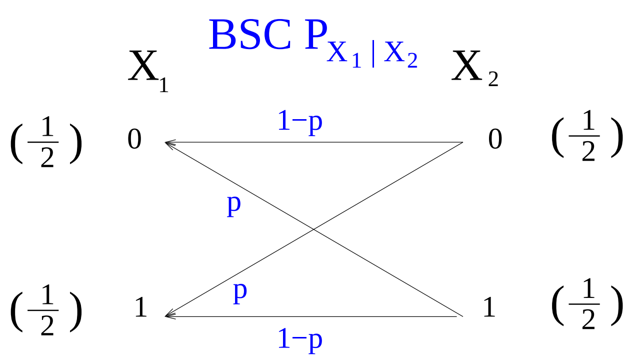
- **Step 1: Data compression:** The terminals communicate over the public channel using compressed data in order to generate "common randomness." These public transmissions are observed by the eavesdropper.

- **Step 2: Secret key construction:** The terminals then process this "common randomness" to extract a secret key of which the eavesdropper has provably little or no knowledge.

Model 1: Two Terminals with Symmetrically Correlated Signals

- Terminals 1 and 2 observe, respectively, n i.i.d. repetitions of the correlated rvs X_1 and X_2, where

- X_1, X_2 are $\{0,1\}$-valued rvs that are "symmetrically" connected by a *virtual* BSC(p), $p < \frac{1}{2}$.



- Have seen that: $C_{SK} = I(X_1 \wedge X_2) = 1 - h_b(p)$ bit/symbol.

- Can assume: $X_1^n = X_2^n \oplus V^n$, where $V^n = (V_1, \cdots, V_n)$ is independent of X_2^n, and is a Bernoulli(p) sequence of rvs.

$\boxed{\textbf{A Useful Fact}}$

P. Elias, 1955

For a BSC $P_{X_1|X_2}$ with $0 < p < \frac{1}{2}$, there exists a binary *linear* block code with parity check matrix $\mathbf{P}$ and codewords of blocklength n, and with

- rate $\cong$ channel capacity $= 1 - h_b(p)$; and
- average error probability of ML decoding

$$1 \;-\; \Pr\{f_{\mathbf{P}}(\mathbf{P}V^n) = V^n\}$$

  vanishing exponentially rapidly with increasing blocklength n, where $f_{\mathbf{P}}(\mathbf{P}V^n)$ is the most likely noise sequence V^n with syndrome $\mathbf{P}V^n$.

Step 1: Slepian-Wolf Data Compression

A.D. Wyner, 1974: Scheme for reconstructing x_1^n at terminal 2

- Standard array for $(n, n-m)$ linear channel code with parity check matrix $\mathbf{P}$:

$$
\begin{array}{cccccc}
c_1^n & c_2^n & \cdots & c_j^n & \cdots & c_{2^{n-m}}^n \\
e_2^n & e_2^n + c_2^n & & e_2^n + c_j^n & & e_2^n + c_{2^{n-m}}^n \\
\vdots & & & & & \vdots \\
e_i^n & e_i^n + c_2^n & & e_i^n + c_j^n = x_1^n & & e_i^n + c_{2^{n-m}}^n \\
\vdots & & & & & \\
e_{2^m}^n & e_{2^m}^n + c_2^n & \cdots & e_{2^m}^n + c_j^n & \cdots & e_{2^m}^n + c_{2^{n-m}}^n
\end{array}
$$

- Terminal 1 transmits $\mathbf{F} =$ the syndrome $\mathbf{P}x_1^n (= \mathbf{P}(x_1^n)^t)$ to terminal 2.
- Terminal 2 computes the ML estimate $\widehat{x}_1^n = \widehat{x}_1^n(x_2^n, \ \mathbf{F})$ as:

$$\widehat{x}_1^n = x_2^n \oplus f_{\mathbf{P}}(\mathbf{P}x_1^n \oplus \mathbf{P}x_2^n),$$

where $f_{\mathbf{P}}(\mathbf{P}x_1^n \oplus \mathbf{P}x_2^n) =$ most likely noise sequence v^n with syndrome

$$\mathbf{P}v^n = \mathbf{P}x_1^n \oplus \mathbf{P}x_2^n.$$

- Thus, terminal 2 reconstructs x_1^n with

$$\Pr\{\widehat{X}_1^n = X_1^n\} = \cdots = \Pr\{f_{\mathbf{P}}(\mathbf{P}V^n) = V^n\} \cong 1.$$

$\boxed{\textbf{Step 2: Secret Key Construction}}$

C. Ye - P.N., '05

- Secret key for terminals 1 and 2

  Terminal 1 sets K_1 = numerical index of x_1^n in coset containing x_1^n;

  Terminal 2 sets K_2 = numerical index of $\widehat{x}_1^n$ in coset containing x_1^n.

- For a systematic channel code: K_1 (resp. K_2) = first $(n - m)$ bits of x_1^n (resp. $\widehat{x}_1^n$).

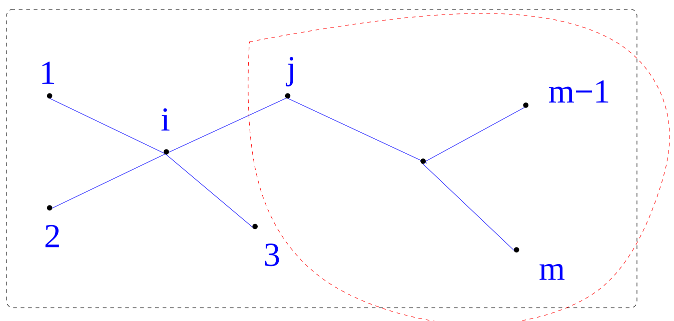
- K_1 or K_2 forms an optimal secret key, since:
  - $\Pr\{K_1 = K_2\} = \Pr\{\widehat{X}_1^n = X_1^n\} \cong 1;$ (common randomness)
  - $I(K_1 \wedge \mathbf{F}) = 0;$ (secrecy)

    as K_1 conditioned on $\mathbf{F} = \mathbf{P}X_1^n \sim$ uniform $\{1, \cdots, 2^{n-m}\}$;
  - $K_1 \sim$ uniform $\{1, \cdots, 2^{n-m}\}$; (uniformity)
  - $\frac{1}{n}H(K_1) = \frac{n-m}{n} \cong 1 - h_b(p).$ (SK-capacity)

Model 2: Markov Chain on a Tree

- Connected graph G with vertex set $= \{1, \cdots, m\}$, edge set E, no circuits (tree).
- $X_1, \cdots, X_m$ are assigned to the vertices $1, \cdots, m$.
- Conditional independence structure determined by G.

G

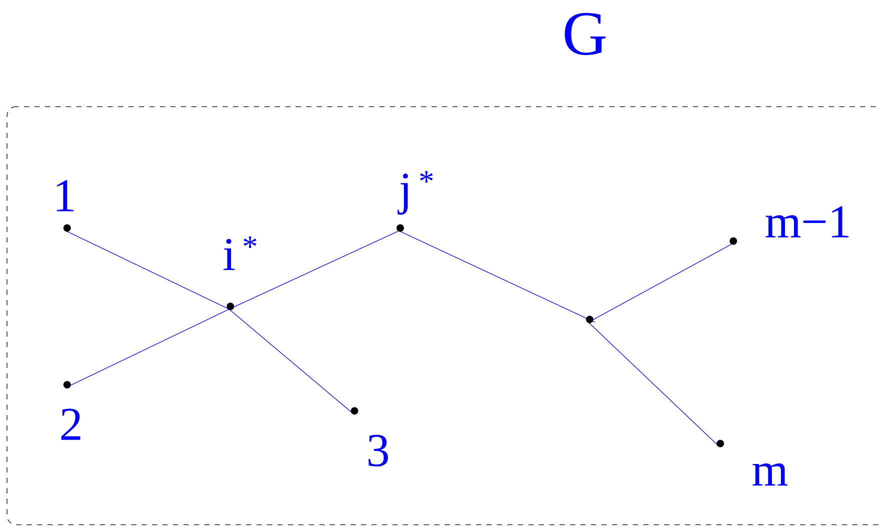


$$P_{X_i|\circ} = P_{X_i|X_j}$$

- If G is a chain, concept reduces to that of a standard Markov chain.
- $X_1, \cdots, X_m$ are $\{0, 1\}$-valued rvs with joint pmf $P_{X_1 \cdots X_m}$ satisfying:

  for $(i, j) \in E$: the rvs X_i, X_j are "symmetrically" connected by a *virtual* BSC(p_{ij}), $p_{ij} < \frac{1}{2}$.

Model 2: Markov Chain on a Tree



I. Csiszár - P.N., '04

$$
\begin{aligned}
C_{SK} &= \min_{(i,j)\in E} I(X_i \wedge X_j) \\
&= I(X_{i^*} \wedge X_{j^*}) = 1 - h_b(p_{max}) \text{ bit/symbol,}
\end{aligned}
$$

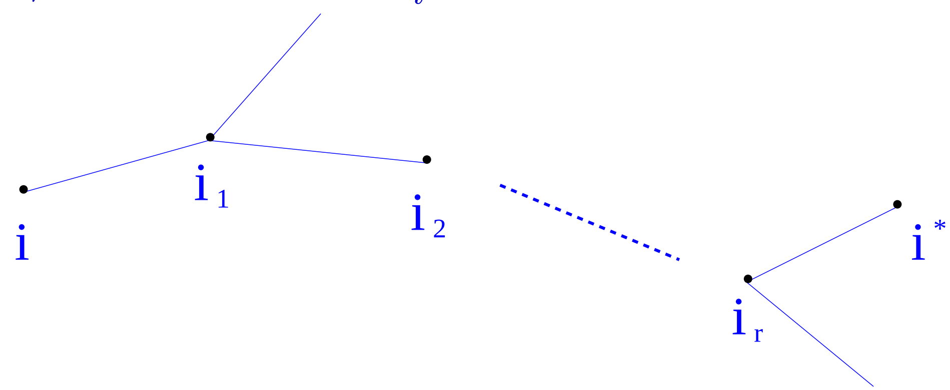
where

$$p_{max} \stackrel{\triangle}{=} p_{i^*j^*} = \max_{(i,j)\in E} p_{ij}.$$

Step 1: Slepian-Wolf Scheme for Reconstructing $x_{i^*}^n$ at All Terminals

C. Ye - P.N., '05

- Consider a linear code of blocklength n with parity check matrix $\mathbf{P}$ for the BSC $P_{X_{i^*}|X_{j^*}} = \text{BSC}(p_{max})$, of rate $\cong 1 - h_b(p_{max})$ and small decoding error probability.
- Each terminal $i = 1, \cdots, m$ transmits the syndrome $\mathbf{P}x_i^n$.
- Each terminal $i \neq i^*$ reconstructs $x_{i^*}^n$ as follows:



$$
\begin{aligned}
\widehat{x}_{i_1}^n &= x_i^n \oplus f_{\mathbf{P}}(\mathbf{P}x_i^n \oplus \mathbf{P}x_{i_1}^n) \\
\widehat{x}_{i_2}^n &= \widehat{x}_{i_1}^n \oplus f_{\mathbf{P}}(\mathbf{P}x_{i_1}^n \oplus \mathbf{P}x_{i_2}^n) \\
&\vdots \\
\widehat{x}_{i^*}^n &= \widehat{x}_{i_r}^n \oplus f_{\mathbf{P}}(\mathbf{P}x_{i_r}^n \oplus \mathbf{P}x_{i^*}^n)
\end{aligned}
$$

- Can show: $\Pr\{\widehat{X}_{i^*}^n = X_{i^*}^n \text{ at every terminal}\} \cong 1$.

Step 2: Secret Key Construction

- Secret key for terminals $1, \cdots, m$

  Terminal i^* sets K_{i^*} = numerical index of $x_{i^*}^n$ in coset containing $x_{i^*}^n$.

  Each terminal $i \neq i^*$ sets K_i = numerical index of its estimate $\widehat{x}_{i^*}^n$ in coset containing $x_{i^*}^n$.

- Any of $K_1, \cdots, K_m$ forms an optimal secret key, since:
  - $\Pr\{K_1 = \cdots = K_m\} \cong 1$; (common randomness)
  - $I(K_1 \wedge \mathbf{F}) = 0$; (secrecy)
  - $H(K_1) = \log$ (cardinality of key space); (uniformity)
  - $\frac{1}{n} H(K_1) \cong 1 - h_b(p_{max})$. (SK-capacity)

Open Problems and Work in Progress

- Model with eavesdropper possessing wiretapped side information.
- Secret key constructions
  - "Good" Slepian-Wolf data compression codes for terminals with arbitrarily correlated signals??
  - Secret key extraction techniques ..... (S. Nitinawarat)

⋮

- Models for the *simultaneous* generation of *multiple* secret keys. (C. Ye)
- Models with real-valued signals observed by the terminals. (S. Nitinawarat)