MULTITERMINAL SECRECY AND TREE PACKING

**With Imre Csiszár, Sirin Nitinawarat, Chunxuan Ye,
Alexander Barg and Alex Reznik**

Information Theoretic Security

- A complementary approach to `computational security` for *secret key* cryptosystems.
- *Unconditional Security*: A quantifiable and provable notion of security, with no assumption of "one-way" functions and no restrictions on the computational power of an adversary.
- Information theoretic perfect secrecy: Shannon, 1949.
- A modified notion: Maurer (1990, 1993), Ahlswede-Csiszár (1993)
  - an adversary does not have access to precisely the same observations as the legitimate users;
  - the legitimate plaintext messages and secret keys are, in effect, "nearly statistically independent" of the observations of the adversary.
- ? New insights: Innate connections with multiterminal data compression and points of contact with combinatorial tree packing algorithms.
- ??? New algorithms: Potential rests on advances in algorithms for multiterminal data compression and a better understanding of connections with combinatorial tree packing of multigraphs.

SECRET KEY GENERATION

Multiterminal Source Model

- The terminals in $\mathcal{M} = \{1, \ldots, m\}$ 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 in a given *subset* $A \subseteq \mathcal{M}$ wish to generate a "secret key" with the cooperation of the remaining terminals, to which end *all* the terminals can communicate among themselves – possibly interactively in multiple rounds – over a public noiseless channel of unlimited capacity.

- A secret key:

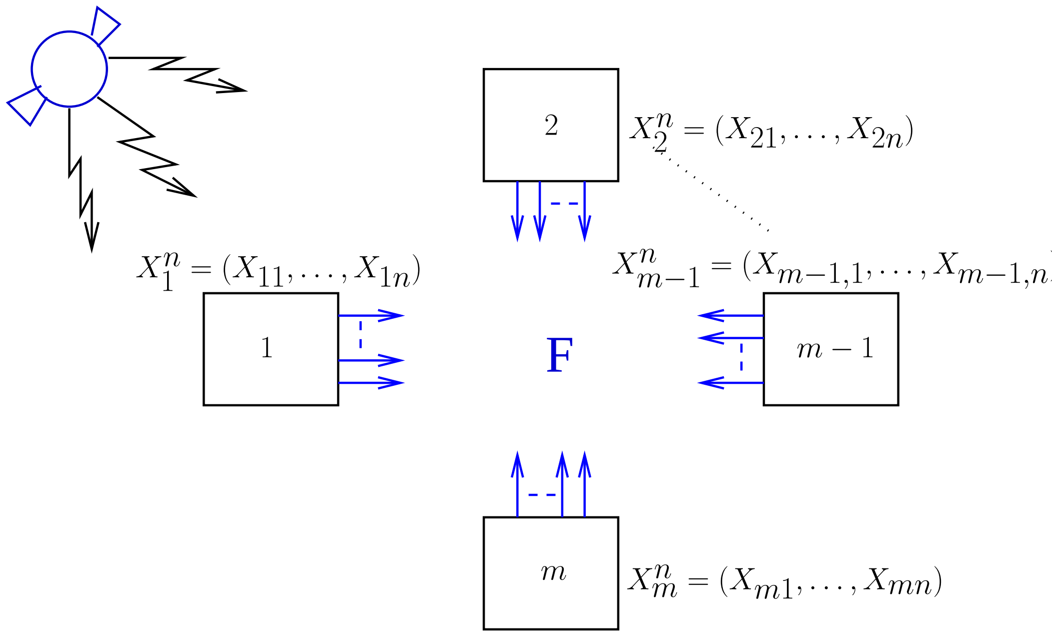
  - random variables (rvs) generated at each terminal in A which agree with probability $\cong 1$; and

  - the rvs are effectively concealed from an eavesdropper with access to the public communication.

- The key generation exploits the correlated nature of the observed signals.

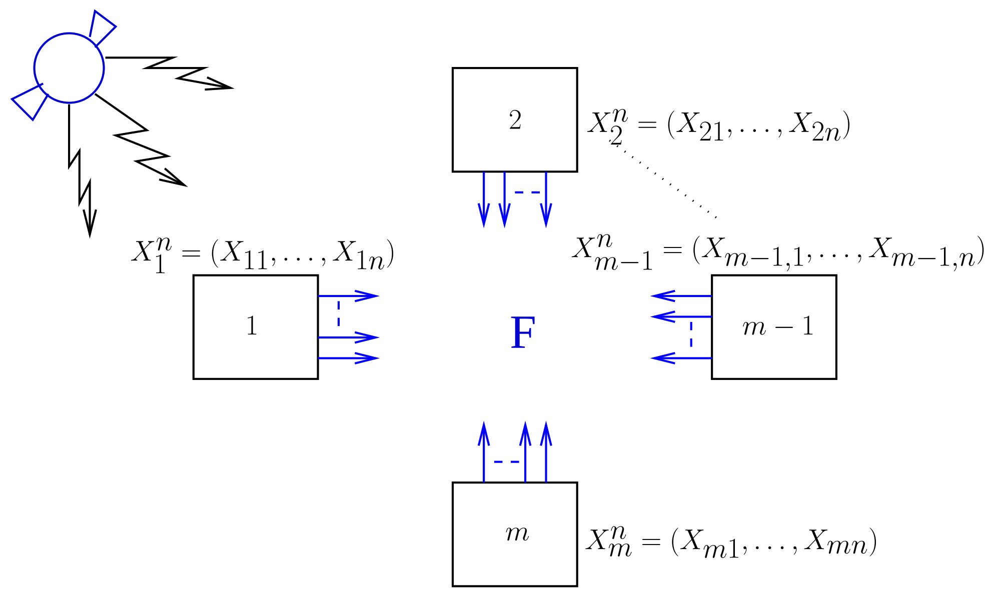
- The secret key thereby generated can be used as a *one-time pad* for secure encrypted communication among the terminals in A.

Multiterminal Source Model



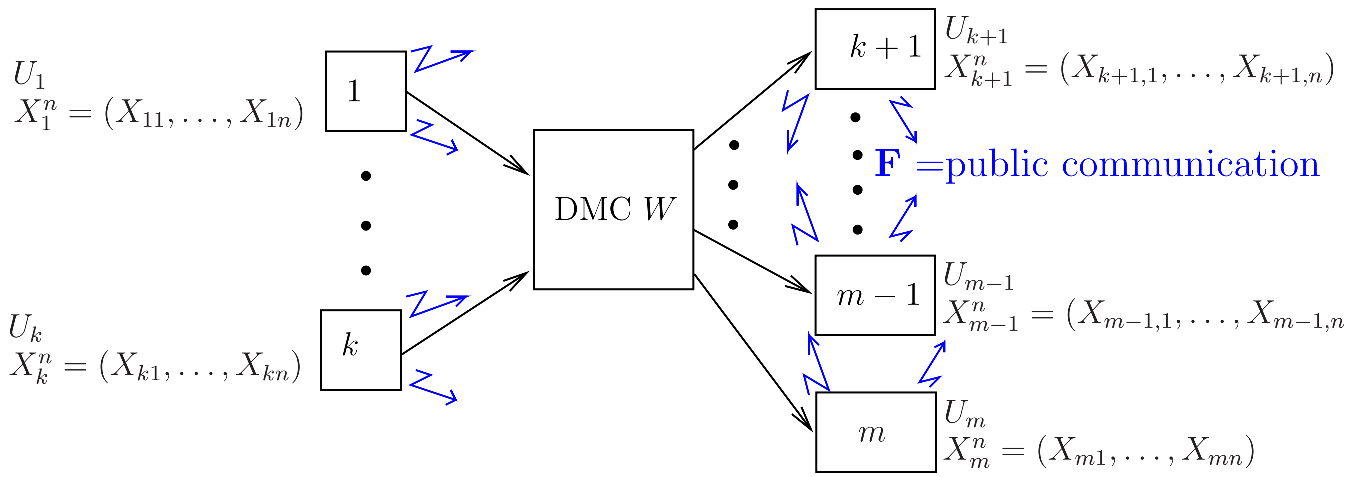
- The m legitimate terminals in $\mathcal{M} = \{1, \ldots, m\}$ cooperate in secret key generation.
- $X_1, \ldots, X_m$ are finite-valued random variables (rvs) with (known) joint distribution $P_{X_1,\ldots,X_m}$.
- Each terminal $i,\ i = 1, \ldots, m$, observes a signal comprising n independent and identically distributed repetitions (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 successive time instants are i.i.d. according to $P_{X_1,\ldots,X_m}$.

Multiterminal Source Model



- *All* the terminals are allowed to communicate over a *noiseless* channel of *unlimited capacity*, possibly interactively in several rounds.
- The communication from any terminal is observed by all the other terminals.
- The communication from a terminal is allowed to be any function of its own signal, and of all previous communication.
- Let **F** denote collectively all the communication.

Multiterminal Channel Model

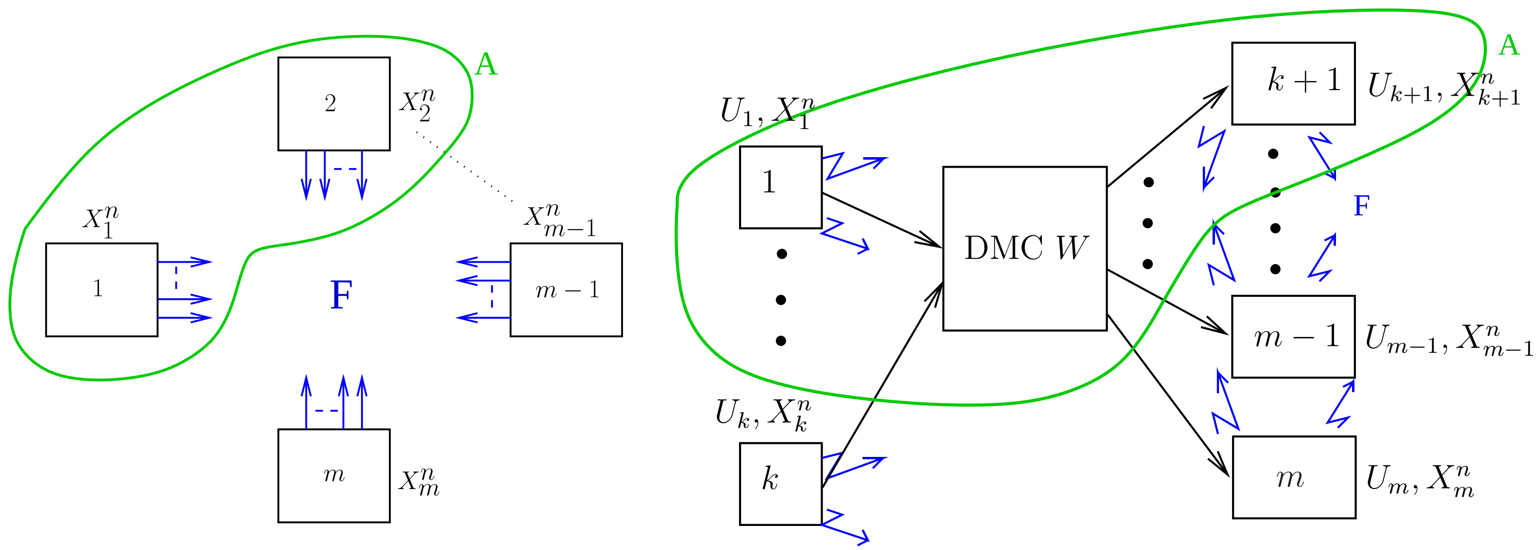


- Terminals $1, \ldots, k$ govern the inputs of a *secure* discrete memoryless channel W, with input terminal i transmitting a signal $X_i^n = (X_{i1}, \ldots, X_{in})$ of length n. Terminals $k+1, \ldots, m$ observe the corresponding output signals, with output terminal i observing X_i^n of length n.

- Following each simultaneous transmission of symbols over the channel W, communication over a *public noiseless channel of unlimited capacity* is allowed between all the terminals, perhaps interactively, and observed by all the terminals. Let $\mathbf{F}$ denote collectively all such public communication.

- Randomization at the terminals is permitted, and is modeled by the rvs $U_i,\ i = 1, \ldots, m$, which are taken to be mutually independent.

The Objective



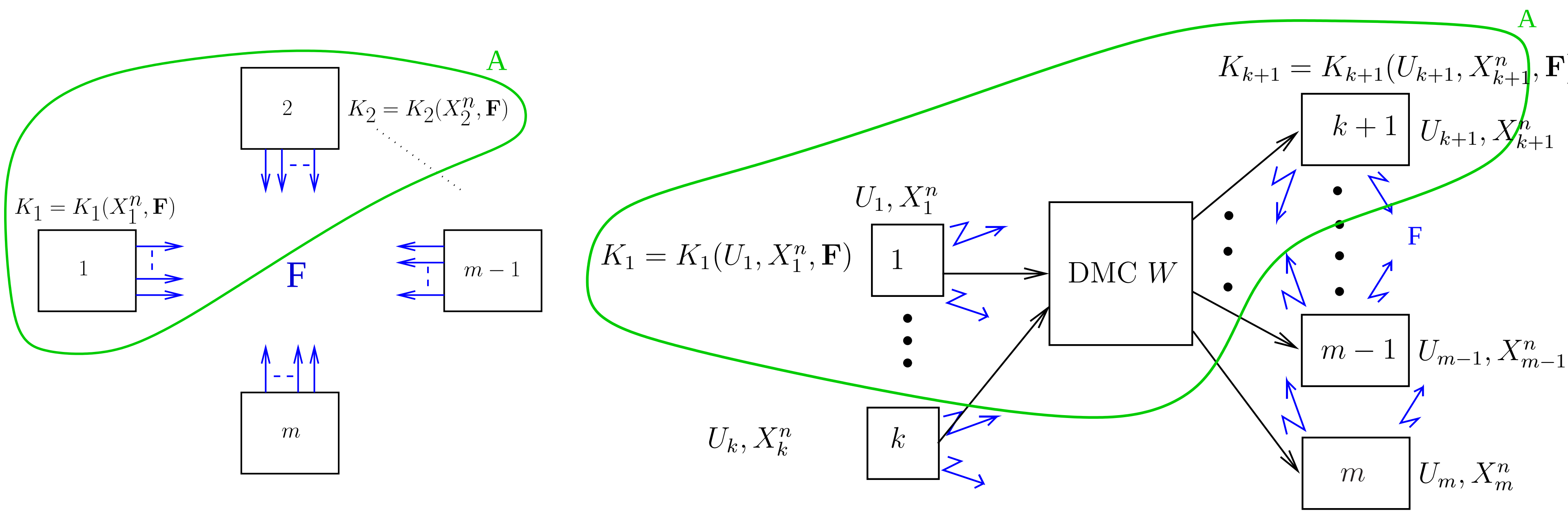
Objective: To generate a *secret key* of the largest "size" for a given set $A \subseteq \{1, \ldots, m\}$ of terminals, i.e., *common randomness* shared by the terminals in A, which is

- of near uniform distribution;
- concealed from an eavesdropper that observes the public communication **F**.

All the terminals $1, \ldots, m$ cooperate in achieving this goal.

Assume: The eavesdropper is passive and cannot wiretap.

What is a Secret Key?

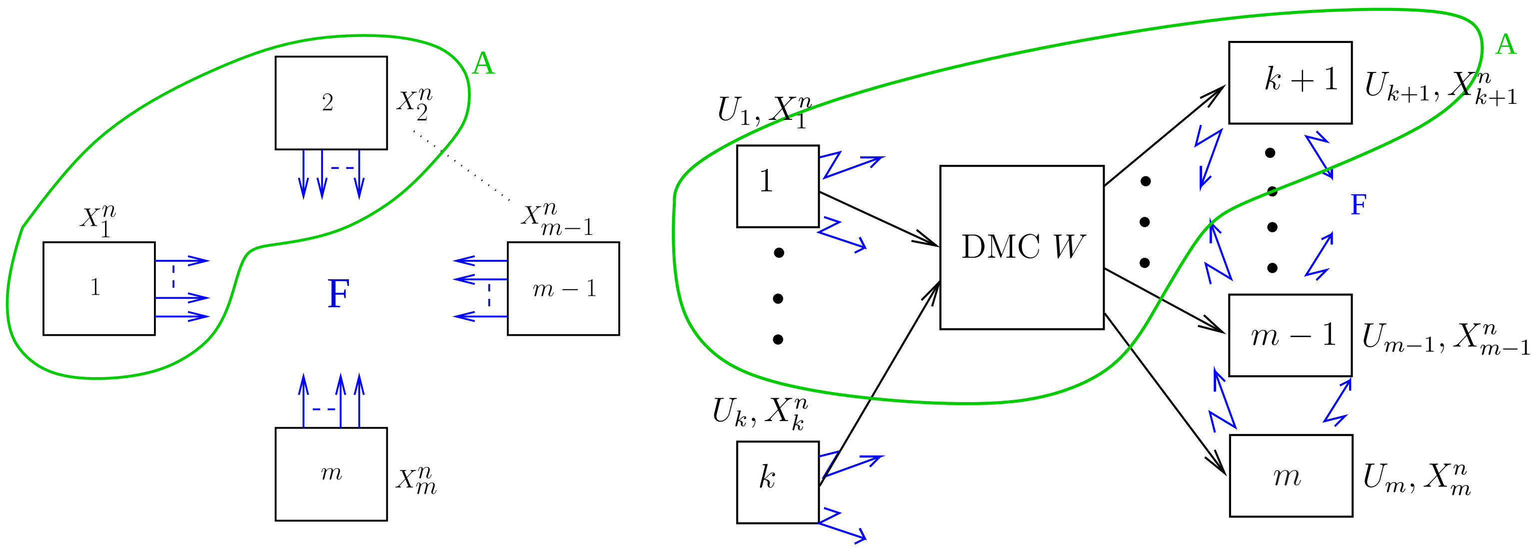


Secret Key (SK): A random variable K is a *SK* for the terminals in A, achievable with communication $\mathbf{F}$, if

- $Pr\{K = K_i,\ i \in A\} \cong 1$ (“common randomness”)
- $I(K \wedge \mathbf{F}) \cong 0$ (“secrecy”)
- $H(K) \cong \log |\text{key space}|$. (“uniformity”)

Thus, a SK is effectively concealed from an eavesdropper with access to **F**, and is nearly uniformly distributed.

Secret Key Capacity

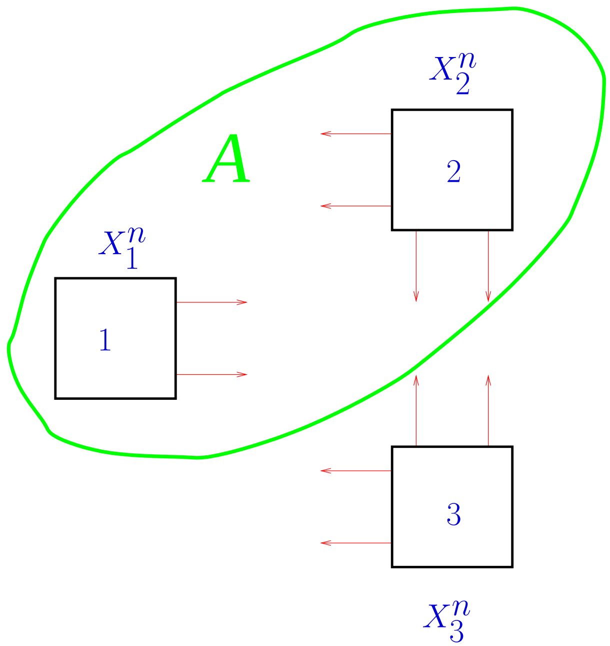


?? What is the *largest rate* $\lim_n \frac{1}{n} \log |\text{key space}|$ of such a SK for A which can be achieved with suitable communication: SK capacity $C_S(A)$?

?? How to construct such a SK?

Hereafter, we shall restrict ourselves to the multiterminal source model.

A Toy Example: $\mathcal{M} = \{1, 2, 3\}, \;\; A = \{1, 2\}$



- X_1^n, X_2^n are $\{0, 1\}$-valued Bernoulli $(\frac{1}{2})$ sequences, and X_1^n is independent of X_2^n.

$$X_{3t} = X_{1t} \oplus X_{2t}, \qquad t = 1, \ldots, n.$$

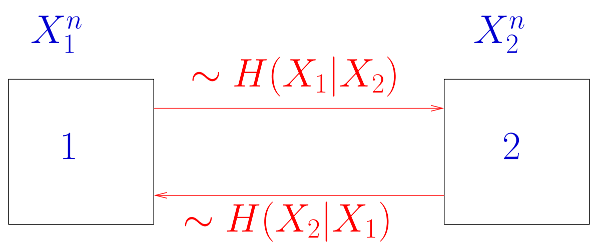
- *Scheme (using $n = 1$):* Terminal 3 communicates publicly $X_{31} = X_{11} \oplus X_{21}$.
- Terminals 1 and 2 respectively infer X_{21} and X_{11}.
- X_{11} is independent of $\mathbf{F} = X_{31} = X_{11} \oplus X_{21}$, and is uniform on $\{0, 1\}$.
- Thus, X_{11} is a *perfect* SK of rate 1.0 (optimal), so that SK capacity $C_S(A) = 1.0$.

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, 2000, 2003, $\cdots$
- Venkatesan - Anantharam 1995, 1997, 1998, 2000, $\cdots$
- Csiszár - Narayan 2000, 2004, 2005, 2007, 2008, 2009.
- Renner - Wolf 2003.
- Muramatsu 2004, 2005.
- Ye - Narayan 2004, 2005.
- Gohari-Anantharam 2007, 2008.
- Ye-Reznik, 2007.
- Nitinawarat-Ye-Barg-Narayan-Reznik, 2008, 2009.

$\vdots$

Secret Key Capacity: $\mathcal{M} = \{1,2\} = A$



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

$$C_S(A) = I(X_1 \wedge X_2).$$

- **An interpretation:**

$$\begin{aligned} C_S(A) &= I(X_1 \wedge X_2) \\ &= H(X_1, X_2) - [H(X_1|X_2) + H(X_2|X_1)] \\ &= \text{Entropy rate of "omniscience"} \; - \\ & \quad \text{Smallest aggregate rate of communication, } R_{CO}(A), \\ & \quad \text{that enables the terminals in } A \text{ to become omniscient.} \end{aligned}$$

Secret Key Capacity

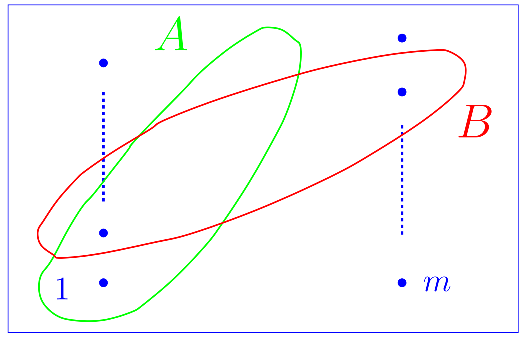
Theorem [I. Csiszár - P. N., '04, '08]:

$$\begin{aligned} C_S(A) &= H(X_1, \ldots, X_m) - \text{Smallest aggregate rate of overall} \\ &\quad \text{interterminal communication, } R_{CO}(A)\text{, that enables} \\ &\quad \textit{all} \text{ the terminals in } A \text{ to become omniscient} \\ &= H(X_1, \ldots, X_m) - \max_{\lambda \in \Lambda(A)} \sum_{B \in \mathcal{B}(A)} \lambda_B H\left(X_B | X_{B^c}\right) \end{aligned}$$

and can be achieved with noninteractive communication.

Remark: $R_{CO}(A)$ is obtained as the solution to a multiterminal data compression problem of omniscience generation that *does not involve any secrecy constraints*.

Interpretation: All the terminals cooperate – through public communication – in enabling the terminals in A to attain omniscience. Then the terminals in A extract a SK from their omniscience by purging the rate of this communication.



Proposition [I. Csiszár - P. N., '04]: The smallest aggregate rate of interterminal communication, $R_{CO}(A)$, that enables all the terminals in A to become omniscient, is

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

where

$$\mathcal{R}_{SW}(A) = \left\{ (R_1, \cdots, R_m) : \sum_{i\in B} R_i \geq H(X_B|X_{B^c}),\ \forall B \subset \mathcal{M},\ B \neq \emptyset,\ A \not\subseteq B \right\},$$

and can be achieved with noninteractive communication. Furthermore

$$R_{CO}(A) \ = \ \max_{\lambda\in\Lambda(A)} \sum_{B\in\mathcal{B}(A)} \lambda_B H\left(X_B|X_{B^c}\right).$$

How Can a Secret Key be Constructed?

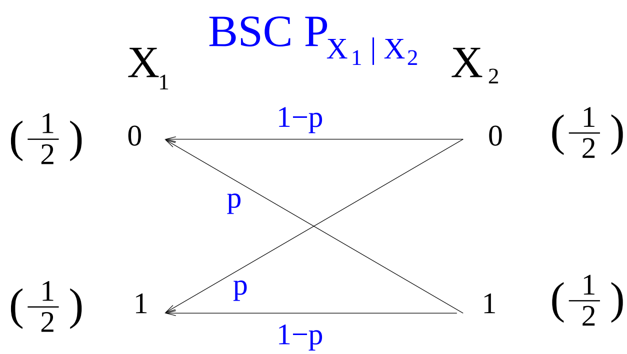
- **Step 1: Data compression:** The terminals communicate over the public channel using compressed data in order to generate omniscience or some form of "common randomness." This public communication is observed by the eavesdropper.

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

Example: 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 with

$$P_{X_1X_2}(x_1,x_2) = \frac{1}{2}(1-p)\delta_{x_1x_2} + \frac{1}{2}p\,(1-\delta_{x_1x_2}), \qquad p < \frac{1}{2}.$$



- $C_S(\{1,2\}) = I(X_1 \wedge X_2) = 1 - h_b(p)$ 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.

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 $\mathbf{P}$ for a channel with noise V^n:

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

- Terminal 1 communicates $\mathbf{F}$ = the syndrome $\mathbf{P}x_1^n$ 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

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

Step 2: Secret Key Construction

C. Ye - P.N., '05

- SK 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.

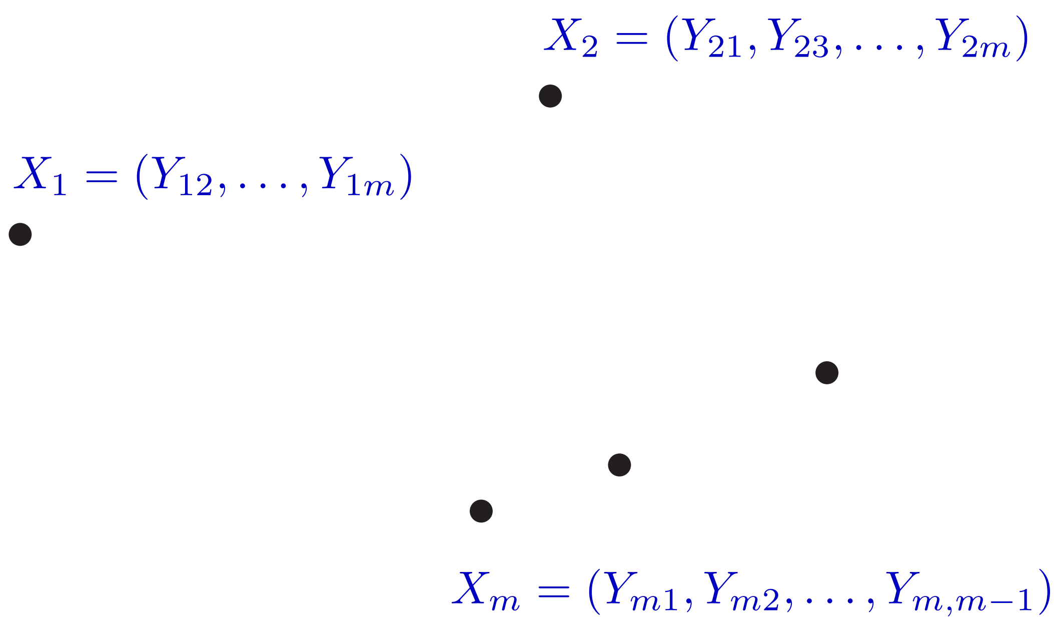
- For a systematic channel code: K_1 (resp. K_2) = first $(n-m)$ bits of x_1^n (resp. $\widehat{x}_1^n$).

- K_1 or K_2 forms an optimal rate SK, 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)

TREE PACKING

Pairwise Independent Network (PIN) Model



A special form of a multiterminal source model in which

- $X_i = (Y_{i\,j}, j \in \{1, \ldots, m\} \backslash \{i\}),\ i = 1, \ldots, m;$
- $Y_{i\,j}$ is correlated with $Y_{j\,i},\ \ 1 \le i \ne j \le m;$
- the pairs $\{(Y_{i\,j}, Y_{j\,i})\}$ are mutually independent across $1 \le i < j \le m.$

Secret Key Capacity for the PIN Model

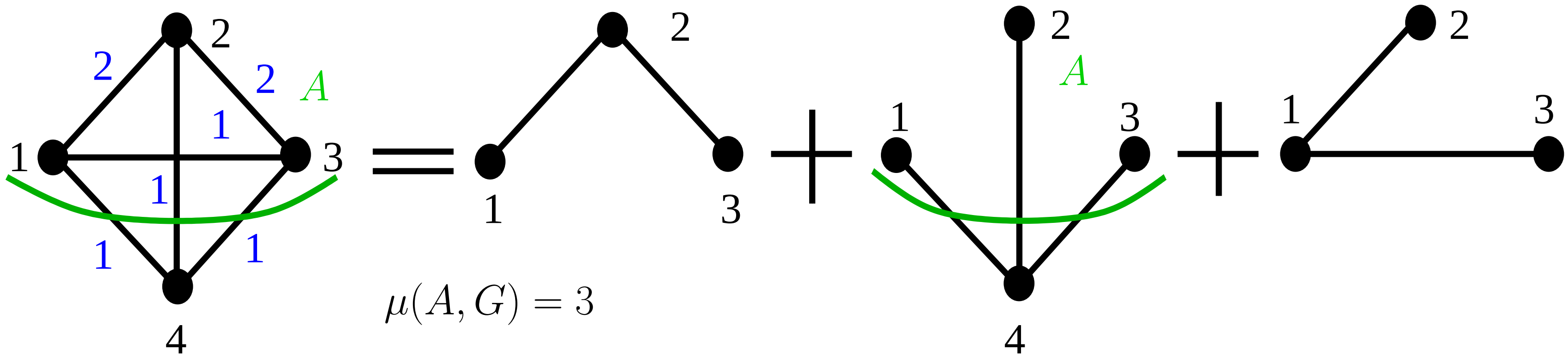
Proposition [Nitinawarat *et al*, '08]: For a PIN model, the SK capacity for a set of terminals $A \subseteq \mathcal{M} = \{1, \ldots, m\}$ is

$$C_S(A) = \min_{\lambda \in \Lambda(A)} \left[\sum_{1 \leq i < j \leq m} \left(\sum_{\substack{B \in \mathcal{B}(A): \\ i \in B,\ j \in B^c}} \lambda_B \right) I(Y_{ij} \wedge Y_{ji}) \right].$$

Remark: $C_S(A)$ depends on the underlying joint probability distribution only through a linear combination of $\{I(Y_{ij} \wedge Y_{ji})\}_{i \neq j}$, i.e., the *best pairwise SK rates*; the corresponding pairwise SKs are mutually independent.

?? Can a SK for the set of terminals A be formed by propagating *independent* and *locally generated pairwise* SKs, for instance, by some form of tree packing in an associated multigraph??

Steiner Tree Packing



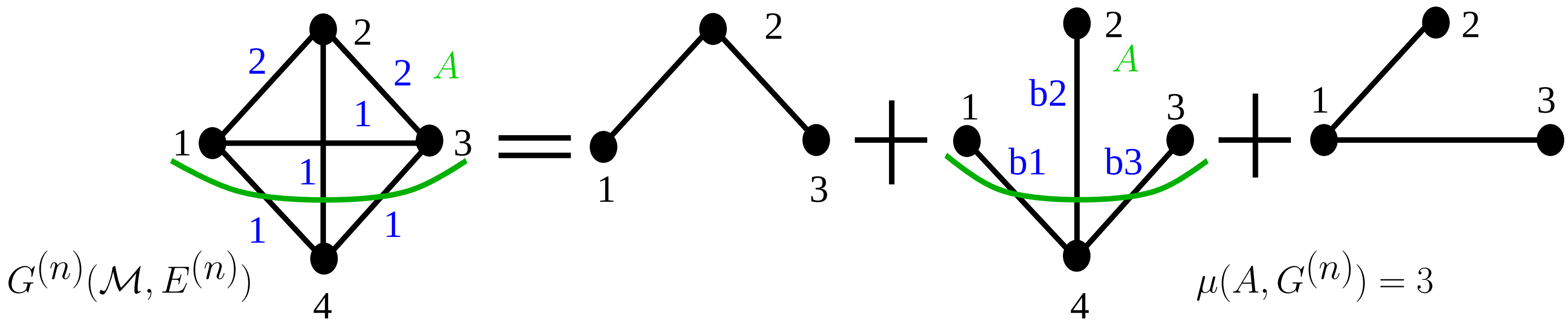
$G(\mathcal{M}, E)$ = multigraph with vertex set $\mathcal{M}$ and edge set E.

Definition

- For $A \subseteq \mathcal{M}$, a *Steiner tree* of G is a subgraph of G which is a tree and whose vertex set contains A.
- A *Steiner tree packing* of G is any collection of *edge*-disjoint Steiner trees of G. Let $\mu(A, G)$ denote the maximum size of such a packing.

How to Generate a Secret Key by Steiner Tree Packing?

- Given a PIN model, calculate $\{I(Y_{ij} \wedge Y_{ji})\}_{i \neq j}$.
- With the given PIN model, associate a multigraph $G^{(n)}(\mathcal{M}, E^{(n)})$ with vertex set $\mathcal{M} = \{1, \ldots, m\}$ and edge set $E^{(n)} = \{e_{ij}^{(n)} = nI(Y_{ij} \wedge Y_{ji})\}_{i \neq j}$.



- *Local SK generation*: For every pair of vertices $(i, j) \in G^{(n)}$, the terminals i, j generate a pairwise SK of size $nI(Y_{ij} \wedge Y_{ji})$ bits; these pairwise SKs are mutually independent.
- *SK propagation by Steiner tree packing*:
  - Claim: *Every Steiner tree corresponds to* 1 *bit of SK for the terminals in* A.
  - A Steiner packing of size p yields p SK bits shared by the terminals in A.

Remark: For m fixed, this algorithm can be implemented in linear time (in n).

Theorem [Nitinawarat *et al*, '08]: For a PIN model, the SK capacity satisfies

$$C_S(A) \geq \sup_n \frac{1}{n} \, \mu(A, G^{(n)}).$$

A consequence of independent interest: Given a multigraph $G = G^{(1)}$,

- the SK capacity of an associated PIN model with

$$I(Y_{ij} \wedge Y_{ji}) = e_{ij}, \quad 1 \leq i < j \leq m$$

  provides a new (information theoretic) upper bound for the maximum rate of Steiner tree packing $\sup_n \frac{1}{n} \, \mu(A, G^{(n)})$;

- this bound is tight when $|A| = 2$ and $|A| = m$ but can be loose otherwise.

When $A = \mathcal{M} = \{1, \ldots, m\}$, a Steiner tree becomes a spanning tree.

Theorem [Nitinawarat *et al*, '08]: For a PIN model,

$$C_S(\mathcal{M}) = \sup_n \frac{1}{n}\, \mu(\mathcal{M}, G^{(n)}).$$

Idea of proof: By a result of Nash-Williams and Tutte,

$$\sup_n \frac{1}{n}\, \mu(\mathcal{M}, G^{(n)}) = \min_{\mathcal{P}:\ \mathcal{P} \text{ a partition of } \mathcal{M}} \frac{1}{|\mathcal{P}| - 1} \Big(\text{No. of edges of } G^{(1)} \text{that cross } \mathcal{P}\Big),$$

which coincides with an upper bound for $C_S(\mathcal{M})$ in [I. Csiszár-P.N., '04].

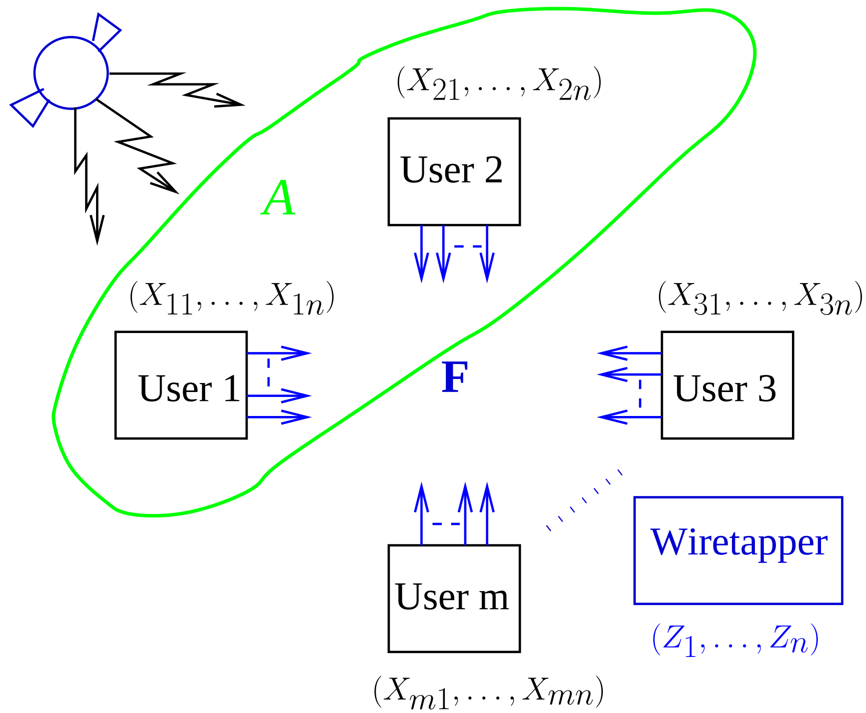
Remarks:

(i) Thus, maximal spanning tree packing attains the SK capacity $C_S(\mathcal{M})$.

(ii) There exists a polynomial-time algorithm (in both m, n) for finding a maximal collection of edge-disjoint spanning trees for $G^{(n)}$ [Gabor-Westermann] and forming an optimal rate SK.

VARIANT MODELS FOR SECRET KEY GENERATION

Multiterminal Source Model with Wiretapper



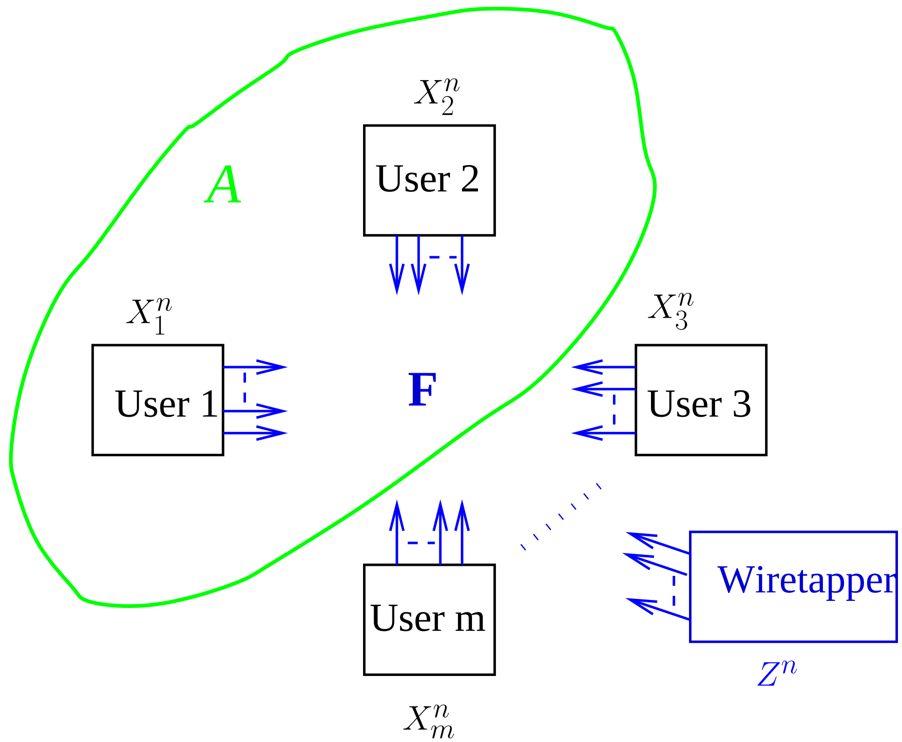
The legitimate user terminals in A wish to generate a secret key K with the cooperation of the remaining legitimate terminals, which is concealed from an eavesdropper with access to the public interterminal communication $\mathbf{F}$ and wiretapped side information $Z^n = (Z_1, \ldots, Z_n)$.

The *secrecy condition* is now strengthened to

$$I(K \wedge \mathbf{F}, Z^n) \cong 0.$$

??? Largest rate of a *wiretap secret key* for A: Unknown in general but for special cases and bounds.

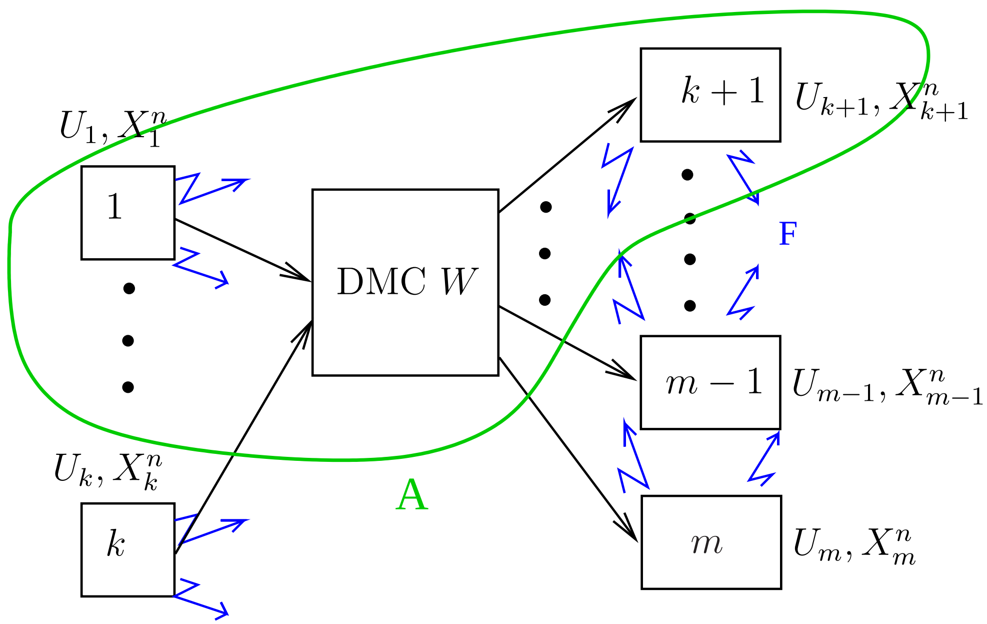
Multiterminal Source Model with Cooperative Wiretapper: Private Key



The wiretapped terminal *cooperates* in the secrecy generation by "revealing" its observations to all the legitimate terminals; the resulting key must be concealed from the eavesdropper which knows $(\mathbf{F}, Z^n)$.

??? Largest rate of a *private key* for A: Known.

Multiterminal Channel Model



??? Largest rate of a *secret key* for A: Unknown in general but for special cases and bounds.

IN CLOSING

A Few Questions

- Information theoretic secrecy generation in a network is intertwined with multiterminal data compression and channel coding for certain network models.
  - What are the *explicit connections* for general network models?
  - What are the corresponding best rates of secret keys?
  - New algorithms for secret key construction?
- Multiuser secrecy generation for the PIN model has connections to the combinatorial problem of tree packing in multigraphs.
  - Tree packing algorithms for global secret key generation?
  - Information theoretic tools for tackling combinatorial tree packing problems?

Idea of Proof of SK Capacity Theorem

Achievability

If L represents "common randomness" for all the terminals in A, achievable with communication F for some (signal) observation length n, then $\frac{1}{n}H(L|\mathbf{F})$ is an achievable SK rate for the terminals in A.

- The terminals communicate publicly using compressed data in order to generate common randomness for the terminals in A equalling
  $L \cong$ omniscience $= (X_1^n, \ldots, X_m^n)$, with $\mathbf{F} = \mathbf{F}_{CO} = \mathbf{F}_{CO}(X_1^n, \ldots, X_m^n)$.
- The terminals in A then process this L to extract a SK of rate

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

  and of which the eavesdropper has provably little or no knowledge.

Converse

Tricky, since interactive communication is not excluded a priori.

Decomposition interpretation:

$$\text{Omniscience} = (X_1^n, \ldots, X_m^n) \cong (\text{Optimum secret key for A}, \ \mathbf{F}_{CO}).$$

Private Key Capacity

Theorem [I. Csiszár - P. N., '04, '08]:

$$\begin{aligned} C_P(A|Z) &= H(X_1, \ldots, X_m, Z) - H(Z) - \text{ Smallest aggregate rate of} \\ &\quad \text{public communication which enables the terminals in } A \text{ to} \\ &\quad \text{become omniscient when all terminals additionally know } Z^n \\ &= H(X_1, \ldots, X_m | Z) - \max_{\lambda \in \Lambda(A|Z)} \sum_{B \in \mathcal{B}(A|Z)} \lambda_B H\left(X_B | X_{B^c}, Z\right) \end{aligned}$$

and can be achieved with noninteractive communication.

Remarks:

- Clearly, WSK capacity $C_W(A|Z) \le$ PK capacity $C_P(A|Z)$ with equality in special cases.
- Better upper bounds on WSK capacity are available due to Renner-Wolf ('03) and Gohari-Anantharam ('07, '08).

Achievability:

- The terminals in A generate common randomness L such that

$$(L, Z^n) \cong (\text{omniscience}, Z^n) = (X_1^n, \ldots, X_m^n, Z^n),$$

  using public interterminal communication $\mathbf{F}_{CO} = \mathbf{F}_{CO}(X_1^n, \ldots, X_m^n, Z^n)$ that is independent of Z^n.

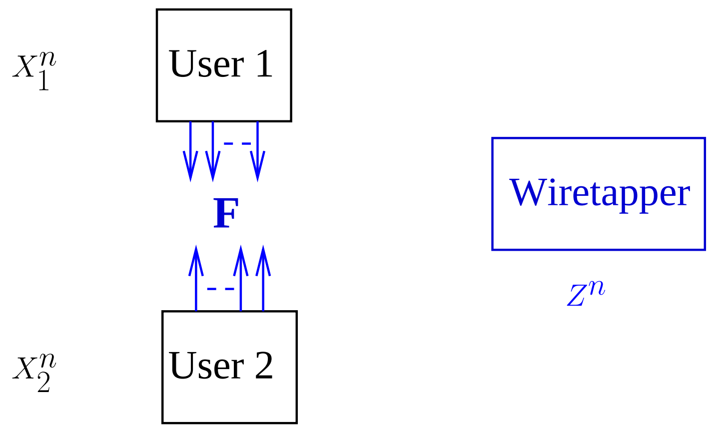
- The terminals in A then extract secrecy of rate

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

Decomposition interpretation:

$$(X_1^n, \ldots, X_m^n, Z^n) \;\cong\; \left(\text{Optimum private key for A},\; \mathbf{F}_{CO},\; Z^n\right).$$

Open Problem: The General Wiretapper Model with $\mathcal{M} = \{1,2\} = A$



Gohari-Anantharam, '07, '08

- Terminals $1, 2$ generate common randomness L using public interterminal communication $\mathbf{F} = \mathbf{F}(X_1^n, X_2^n)$, such that

$$(L, Z^n) \cong (\text{omniscience}, Z^n) = (X_1^n, X_2^n, Z^n).$$

  Note that that $\mathbf{F}$ is *not* a function of Z^n.

- The "nonsingle-letter" characterization of WSK capacity is

$$C_W(A|Z) = \lim_n \max_{L,\mathbf{F}} \frac{1}{n} H(L|\mathbf{F}, Z^n) = \cdots = H(X_1, X_2|Z) - \lim_n \min_{\mathbf{F}} \frac{1}{n} H(\mathbf{F}|Z^n).$$

Question: If L is the *Slepian-Wolf codeword* for the *joint source* (X_1^n, X_2^n) with "decoder side information" Z^n, what is $\lim_n \min_{\mathbf{F}} \frac{1}{n} H(\mathbf{F}|Z^n)$ where $\mathbf{F}$ is the interterminal communication needed to form L by distributed processing?