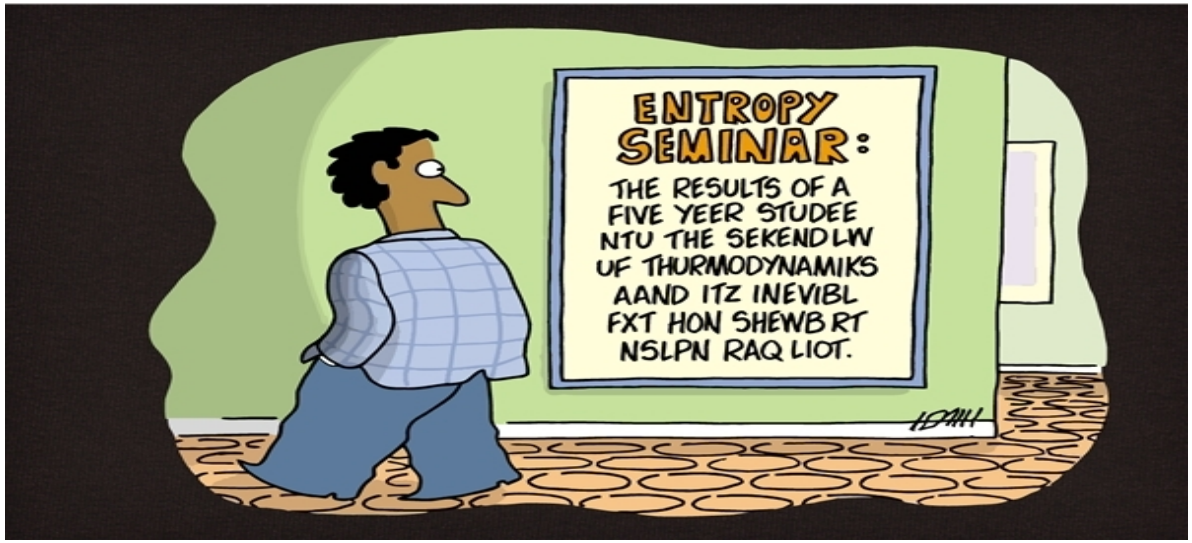
Codes, entropies and groups

Terence Chan†

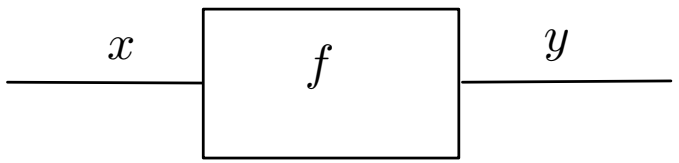
† ITR, University of South Australia



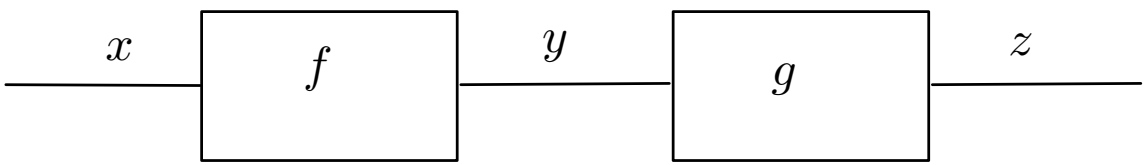
Tools

Tools (Normal Factor Graphs[1])

- Edges - variables (e.g., $x, y \ldots$)
- Boxes - functions (e.g., $f, g \ldots$)
- SUM over internal variables



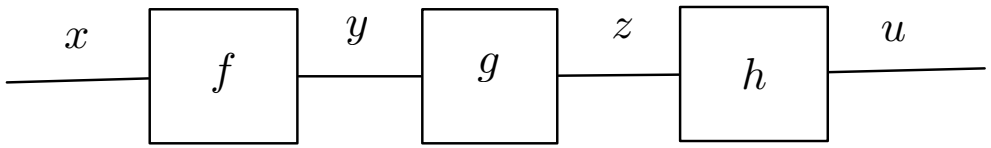
$$f(x, y)$$



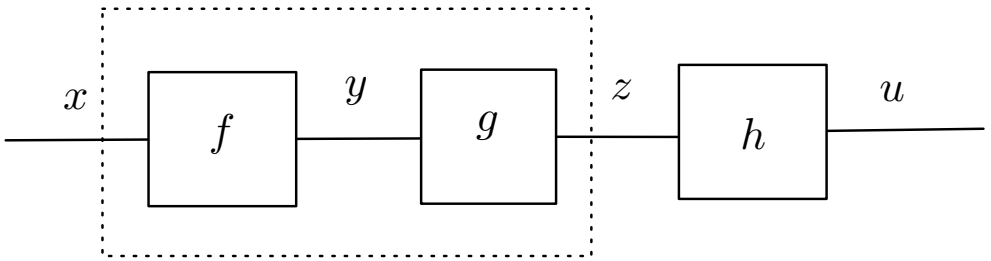
$$\sum_{y} f(x, y) g(y, z)$$

[1] See works by Forney, Al-Bashabsheh, Mao, et al.

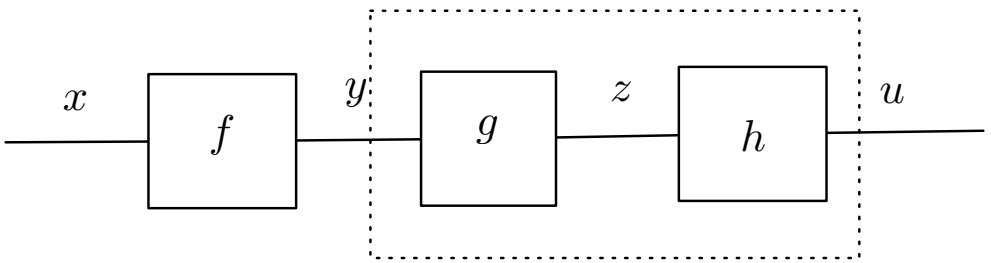
Distributive Law



$$\sum_{y,z} f(x,y)g(y,z)h(z,u)$$

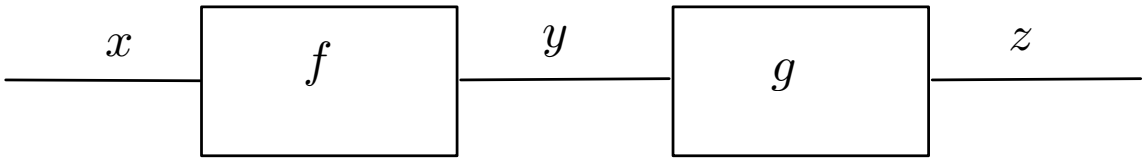


$$\sum_{z}\left(\sum_{y} f(x,y)g(y,z)\right)h(z,u)$$

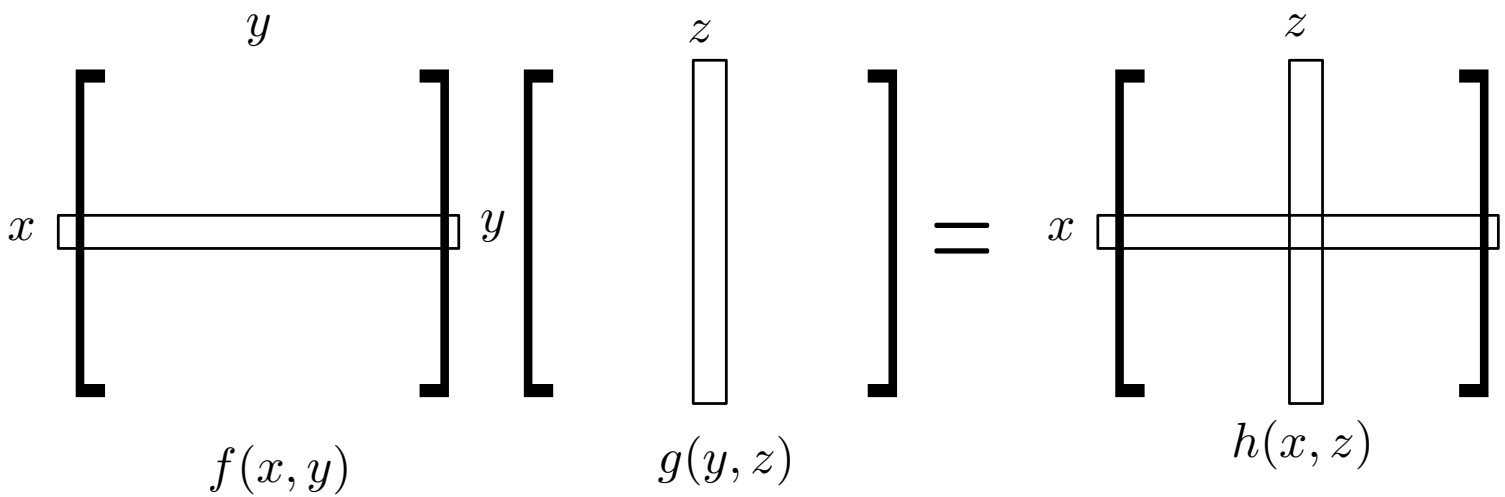


$$\sum_{y} f(x,y)\left(\sum_{z} g(y,z)h(z,u)\right)$$

Matrix multiplication



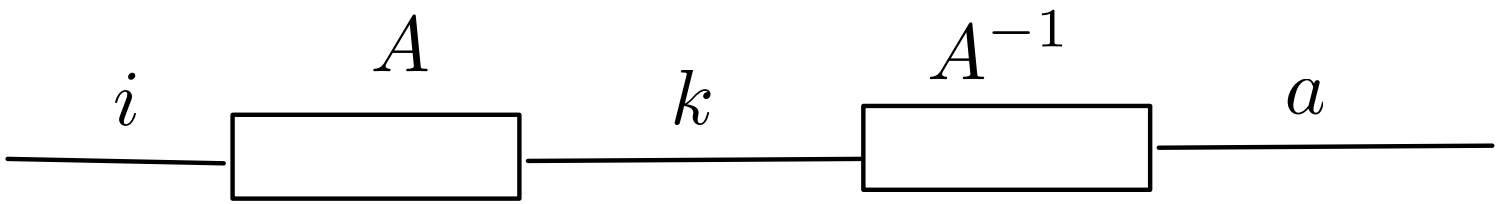
$$h(x,z) = \sum_y f(x,y)g(y,z)$$

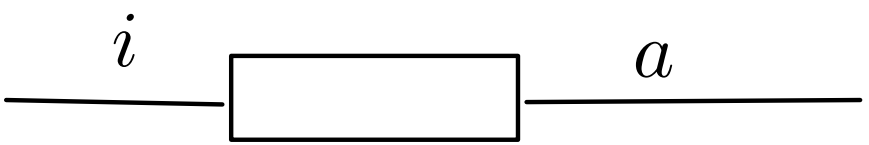


- For identify matrix:

$$h(x,z) = \begin{cases} 1 & \text{if } x = z \\ 0 & \text{otherwise.} \end{cases}$$

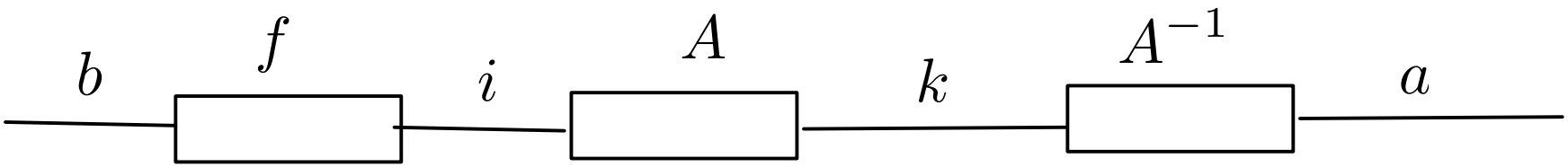
Matrix multiplication

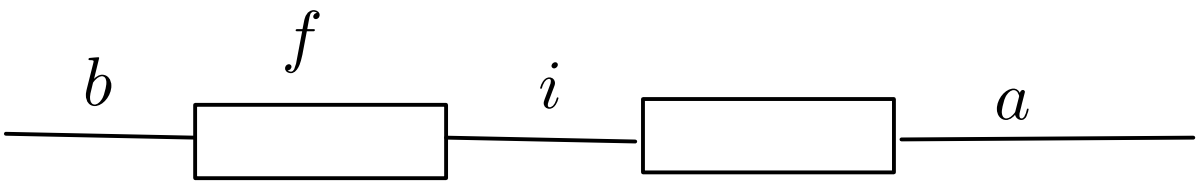




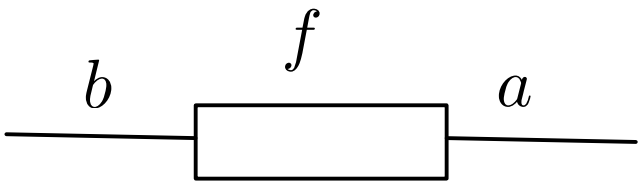
$$[i = a]$$

Matrix multiplication



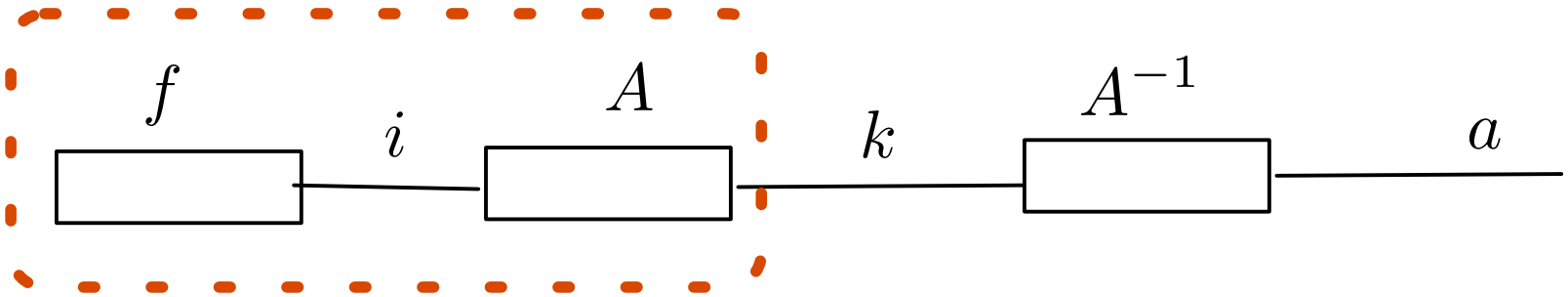


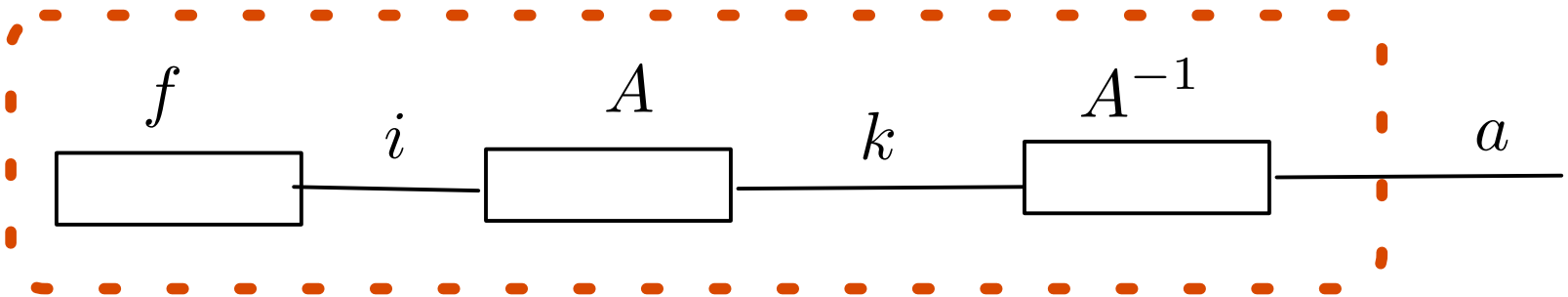
$[i = a]$



Transform

$$f(i) \Rightarrow \hat{f}(k)$$



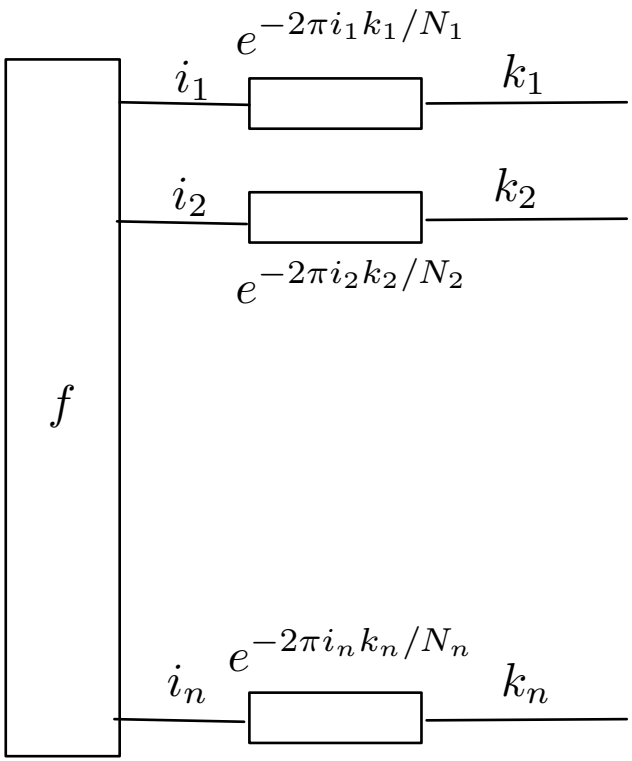


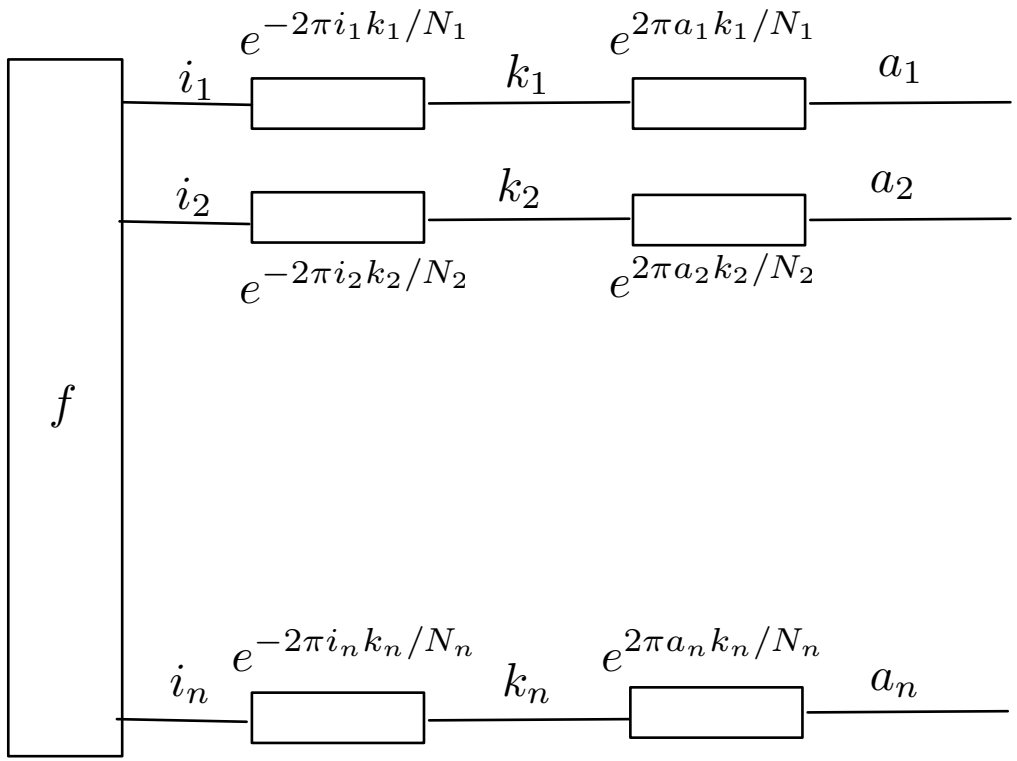
$$\hat{f}(k) \Rightarrow f(a)$$

Example: Fourier Transform

- i_j and k_j in the set $\{1, \ldots, N_j\}$

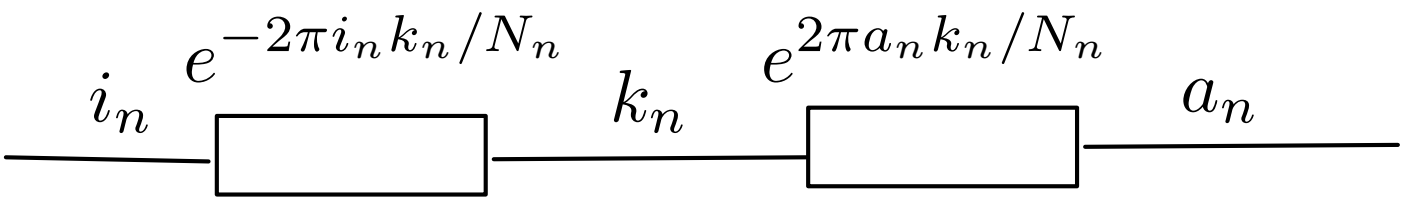
$$\hat{f}(k_1, \ldots, k_n) = \sum_{i_1, \ldots, i_n} f(i_1, \ldots, i_n) \prod_{j=1}^{n} e^{2\pi a_j k_j / N_j}$$

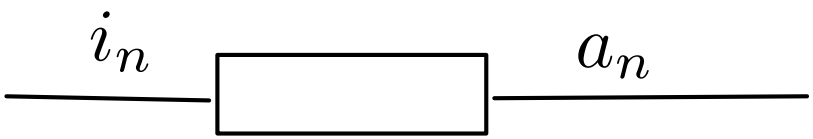




$$f(a_1, \ldots, a_n) = \sum_{k_1, \ldots, k_n} \hat{f}(k_1, \ldots, k_n) \prod_{j=1}^{n} e^{2\pi a_j k_j / N_j}$$

Example: Fourier Transform



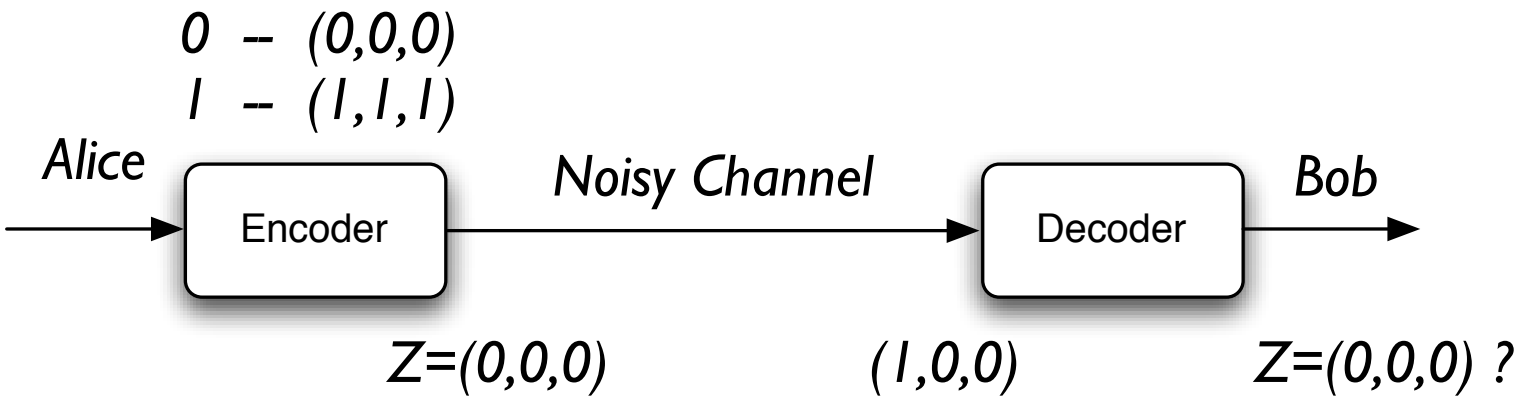


$$\sum_{k_n} e^{-2\pi i_n k_n / N_n} e^{2\pi a_n k_n / N_n} = [i_n = a_n \, ?]$$

Codes

What are codes?

- (Error control) coding is a technique to protect transmitted data against errors



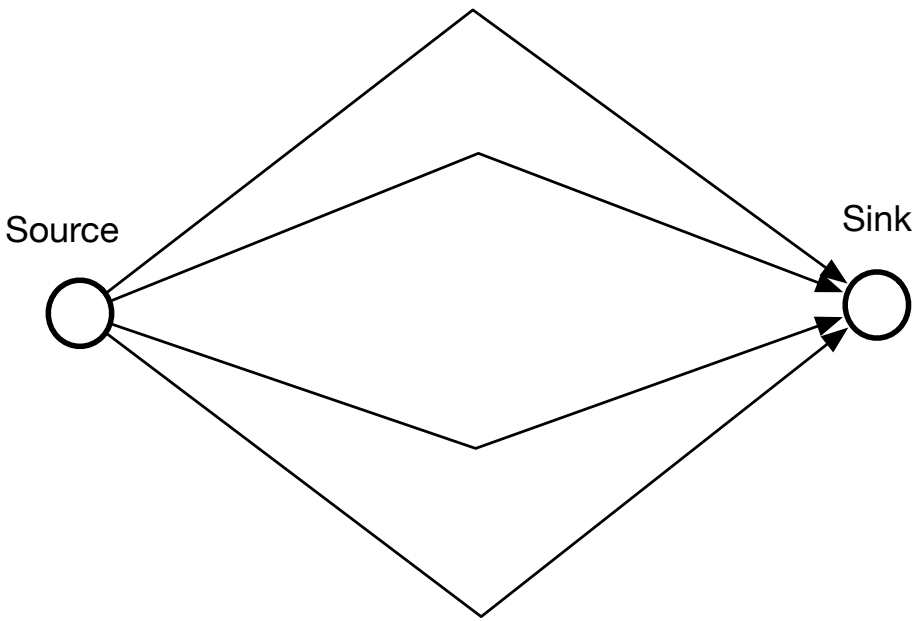
- Codebook size vs. Error correcting capability

$$\text{Error probability} = \Pr(\text{more than one symbol error})$$

- Code – a set of random variables $(Z_1, \ldots, Z_n)$
  - Z_i – the i^{th} codeword symbol.

Tamper-proof transmission

- Transmitter and receiver connected via n parallel links
- Adversary – obstruct data transmission
  - Replacing the messages transmitted on the attacked links with any other messages.
  - Message transmitted on untampered link received without error.



Tamper-proof transmission

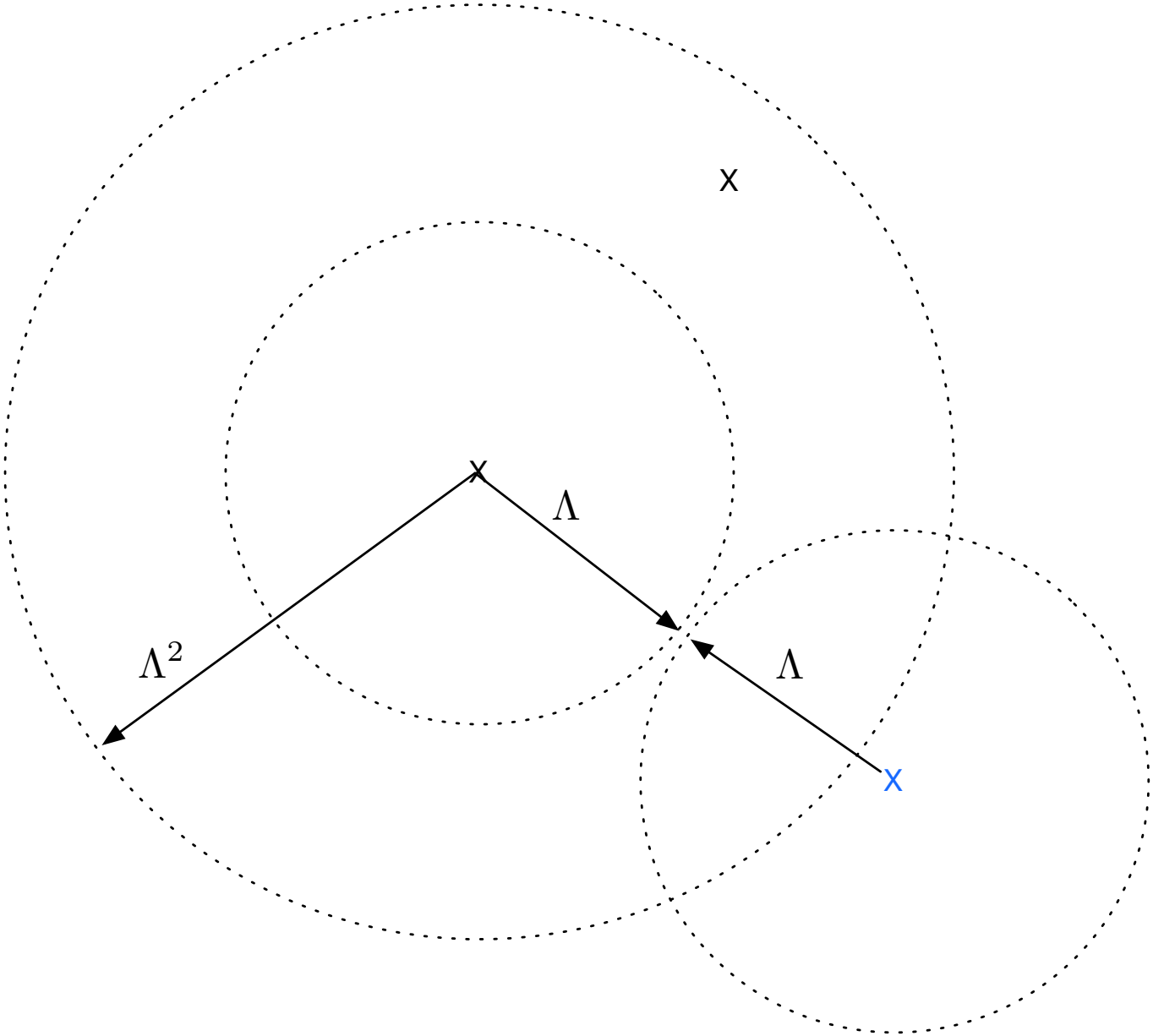
Find the highest rate code that resilient to attacks:

- Adversary's *tampering pattern* Λ – the possible link subsets that an adversary can attack.
- If the adversary can attack up to any t links, then Λ contains all subsets of sizes up to t.
- Codebook size - $H(Z_1, \ldots, Z_n)$
- Code is resilient if

$$H(Z_1, \ldots, Z_n | Z_i, i \in \alpha^c) = 0$$

for all $\alpha \subseteq \Lambda^2$ where $\Lambda^2 \triangleq \{\mathcal{B} \cup \mathcal{C} : \mathcal{B}, \mathcal{C} \in \Lambda\}$.

Tamper-proof transmission



Distributed Storage

- Data encoded into n pieces $Z_1, \ldots, Z_n$,
- each stored in a data centre (DC)
- In case of data centre failures, the stored data can be restored from other DC
- Ξ – failure pattern,
- Design a storage code such that data can be restored if a set $\mathcal{A} \in \Xi$ of data centres fail.

Distributed Storage

Find the most efficient storage code (resilient to failures)

- Code size – $H(Z_1, \ldots, Z_n)$
- Robustness if

$$H(Z_1, \ldots, Z_n | Z_j, j \in \alpha^c) = 0$$

  for all $\alpha \in \Xi$
- Extension to subset recovery

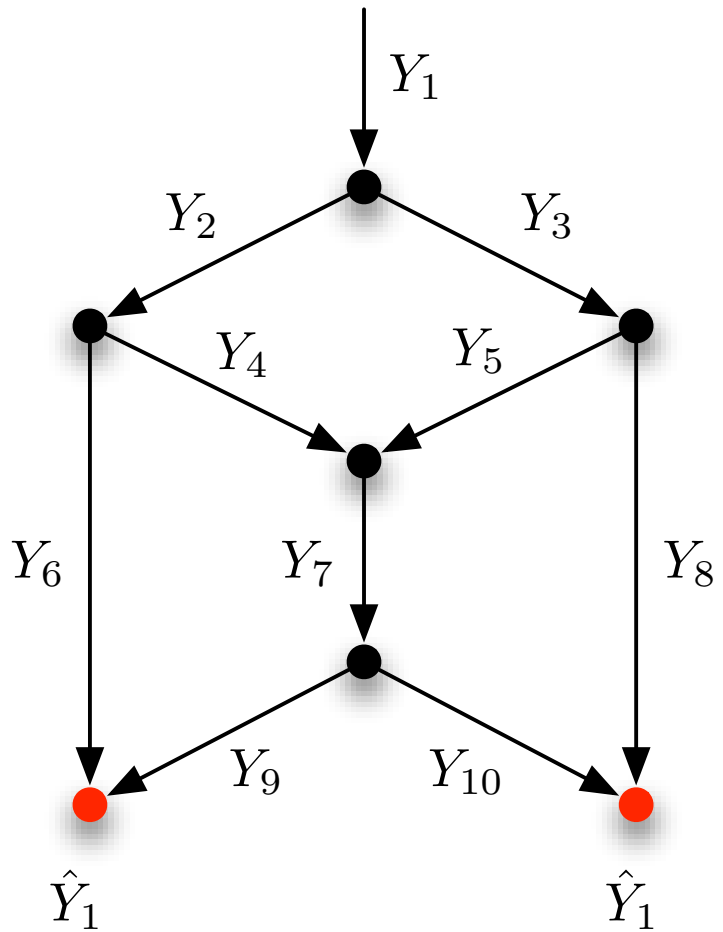
Network Coding

- Network code - specified by a set of random variables
- Source variables
- Link variables
- Topological constraint:

$$H(Y_7|Y_4, Y_5) = 0$$

- Decoding constraint:

$$H(Y_1|Y_6, Y_9) = 0$$



Secret Sharing

- Dealer share a secret with $n-1$ participants, indexed by the set $\{2,\ldots,n\}$. (Dealer is player 1)
- only specified legitimate groups of participants can reconstruct the secret data
- Ω – *access structure*, only participants indexed by $\mathcal{A} \in \Omega$ can access the secret.
- A secret sharing scheme is a random vector $(Z_1,\ldots,Z_n)$ such that
  1. Z_1 is the secret;
  2. Z_j is the share held by participant j;
  3. $H(Z_1|Z_j, j \in \mathcal{A}) = 0$ if $\mathcal{A} \in \Omega$;
  4. Z_1 and $(Z_j : j \in \mathcal{A})$ are independent whenever $\mathcal{A} \notin \Omega$.

Fundamental questions ...

These are codes

- specified by random variables
- satisfied functional dependency constraint

The basic questions are ...

- How to find an efficient code?
- Bounds on the rate of codes?
- Necessary condition for the existence of a code?
- In particular, assume a finite regime – alphabet sizes are fixed

Codes

- Let $\mathcal{Z}_1, \ldots, \mathcal{Z}_n$ be a set of non-empty sets, of sizes $N_1, \ldots, N_n$
- Assume WLOG $\mathcal{Z}_i = \{0, \ldots, N_i - 1\}$
- A code $\mathcal{C}$ is a non-empty subset of $\prod_{i=1}^{n} \mathcal{Z}_i$ (or simply $\mathcal{Z}^{\mathcal{N}}$).
- For any codewords, $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n) \in \mathcal{Z}^{\mathcal{N}}$, their
  - *difference* – $(\mathbf{a} - \mathbf{b}) \triangleq (a_1 - b_1, \ldots, a_n - b_n)$
  - *support* – $S(\mathbf{a}, \mathbf{b}) \triangleq \{j \in \{1, \ldots, n\} : a_j - b_j \neq 0\}$.
  - *distance* – $|S(\mathbf{a}, \mathbf{b})|$
- The minimum distance of a code $\mathcal{C}$ is defined as

$$\min_{\mathbf{a},\mathbf{b} \in \mathcal{C} : \mathbf{a} \neq \mathbf{b}} |S(\mathbf{a}, \mathbf{b})|.$$

Example

- Suppose $\mathcal{C} = \{(0,1,1),(0,2,1),(1,2,1)\}$ where $N_i = \{0,1,2\}$
- Consider the pair of codewords $(0,1,1)$ and $(0,2,1)$
  - Difference is $(0,2,0)$
  - Support is the subset $\{1\}$ and the distance is 1.
- Consider the pair of codewords $(0,1,1)$ and $(1,2,1)$
  - Difference is $(2,2,0)$
  - Support is the subset $\{1,2\}$ and the distance is 2.
- Denote the support be a *binary vector* of length n
- E.g., $(1,0,0)$ and $(1,1,0)$

Enumerators

- *Difference enumerator (FE)*

$$F(\mathbf{a}) = |\{(\mathbf{b},\mathbf{c}) : \mathbf{b},\mathbf{c} \in \mathcal{C} \text{ and } \mathbf{b}-\mathbf{c}=\mathbf{a}\}|.$$

- *Support enumerator (SE)*

$$\xi_f(\mathbf{r}) = \sum_{\mathbf{a}:a_i \neq 0 \text{ iff } i \in \mathbf{r}} F(\mathbf{a}).$$

- *Distance enumerator (DE)*

$$\gamma_f(i) = \sum_{\mathbf{r}:|\mathbf{r}|=i} \xi_f(\mathbf{r}).$$

- Sometimes, normalised with the factor $1/|\mathcal{C}|^2$

Necessary condition

Theorem (Necessary condition)

Support enumerator will satisfy the following conditions:

$$\xi_f(\mathbf{r}) \geq 0$$

$$\sum_{\mathbf{r}} \xi_f(\mathbf{r}) \prod_{j=1}^{n} \kappa_{N_j}(r_j, s_j) \geq 0$$

where $\mathbf{r} = (r_1, \ldots, r_n), \mathbf{s} = (s_1, \ldots, s_n) \subseteq \mathcal{N}$, *and*

$$\kappa_{N_j}(s, r) = \begin{cases} 1 & \text{if } r_j = 0 \\ N_j - 1 & \text{if } s_j = 0 \text{ and } r_j = 1 \\ -1 & \text{otherwise,} \end{cases}$$

Proving necessary condition

- For each code $\mathcal{C}$, it is associated with an "indicator function" f defined as follows

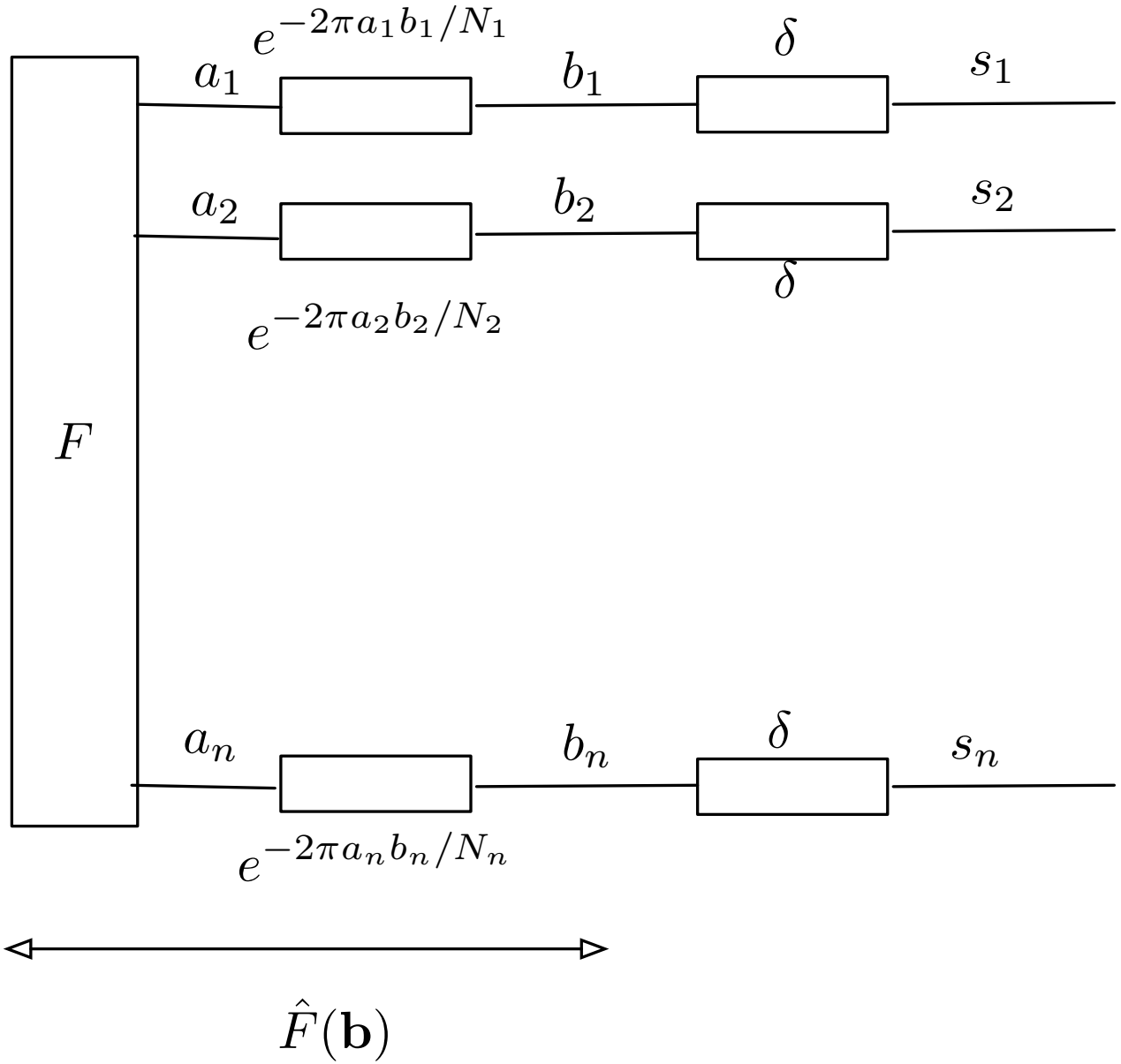
$$f(z_1, \ldots, z_n) = \begin{cases} 1 & \text{if } (z_1, \ldots, z_n) \in \mathcal{C} \\ 0 & \text{otherwise.} \end{cases}$$

- Then $F(\mathbf{a}) = \sum_{\mathbf{b}} f(\mathbf{b}) f(\mathbf{b} + \mathbf{a})$.

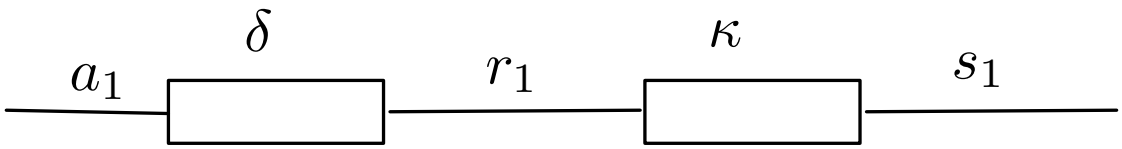
Theorem (Nonnegativity)

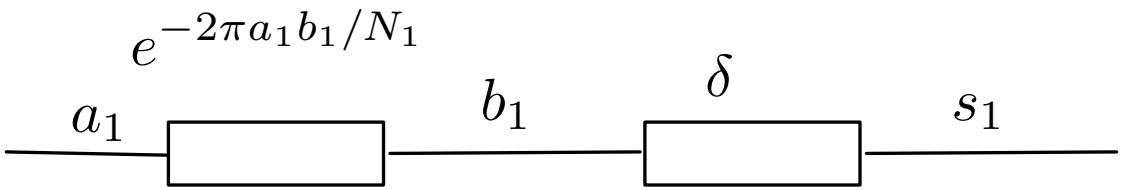
- $F(\mathbf{a}) = \sum_{\mathbf{b}} f(\mathbf{b}) f(\mathbf{b} + \mathbf{a}) \geq 0.$
- $\hat{F}(k_1, \ldots, k_n) \triangleq \sum_{a_1, \ldots, a_n} F(a_1, \ldots, a_n) \prod_{j=1}^{n} e^{-2\pi a_j k_j / N_j} \geq 0$

Proving necessary condition



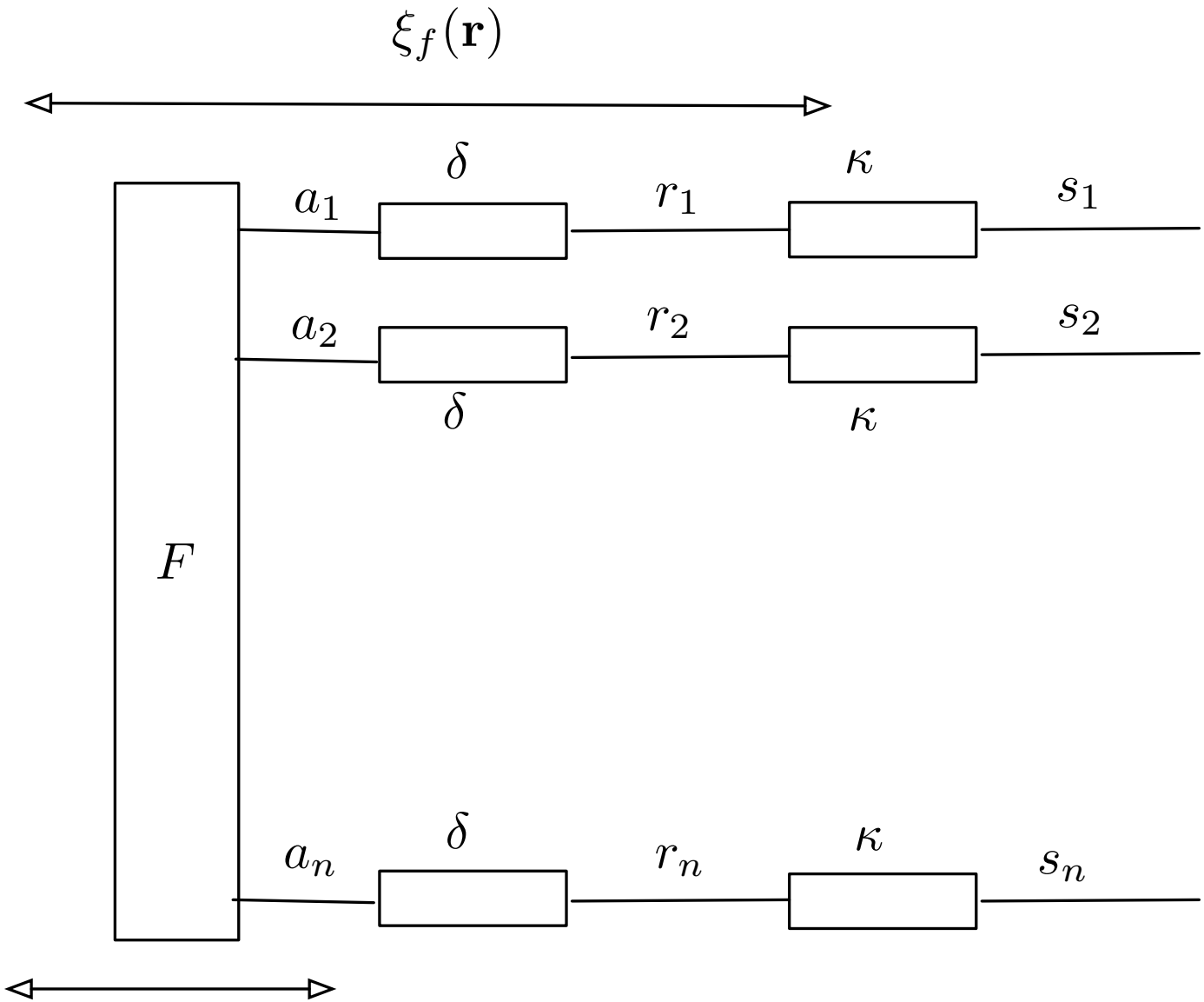
Proving necessary condition





$$\delta_{N_j}(a_j, r_j) = \begin{cases} 1 & \text{if } a_j = r_j = 0 \\ 1 & \text{if } a_j, r_j \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Proving necessary condition



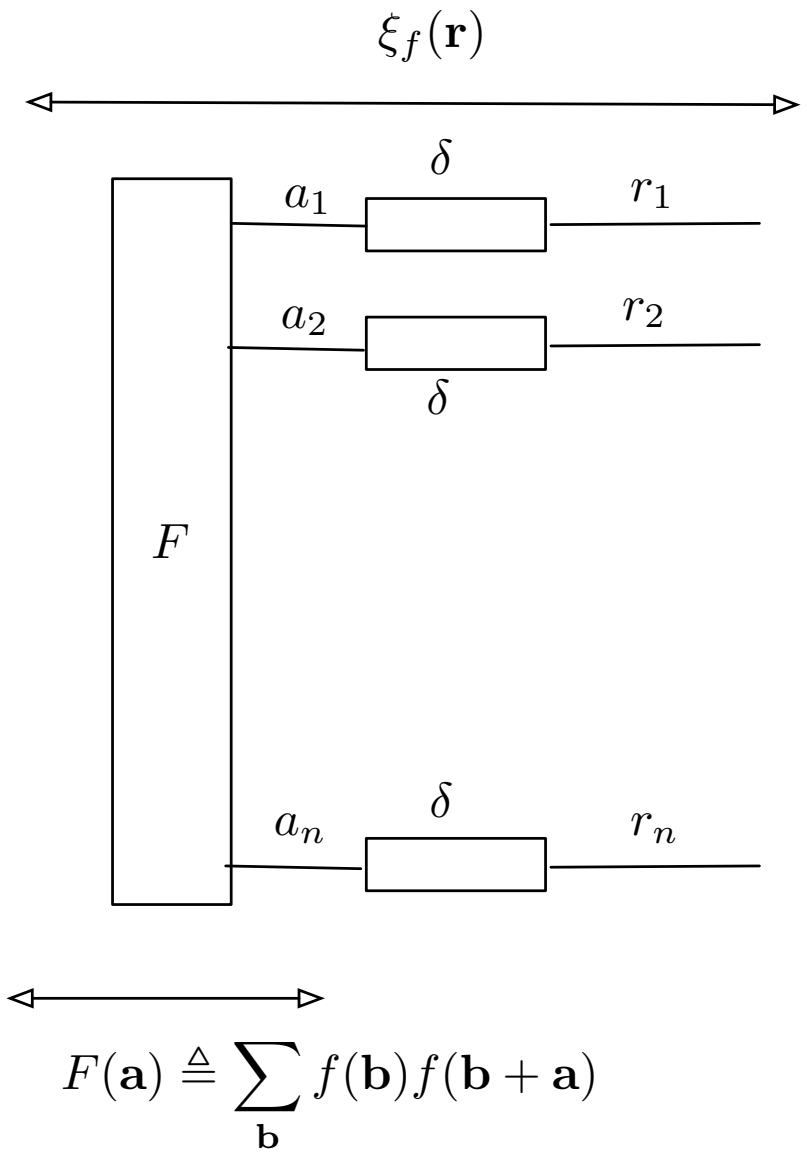
$$F(\mathbf{a}) = \sum_{\mathbf{b}} f(\mathbf{b}) f(\mathbf{b}+\mathbf{a})$$

Proving necessary condition

- $\xi_f(\mathbf{r}) = \sum_{\mathbf{a}:a_i \neq 0 \text{ iff } i \in \mathbf{r}} F(\mathbf{a})$.
- $\xi_f(\mathbf{r}) = \sum_{\mathbf{a}} F(\mathbf{a}) \prod_{j=1}^{n} \delta_{N_j}(a_j, r_j)$ where

$$\delta_{N_j}(a_j, r_j) = \begin{cases} 1 & \text{if } a_j = r_j = 0 \\ 1 & \text{if } a_j, r_j \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

- δ_{N_j} is a $N_j \times 2$ matrix
- $r_j \in \{0, 1\}$



Coding constraint

Suppose $\mathcal{C} \subseteq \mathcal{Z}^{\mathcal{N}}$ is a code. Let

$$\psi_f(\mathbf{s}) = \sum_{\mathbf{r}:\mathbf{r}\subseteq\mathbf{s}} \xi_f(\mathbf{r})$$

- $\psi_f(\mathcal{N})$ is the number of codewords in $\mathcal{C}$
- If $\mathbf{Z}_{\mathcal{B}}$ is a function of $\mathbf{Z}_{\mathcal{A}}$, then $\psi_f(\mathcal{N} - (\mathcal{A} \cup \mathcal{B})) = \psi_f(\mathcal{N} - \mathcal{A})$
- If $\mathbf{Z}_{\mathcal{A}}$ and $\mathbf{Z}_{\mathcal{B}}$ are independent, then $\psi_f(\mathcal{N} - (\mathcal{A} \cup \mathcal{B}))\psi_f(\mathcal{N}) = \psi_f(\mathcal{N} - \mathcal{A})\psi_f(\mathcal{N} - \mathcal{B})$.

Delsarte's LP bound

Theorem (LP bound)

Let $\mathcal{C}$ be a code such that the minimum Hamming distance of $\mathcal{C}$ is at least d. Then $|\mathcal{C}|^2$ is upper bounded by the maximum of the following optimisation problem:

$$
\begin{aligned}
&\textbf{maximize} && \textstyle\sum_{\mathbf{r}} \xi_f(\mathbf{r}) && \\
&\textbf{subject to} && \xi_f(\mathbf{r}) \geq 0 && \forall \mathbf{r} \\
& && \sum_{\mathbf{s}} \xi_f(\mathbf{s}) \prod_{j=1}^{n} \kappa_{N_j}(s_j, r_j) \geq 0 && \forall \mathbf{r} \\
& && |\xi_f(\mathbf{r})| = 0 && \forall \mathbf{r} : \ 1 \leq |\mathbf{r}| \leq d-1 \, .
\end{aligned}
$$

Bounds

$\mathbf{R}_{\text{tamper}}(\Lambda)$ is upper-bounded by the optimum of the following linear programming problem:

$$
\begin{aligned}
\textbf{maximize} \quad & \psi_f(\mathcal{N}) \\
\textbf{subject to} \quad & \xi_f(\ell) \geq 0 && \forall \ell \in \{0,1\}^n \\
& \sum_{\ell \in \{0,1\}^n} \xi_f(\ell) \prod_{j=1}^{n} \kappa_{N_j}(\ell_j, \nu_j) \geq 0 && \forall \nu \in \{0,1\}^n \\
& \psi_f(\mathbf{r}) = \sum_{\ell \leq \mathbf{r}} \xi_f(\ell) && \forall \mathbf{r} \in \{0,1\}^n \\
& \psi_f(\mathcal{A}) = 1 && \forall \mathcal{A} \in \Lambda^2 \,.
\end{aligned}
$$

Bounds

$\mathbf{R}_{\text{secret}}(\Omega)$ is upper-bounded by the optimum of the following optimisation:

maximize $\min_{j \in \mathcal{N}-\{1\}} \dfrac{\log \psi_f(\mathcal{N}) - \log \psi_f(\mathcal{N} - \{1\})}{\log \psi_f(\mathcal{N}) - \log \psi_f(\mathcal{N} - \{j\})}$

subject to

$$\xi_f(\ell) \geq 0$$

$$\sum_{\ell \in \{0,1\}^n} \xi_f(\ell) \prod_{j=1}^{n} \kappa_{N_j}(\ell_j, \nu_j) \geq 0$$

$$\psi_f(\mathbf{r}) = \sum_{\ell \leq \mathbf{r}} \xi_f(\ell)$$

$$\psi_f(\mathcal{N} - (\mathcal{A} \cup \{1\})) = \psi_f(\mathcal{N} - \mathcal{A})$$

$$\psi_f(\mathcal{N} - (\mathcal{A} \cup \{1\}))\psi_F(\mathcal{N}) = \psi_f(\mathcal{N} - \mathcal{A})\psi_f(\mathcal{N} - \{1\})$$

Renyi entropy

Renyi entropy

Definition

Let Z be a random variable with probability distribution $f(z)$. Then its Renyi entropy of order α for $\alpha \geq 0$ and $\alpha \neq 1$ is defined as

$$H_\alpha(Z) \triangleq \frac{1}{1-\alpha} \log \left(\sum_{z:f(z)>0} f(z)^\alpha \right).$$

When $\alpha = 1$, $H_1(Z) \triangleq \lim_{\alpha \to 1} H_\alpha(Z)$.

Examples

$$H_2(Z) = -\log \left(\sum_z f(z)^2 \right)$$

$$H_1(Z) = -\sum_z f(z) \log f(x)$$

$$H_0(Z) = \log \left| \{z : f(z) > 0\} \right|.$$

Renyi Entropy - Interpretation

- Let X and Y be two independent random variables, identically distributed as Z.
- Then $H_2(Z) = -\log \Pr(X = Y)$.
- Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be two independent sets of random variables with the same probability distribution f. Then for any $\mathbf{s} \subseteq \mathcal{N}$,

$$\psi_f(\mathbf{s}) \triangleq \Pr(S(\mathbf{X}, \mathbf{Y}) \subseteq \mathbf{s}) = 2^{-H_2(X_{\bar{\mathbf{s}}})}.$$

Extension

- Let f be a probability mass function for random variables $(Z_1, \ldots, Z_n)$.
- Let

$$F(\mathbf{a}) = \sum_{\mathbf{b}} f(\mathbf{b}) f(\mathbf{b} + \mathbf{a})$$

$$\xi_f(\mathbf{r}) = \sum_{\mathbf{a}} F(\mathbf{a}) \prod_{j=1}^{n} \delta_{N_j}(a_j, r_j)$$

- Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be two independent sets of random variables with the same probability distribution f.
- Then for any $\mathbf{s} \subseteq \mathcal{N}$,

$$\xi_f(\mathbf{s}) = \Pr(S(\mathbf{X}, \mathbf{Y}) = \mathbf{s}).$$

Extension

Theorem

$$\xi_f(\mathbf{r}) \geq 0$$

$$\sum_{\mathbf{r}} \xi_f(\mathbf{r}) \prod_{j=1}^{n} \kappa_{N_j}(r_j, s_j) \geq 0$$

for all $\mathbf{r} = (r_1, \ldots, r_n), \mathbf{s} = (s_1, \ldots, s_n) \subseteq \mathcal{N}$.

Mobius Transform

Theorem (Relation to Renyi entropy)

$$\sum_{\mathbf{r}:\mathbf{r}\subseteq\mathbf{s}} \xi_f(\mathbf{r}) = \psi_f(\mathbf{s}),$$

$$\sum_{\mathbf{s}:\mathbf{s}\subseteq\mathbf{v}} (-1)^{|\mathbf{v}\setminus\mathbf{s}|}\psi_f(\mathbf{s}) = \xi_f(\mathbf{v})$$

for all $\mathbf{r}, \mathbf{s}, \mathbf{v} \subseteq \mathcal{N}$

$$\psi_f(\mathbf{s}) \triangleq \Pr(S(\mathbf{X}, \mathbf{Y}) \subseteq \mathbf{s}) = 2^{-H_2(X_{\bar{\mathbf{s}}})}.$$

Renyi information inequalities

Theorem

Let f be a probability distribution of a set of discrete random variables $(Z_1, \ldots, Z_n)$. Then for all $\mathbf{r} \subseteq \mathcal{N}$,

$$\xi_f(\mathbf{r}) = \sum_{\mathbf{s}:\mathbf{s}\subseteq\mathbf{r}} (-1)^{|\mathbf{r}\setminus\mathbf{s}|} 2^{-H_2(Z_{\bar{\mathbf{s}}})} \geq 0,$$

$$\sum_{\mathbf{r}} \xi_f(\mathbf{r}) \prod_{j=1}^{n} \kappa_{N_j}(r_j, s_j) = \sum_{\mathbf{u}:\mathbf{u}\subseteq\mathbf{r}} (-1)^{|\mathbf{u}|} 2^{-H_2(Z_{\bar{\mathbf{u}}\cap\mathbf{r}})} \prod_{j:j\in\mathbf{r}\setminus\mathbf{u}} 2^{H_0(Z_j)} \geq 0.$$

Corollary

Theorem

Let $\{Z_1, \ldots, Z_n\}$ be a set of marginally uniform random variables. Then for all $\mathbf{r} \subseteq \mathcal{N}$,

$$\sum_{\mathbf{s}:\mathbf{s}\subseteq\mathbf{r}} (-1)^{|\mathbf{r}\setminus\mathbf{s}|} 2^{-H_2(Z_{\bar{\mathbf{s}}})} \geq 0,$$

$$\sum_{\mathbf{u}:\mathbf{u}\subseteq\mathbf{r}} (-1)^{|\mathbf{u}|} 2^{-H_2(Z_{\mathbf{r}\setminus\mathbf{u}}) + \sum_{j:j\in\mathbf{r}\setminus\mathbf{u}} H_2(Z_j)} \geq 0.$$

Dualities

Theorem (Dualities)

- *Let f be a probability distribution of a set of marginally uniform discrete random variables $(Z_1, \ldots, Z_n)$.*
- *Let $\rho(\mathbf{r}) \triangleq H_2(Z_j, j \in \mathbf{r})$ be the collision (or extension) entropy function*
- *Let $\mu(\mathbf{r}) \triangleq \sum_{i \in \mathbf{r}} \rho(i) + \rho(\bar{\mathbf{r}}) - \rho(\mathcal{N})$ be its induced dual.*
- *Then for all $\mathbf{r} \subseteq \mathcal{N}$,*

$$\sum_{\mathbf{s}:\mathbf{s}\subseteq\mathbf{r}} (-1)^{|\mathbf{r}\setminus\mathbf{s}|} 2^{\rho(\mathcal{N})-\rho(\bar{\mathbf{s}})} \geq 0$$

$$\sum_{\mathbf{s}:\mathbf{s}\subseteq\mathbf{r}} (-1)^{|\mathbf{r}\setminus\mathbf{s}|} 2^{\mu(\mathcal{N})-\mu(\bar{\mathbf{s}})} \geq 0$$

Group induced random variables

Theorem

Let G be a finite group and $G_1, \ldots, G_n$ be its subgroups. There exists random variables $U_1, \ldots, U_n$ such that

$$H_0(U_i, i \in \alpha) = H_2(U_i, i \in \alpha) = \log |G| - \log | \cap_{i \in \alpha} G_i|.$$

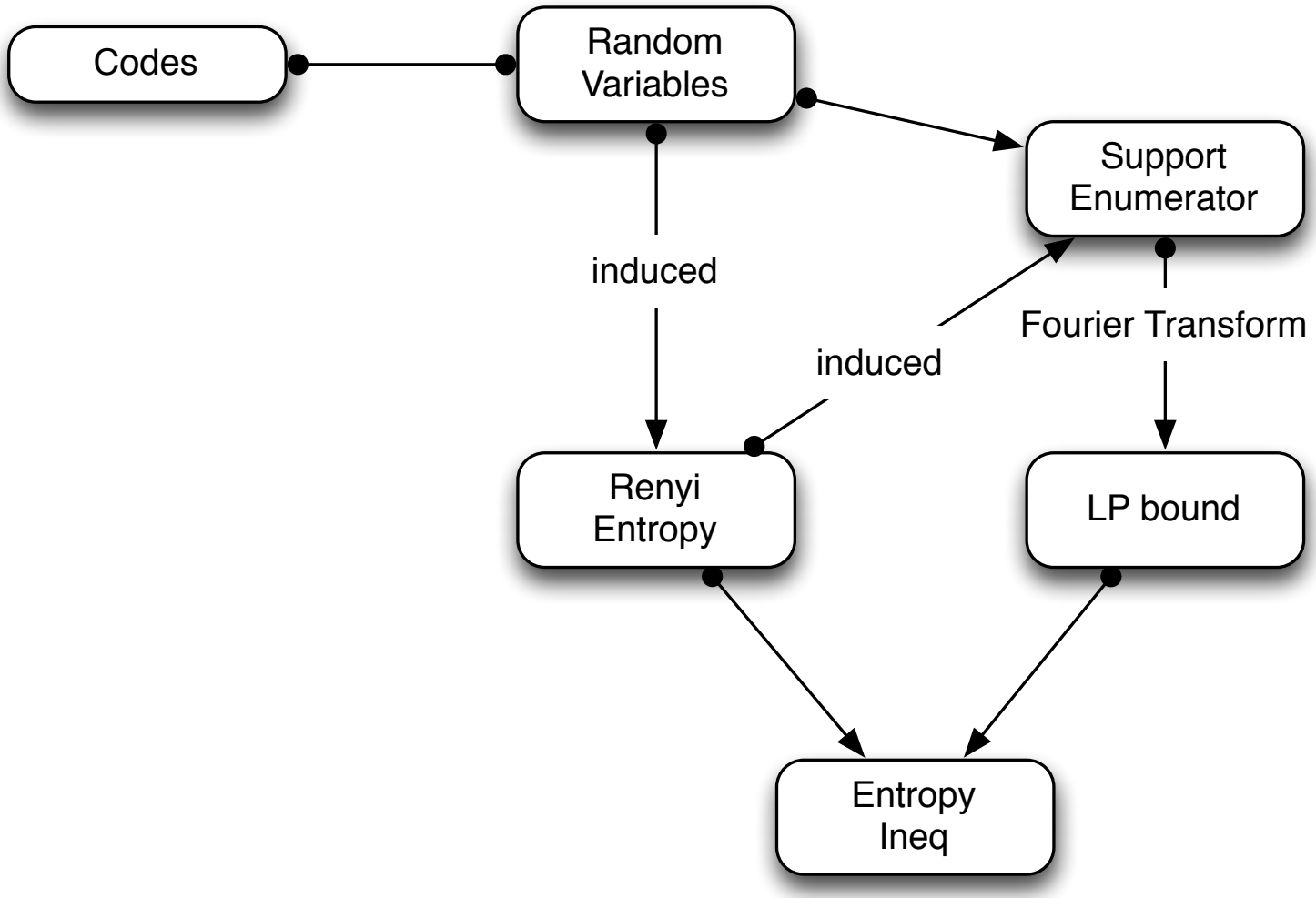
Corollary

$$\sum_{\mathbf{s}: \mathbf{s} \supseteq \bar{\mathbf{r}}} (-1)^{|\mathbf{s}-\mathbf{r}|} | \cap_{i \in \mathbf{s}} G_i| \geq 0$$

$$\sum_{\mathbf{s}: \mathbf{s} \subseteq \mathbf{r}} \left(\frac{-1}{|G|} \right)^{|\mathbf{s}|} \frac{| \cap_{i \in \mathbf{r} \setminus \mathbf{s}} G_i|}{\prod_{j: j \in \mathbf{r} \setminus \mathbf{s}} |G_j|} \geq 0$$

for all $\mathbf{r} \subseteq \mathcal{N}$.

Conclusion


Codes
Random Variables
Support Enumerator
induced
induced
Fourier Transform
Renyi Entropy
LP bound
Entropy Ineq

Thank You !!