Less Noisy Domination by Symmetric Channels

Anuran Makur and Yury Polyanskiy

EECS Department, Massachusetts Institute of Technology

ISIT 2017

Outline

Preliminaries

- probability distributions – *row* vectors

Preliminaries

- probability distributions – *row* vectors
- channels (conditional distributions) – *row stochastic* matrices

Preliminaries

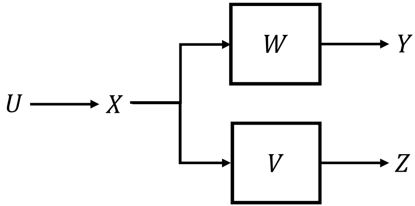
- probability distributions – *row* vectors
- channels (conditional distributions) – *row stochastic* matrices

Definition (Less Noisy Preorder [Körner-Marton 1977])

$P_{Y|X} = W$ is less noisy than $P_{Z|X} = V$, denoted $W \succeq_{\text{ln}} V$, if and only if

$$I(U;Y) \geq I(U;Z)$$

for every joint distribution $P_{U,X}$ such that $U \to X \to (Y,Z)$.



Preliminaries

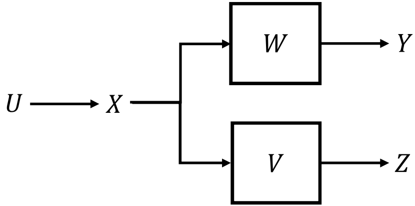
- probability distributions – *row* vectors
- channels (conditional distributions) – *row stochastic* matrices

Definition (Less Noisy Preorder [Körner-Marton 1977])

$P_{Y|X} = W$ is less noisy than $P_{Z|X} = V$, denoted $W \succeq_{\text{ln}} V$, if and only if

$$D(P_X W || Q_X W) \geq D(P_X V || Q_X V)$$

for every pair of input distributions P_X and Q_X.



Main Results

1. Test $\succeq_{\ln}$ using different divergence measure?

Main Results

1. Test $\succeq_{\ln}$ using different divergence measure?
   Yes, χ^2-divergence

Main Results

1. Test $\succeq_{\ln}$ using different divergence measure?
   Yes, χ^2-divergence

2. Sufficient conditions for $\succeq_{\ln}$ domination by symmetric channels?

Main Results

1. Test $\succeq_{\ln}$ using different divergence measure?
   Yes, χ^2-divergence

2. Sufficient conditions for $\succeq_{\ln}$ domination by symmetric channels?
   Yes
   - degradation criterion for general channels
   - stronger criterion for additive noise channels

Main Results

1. Test $\succeq_{\mathsf{ln}}$ using different divergence measure?
   Yes, χ^2-divergence
2. Sufficient conditions for $\succeq_{\mathsf{ln}}$ domination by symmetric channels?
   Yes
   - degradation criterion for general channels
   - stronger criterion for additive noise channels
3. Why $\succeq_{\mathsf{ln}}$ domination by symmetric channels?

Main Results

1. Test $\succeq_{\ln}$ using different divergence measure?
   Yes, χ^2-divergence

2. Sufficient conditions for $\succeq_{\ln}$ domination by symmetric channels?
   Yes
   - degradation criterion for general channels
   - stronger criterion for additive noise channels

3. Why $\succeq_{\ln}$ domination by symmetric channels?
   - just because we ♥ IT
   - $\succeq_{\ln}$ domination $\Rightarrow$ log-Sobolev inequality
   - secrecy capacity

Motivation: Strong Data Processing Inequality

Data Processing Inequality:

For any channel V,

$$\forall P_X, Q_X, \;\; D(P_X || Q_X) \geq D(P_X V || Q_X V)$$

Motivation: Strong Data Processing Inequality

Strong Data Processing Inequality [Ahlswede-Gács 1976]**:**
For any channel V,

$$\forall P_X, Q_X, \ \eta_{\mathsf{KL}}(V)\, D(P_X || Q_X) \geq D(P_X V || Q_X V)$$

where $\eta_{\mathsf{KL}}(V)$ – contraction coefficient:

$$\eta_{\mathsf{KL}}(V) \triangleq \sup_{P_X, Q_X} \frac{D(P_X V || Q_X V)}{D(P_X || Q_X)} \in [0, 1] \,.$$

Motivation: Strong Data Processing Inequality

Strong Data Processing Inequality [Ahlswede-Gács 1976]**:**
For any channel V,

$$\forall P_X, Q_X, \ \eta_{\text{KL}}(V)\, D(P_X||Q_X) \geq D(P_X V||Q_X V)$$

where $\eta_{\text{KL}}(V)$ – contraction coefficient.

Relation to Erasure Channels [Polyanskiy-Wu 2016]**:**

- **Definition:** q-ary erasure channel q-$EC(1-\eta)$
  erases input w.p. $1-\eta$, and reproduces input w.p. η.

Strong Data Processing Inequality [Ahlswede-Gács 1976]**:**
For any channel V,

$$\forall P_X, Q_X, \ \eta_{\mathsf{KL}}(V)\, D(P_X || Q_X) \geq D(P_X V || Q_X V)$$

where $\eta_{\mathsf{KL}}(V)$ – contraction coefficient.

Relation to Erasure Channels [Polyanskiy-Wu 2016]**:**

- **Definition:** q-ary erasure channel $q\text{-}EC(1-\eta)$ erases input w.p. $1-\eta$, and reproduces input w.p. η.
- **Prop** [Polyanskiy-Wu 2016]**:**

$$q\text{-}EC(1-\eta) \succeq_{\mathsf{ln}} V \ \Leftrightarrow \ \forall P_X, Q_X, \eta D(P_X || Q_X) \geq D(P_X V || Q_X V) \,.$$

Motivation: Strong Data Processing Inequality

Strong Data Processing Inequality [Ahlswede-Gács 1976]**:**
For any channel V,

$$\forall P_X, Q_X, \ \eta_{\mathsf{KL}}(V)\, D(P_X||Q_X) \geq D(P_X V||Q_X V)$$

where $\eta_{\mathsf{KL}}(V)$ – contraction coefficient.

Relation to Erasure Channels [Polyanskiy-Wu 2016]**:**

- **Definition:** q-ary erasure channel $q\text{-}EC(1-\eta)$ erases input w.p. $1-\eta$, and reproduces input w.p. η.
- **Prop** [Polyanskiy-Wu 2016]**:**

$$q\text{-}EC(1-\eta) \succeq_{\mathsf{ln}} V \ \Leftrightarrow \ \forall P_X, Q_X, \, \eta D(P_X||Q_X) \geq D(P_X V||Q_X V)\,.$$

SDPI $\Leftrightarrow$ $\succeq_{\mathsf{ln}}$ domination by erasure channel

Main Question

Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

Main Question

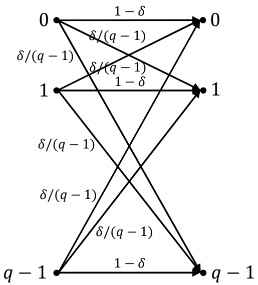
Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

Definition (q-ary Symmetric Channel)

Channel matrix:

$$W_\delta \triangleq \begin{bmatrix} 1-\delta & \frac{\delta}{q-1} & \cdots & \frac{\delta}{q-1} \\ \frac{\delta}{q-1} & 1-\delta & \cdots & \frac{\delta}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta}{q-1} & \frac{\delta}{q-1} & \cdots & 1-\delta \end{bmatrix}$$

where $\delta \in [0,1]$ – total crossover probability.



Main Question

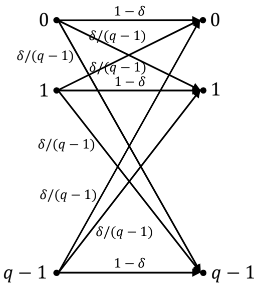
Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

Definition (q-ary Symmetric Channel)

Channel matrix:

$$W_\delta \triangleq \begin{bmatrix} 1-\delta & \frac{\delta}{q-1} & \cdots & \frac{\delta}{q-1} \\ \frac{\delta}{q-1} & 1-\delta & \cdots & \frac{\delta}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta}{q-1} & \frac{\delta}{q-1} & \cdots & 1-\delta \end{bmatrix}$$

where $\delta \in [0,1]$ – total crossover probability.



- For every channel V, $W_0 \succeq_{\text{ln}} V$ and $V \succeq_{\text{ln}} W_{(q-1)/q}$.

Main Question

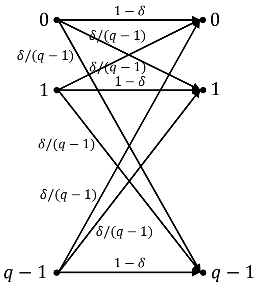
Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

Definition (q-ary Symmetric Channel)

Channel matrix:

$$W_\delta \triangleq \begin{bmatrix} 1-\delta & \frac{\delta}{q-1} & \cdots & \frac{\delta}{q-1} \\ \frac{\delta}{q-1} & 1-\delta & \cdots & \frac{\delta}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta}{q-1} & \frac{\delta}{q-1} & \cdots & 1-\delta \end{bmatrix}$$

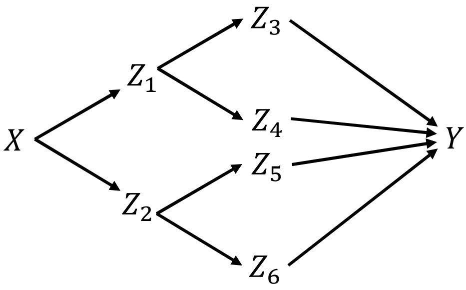
where $\delta \in [0,1]$ – total crossover probability.



- For every channel V, $W_0 \succeq_{\text{ln}} V$ and $V \succeq_{\text{ln}} W_{(q-1)/q}$.
- $\forall \epsilon, \delta \in (0,1)$, $W_\delta \not\succeq_{\text{ln}} q\text{-}EC(\epsilon)$.

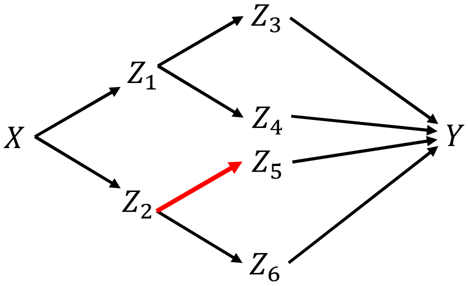
Less Noisy Channels in Networks

Consider general Bayesian network:



Less Noisy Channels in Networks

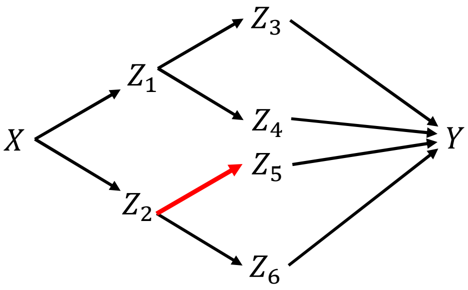
Consider general Bayesian network:



Conjecture:
Replace $P_{Z_5|Z_2}$ with less noisy channel $\Rightarrow$ $P_{Y|X}$ becomes less noisy.

Less Noisy Channels in Networks

Consider general Bayesian network:

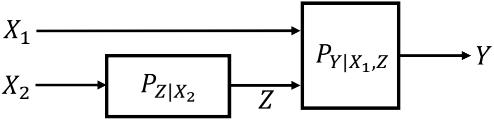


Conjecture:
Replace $P_{Z_5|Z_2}$ with less noisy channel $\Rightarrow$ $P_{Y|X}$ becomes less noisy.

Motivation: Results of [Polyanskiy-Wu 2016] on SDPIs in networks.

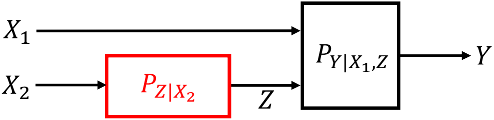
Less Noisy Channels in Networks

Consider Bayesian network with binary r.v.s



Less Noisy Channels in Networks

Consider Bayesian network with binary r.v.s

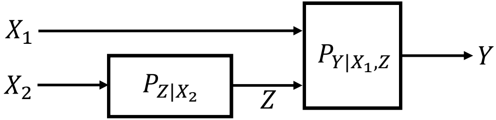


where we replace $P_{Z|X_2}$ with less noisy channel.

Can this decrease $I(X_1, X_2; Y)$**?**

Less Noisy Channels in Networks

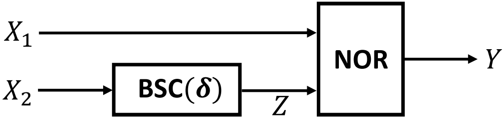
Consider Bayesian network with binary r.v.s



where we replace $P_{Z|X_2}$ with less noisy channel.

Can this decrease $I(X_1, X_2; Y)$? YES

Example: Let $X_1 \sim \text{Ber}\left(\frac{1}{2}\right)$ and $X_2 = 1$ *a.s.*, and let $I(\delta) = I(X_1, X_2; Y)$.



For $\delta > 0$, $\text{BSC}(0) \succeq_{\text{ln}} \text{BSC}(\delta)$, but $h(\delta/2) - h(\delta)/2 = I(\delta) > I(0) = 0$.

Outline

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V) \,.$$

Recall χ^2-divergence between P_X and Q_X:

$$\chi^2(P_X || Q_X) \triangleq \sum_{x \in \mathcal{X}} \frac{(P_X(x) - Q_X(x))^2}{Q_X(x)} \,.$$

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V) \, .$$

Proof: ($\Rightarrow$) Fix any P_X, Q_X. Recall local approximation:

$$\lim_{\lambda \to 0^+} \frac{2}{\lambda^2} D(\lambda P_X + (1-\lambda) Q_X || Q_X) = \chi^2(P_X || Q_X) \, .$$

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right) .$$

Proof: ($\Rightarrow$) Fix any P_X, Q_X. Recall local approximation:

$$\lim_{\lambda \to 0^+} \frac{2}{\lambda^2} D\left(\lambda P_X + (1-\lambda) Q_X || Q_X\right) = \chi^2\left(P_X || Q_X\right) .$$

$W \succeq_{\text{ln}} V$ implies

$$D\left(\lambda P_X W + (1-\lambda) Q_X W || Q_X W\right) \geq D(\lambda P_X V + (1-\lambda) Q_X V || Q_X V)$$

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V) \, .$$

Proof: ($\Rightarrow$) Fix any P_X, Q_X. Recall local approximation:

$$\lim_{\lambda \to 0^+} \frac{2}{\lambda^2} D(\lambda P_X + (1-\lambda) Q_X || Q_X) = \chi^2(P_X || Q_X) \, .$$

$W \succeq_{\text{ln}} V$ implies

$$D(\lambda P_X W + (1-\lambda) Q_X W || Q_X W) \geq D(\lambda P_X V + (1-\lambda) Q_X V || Q_X V)$$
$$\chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V)$$

after taking limits.

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right) .$$

Proof: $(\Leftarrow)$ Fix any P_X, Q_X. Recall integral representation:

$$D\left(P_X || Q_X\right) = \int_0^\infty \chi^2\left(P_X || Q_X^t\right) \, dt$$

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$ [Choi-Ruskai-Seneta 1994].

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right) .$$

Proof: ($\Leftarrow$) Fix any P_X, Q_X. Recall integral representation:

$$D\left(P_X || Q_X\right) = \int_0^\infty \chi^2\left(P_X || Q_X^t\right) \, dt$$

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$ [Choi-Ruskai-Seneta 1994].

$$\chi^2\left(P_X W || Q_X^t W\right) \geq \chi^2\left(P_X V || Q_X^t V\right)$$

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right) .$$

Proof: ($\Leftarrow$) Fix any P_X, Q_X. Recall integral representation:

$$D\left(P_X || Q_X\right) = \int_0^\infty \chi^2\left(P_X || Q_X^t\right) \, dt$$

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$ [Choi-Ruskai-Seneta 1994].

$$\chi^2\left(P_X W || Q_X^t W\right) \geq \chi^2\left(P_X V || Q_X^t V\right)$$

$$\int_0^\infty \chi^2\left(P_X W || Q_X^t W\right) \, dt \geq \int_0^\infty \chi^2\left(P_X V || Q_X^t V\right) \, dt$$

χ^2-Divergence Characterization of Less Noisy

Theorem 1 (χ^2-Divergence Characterization of $\succeq_{\text{ln}}$)

Given channels W and V, $W \succeq_{\text{ln}} V$ if and only if

$$\forall P_X, Q_X, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V) \, .$$

Proof: ($\Leftarrow$) Fix any P_X, Q_X. Recall integral representation:

$$D(P_X || Q_X) = \int_0^\infty \chi^2\left(P_X || Q_X^t\right) \, dt$$

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$ [Choi-Ruskai-Seneta 1994].

$$\chi^2\left(P_X W || Q_X^t W\right) \geq \chi^2\left(P_X V || Q_X^t V\right)$$

$$\int_0^\infty \chi^2\left(P_X W || Q_X^t W\right) \, dt \geq \int_0^\infty \chi^2\left(P_X V || Q_X^t V\right) \, dt$$

$$D(P_X W || Q_X W) \geq D(P_X V || Q_X V)$$

Löwner and Spectral Characterizations of Less Noisy

Theorem 1 (Equivalent Characterizations of $\succeq_{\text{ln}}$)

Given channels W and V,

$$W \succeq_{\text{ln}} V \;\Leftrightarrow\; \forall P_X, Q_X,\ \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V)$$
$$\Leftrightarrow\; \forall P_X,\ W \text{diag}(P_X W)^{-1} W^T \succeq_{\text{PSD}} V \text{diag}(P_X V)^{-1} V^T$$
$$\Leftrightarrow\; \forall P_X,\ \rho\Big(\big(W \text{diag}(P_X W)^{-1} W^T\big)^{\dagger} V \text{diag}(P_X V)^{-1} V^T\Big) = 1$$

where $\succeq_{\text{PSD}}$ – Löwner (PSD) partial order,
$A^{\dagger}$ – Moore-Penrose pseudoinverse of A,
and $\rho(\cdot)$ – spectral radius.

Outline

Condition for Degradation by Symmetric Channels

Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

Condition for Degradation by Symmetric Channels

Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

- **Definition (Degradation)** [Bergmans 1973]**:** V is degraded version of W, denoted $W \succeq_{\text{deg}} V$, if $V = WA$ for some channel A.

Condition for Degradation by Symmetric Channels

Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\ln} V$?

- **Definition (Degradation)** [Bergmans 1973]**:** V is degraded version of W, denoted $W \succeq_{\deg} V$, if $V = WA$ for some channel A.
- **Prop:** $W \succeq_{\deg} V \Rightarrow W \succeq_{\ln} V$.

Condition for Degradation by Symmetric Channels

Given channel V, find q-ary symmetric channel W_δ with largest $\delta \in \left[0, \frac{q-1}{q}\right]$ such that $W_\delta \succeq_{\text{ln}} V$?

- **Definition (Degradation)** [Bergmans 1973]**:** V is degraded version of W, denoted $W \succeq_{\text{deg}} V$, if $V = WA$ for some channel A.
- **Prop:** $W \succeq_{\text{deg}} V \Rightarrow W \succeq_{\text{ln}} V$.

Theorem 2 (Degradation by Symmetric Channels)

For channel V with common input and output alphabet, and minimum probability $\nu = \min\left\{[V]_{i,j} : 1 \leq i, j \leq q\right\}$,

$$0 \leq \delta \leq \frac{\nu}{1 - (q-1)\nu + \frac{\nu}{q-1}} \Rightarrow W_\delta \succeq_{\text{deg}} V \, .$$

Condition for Degradation by Symmetric Channels

Theorem 2 (Degradation by Symmetric Channels)

For channel V with common input and output alphabet, and minimum probability $\nu = \min\{[V]_{i,j} : 1 \leq i,j \leq q\}$,

$$0 \leq \delta \leq \frac{\nu}{1-(q-1)\nu + \frac{\nu}{q-1}} \ \Rightarrow \ W_\delta \succeq_{\text{deg}} V \,.$$

Remark: Condition is tight when no further information about V known. For example, suppose

$$V = \begin{bmatrix} \nu & 1-(q-1)\nu & \nu & \cdots & \nu \\ 1-(q-1)\nu & \nu & \nu & \cdots & \nu \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-(q-1)\nu & \nu & \nu & \cdots & \nu \end{bmatrix}.$$

Then, $0 \leq \delta \leq \nu/\left(1-(q-1)\nu + \frac{\nu}{q-1}\right) \ \Leftrightarrow \ W_\delta \succeq_{\text{deg}} V$.

Additive Noise Channels

- Fix Abelian group $(\mathcal{X}, \oplus)$ with order q as alphabet.

Additive Noise Channels

- Fix Abelian group $(\mathcal{X}, \oplus)$ with order q as alphabet.
- Additive noise channel:

$$Y = X \oplus Z, \qquad X \perp\!\!\!\perp Z$$

where $X, Y, Z \in \mathcal{X}$ are input, output, and noise r.v.s.

Additive Noise Channels

- Fix Abelian group $(\mathcal{X}, \oplus)$ with order q as alphabet.
- Additive noise channel:

$$Y = X \oplus Z, \qquad X \perp\!\!\!\perp Z$$

  where $X, Y, Z \in \mathcal{X}$ are input, output, and noise r.v.s.
- Channel probabilities given by noise pmf P_Z:

$$\forall x, y \in \mathcal{X}, \;\; P_{Y|X}(y|x) = P_Z(-x \oplus y)\,.$$

Additive Noise Channels

- Fix Abelian group $(\mathcal{X}, \oplus)$ with order q as alphabet.
- Additive noise channel:

$$Y = X \oplus Z, \qquad X \perp\!\!\!\perp Z$$

  where $X, Y, Z \in \mathcal{X}$ are input, output, and noise r.v.s.
- Channel probabilities given by noise pmf P_Z:

$$\forall x, y \in \mathcal{X}, \;\; P_{Y|X}(y|x) = P_Z(-x \oplus y) \, .$$

- P_Y is convolution of P_X and P_Z:

$$\forall y \in \mathcal{X}, \;\; P_Y(y) = (P_X \star P_Z)(y) \triangleq \sum_{x \in \mathcal{X}} P_X(x) P_Z(-x \oplus y) \, .$$

Additive Noise Channels

- Fix Abelian group $(\mathcal{X}, \oplus)$ with order q as alphabet.
- Additive noise channel:

$$Y = X \oplus Z, \qquad X \perp\!\!\!\perp Z$$

  where $X, Y, Z \in \mathcal{X}$ are input, output, and noise r.v.s.
- Channel probabilities given by noise pmf P_Z:

$$\forall x, y \in \mathcal{X}, \;\; P_{Y|X}(y|x) = P_Z(-x \oplus y) \, .$$

- P_Y is convolution of P_X and P_Z:

$$\forall y \in \mathcal{X}, \;\; P_Y(y) = (P_X \star P_Z)(y) \triangleq \sum_{x \in \mathcal{X}} P_X(x) P_Z(-x \oplus y) \, .$$

- q-ary symmetric channel: $P_Z = \left(1 - \delta, \frac{\delta}{q-1}, \ldots, \frac{\delta}{q-1}\right)$ for $\delta \in [0, 1]$

$$(\cdot \star P_Z) = W_\delta$$

More Noisy and Degradation Regions

- Fix q-ary symmetric channel W_{δ} with $\delta \in [0, 1]$.

More Noisy and Degradation Regions

- Fix q-ary symmetric channel W_δ with $\delta \in [0,1]$.
- More noisy region of W_δ is

$$\textit{more-noisy}\,(W_\delta) \triangleq \{P_Z : W_\delta \succeq_{\mathsf{ln}} (\cdot \star P_Z)\}\,.$$

More Noisy and Degradation Regions

- Fix q-ary symmetric channel W_δ with $\delta \in [0,1]$.
- More noisy region of W_δ is

$$\textit{more-noisy}\,(W_\delta) \triangleq \{P_Z : W_\delta \succeq_{\mathsf{ln}} (\cdot \star P_Z)\}\,.$$

- Degradation region of W_δ is

$$\textit{degrade}\,(W_\delta) \triangleq \{P_Z : W_\delta \succeq_{\mathsf{deg}} (\cdot \star P_Z)\}\,.$$

Domination Structure of Additive Noise Channels

Theorem 3 (More Noisy and Degradation Regions)

For W_δ with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $q \geq 2$,

$$\begin{aligned} \textit{degrade}\,(W_\delta) &= \textit{co}\,(\text{rows of } W_\delta) \\ &\subseteq \textit{co}\,(\text{rows of } W_\delta \text{ and } W_\gamma) \\ &\subseteq \textit{more-noisy}\,(W_\delta) \\ &\subseteq \{P_Z : \|P_Z - \mathbf{u}\|_{\ell^2} \leq \|w_\delta - \mathbf{u}\|_{\ell^2}\} \end{aligned}$$

where $co\,(\cdot)$ – convex hull, $\gamma = (1-\delta)/\left(1 - \delta + \frac{\delta}{(q-1)^2}\right)$, $\mathbf{u}$ – uniform pmf, and w_δ – first row of W_δ.

Domination Structure of Additive Noise Channels

Theorem 3 (More Noisy and Degradation Regions)

For W_δ with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $q \geq 2$,

$$\begin{aligned} \textit{degrade}\,(W_\delta) &= \textit{co}\,(\text{rows of } W_\delta) \\ &\subseteq \textit{co}\,(\text{rows of } W_\delta \text{ and } W_\gamma) \\ &\subseteq \textit{more-noisy}\,(W_\delta) \\ &\subseteq \{P_Z : \|P_Z - \mathbf{u}\|_{\ell^2} \leq \|w_\delta - \mathbf{u}\|_{\ell^2}\} \end{aligned}$$

where $co\,(\cdot)$ – convex hull, $\gamma = (1-\delta)/\left(1-\delta+\frac{\delta}{(q-1)^2}\right)$, $\mathbf{u}$ – uniform pmf, and w_δ – first row of W_δ.

Furthermore, *more-noisy* (W_δ) is closed, convex, and invariant under permutations of $(\mathcal{X}, \oplus)$.

Illustration of the $q = 3$ case:

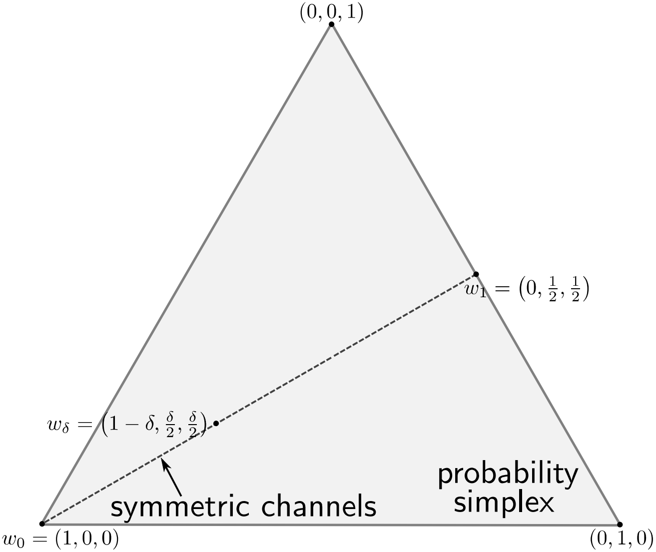
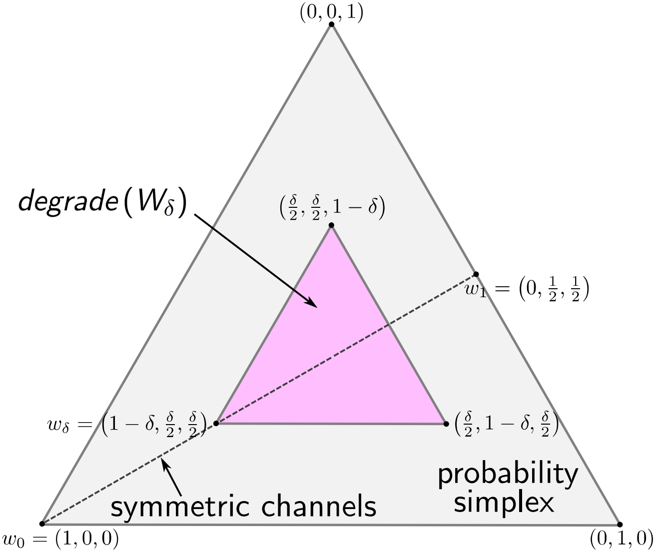
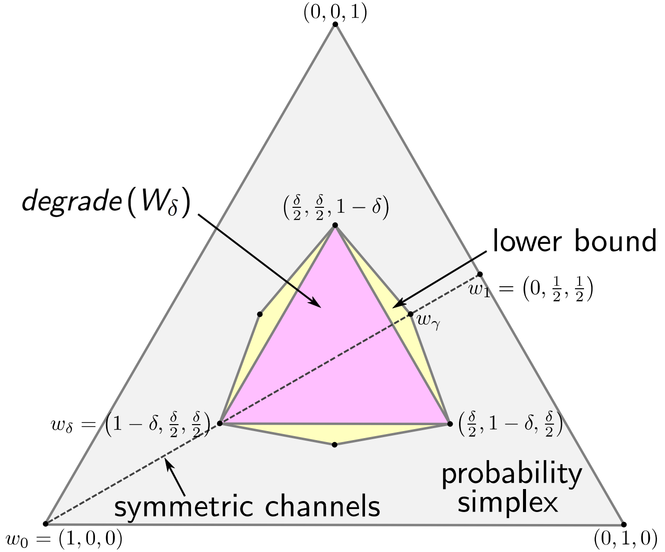


Illustration of the $q = 3$ case:



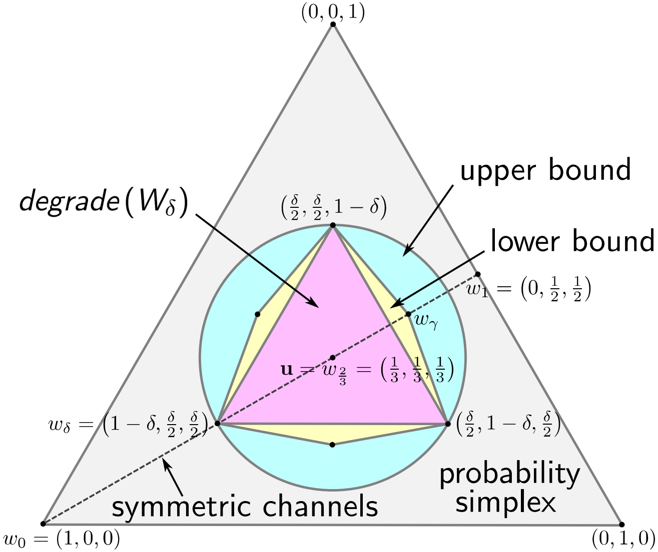
Domination Structure of Additive Noise Channels

Illustration of the $q = 3$ case:



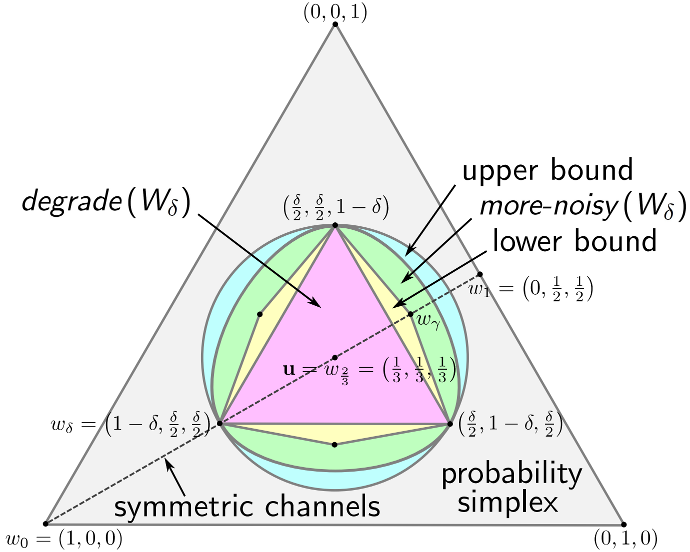
Domination Structure of Additive Noise Channels

Illustration of the $q = 3$ case:



Domination Structure of Additive Noise Channels

Illustration of the $q = 3$ case:



Domination Structure of Additive Noise Channels

Theorem 3 (More Noisy and Degradation Regions)

For W_δ with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $q \geq 2$,

$$\begin{aligned} \mathit{degrade}(W_\delta) &= \mathit{co}(\text{rows of } W_\delta) \\ &\subseteq \mathit{co}(\text{rows of } W_\delta \text{ and } W_\gamma) \\ &\subseteq \mathit{more\text{-}noisy}(W_\delta) \\ &\subseteq \{P_Z : \|P_Z - \mathbf{u}\|_{\ell^2} \leq \|w_\delta - \mathbf{u}\|_{\ell^2}\} \end{aligned}$$

where $\mathit{co}(\cdot)$ – convex hull, $\gamma = (1-\delta)/\left(1 - \delta + \frac{\delta}{(q-1)^2}\right)$, $\mathbf{u}$ – uniform pmf, and w_δ – first row of W_δ.

Furthermore, *more-noisy*(W_δ) is closed, convex, and invariant under permutations of $(\mathcal{X}, \oplus)$.

Proof Sketch: co(rows of W_δ and W_γ) $\subseteq$ $\textit{more-noisy}(W_\delta)$

- $\textit{more-noisy}(W_\delta)$ is convex, invariant under permutations of $(\mathcal{X}, \oplus)$
  $\Rightarrow$ suffices to prove $W_\delta \succeq_{\text{ln}} W_\gamma$.

Proof Sketch: $co\left(\text{rows of } W_\delta \text{ and } W_\gamma\right) \subseteq \textit{more-noisy}\left(W_\delta\right)$

- $\textit{more-noisy}\left(W_\delta\right)$ is convex, invariant under permutations of $(\mathcal{X}, \oplus)$
  $\Rightarrow$ suffices to prove $W_\delta \succeq_{\text{ln}} W_\gamma$.
- By Theorem 1,

$$\forall P_X, \quad W_\delta \operatorname{diag}(P_X W_\delta)^{-1} W_\delta^T \succeq_{\text{PSD}} W_\gamma \operatorname{diag}(P_X W_\gamma)^{-1} W_\gamma^T$$

Proof Sketch: $\textit{co}\,(\text{rows of } W_\delta \text{ and } W_\gamma) \subseteq \textit{more-noisy}\,(W_\delta)$

- $\textit{more-noisy}\,(W_\delta)$ is convex, invariant under permutations of $(\mathcal{X}, \oplus)$
  $\Rightarrow$ suffices to prove $W_\delta \succeq_{\text{ln}} W_\gamma$.

- By Theorem 1,

$$\forall P_X, \quad W_\delta \operatorname{diag}(P_X W_\delta)^{-1} W_\delta^T \succeq_{\text{PSD}} W_\gamma \operatorname{diag}(P_X W_\gamma)^{-1} W_\gamma^T$$
$$\Leftrightarrow \quad 1 \geq \|A\|_{\text{op}}$$

where $\|\cdot\|_{\text{op}}$ – operator norm, and A is symmetric PSD:

$$A \triangleq \operatorname{diag}(w_\gamma)^{-\frac{1}{2}} W_\gamma W_\delta^{-1} \operatorname{diag}(w_\delta) W_\delta^{-1} W_\gamma \operatorname{diag}(w_\gamma)^{-\frac{1}{2}}$$

with w_δ – first row of W_δ, and w_γ – first row of W_γ.

Proof Sketch: $co\,(\text{rows of } W_\delta \text{ and } W_\gamma) \subseteq \textit{more-noisy}\,(W_\delta)$

- $\textit{more-noisy}\,(W_\delta)$ is convex, invariant under permutations of $(\mathcal{X}, \oplus)$
  $\Rightarrow$ suffices to prove $W_\delta \succeq_{\text{ln}} W_\gamma$.

- By Theorem 1,

$$\forall P_X, \quad W_\delta \operatorname{diag}(P_X W_\delta)^{-1} W_\delta^T \succeq_{\text{PSD}} W_\gamma \operatorname{diag}(P_X W_\gamma)^{-1} W_\gamma^T$$
$$\Leftrightarrow \quad 1 \geq \|A\|_{\text{op}}$$

  where $\|\cdot\|_{\text{op}}$ – operator norm, and A is symmetric PSD:

$$A \triangleq \operatorname{diag}(w_\gamma)^{-\frac{1}{2}} W_\gamma W_\delta^{-1} \operatorname{diag}(w_\delta) W_\delta^{-1} W_\gamma \operatorname{diag}(w_\gamma)^{-\frac{1}{2}}$$

  with w_δ – first row of W_δ, and w_γ – first row of W_γ.

- A has left eigenvector $\sqrt{w_\gamma} > 0$ with eigenvalue 1:

$$\sqrt{w_\gamma} A = \sqrt{w_\gamma}\,.$$

Proof Sketch: $co\,($rows of W_δ and $W_\gamma) \subseteq$ $\textit{more-noisy}\,(W_\delta)$

- $A \geq 0$ (entry-wise) $\Rightarrow$ largest eigenvalue of A is 1 by Perron-Frobenius theorem, because $\sqrt{w_\gamma} > 0$.

Proof Sketch: $co(\text{rows of } W_\delta \text{ and } W_\gamma) \subseteq \textit{more-noisy}(W_\delta)$

- $A \geq 0$ (entry-wise) $\Rightarrow$ largest eigenvalue of A is 1 by Perron-Frobenius theorem, because $\sqrt{w_\gamma} > 0$.
- Since A – symmetric PSD, $A \succeq 0 \Rightarrow \|A\|_{\text{op}} \leq 1$.
  $\Rightarrow$ Suffices to prove $A \geq 0$.

Proof Sketch: co (rows of W_δ and W_γ) $\subseteq$ $more\text{-}noisy(W_\delta)$

- $A \geq 0$ (entry-wise) $\Rightarrow$ largest eigenvalue of A is 1 by Perron-Frobenius theorem, because $\sqrt{w_\gamma} > 0$.
- Since A – symmetric PSD, $A \geq 0 \Rightarrow \|A\|_{\text{op}} \leq 1$.
  $\Rightarrow$ Suffices to prove $A \geq 0$.
- Verify that:

$$\min_{i,j} [A]_{i,j} \geq 0 \quad \Leftrightarrow \quad \delta \leq \gamma \leq \frac{1-\delta}{1-\delta+\frac{\delta}{(q-1)^2}} .$$

■

Outline

Log-Sobolev Inequalities

- Consider irreducible Markov chain V with uniform stationary pmf $\mathbf{u}$ on state space of size q.

Log-Sobolev Inequalities

- Consider irreducible Markov chain V with uniform stationary pmf $\mathbf{u}$ on state space of size q.
- Dirichlet form $\mathcal{E}_V : \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}^+$

$$\mathcal{E}_V(f, f) \triangleq \frac{1}{q} f^T \left(I - \frac{V + V^T}{2} \right) f$$

Log-Sobolev Inequalities

- Consider irreducible Markov chain V with uniform stationary pmf $\mathbf{u}$ on state space of size q.
- Dirichlet form $\mathcal{E}_V : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}^+$

$$\mathcal{E}_V(f,f) \triangleq \frac{1}{q} f^T \left(I - \frac{V + V^T}{2} \right) f$$

- Log-Sobolev inequality with constant $\alpha \in \mathbb{R}^+$:
  For every $f \in \mathbb{R}^q$ such that $f^T f = q$,

$$D\left(f^2\mathbf{u} \,||\, \mathbf{u}\right) = \frac{1}{q}\sum_{i=1}^{q} f_i^2 \log\left(f_i^2\right) \leq \frac{1}{\alpha}\mathcal{E}_V(f,f) \, .$$

Log-Sobolev Inequalities

- Consider irreducible Markov chain V with uniform stationary pmf $\mathbf{u}$ on state space of size q.
- Dirichlet form $\mathcal{E}_V : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}^+$

$$\mathcal{E}_V(f,f) \triangleq \frac{1}{q} f^T \left(I - \frac{V + V^T}{2} \right) f$$

- Log-Sobolev inequality with constant $\alpha \in \mathbb{R}^+$:
  For every $f \in \mathbb{R}^q$ such that $f^T f = q$,

$$D\left(f^2 \mathbf{u} \,||\, \mathbf{u}\right) = \frac{1}{q} \sum_{i=1}^{q} f_i^2 \log\left(f_i^2\right) \leq \frac{1}{\alpha} \mathcal{E}_V(f,f) \, .$$

- Log-Sobolev constant – largest α satisfying log-Sobolev inequality.

Comparison of Dirichlet Forms

- Standard Dirichlet form:

$$\mathcal{E}_{\text{std}}(f, f) \triangleq \mathbb{VAR}_{\mathbf{u}}(f) = \sum_{i=1}^{q} \frac{1}{q} f_i^2 - \left(\sum_{i=1}^{q} \frac{1}{q} f_i \right)^2$$

Comparison of Dirichlet Forms

- For standard Dirichlet form, $\mathcal{E}_{\mathsf{std}}(f, f) \triangleq \mathbb{VAR}_{\mathbf{u}}(f)$, log-Sobolev constant known [Diaconis-Saloff-Coste 1996]:

$$D\left(f^2\mathbf{u} \,||\, \mathbf{u}\right) \leq \frac{q \log(q-1)}{(q-2)} \mathcal{E}_{\mathsf{std}}(f, f)$$

for all $f \in \mathbb{R}^q$ with $f^T f = q$.

Comparison of Dirichlet Forms

- For standard Dirichlet form, $\mathcal{E}_{\mathsf{std}}(f,f) \triangleq \mathbb{VAR}_{\mathbf{u}}(f)$, log-Sobolev constant known [Diaconis-Saloff-Coste 1996]:

$$D\left(f^2\mathbf{u}\,||\,\mathbf{u}\right) \leq \frac{q\log(q-1)}{(q-2)}\,\mathcal{E}_{\mathsf{std}}(f,f)$$

for all $f \in \mathbb{R}^q$ with $f^T f = q$.

Theorem 4 (Domination of Dirichlet Forms)

For channels W_δ and V with $\delta \in \left[0, \frac{q-1}{q}\right]$ and stationary pmf $\mathbf{u}$,

$$W_\delta \succeq_{\mathsf{ln}} V \;\;\Rightarrow\;\; \mathcal{E}_V \geq \frac{q\delta}{q-1}\,\mathcal{E}_{\mathsf{std}} \text{ pointwise}\,.$$

Comparison of Dirichlet Forms

- For standard Dirichlet form, $\mathcal{E}_{\text{std}}(f,f) \triangleq \mathbb{VAR}_{\mathbf{u}}(f)$, log-Sobolev constant known [Diaconis-Saloff-Coste 1996]:

$$D\left(f^2\mathbf{u}\,||\,\mathbf{u}\right) \leq \frac{q\log(q-1)}{(q-2)}\,\mathcal{E}_{\text{std}}(f,f)$$

for all $f \in \mathbb{R}^q$ with $f^T f = q$.

Theorem 4 (Domination of Dirichlet Forms)

For channels W_δ and V with $\delta \in \left[0, \frac{q-1}{q}\right]$ and stationary pmf $\mathbf{u}$,

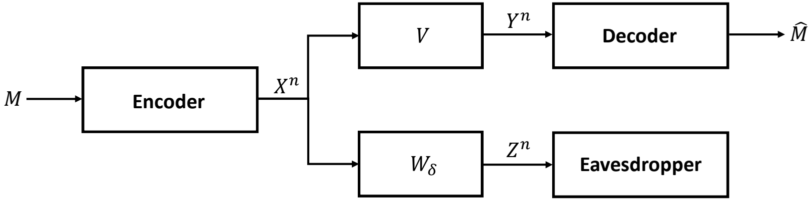
$$W_\delta \succeq_{\text{ln}} V \;\;\Rightarrow\;\; \mathcal{E}_V \geq \frac{q\delta}{q-1}\,\mathcal{E}_{\text{std}} \;\; \text{pointwise}\,.$$

- $W_\delta \succeq_{\text{ln}} V \Rightarrow$ log-Sobolev inequality for V,

$$D(f^2\mathbf{u}\,||\,\mathbf{u}) \leq \frac{(q-1)\log(q-1)}{\delta\,(q-2)}\,\mathcal{E}_V(f,f)$$

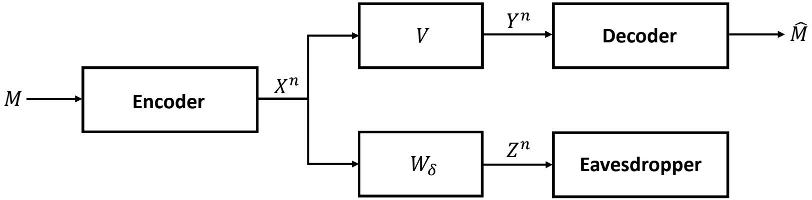
for every $f \in \mathbb{R}^q$ satisfying $f^T f = q$.

Interpretation via Wyner's Wiretap Channel



- V – *main channel*, W_δ – *eavesdropper channel*

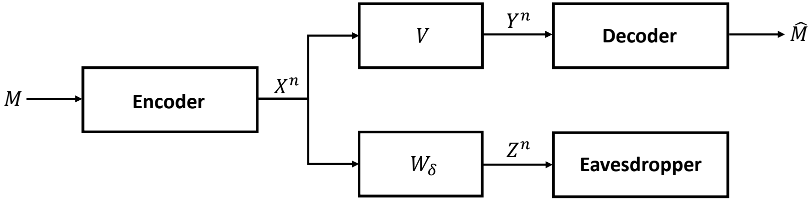
Interpretation via Wyner's Wiretap Channel



- V – *main channel*, W_δ – *eavesdropper channel*
- Secrecy capacity – maximum rate to legal receiver such that $\mathbb{P}(M \neq \hat{M}) \to 0$ and $\frac{1}{n} I(M; Z^n) \to 0$

$$C_S = \max_{P_{U,X}} I(U;Y) - I(U;Z) \qquad \text{[Csiszár-Körner 1978]}$$

Interpretation via Wyner's Wiretap Channel

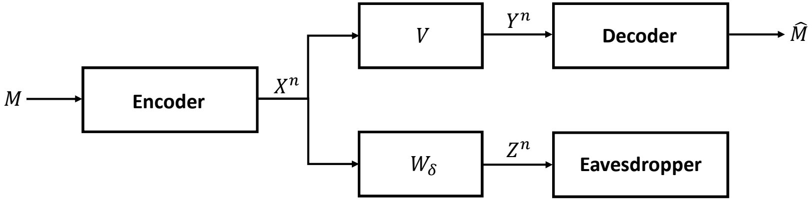


- *V – main channel, W_δ – eavesdropper channel*
- Secrecy capacity – maximum rate to legal receiver such that $\mathbb{P}(M \neq \hat{M}) \to 0$ and $\frac{1}{n} I(M; Z^n) \to 0$

$$C_S = \max_{P_{U,X}} I(U;Y) - I(U;Z) \qquad \text{[Csiszár-Körner 1978]}$$

- **Prop** [Csiszár-Körner 1978]**:** $C_S = 0 \Leftrightarrow W_\delta \succeq_{\text{ln}} V$.

Interpretation via Wyner's Wiretap Channel



- *V – main channel, W_δ – eavesdropper channel*
- Secrecy capacity – maximum rate to legal receiver such that $\mathbb{P}(M \neq \hat{M}) \to 0$ and $\frac{1}{n} I(M; Z^n) \to 0$

$$C_S = \max_{P_{U,X}} I(U;Y) - I(U;Z) \qquad \text{[Csiszár-Körner 1978]}$$

- **Prop** [Csiszár-Körner 1978]**:** $C_S = 0 \Leftrightarrow W_\delta \succeq_{\text{ln}} V$.
- Finding maximally noisy $W_\delta \succeq_{\text{ln}} V$ establishes minimal noise on $P_{Z|X}$ so that secret communication feasible.

Thank You!