Identification Without Randomization

Mohammad Javad Salariseddigh

July 16, 2020

Outline

1. Motivation
2. Transmission
3. Identification
4. Conclusions

Outline

Motivation

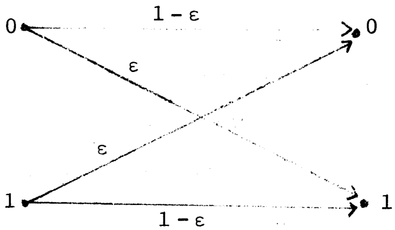


Figure 1: The Binary Symmetric Channel (BSC)

JaJa 1985

Identification for $BSC_{\epsilon \neq 0.5}$ with rate arbitrarily close to 1 is possible by exploiting Gilbert's bound where:

- $d_H = n\delta$ ($\delta \to 0$)
- Radius of Hamming spheres: $n(\epsilon + \eta)$ where $\epsilon < \frac{1}{n}, \eta \ll \epsilon$

Motivation cont

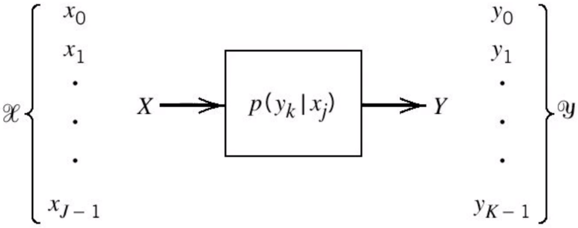


Figure 2: The Discrete Memoryless Channel (DMC)

Ahlswede 1989

For general DMC $\mathcal{D}$ with all distinct rows in stochastic matrix $\rightarrow$

$$C_{NRI}(\mathcal{D}) = \log |\mathcal{X}|$$

Motivation cont

- Arbitrary Varying Channel (AVC) and Compound Channel (CC) are suitable models for **Jamming**
- Pessimistic assumption $\rightarrow$ Jammer knows input sequence
- Randomization in encoding would be superfluous

$\Longrightarrow$ Identification without randomization (NRI) ✓

NRI $\neq$ Transmission

For general channel $\mathcal{W}$, the NRI capacity $C_{NRI}(\mathcal{W})$ is quite different from the transmission capacity $C(\mathcal{W})$

Outline

1. Motivation
2. Transmission
3. Identification
4. Conclusions

Transmission

(Shannon 1948):

Alice: $i \in [\![M]\!] \Longrightarrow \boxed{\varphi_n} \xrightarrow{X^n} \boxed{\mathcal{D}} \xrightarrow{Y^n} \boxed{\psi_n} \Longrightarrow \hat{i} \approx_\lambda i$:Bob

Definition

Transmission (n, M, λ) code for $\mathcal{D}$ is a system $\{(\boldsymbol{u_i}, \mathcal{D}_i)\}_{i\in[\![M]\!]}$:

1. $\boldsymbol{u_i} \in \mathcal{X}^n, \mathcal{D}_i \subset \mathcal{Y}^n$
2. $W^n(\mathcal{D}_i|\boldsymbol{u_i}) \geq 1 - \lambda$
3. $\mathcal{D}_i \bigcap_{i\neq j} \mathcal{D}_j = \emptyset$

Transmission

(Shannon 1948):

$$\text{Alice:}\quad i \in [\![M]\!] \Longrightarrow \boxed{\varphi_n} \xrightarrow{X^n} \boxed{\mathcal{D}} \xrightarrow{Y^n} \boxed{\psi_n} \Longrightarrow \hat{i} \approx_\lambda i \quad \text{:Bob}$$

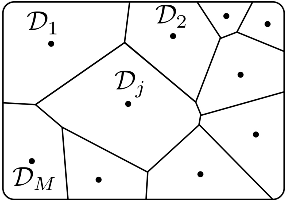


Figure 3: Decoding system partitions output space into M subsets

Capacity Theorem

$$\lim_{n\to\infty} n^{-1} \log M_{max}(n,\lambda) = C_T \qquad \forall \lambda \in (0,1)$$

NRS-Codes

(Ahlswede 1980):

$$\text{Alice:} \quad i \in [\![M]\!] \Longrightarrow \boxed{\varphi_n} \xrightarrow{X^n} \boxed{\mathcal{A}} \xrightarrow{Y^n} \boxed{\boxed{\psi_n} \Longrightarrow \hat{i} \approx_\lambda i} \quad \text{:Bob}$$

Definition

(n, M, λ) NRS-code for AVC $\mathcal{A}$ is a system $\{\boldsymbol{u}, \boldsymbol{u}', \mathcal{D}(\boldsymbol{u}, \boldsymbol{u}'), \mathcal{D}(\boldsymbol{u}', \boldsymbol{u})\}_{\substack{\boldsymbol{u}, \boldsymbol{u}' \in \mathcal{U} \\ \boldsymbol{u} \neq \boldsymbol{u}'}}$:

1. $\mathcal{U} \subset \mathcal{X}^n, \ \mathcal{D}(\boldsymbol{u}, \boldsymbol{u}'), \ \mathcal{D}(\boldsymbol{u}', \boldsymbol{u}) \subset \mathcal{Y}^n, \ |\mathcal{U}| = M$
2. $\mathcal{D}(\boldsymbol{u}, \boldsymbol{u}') \cap \mathcal{D}(\boldsymbol{u}', \boldsymbol{u}) = \emptyset$ separation property
3. $W^n(\mathcal{D}(\boldsymbol{u}, \boldsymbol{u}') | \boldsymbol{u}, s^n) \geq 1 - \lambda(s^n)$

$$\forall u, u' \underset{\boldsymbol{u} \neq \boldsymbol{u}'}{\in} \mathcal{U}, \quad s^n \in \mathcal{S}^n$$

NRS-Codes

(Ahlswede 1980):

Alice: $i \in [\![M]\!] \Longrightarrow \boxed{\varphi_n} \xrightarrow{X^n} \boxed{\mathcal{A}} \xrightarrow{Y^n} \boxed{\psi_n} \Longrightarrow \hat{i} \approx_\lambda i$:Bob

Why NRS-codes?

- We can associate (n, M, λ) SP-code (resp. NRS-code) to $(n, M, \lambda_1, \lambda_2)$ ID-code (resp. NRI-code) to prove soft-converse of identification capacity

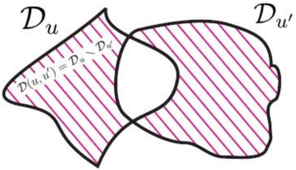


Figure 4: SP-codes association with ID-codes

Transmission

(Shannon 1948):

$$\text{Alice:}\quad i \in [\![M]\!] \Longrightarrow \boxed{\varphi_n} \xrightarrow{X^n} \boxed{\mathcal{D}} \xrightarrow{Y^n} \boxed{\boxed{\psi_n} \Longrightarrow \hat{i} \approx_\lambda i} \quad \text{:Bob}$$

Randomized Encoding

1. $Q_i W^n(\mathcal{D}_i) = \sum_{x^n} Q_i(x^n) W(\mathcal{D}_i | x^n) \geq 1 - \lambda$
   $\Rightarrow \sum_{x^n} Q_i(x^n) W(\mathcal{D}_i^c | x^n) < \lambda$
2. $\exists\ u_i \in \mathcal{X}^n$ such that:

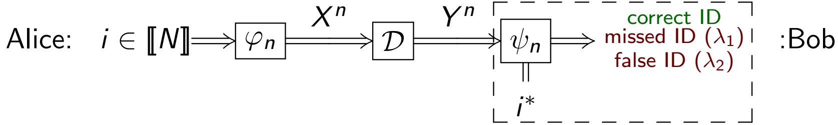
$$W(\mathcal{D}_i^c | u_i) \leq \sum_{x^n} Q_i(x^n) W(\mathcal{D}_i^c | x^n) < \lambda$$

$\Longrightarrow$ Transmission can not benefit randomization

Outline

Randomized Identification

(Ahlswede & Dueck 1989):



Definition

$(n, N, \lambda_1, \lambda_2)$ ID-code for DMC $\mathcal{D}$ is a system $\{(Q_i, \mathcal{D}_i)\}_{i \in [\![N]\!]}$:

1. $Q_i \in \mathcal{P}(\mathcal{X}^n), \mathcal{D}_i \subset \mathcal{Y}^n$
2. $Q_i W^n(\mathcal{D}_i) > 1 - \lambda_1$ correctedness property
3. $Q_j W^n(\mathcal{D}_i) \underset{i \neq j}{<} \lambda_2$ disjointedness property

Randomized Identification

(Ahlswede & Dueck 1989):

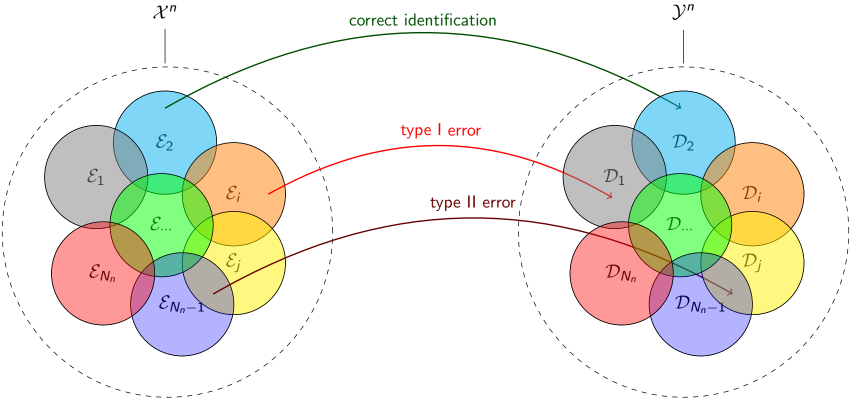
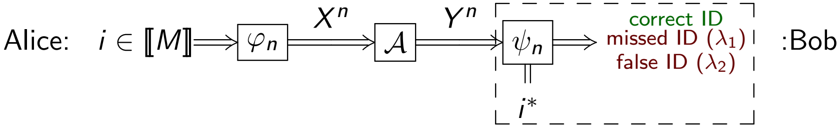


Figure 5: A visual representation of ID-codes

NRI-codes

(Ahlswede & Cai 1999):



Definition

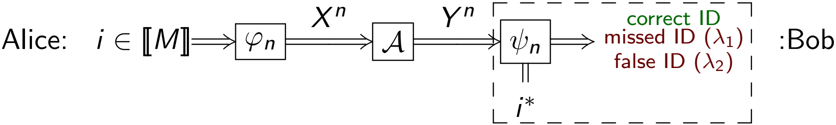
$(n, M, \lambda_1, \lambda_2)$ NRI-code for AVC $\mathcal{A}$ is a system $\{(\boldsymbol{u}, \mathcal{D}_u)\}_{\boldsymbol{u} \in \mathcal{U}}$:

1. $\mathcal{U} \subset \mathcal{X}^n, \ \mathcal{D}_{\boldsymbol{u}} \subset \mathcal{Y}^n, \ |\mathcal{U}| = M$
2. $W^n(\mathcal{D}_{\boldsymbol{u}} | \boldsymbol{u}, s^n) > 1 - \lambda_1$
3. $W^n(\mathcal{D}_{\boldsymbol{u}} | \boldsymbol{u}', s^n) < \lambda_2$

$$\forall u, u' \underset{\boldsymbol{u} \neq \boldsymbol{u}'}{\in} \mathcal{U}, \quad s^n \in \mathcal{S}^n$$

NRA-codes

(Ahlswede & Cai 1999):



Definition

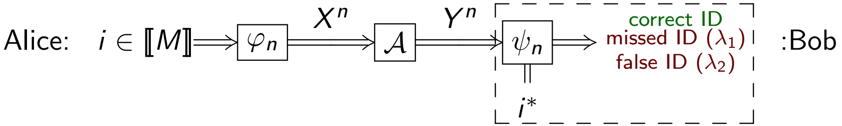
$(n, M, \lambda_1, \lambda_2, \overline{L})$ NRA-code for AVC $\mathcal{A}$ is a system $\{(\boldsymbol{u}, \mathcal{D}_{\boldsymbol{u}})\}_{\boldsymbol{u} \in \mathcal{U}}$:

1. $\mathcal{U} \subset \mathcal{X}^n, \; \mathcal{D}_u \subset \mathcal{Y}^n, \; |\mathcal{U}| = M$
2. $W^n(\mathcal{D}_{\boldsymbol{u}} | \boldsymbol{u}, s^n) > 1 - \lambda_1$
3. $\mathcal{D}_{\boldsymbol{u}} \cap \mathcal{D}_{\boldsymbol{u}'} = A(\boldsymbol{u}, \boldsymbol{u}') \,\dot{\bigcup}\, A(\boldsymbol{u}', \boldsymbol{u})$
4. $W^n(A(\boldsymbol{u}', \boldsymbol{u}) | \boldsymbol{u}, s^n) < \lambda_2$
5. $\overline{L}_{\mathcal{U}} \leq \overline{L}$

$$\forall u, u' \underset{\boldsymbol{u} \neq \boldsymbol{u}'}{\in} \mathcal{U}$$

NRA-codes

(Ahlswede & Cai 1999):



Definition (Cont)

1. $\overline{L}_{\mathcal{U}} = \max_{\boldsymbol{u} \in \mathcal{U}, s^n \in \mathcal{S}^n} L(\boldsymbol{u}, s^n)$ [worst case average list size]
2. $\overline{L}(\boldsymbol{u}, s^n) = \sum_{y^n \in D_{\boldsymbol{u}}} L(y^n) W^n(y^n | \boldsymbol{u}, s^n)$
3. $L(y^n) = |\{\boldsymbol{u}' \in \mathcal{U} : y^n \in D_{\boldsymbol{u}'}\}|$

- NRA-codes $\equiv$ list code + separation property or
- separation code with worst case average list size

NRI Capacity of CC

(Ahlswede & Cai 1999):

- Each member in $\mathcal{V} = \left\{V(.|.,s) : s \in \mathcal{S}, |\mathcal{S}| < \infty\right\}$ induces a partition $\left\{\mathcal{X}(1|s), \cdots, \mathcal{X}(j_s|s)\right\}$ in $\mathcal{X}$
- x, x' lie in the same subset $\iff$ $V(.|x,s) = V(.|x',s)$

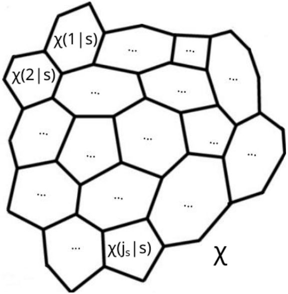


Figure 6: States partition input space

NRI Capacity of CC

(Ahlswede & Cai 1999):

- An RV X taking value in $\mathcal{X}$ induces RV $\hat{X} = \hat{X}(s) \quad \forall s \in \mathcal{S}$
- $\hat{X}(s) = l \iff X \in \mathcal{X}(l|s)$

Capacity Theorem

$$C_{NRI}(\mathcal{V}) = \max_{X} \min_{s \in \mathcal{S}} H(\hat{X}(s))$$

NRI Capacity of AVC

(Ahlswede & Cai 1999):

- Let $P \in \mathcal{P}(\mathcal{X})$ and $\overline{\overline{\mathcal{A}}}$ be row-convex closure of $\mathcal{A}$
- Let $\mathcal{Q}(P, \mathcal{A}) = \left\{(X, X', Y) : P_{Y|X}, P_{Y|X'} \in \overline{\overline{\mathcal{A}}}, P_X = P_{X'} = P, X \to X' \to Y\right\}$
- Set $\hat{I}(P, \mathcal{W}) = \min_{(X,X',Y) \in \mathcal{Q}(\mathcal{P},\mathcal{W})} I(X' \wedge XY)$

Theorem [A Lower Bound for $C_{NRI}(\mathcal{A})$]

$$C_{NRI}(\mathcal{A}) \geq \max_P \hat{I}(P, \mathcal{A})$$

NRI-codes in a Gaussian Channel

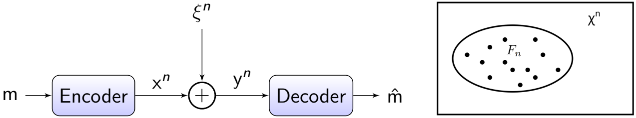


Figure 6: Discrete-time Gaussian Channel with Cost Constraint

$$\mathcal{G}: y_t = x_t + \xi_t \quad , \quad \xi_t \sim_{i.i.d} \mathcal{N}(0, \sigma^2) \quad , \quad t \in [1:n] \qquad (1)$$

Induced Input Space

- Input constraint induces a new input space to which all code-words must belong:

$$F_n \triangleq \left\{ x^n; \frac{1}{n} \sum_{1}^{n} c(x_t) \leq P \right\} \subset \mathcal{X}^n$$

NRI-codes in a Gaussian Channel

Coding Method

1. Let $p^* \in \mathcal{P}(\mathcal{X})$ such that:

$$\mathbb{E}_{p^*}\left\{c(X)\right\} = \sum_{x\in\mathcal{X}} p^*(x)c(x) \leq P - \delta \tag{2}$$

for $X \sim p^*(x)$, where $\delta > 0$ is arbitrarily small (**LLN**)

2. Let $U_1, ..., U_M$ be i.i.d. RV's with distribution p^* such that

$$\Pr(U_i = x^n) = p^*(x^n) = \prod_{t=1}^{n} p^*(x_t) \tag{3}$$

3. $V_i = \begin{cases} U_i & \|U_i - U_j\|_2 \geq n\epsilon \quad \forall i \neq j \\ x_0^n & \|U_i - U_j\|_2 \leq n\epsilon \quad \exists j \neq i, \quad c(x_0) = 0 \end{cases}$

NRI-codes in a Gaussian Channel

Dictionary

- $\mathcal{M} = \{i \ ; \ V_i \neq \mathbf{x}_0 \ ; \ 1 \leq i \leq M\}$ is set of code-words where $M \sim 2^{nC_{NRI}(\mathcal{G})}$

Decoding sets

- $B(V_i) = \left\{ y^n \ ; \ \frac{1}{n} \log \frac{W(y^n|V_i)}{q^*(y^n)} \geq C - \gamma \right\} , \qquad q^* = p^* W$

Further Issues

- Derive $\Pr\left(|\mathcal{M}| \leq \frac{1}{2} M\right)$, study error analysis

Outline

Conclusions and Outlook

1. Randomized encoding for transmission is not necessary
2. Randomized encoding for identification effects type II error
3. $C_{NRI}(\mathcal{A})$ for general AVC $\mathcal{A}$ is still **unknown**
4. For every AVC $\mathcal{A}$ under maximal error probability criterion, randomization in decoding does not provide higher capacity than $C(\mathcal{A})$, $C_{NRI}(\mathcal{A})$ and $C_I(\mathcal{A})$ reps.
5. NRS $\prec$ NRA $\prec$ NRI (strongest in property)
6. NRI coding method for Gaussian channel mimics similar spirit as of DMC technique (Hamming distance property)

Discussion

Thank you! Questions?

References

- R. Ahlswede, G. Dueck. Identification via channels. *IEEE Trans. Inform. Theory*, Vol. 35, No. 1, P. 15-29, Jan. 1989.
- R. Ahlswede and Ning Cai. Identification Without Randomization. *IEEE Transactions on Information Theory*, P. 356–361, 2019.
- R. Ahlswede. Identification without randomization. *Discrete Applied Mathematics*, P. 1348 - 1388, 2008.
- R. Ahlswede and G. Dueck. Identification in the presence of feedback-a discovery of new capacity formulas. *IEEE Transactions on Information Theory*. P. 30-60, Jan 89.
- R. Ahlswede. A method of coding and its application to arbitrarily varying channels *Journal of combinatorics, information and system sciences*. P. 10-35, 1999.
- S. Verdu ; V.K. Wei All Authors. Explicit construction of optimal constant-weight codes for identification via channels.*IEEE Transactions on Information Theory*, P. 30-36, Jan 93.

NRI Capacity of AVC I

- For every fixed $x \in \mathcal{X}$ define

$$\mathcal{A}_1(x) = \{A(.|x, s) : s \in \mathcal{S}\}$$

as set of PDs on $\mathcal{Y}$ where $\mathcal{A}_1 = \{A(.|.,\ s) : s \in \mathcal{S}\}$

- Define $\overline{A}(x)$ as **convex closure** of $\mathcal{A}_1(x)$ i.e. of entries in form

$$\sum_{s \in \tilde{\mathcal{S}}} P(s)A(y|x, s)$$

NRI Capacity of AVC II

- Define **row-convex closure** of $\mathcal{A}$ denote by $\overline{\overline{\mathcal{A}}}$ as follows:

$$\overline{\overline{\mathcal{A}}} = \left\{ (A(y|x))_{x \in \mathcal{X}, y \in \mathcal{Y}} : A(.|x) \in \overline{A}(x) \right\}$$

$\overline{\overline{\mathcal{A}}}$ has entries of form:

$$\sum_{s \in \tilde{\mathcal{S}}} P(s|x) A(y|x, s)$$

$P(s|x)$ means that coefficient are conditioned on choice of x, i.e., for every different x there would be in general a complete different set of coefficients than that of required for defining entries of $\overline{A}(x)$

NRI-codes in a Gaussian Channel

Cost Constraints

1. Average power constraint:

$$\frac{1}{n}\sum_{1}^{n}|x_t|^2 \leq P \iff \|x^n\|_2 \leq \sqrt{nP}$$

2. Peak power constraint:

$$\max_{1\leq t\leq n}|x_t| \leq A \iff \|x^n\|_\infty \leq A$$

NRI-codes in a Gaussian Channel

Error Analysis

1. $\mathcal{A}_1 = \{\phi^n(V_i) > A\}$
2. $\mathcal{A}_2 = \{Y^n \notin B(V_i)\}$
3. $\mathcal{A}_3 = \{Y^n \in B(V_j), \quad \exists j \neq i\}$