New Results on the Less Noisy Preorder over Channels

Anuran Makur and Yury Polyanskiy

EECS Department, Massachusetts Institute of Technology

10 November 2016

Outline

1. Introduction
   - Preliminaries
   - Channel Preorders in Information Theory
   - Relation to Strong Data Processing Inequalities
   - Symmetric Channels
   - Main Questions
2. Equivalent Characterizations of Less Noisy Preorder
3. Condition for Domination by a Symmetric Channel
4. Comparison of Additive Noise Channels
5. Less Noisy Domination and Log-Sobolev Inequalities

Preliminiaries

- $\mathbb{R}_{\text{sto}}^{q\times r}$ = set of $q \times r$ *row* stochastic matrices
- $\mathcal{P}_q$ = probability simplex of *row* vectors in $\mathbb{R}^q$

Preliminiaries

- $\mathbb{R}_{\mathsf{sto}}^{q \times r}$ = set of $q \times r$ *row* stochastic matrices
- $\mathcal{P}_q$ = probability simplex of *row* vectors in $\mathbb{R}^q$
- Alphabets $\mathcal{X} = [q] \triangleq \{0, 1, \ldots, q-1\}$ and $\mathcal{Y} = [r]$
- Input and output random variables $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$

Preliminiaries

- $\mathbb{R}^{q\times r}_{\text{sto}}$ = set of $q \times r$ *row* stochastic matrices
- $\mathcal{P}_q$ = probability simplex of *row* vectors in $\mathbb{R}^q$
- Alphabets $\mathcal{X} = [q] \triangleq \{0, 1, \ldots, q-1\}$ and $\mathcal{Y} = [r]$
- Input and output random variables $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$
- A channel is the set of conditional distributions $W_{Y|X}$ that associates each $x \in \mathcal{X}$ with a conditional pmf $W_{Y|X}(\cdot|x) \in \mathcal{P}_r$
- Represent a channel $W_{Y|X}$ with a stochastic matrix $W \in \mathbb{R}^{q\times r}_{\text{sto}}$ so that $P_Y = P_X W$

Channel Preorders in Information Theory

- **Less Noisy** [KM77]
  $W \in \mathbb{R}^{q \times r}_{\text{sto}}$ is less noisy than $V \in \mathbb{R}^{q \times s}_{\text{sto}}$:

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

  if $D(P_X W || Q_X W) \geq D(P_X V || Q_X V)$ for every $P_X, Q_X \in \mathcal{P}_q$.

Channel Preorders in Information Theory

- **Less Noisy** [KM77]
  $W \in \mathbb{R}^{q \times r}_{\mathsf{sto}}$ is less noisy than $V \in \mathbb{R}^{q \times s}_{\mathsf{sto}}$:

$$W \succeq_{\mathsf{ln}} V$$

  if $D(P_X W || Q_X W) \geq D(P_X V || Q_X V)$ for every $P_X, Q_X \in \mathcal{P}_q$.

- **Degradation** [Ber73]
  $V \in \mathbb{R}^{q \times s}_{\mathsf{sto}}$ is a degraded version of $W \in \mathbb{R}^{q \times r}_{\mathsf{sto}}$:

$$W \succeq_{\mathsf{deg}} V$$

  if $V = WA$ for some channel $A \in \mathbb{R}^{r \times s}_{\mathsf{sto}}$.

Channel Preorders in Information Theory

- **Less Noisy** [KM77]
  $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ is less noisy than $V \in \mathbb{R}_{\text{sto}}^{q \times s}$:

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

  if $D(P_X W || Q_X W) \geq D(P_X V || Q_X V)$ for every $P_X, Q_X \in \mathcal{P}_q$.

- **Degradation** [Ber73]
  $V \in \mathbb{R}_{\text{sto}}^{q \times s}$ is a degraded version of $W \in \mathbb{R}_{\text{sto}}^{q \times r}$:

$$W \succeq_{\text{deg}} V$$

  if $V = WA$ for some channel $A \in \mathbb{R}_{\text{sto}}^{r \times s}$.

Observation: $W \succeq_{\text{deg}} V \Rightarrow W \succeq_{\text{ln}} V$

Less Noisy Preorder and Strong Data Processing

Data Processing Inequality: For any channel $W \in \mathbb{R}_{\text{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\mathsf{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0, 1]$ is a channel dependent contraction coefficient.

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\text{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0,1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\text{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1 - \eta$.

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\mathsf{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0, 1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\mathsf{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1 - \eta$.

$$E_{1-\eta} \succeq_{\mathsf{ln}} W$$

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\text{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0, 1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\text{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1 - \eta$.

$$E_{1-\eta} \succeq_{\text{ln}} W$$
$$\Leftrightarrow \; \forall P_X, Q_X, \; D(P_X E_{1-\eta} || Q_X E_{1-\eta}) \geq D(P_X W || Q_X W)$$

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\mathsf{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0, 1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\mathsf{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1 - \eta$.

$$\begin{aligned} & E_{1-\eta} \succeq_{\mathsf{ln}} W \\ \Leftrightarrow \; & \forall P_X, Q_X, \; D(P_X E_{1-\eta} || Q_X E_{1-\eta}) \geq D(P_X W || Q_X W) \\ \Leftrightarrow \; & \forall P_X, Q_X, \; D(\eta P_X + (1-\eta)\delta_{\mathsf{e}} || \eta Q_X + (1-\eta)\delta_{\mathsf{e}}) \geq D(P_X W || Q_X W) \end{aligned}$$

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\text{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0,1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\text{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1-\eta$.

$$\begin{aligned}
& E_{1-\eta} \succeq_{\text{ln}} W \\
\Leftrightarrow \; & \forall P_X, Q_X, \; D(P_X E_{1-\eta} || Q_X E_{1-\eta}) \geq D(P_X W || Q_X W) \\
\Leftrightarrow \; & \forall P_X, Q_X, \; D(\eta P_X + (1-\eta)\delta_{\text{e}} || \eta Q_X + (1-\eta)\delta_{\text{e}}) \geq D(P_X W || Q_X W) \\
\Leftrightarrow \; & \forall P_X, Q_X, \; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)
\end{aligned}$$

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\text{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0,1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\text{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1 - \eta$.

$$\begin{aligned} & E_{1-\eta} \succeq_{\text{ln}} W \\ \Leftrightarrow \; & \forall P_X, Q_X, \; D(P_X E_{1-\eta} || Q_X E_{1-\eta}) \geq D(P_X W || Q_X W) \\ \Leftrightarrow \; & \forall P_X, Q_X, \; D(\eta P_X + (1-\eta)\delta_{\text{e}} || \eta Q_X + (1-\eta)\delta_{\text{e}}) \geq D(P_X W || Q_X W) \\ \Leftrightarrow \; & \forall P_X, Q_X, \; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W) \end{aligned}$$

SDPI = $\succeq_{\text{ln}}$ domination by erasure channel

Less Noisy Preorder and Strong Data Processing

Strong Data Processing Inequality [AG76]: For any channel $W \in \mathbb{R}_{\text{sto}}^{q \times r}$,

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)$$

where $\eta \in [0, 1]$ is a channel dependent contraction coefficient.

Relation to Less Noisy Preorder [PW17]:
Let $E_{1-\eta} \in \mathbb{R}_{\text{sto}}^{q \times (q+1)}$ be an erasure channel with erasure probability $1 - \eta$.

$$\begin{aligned}
& E_{1-\eta} \succeq_{\text{ln}} W \\
\Leftrightarrow \; & \forall P_X, Q_X, \; D(P_X E_{1-\eta} || Q_X E_{1-\eta}) \geq D(P_X W || Q_X W) \\
\Leftrightarrow \; & \forall P_X, Q_X, \; D(\eta P_X + (1-\eta)\delta_{\text{e}} || \eta Q_X + (1-\eta)\delta_{\text{e}}) \geq D(P_X W || Q_X W) \\
\Leftrightarrow \; & \forall P_X, Q_X, \; \eta D(P_X || Q_X) \geq D(P_X W || Q_X W)
\end{aligned}$$

SDPI $= \succeq_{\text{ln}}$ domination by erasure channel

When does a q-ary symmetric channel dominate another channel?

Symmetric Channels

Definition (Symmetric Channel)

Given the alphabet $\mathcal{X} = \mathcal{Y} = [q]$, the q-ary symmetric channel is given by the stochastic matrix:

$$W_\delta \triangleq \begin{bmatrix} 1-\delta & \frac{\delta}{q-1} & \cdots & \frac{\delta}{q-1} \\ \frac{\delta}{q-1} & 1-\delta & \cdots & \frac{\delta}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta}{q-1} & \frac{\delta}{q-1} & \cdots & 1-\delta \end{bmatrix} \in \mathbb{R}_{\mathsf{sto}}^{q \times q}$$

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

Symmetric Channels

Definition (Symmetric Channel)

Given the alphabet $\mathcal{X} = \mathcal{Y} = [q]$, the *q-ary symmetric channel* is given by the stochastic matrix:

$$W_\delta \triangleq \begin{bmatrix} 1-\delta & \frac{\delta}{q-1} & \cdots & \frac{\delta}{q-1} \\ \frac{\delta}{q-1} & 1-\delta & \cdots & \frac{\delta}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta}{q-1} & \frac{\delta}{q-1} & \cdots & 1-\delta \end{bmatrix} \in \mathbb{R}_{\text{sto}}^{q \times q}$$

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

Properties:

- $\{W_\delta \in \mathbb{R}_{\text{sym}}^{q \times q} : \delta \in \mathbb{R}\}$ are symmetric, circulant, doubly stochastic matrices that are jointly diagonalized by the DFT matrix.

Symmetric Channels

Definition (Symmetric Channel)

Given the alphabet $\mathcal{X} = \mathcal{Y} = [q]$, the q-ary symmetric channel is given by the stochastic matrix:

$$W_\delta \triangleq \begin{bmatrix} 1-\delta & \frac{\delta}{q-1} & \cdots & \frac{\delta}{q-1} \\ \frac{\delta}{q-1} & 1-\delta & \cdots & \frac{\delta}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta}{q-1} & \frac{\delta}{q-1} & \cdots & 1-\delta \end{bmatrix} \in \mathbb{R}_{\mathsf{sto}}^{q\times q}$$

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

Properties:

- $\{W_\delta \in \mathbb{R}_{\mathsf{sym}}^{q\times q} : \delta \in \mathbb{R}\}$ are symmetric, circulant, doubly stochastic matrices that are jointly diagonalized by the DFT matrix.
- $\{W_\delta \in \mathbb{R}_{\mathsf{sym}}^{q\times q} : \delta \in \mathbb{R}\backslash\{\frac{q-1}{q}\}\}$ with the operation of matrix multiplication is an Abelian group.

Main Questions

1. Can we test $\succeq_{\ln}$ using a different divergence measure?

Main Questions

1. Can we test $\succeq_{\mathsf{ln}}$ using a different divergence measure?
   Yes

Main Questions

1. Can we test $\succeq_{\text{ln}}$ using a different divergence measure?
   Yes

2. Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$, is there a simple sufficient condition for $W_\delta \succeq_{\text{ln}} V$?

Main Questions

1. Can we test $\succeq_{\text{ln}}$ using a different divergence measure?
   Yes

2. Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$, is there a simple sufficient condition for $W_\delta \succeq_{\text{ln}} V$?
   Yes (via degradation $W_\delta \succeq_{\text{deg}} V$)

Main Questions

1. Can we test $\succeq_{\text{ln}}$ using a different divergence measure?
   Yes

2. Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$, is there a simple sufficient condition for $W_\delta \succeq_{\text{ln}} V$?
   Yes (via degradation $W_\delta \succeq_{\text{deg}} V$)

3. Can we say anything stronger about $\succeq_{\text{ln}}$ domination by a symmetric channel when $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ is an additive noise channel?

Main Questions

1. Can we test $\succeq_{\text{ln}}$ using a different divergence measure?
   Yes

2. Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$, is there a simple sufficient condition for $W_\delta \succeq_{\text{ln}} V$?
   Yes (via degradation $W_\delta \succeq_{\text{deg}} V$)

3. Can we say anything stronger about $\succeq_{\text{ln}}$ domination by a symmetric channel when $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ is an additive noise channel?
   Yes (picture!)

Main Questions

1. Can we test $\succeq_{\text{ln}}$ using a different divergence measure?
   Yes

2. Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$, is there a simple sufficient condition for $W_\delta \succeq_{\text{ln}} V$?
   Yes (via degradation $W_\delta \succeq_{\text{deg}} V$)

3. Can we say anything stronger about $\succeq_{\text{ln}}$ domination by a symmetric channel when $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ is an additive noise channel?
   Yes (picture!)

4. Why do we care about $\succeq_{\text{ln}}$ domination by symmetric channels?

Main Questions

1. Can we test $\succeq_{\text{ln}}$ using a different divergence measure?
   Yes

2. Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$, is there a simple sufficient condition for $W_\delta \succeq_{\text{ln}} V$?
   Yes (via degradation $W_\delta \succeq_{\text{deg}} V$)

3. Can we say anything stronger about $\succeq_{\text{ln}}$ domination by a symmetric channel when $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ is an additive noise channel?
   Yes (picture!)

4. Why do we care about $\succeq_{\text{ln}}$ domination by symmetric channels?
   Dirichlet form domination $\Rightarrow$ Log-Sobolev Inequality

Outline

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right).$$

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right).$$

Recall that for any two pmfs $P_X, Q_X \in \mathcal{P}_q$, their χ^2-divergence is given by:

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

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q\times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q\times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right).$$

Proof: ($\Rightarrow$) Recall that for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$ [PW16]:

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

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \ \ \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V).$$

Proof: ($\Rightarrow$) Recall that for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$ [PW16]:

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

So, for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$, $W \succeq_{\text{ln}} V$ implies:

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

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q\times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q\times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right).$$

Proof: ($\Rightarrow$) Recall that for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$ [PW16]:

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

So, for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$, $W \succeq_{\text{ln}} V$ implies:

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

after taking limits.

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V) .$$

Proof: ($\Rightarrow$) Recall that for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$ [PW16]:

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

So, for any $P_X \in \mathcal{P}_q$ and $Q_X \in \mathcal{P}_q^{\circ}$, $W \succeq_{\text{ln}} V$ 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. Continuity completes this direction.

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2\left(P_X W || Q_X W\right) \geq \chi^2\left(P_X V || Q_X V\right).$$

Proof: ($\Leftarrow$) Recall that for any $P_X, Q_X \in \mathcal{P}_q$ [PW17]:

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

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$.

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q\times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q\times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V) .$$

Proof: ($\Leftarrow$) Recall that for any $P_X, Q_X \in \mathcal{P}_q$ [PW17]:

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

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$.
Hence, for every $P_X, Q_X \in \mathcal{P}_q$, we have:

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

χ^2-Divergence Characterization of $\succeq_{\text{ln}}$

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

Given the channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, $W \succeq_{\text{ln}} V$ if and only if:

$$\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V).$$

Proof: ($\Leftarrow$) Recall that for any $P_X, Q_X \in \mathcal{P}_q$ [PW17]:

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

where $Q_X^t = \frac{t}{1+t} P_X + \frac{1}{t+1} Q_X$ for $t \in [0, \infty)$.
Hence, for every $P_X, Q_X \in \mathcal{P}_q$, we have:

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

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

which means that $W \succeq_{\text{ln}} V$.

Löwner and Spectral Characterizations of $\succeq_{\text{ln}}$

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

Given channels $W \in \mathbb{R}_{\text{sto}}^{q \times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q \times s}$, the following are equivalent:

- $W \succeq_{\text{ln}} V$
- $\forall P_X, Q_X \in \mathcal{P}_q, \;\; \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V)$
- $\forall P_X \in \mathcal{P}_q^{\circ}, \;\; W \text{diag}(P_X W)^{-1} W^T \succeq_{\text{PSD}} V \text{diag}(P_X V)^{-1} V^T$
- $\forall P_X \in \mathcal{P}_q^{\circ}, \;\; \rho\left(\left(W \text{diag}(P_X W)^{-1} W^T\right)^{\dagger} V \text{diag}(P_X V)^{-1} V^T\right) = 1$

where $X^{\dagger}$ denotes the *Moore-Penrose pseudoinverse* of any matrix X, and $\rho(X)$ denotes the *spectral radius* of any square matrix X.

Löwner and Spectral Characterizations of $\succeq_{\text{ln}}$

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

Given channels $W \in \mathbb{R}_{\text{sto}}^{q\times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q\times s}$, the following are equivalent:

- $W \succeq_{\text{ln}} V$
- $\forall P_X, Q_X \in \mathcal{P}_q,\ \ \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V)$
- $\forall P_X \in \mathcal{P}_q^{\circ},\ \ W \text{diag}(P_X W)^{-1} W^T \succeq_{\text{PSD}} V \text{diag}(P_X V)^{-1} V^T$
- $\forall P_X \in \mathcal{P}_q^{\circ},\ \ \rho\left(\left(W \text{diag}(P_X W)^{-1} W^T\right)^{\dagger} V \text{diag}(P_X V)^{-1} V^T\right) = 1$

where $X^{\dagger}$ denotes the *Moore-Penrose pseudoinverse* of any matrix X, and $\rho(X)$ denotes the *spectral radius* of any square matrix X.

Löwner Characterization: For every $P_X \in \mathcal{P}_q$ and every $Q_X \in \mathcal{P}_q^{\circ}$,

$$\chi^2(P_X W || Q_X W) = (P_X - Q_X)\, W \text{diag}(Q_X W)^{-1} W^T (P_X - Q_X)^T .$$

Löwner and Spectral Characterizations of $\succeq_{\text{ln}}$

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

Given channels $W \in \mathbb{R}_{\text{sto}}^{q\times r}$ and $V \in \mathbb{R}_{\text{sto}}^{q\times s}$, the following are equivalent:

- $W \succeq_{\text{ln}} V$
- $\forall P_X, Q_X \in \mathcal{P}_q,\ \ \chi^2(P_X W || Q_X W) \geq \chi^2(P_X V || Q_X V)$
- $\forall P_X \in \mathcal{P}_q^{\circ},\ \ W \text{diag}(P_X W)^{-1} W^T \succeq_{\text{PSD}} V \text{diag}(P_X V)^{-1} V^T$
- $\forall P_X \in \mathcal{P}_q^{\circ},\ \ \rho\left(\left(W \text{diag}(P_X W)^{-1} W^T\right)^{\dagger} V \text{diag}(P_X V)^{-1} V^T\right) = 1$

where $X^{\dagger}$ denotes the *Moore-Penrose pseudoinverse* of any matrix X, and $\rho(X)$ denotes the *spectral radius* of any square matrix X.

Löwner Characterization: For every $P_X \in \mathcal{P}_q$ and every $Q_X \in \mathcal{P}_q^{\circ}$,

$$\chi^2(P_X W || Q_X W) = (P_X - Q_X)\, W \text{diag}(Q_X W)^{-1} W^T (P_X - Q_X)^T .$$

Spectral Characterization: Exercise in matrix analysis.

Outline

Condition for Degradation by Symmetric Channels

Theorem (Degradation by Symmetric Channels)

Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ with $q \geq 2$ and minimum probability $\nu = \min\{[V]_{i,j} : 1 \leq i,j \leq q\}$, we have:

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

where $W_\delta \in \mathbb{R}_{\text{sto}}^{q \times q}$ is a symmetric channel.

Proof Sketch

Consider the channels $W_{(q-1)\nu}, V \in \mathbb{R}^{q \times q}_{\text{sto}}$:

$$W_{(q-1)\nu} = \begin{bmatrix} - & w_1 & - \\ & \vdots & \\ - & w_q & - \end{bmatrix}, \qquad V = \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_q & - \end{bmatrix}$$

where $w_i = (\nu, \ldots, \nu, 1 - (q-1)\nu, \nu, \ldots, \nu)$ has $1 - (q-1)\nu$ in the ith position, and V has minimum entry ν.

Proof Sketch

Consider the channels $W_{(q-1)\nu}, V \in \mathbb{R}_{\text{sto}}^{q \times q}$:

$$W_{(q-1)\nu} = \begin{bmatrix} - & w_1 & - \\ & \vdots & \\ - & w_q & - \end{bmatrix}, \qquad V = \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_q & - \end{bmatrix}$$

where $w_i = (\nu, \dots, \nu, 1-(q-1)\nu, \nu, \dots, \nu)$ has $1-(q-1)\nu$ in the ith position, and V has minimum entry ν.

Using majorization: $\forall i \in \{1, \dots, q\}$, $v_i = \sum_{j=1}^{q} p_{i,j} \, w_j$

for some convex weights $p_{i,j} \geq 0$ such that $\sum_{j=1}^{q} p_{i,j} = 1$.

Proof Sketch

Consider the channels $W_{(q-1)\nu}, V \in \mathbb{R}_{\text{sto}}^{q \times q}$:

$$W_{(q-1)\nu} = \begin{bmatrix} - & w_1 & - \\ & \vdots & \\ - & w_q & - \end{bmatrix}, \qquad V = \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_q & - \end{bmatrix}$$

where $w_i = (\nu, \dots, \nu, 1-(q-1)\nu, \nu, \dots, \nu)$ has $1-(q-1)\nu$ in the ith position, and V has minimum entry ν.

Using majorization: $\forall i \in \{1, \dots, q\}$, $v_i = \sum_{j=1}^{q} p_{i,j}\, w_j$

for some convex weights $p_{i,j} \geq 0$ such that $\sum_{j=1}^{q} p_{i,j} = 1$.
Stack rows of V and observe that:

$$V = \sum_{1 \leq j_1, \dots, j_q \leq q} \left(\prod_{i=1}^{q} p_{i,j_i} \right) \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix}$$

where $\left\{ \prod_{i=1}^{q} p_{i,j_i} : 1 \leq j_1, \dots, j_q \leq q \right\}$ form a product pmf.

Proof Sketch

Suffices to find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$\exists M_{j_1,\ldots,j_q} \in \mathbb{R}_{\text{sto}}^{q\times q}, \quad W_\delta M_{j_1,\ldots,j_q} = \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix}$$

Proof Sketch

Suffices to find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$\exists M_{j_1,\ldots,j_q} \in \mathbb{R}^{q \times q}_{\mathsf{sto}}, \quad W_\delta M_{j_1,\ldots,j_q} = \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix}$$

or equivalently, find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$W_\delta^{-1} \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix} \in \mathbb{R}^{q \times q}_{\mathsf{sto}}.$$

Proof Sketch

Suffices to find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$\exists M_{j_1,\ldots,j_q} \in \mathbb{R}_{\mathsf{sto}}^{q\times q}, \quad W_\delta M_{j_1,\ldots,j_q} = \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix}$$

or equivalently, find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$W_\delta^{-1} \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix} \in \mathbb{R}_{\mathsf{sto}}^{q\times q}.$$

The rows of such matrices always sum to unity.

Proof Sketch

Suffices to find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$\exists M_{j_1,\ldots,j_q} \in \mathbb{R}^{q\times q}_{\mathsf{sto}}, \quad W_\delta M_{j_1,\ldots,j_q} = \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix}$$

or equivalently, find $\delta \in \left[0, \frac{q-1}{q}\right]$ such that for every $1 \leq j_1, \ldots, j_q \leq q$:

$$W_\delta^{-1} \begin{bmatrix} - & w_{j_1} & - \\ & \vdots & \\ - & w_{j_q} & - \end{bmatrix} \in \mathbb{R}^{q\times q}_{\mathsf{sto}}.$$

The rows of such matrices always sum to unity. Finally, we require the minimum possible entry of such matrices to be non-negative, which gives:

$$\delta \leq \frac{\nu}{1 - (q-1)\,\nu + \frac{\nu}{q-1}}.$$

Tightness of Condition for Degradation

Theorem (Degradation by Symmetric Channels)

Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ with $q \geq 2$ and minimum probability $\nu = \min\{[V]_{i,j} : 1 \leq i, j \leq q\}$, we have:

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

Tightness of Condition for Degradation

Theorem (Degradation by Symmetric Channels)

Given a channel $V \in \mathbb{R}_{\text{sto}}^{q \times q}$ with $q \geq 2$ and minimum probability $\nu = \min\{[V]_{i,j} : 1 \leq i,j \leq q\}$, we have:

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

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

$$V = \begin{bmatrix} - & w_2 & - \\ - & w_1 & - \\ & \vdots & \\ - & w_1 & - \end{bmatrix} \in \mathbb{R}_{\text{sto}}^{q \times q}.$$

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

Outline

Additive Noise Channels

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

Additive Noise Channels

- Fix a finite Abelian group $(\mathcal{X}, \oplus)$ with order q as the alphabet.
- An additive noise channel if defined by:

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

where $X, Y, Z \in \mathcal{X}$ are the input, output, and noise random variables.

Additive Noise Channels

- Fix a finite Abelian group $(\mathcal{X}, \oplus)$ with order q as the alphabet.
- An additive noise channel if defined by:

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

  where $X, Y, Z \in \mathcal{X}$ are the input, output, and noise random variables.
- It is characterized by a noise pmf $P_Z \in \mathcal{P}_q$.

Additive Noise Channels

- Fix a finite Abelian group $(\mathcal{X}, \oplus)$ with order q as the alphabet.
- An additive noise channel if defined by:

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

where $X, Y, Z \in \mathcal{X}$ are the input, output, and noise random variables.
- It is characterized by a noise pmf $P_Z \in \mathcal{P}_q$.
- The channel transition probability matrix is a doubly stochastic $\mathcal{X}$-circulant matrix $\text{circ}_{\mathcal{X}}(P_Z) \in \mathbb{R}^{q \times q}_{\text{sto}}$ defined entry-wise as:

$$\forall x, y \in \mathcal{X}, \; [\text{circ}_{\mathcal{X}}(P_Z)]_{x,y} \triangleq P_Z(-x \oplus y) = P_{Y|X}(y|x).$$

Additive Noise Channels

- Fix a finite Abelian group $(\mathcal{X}, \oplus)$ with order q as the alphabet.
- An additive noise channel if defined by:

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

where $X, Y, Z \in \mathcal{X}$ are the input, output, and noise random variables.
- It is characterized by a noise pmf $P_Z \in \mathcal{P}_q$.
- The channel transition probability matrix is a doubly stochastic $\mathcal{X}$-circulant matrix $\text{circ}_{\mathcal{X}}(P_Z) \in \mathbb{R}_{\text{sto}}^{q \times q}$ defined entry-wise as:

$$\forall x, y \in \mathcal{X}, \;\; [\text{circ}_{\mathcal{X}}(P_Z)]_{x,y} \triangleq P_Z(-x \oplus y) = P_{Y|X}(y|x).$$

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

$$\text{circ}_{\mathcal{X}}(P_Z) = W_\delta$$

Additive Noise Channels

- Fix a finite Abelian group $(\mathcal{X}, \oplus)$ with order q as the alphabet.
- An additive noise channel if defined by:

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

where $X, Y, Z \in \mathcal{X}$ are the input, output, and noise random variables.

- It is characterized by a noise pmf $P_Z \in \mathcal{P}_q$.
- The channel transition probability matrix is a doubly stochastic $\mathcal{X}$-circulant matrix $\text{circ}_{\mathcal{X}}(P_Z) \in \mathbb{R}^{q \times q}_{\text{sto}}$ defined entry-wise as:

$$\forall x, y \in \mathcal{X}, \ \ [\text{circ}_{\mathcal{X}}(P_Z)]_{x,y} \triangleq P_Z(-x \oplus y) = P_{Y|X}(y|x).$$

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

$$\text{circ}_{\mathcal{X}}(P_Z) = W_\delta$$

- Cyclic group example: $\mathbb{Z}/q\mathbb{Z}$ ($\mathcal{X} = [q]$ and $\oplus$ is addition modulo q)

$$(\mathbb{Z}/q\mathbb{Z})\text{-circulant matrix} \Rightarrow \text{circulant matrix}$$

Less Noisy Domination and Degradation Regions

- Given a symmetric channel $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{q \times q}$ for $\delta \in [0, 1]$.

Less Noisy Domination and Degradation Regions

- Given a symmetric channel $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{q \times q}$ for $\delta \in [0, 1]$.
- The less noisy domination region of W_δ is:

$$\mathcal{L}_{W_\delta}^{\mathsf{add}} \triangleq \{v \in \mathcal{P}_q : W_\delta \succeq_{\mathsf{ln}} \mathsf{circ}_{\mathcal{X}}(v)\} .$$

Less Noisy Domination and Degradation Regions

- Given a symmetric channel $W_\delta \in \mathbb{R}_{\text{sto}}^{q \times q}$ for $\delta \in [0, 1]$.
- The less noisy domination region of W_δ is:

$$\mathcal{L}_{W_\delta}^{\text{add}} \triangleq \left\{ v \in \mathcal{P}_q : W_\delta \succeq_{\text{ln}} \text{circ}_{\mathcal{X}}(v) \right\}.$$

- The degradation region of W_δ is:

$$\mathcal{D}_{W_\delta}^{\text{add}} \triangleq \left\{ v \in \mathcal{P}_q : W_\delta \succeq_{\text{deg}} \text{circ}_{\mathcal{X}}(v) \right\}.$$

Less Noisy Domination and Degradation Regions

- Given a symmetric channel $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{q\times q}$ for $\delta \in [0,1]$.
- The less noisy domination region of W_δ is:

$$\mathcal{L}_{W_\delta}^{\mathsf{add}} \triangleq \left\{ v \in \mathcal{P}_q : W_\delta \succeq_{\mathsf{ln}} \mathsf{circ}_{\mathcal{X}}(v) \right\}.$$

- The degradation region of W_δ is:

$$\mathcal{D}_{W_\delta}^{\mathsf{add}} \triangleq \left\{ v \in \mathcal{P}_q : W_\delta \succeq_{\mathsf{deg}} \mathsf{circ}_{\mathcal{X}}(v) \right\}.$$

- $\mathcal{D}_{W_\delta}^{\mathsf{add}} \subseteq \mathcal{L}_{W_\delta}^{\mathsf{add}}$.

Domination Structure of Additive Noise Channels

Theorem (Less Noisy Domination and Degradation Regions)

Given $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{q\times q}$ with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $q \geq 2$, we have:

$$\begin{aligned}\mathcal{D}_{W_\delta}^{\mathsf{add}} &= \mathrm{conv}\left(\text{rows of } W_\delta\right) \\ &\subseteq \mathrm{conv}\left(\text{rows of } W_\delta \text{ and } W_\gamma\right) \\ &\subseteq \mathcal{L}_{W_\delta}^{\mathsf{add}} \subseteq \left\{v \in \mathcal{P}_q : \left\|v - \mathbf{u}\right\|_{\ell^2} \leq \left\|w_\delta - \mathbf{u}\right\|_{\ell^2}\right\}\end{aligned}$$

where w_δ is the first row of W_δ, $\gamma = (1-\delta)/\left(1 - \delta + \frac{\delta}{(q-1)^2}\right)$, and $\mathbf{u} \in \mathcal{P}_q$ is the uniform pmf.

Domination Structure of Additive Noise Channels

Theorem (Less Noisy Domination and Degradation Regions)

Given $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{q\times q}$ with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $q \geq 2$, we have:

$$\begin{aligned}\mathcal{D}_{W_\delta}^{\mathsf{add}} &= \operatorname{conv}\left(\text{rows of } W_\delta\right) \\ &\subseteq \operatorname{conv}\left(\text{rows of } W_\delta \text{ and } W_\gamma\right) \\ &\subseteq \mathcal{L}_{W_\delta}^{\mathsf{add}} \subseteq \left\{v \in \mathcal{P}_q : \left\|v - \mathbf{u}\right\|_{\ell^2} \leq \left\|w_\delta - \mathbf{u}\right\|_{\ell^2}\right\}\end{aligned}$$

where w_δ is the first row of W_δ, $\gamma = (1-\delta)/\left(1-\delta+\frac{\delta}{(q-1)^2}\right)$, and $\mathbf{u} \in \mathcal{P}_q$ is the uniform pmf.

Furthermore, $\mathcal{L}_{W_\delta}^{\mathsf{add}}$ is a closed and convex set that is symmetric with respect to the regular permutation representation $\{P_x \in \mathbb{R}^{q\times q} : x \in \mathcal{X}\}$ of $(\mathcal{X}, \oplus)$ (i.e. $v \in \mathcal{L}_{W_\delta}^{\mathsf{add}} \Rightarrow vP_x \in \mathcal{L}_{W_\delta}^{\mathsf{add}}$ for every $x \in \mathcal{X}$).

Domination Structure of Additive Noise Channels

Theorem (Less Noisy Domination and Degradation Regions)

Given $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{q\times q}$ with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $q \geq 2$, we have:

$$\begin{aligned}\mathcal{D}_{W_\delta}^{\mathsf{add}} &= \text{conv}\left(\text{rows of } W_\delta\right) \\ &\subseteq \text{conv}\left(\text{rows of } W_\delta \text{ and } W_\gamma\right) \\ &\subseteq \mathcal{L}_{W_\delta}^{\mathsf{add}} \subseteq \left\{v \in \mathcal{P}_q : \|v - \mathbf{u}\|_{\ell^2} \leq \|w_\delta - \mathbf{u}\|_{\ell^2}\right\}\end{aligned}$$

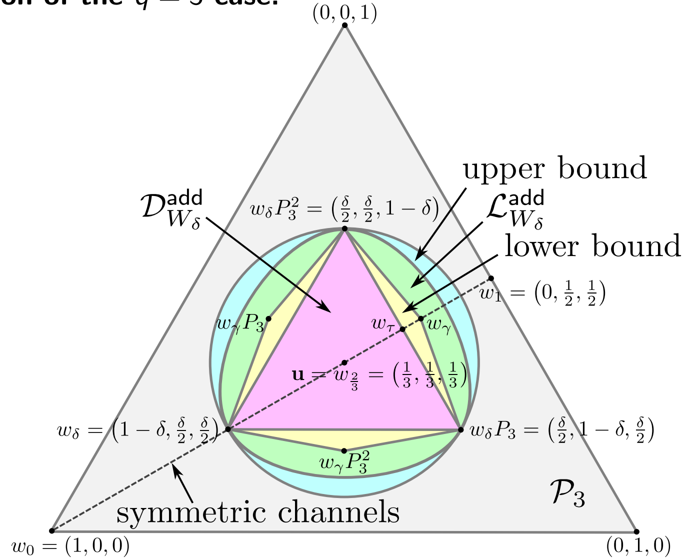
where w_δ is the first row of W_δ, $\gamma = (1-\delta)/\left(1 - \delta + \frac{\delta}{(q-1)^2}\right)$, and $\mathbf{u} \in \mathcal{P}_q$ is the uniform pmf.

Furthermore, $\mathcal{L}_{W_\delta}^{\mathsf{add}}$ is a closed and convex set that is symmetric with respect to the regular permutation representation $\{P_x \in \mathbb{R}^{q\times q} : x \in \mathcal{X}\}$ of $(\mathcal{X}, \oplus)$ (i.e. $v \in \mathcal{L}_{W_\delta}^{\mathsf{add}} \Rightarrow vP_x \in \mathcal{L}_{W_\delta}^{\mathsf{add}}$ for every $x \in \mathcal{X}$).

Remark: The first set inclusion is strict for $\delta \in \left(0, \frac{q-1}{q}\right)$ and $q \geq 3$.

Domination Structure of Additive Noise Channels

Illustration of the $q = 3$ case:



Outline

Some Definitions

- Consider an irreducible Markov chain on $W \in \mathbb{R}_{\text{sto}}^{q \times q}$ on a state space $\mathcal{X} = [q]$ with uniform stationary distribution $\mathbf{u} \in \mathcal{P}_q$.

Some Definitions

- Consider an irreducible Markov chain on $W \in \mathbb{R}_{\text{sto}}^{q \times q}$ on a state space $\mathcal{X} = [q]$ with uniform stationary distribution $\mathbf{u} \in \mathcal{P}_q$.
- Define a continuous-time Markov semigroup with *generator* $L = W - I$ and uniform stationary distribution:

$$\forall t \geq 0, \quad H_t \triangleq \exp\left(-t\left(I - W\right)\right) \in \mathbb{R}_{\text{sto}}^{q \times q}.$$

Some Definitions

- Consider an irreducible Markov chain on $W \in \mathbb{R}_{\text{sto}}^{q \times q}$ on a state space $\mathcal{X} = [q]$ with uniform stationary distribution $\mathbf{u} \in \mathcal{P}_q$.
- Define a continuous-time Markov semigroup with *generator* $L = W - I$ and uniform stationary distribution:

$$\forall t \geq 0, \;\; H_t \triangleq \exp\left(-t\left(I - W\right)\right) \in \mathbb{R}_{\text{sto}}^{q \times q}.$$

- Let $\mathcal{L}^2\left(\mathcal{X}, \mathbf{u}\right)$ (*column* vectors in $\mathbb{R}^q$) be the Hilbert space of functions on $\mathcal{X}$ endowed with the inner product:

$$\forall f, g \in \mathcal{L}^2\left(\mathcal{X}, \mathbf{u}\right), \;\; \langle f, g \rangle_{\mathbf{u}} \triangleq \frac{1}{q} \sum_{x \in \mathcal{X}} f(x) g(x) = \frac{f^T g}{q}.$$

Log-Sobolev Inequalities

Dirichlet Form:

Define the Dirichlet form $\mathcal{E}_W : \mathcal{L}^2(\mathcal{X}, \pi) \times \mathcal{L}^2(\mathcal{X}, \pi) \rightarrow \mathbb{R}^+$:

$$\mathcal{E}_W(f, f) \triangleq \langle (I - W) f, f \rangle_{\mathbf{u}} = \frac{1}{q} f^T \left(I - \frac{W + W^T}{2} \right) f.$$

Dirichlet Form:

Define the Dirichlet form $\mathcal{E}_W : \mathcal{L}^2(\mathcal{X}, \pi) \times \mathcal{L}^2(\mathcal{X}, \pi) \rightarrow \mathbb{R}^+$:

$$\mathcal{E}_W(f, f) \triangleq \langle (I - W) f, f \rangle_{\mathbf{u}} = \frac{1}{q} f^T \left(I - \frac{W + W^T}{2} \right) f.$$

Log-Sobolev Inequality: [DSC96]

The LSI for the Markov semigroup H_t with constant $\alpha \in \mathbb{R}$ states that for every $f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u})$ such that $\|f\|_{\mathbf{u}} = 1$,

$$D\left(f^2 \mathbf{u} \,||\, \mathbf{u}\right) = \frac{1}{q} \sum_{x \in \mathcal{X}} f^2(x) \log\left(f^2(x)\right) \leq \frac{1}{\alpha} \mathcal{E}_W(f, f).$$

Dirichlet Form:

Define the Dirichlet form $\mathcal{E}_W : \mathcal{L}^2(\mathcal{X}, \pi) \times \mathcal{L}^2(\mathcal{X}, \pi) \rightarrow \mathbb{R}^+$:

$$\mathcal{E}_W(f, f) \triangleq \langle (I - W) f, f \rangle_{\mathbf{u}} = \frac{1}{q} f^T \left(I - \frac{W + W^T}{2} \right) f.$$

Log-Sobolev Inequality: [DSC96]

The LSI for the Markov semigroup H_t with constant $\alpha \in \mathbb{R}$ states that for every $f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u})$ such that $\|f\|_{\mathbf{u}} = 1$,

$$D\left(f^2 \mathbf{u} \,||\, \mathbf{u}\right) = \frac{1}{q} \sum_{x \in \mathcal{X}} f^2(x) \log\left(f^2(x)\right) \leq \frac{1}{\alpha} \mathcal{E}_W(f, f).$$

Log-Sobolev Constant:

The largest constant α in the LSI is called the log-Sobolev constant:

$$\alpha(W) \triangleq \inf_{\substack{f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}): \\ \|f\|_{\mathbf{u}} = 1 \\ D(f^2 \mathbf{u} || \mathbf{u}) \neq 0}} \frac{\mathcal{E}_W(f, f)}{D\left(f^2 \mathbf{u} \,||\, \mathbf{u}\right)}.$$

Some Consequences:

Log-Sobolev Inequalities

Some Consequences:

- Continuous case: [DSC96]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall t \geq 0,\ D\left(\mu H_t || \mathbf{u}\right) \leq e^{-2\alpha(W)t} D\left(\mu || \mathbf{u}\right).$$

Log-Sobolev Inequalities

Some Consequences:

- Continuous case: [DSC96]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall t \geq 0,\ D\left(\mu H_t || \mathbf{u}\right) \leq e^{-2\alpha(W)t} D\left(\mu || \mathbf{u}\right).$$

- Discrete case: [Mic97]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall n \in \mathbb{N},\ D\left(\mu W^n || \mathbf{u}\right) \leq \left(1 - \alpha\left(WW^T\right)\right)^n D\left(\mu || \mathbf{u}\right).$$

Log-Sobolev Inequalities

Some Consequences:

- Continuous case: [DSC96]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall t \geq 0,\ D\left(\mu H_t || \mathbf{u}\right) \leq e^{-2\alpha(W)t} D\left(\mu || \mathbf{u}\right).$$

- Discrete case: [Mic97]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall n \in \mathbb{N},\ D\left(\mu W^n || \mathbf{u}\right) \leq \left(1 - \alpha\left(WW^T\right)\right)^n D\left(\mu || \mathbf{u}\right).$$

Computing log-Sobolev Constants:

- Difficult in general ☹

Log-Sobolev Inequalities

Some Consequences:

- Continuous case: [DSC96]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall t \geq 0,\ D\left(\mu H_t || \mathbf{u}\right) \leq e^{-2\alpha(W)t} D\left(\mu || \mathbf{u}\right).$$

- Discrete case: [Mic97]

$$\text{LSI} \Rightarrow \forall \mu \in \mathcal{P}_q, \forall n \in \mathbb{N},\ D\left(\mu W^n || \mathbf{u}\right) \leq \left(1 - \alpha\left(WW^T\right)\right)^n D\left(\mu || \mathbf{u}\right).$$

Computing log-Sobolev Constants:

- Difficult in general ☹
- Easy for q-ary symmetric channels ☺

$$\alpha(W_\delta) = \begin{cases} \frac{(q-2)\delta}{(q-1)\log(q-1)} & , \quad q > 2 \\ \delta & , \quad q = 2 \end{cases}$$

$$\alpha\left(W_\delta W_\delta^T\right) = \begin{cases} \frac{(q-2)(2q-2-q\delta)\delta}{(q-1)^2\log(q-1)} & , \quad q > 2 \\ 2\delta(1-\delta) & , \quad q = 2 \end{cases}$$

Comparison of Dirichlet Forms

How do we prove an LSI for an irreducible channel $V \in \mathbb{R}^{q \times q}_{\textbf{sto}}$?

Comparison of Dirichlet Forms

How do we prove an LSI for an irreducible channel $V \in \mathbb{R}_{\textbf{sto}}^{q\times q}$?

Idea: Domination of Dirichlet forms

$$\forall f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}), \;\; \mathcal{E}_V(f,f) \geq \mathcal{E}_{W_\delta}(f,f)$$

for some $\delta \in \left[0, \frac{q-1}{q}\right]$

Comparison of Dirichlet Forms

How do we prove an LSI for an irreducible channel $V \in \mathbb{R}_{\textbf{sto}}^{q \times q}$?

Idea: Domination of Dirichlet forms

$$\forall f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}), \;\; \mathcal{E}_V(f, f) \geq \mathcal{E}_{W_\delta}(f, f)$$

for some $\delta \in \left[0, \frac{q-1}{q}\right]$ implies an LSI:

$$D\left(f^2 \mathbf{u} \,||\, \mathbf{u}\right) \leq \frac{1}{\alpha(W_\delta)} \mathcal{E}_{W_\delta}(f, f) \leq \frac{1}{\alpha(W_\delta)} \mathcal{E}_V(f, f)$$

for every $f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u})$ with $\|f\|_{\mathbf{u}} = 1$.

Comparison of Dirichlet Forms

How do we prove an LSI for an irreducible channel $V \in \mathbb{R}_{\textbf{sto}}^{q\times q}$?

Idea: Domination of Dirichlet forms

$$\forall f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}), \;\; \mathcal{E}_V(f,f) \geq \mathcal{E}_{W_\delta}(f,f)$$

for some $\delta \in \left[0, \frac{q-1}{q}\right]$ implies an LSI:

$$D\left(f^2\mathbf{u} \,||\, \mathbf{u}\right) \leq \frac{1}{\alpha(W_\delta)} \mathcal{E}_{W_\delta}(f,f) \leq \frac{1}{\alpha(W_\delta)} \mathcal{E}_V(f,f)$$

for every $f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u})$ with $\|f\|_{\mathbf{u}} = 1$.

Connection to channel comparison:

Less noisy domination $\Rightarrow$ Dirichlet form domination

Comparison of Dirichlet Forms

Theorem (Domination of Dirichlet Forms)

Let $W, V \in \mathbb{R}_{\text{sto}}^{q \times q}$ be channels with uniform stationary distribution.

- If $W \succeq_{\text{ln}} V$, then:

$$\forall f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}), \ \ \mathcal{E}_{VV^T}(f, f) \geq \mathcal{E}_{WW^T}(f, f) .$$

- If W is positive semidefinite, V is normal (i.e. $V^T V = VV^T$), and $W \succeq_{\text{ln}} V$, then:

$$\forall f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}), \ \ \mathcal{E}_{V}(f, f) \geq \mathcal{E}_{W}(f, f) .$$

- If $W = W_\delta \in \mathbb{R}_{\text{sto}}^{q \times q}$ is any q-ary symmetric channel with $\delta \in \left[0, \frac{q-1}{q}\right]$ and $W_\delta \succeq_{\text{ln}} V$, then:

$$\forall f \in \mathcal{L}^2(\mathcal{X}, \mathbf{u}), \ \ \mathcal{E}_{V}(f, f) \geq \mathcal{E}_{W_\delta}(f, f) .$$

Outline

Conclusion: Deriving Log-Sobolev Inequalities

Consider the irreducible channel:

$$V = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 1/3 & 5/12 & 1/4 \end{bmatrix}$$

with $q = 3$, stationary pmf $\mathbf{u} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, and minimum entry $\nu = \frac{1}{6}$.

Conclusion: Deriving Log-Sobolev Inequalities

Consider the irreducible channel:

$$V = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 1/3 & 5/12 & 1/4 \end{bmatrix}$$

with $q = 3$, stationary pmf $\mathbf{u} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, and minimum entry $\nu = \frac{1}{6}$.

- Generate 3-ary symmetric channel $W_\delta \in \mathbb{R}^{3\times 3}_{\mathsf{sto}}$ such that $W_\delta \succeq_{\mathsf{ln}} V$:

$$\delta = \frac{\nu}{1 - (q-1)\nu + \frac{\nu}{q-1}} = \frac{2}{9}.$$

Conclusion: Deriving Log-Sobolev Inequalities

Consider the irreducible channel:

$$V = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 1/3 & 5/12 & 1/4 \end{bmatrix}$$

with $q = 3$, stationary pmf $\mathbf{u} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, and minimum entry $\nu = \frac{1}{6}$.

- Generate 3-ary symmetric channel $W_\delta \in \mathbb{R}^{3\times 3}_{\mathsf{sto}}$ such that $W_\delta \succeq_{\mathsf{ln}} V$:

$$\delta = \frac{\nu}{1 - (q-1)\nu + \frac{\nu}{q-1}} = \frac{2}{9}.$$

- Compute the log-Sobolev constant of W_δ:

$$\alpha(W_\delta) = \frac{(q-2)\delta}{(q-1)\log(q-1)} = \frac{1}{9\log(2)} \approx 0.1603\,.$$

Conclusion: Deriving Log-Sobolev Inequalities

Consider the irreducible channel:

$$V = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 1/3 & 5/12 & 1/4 \end{bmatrix}$$

with $q = 3$, stationary pmf $\mathbf{u} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, and minimum entry $\nu = \frac{1}{6}$.

- Generate 3-ary symmetric channel $W_\delta \in \mathbb{R}^{3\times 3}_{\mathsf{sto}}$ such that $W_\delta \succeq_{\mathsf{ln}} V$:

$$\delta = \frac{\nu}{1 - (q-1)\nu + \frac{\nu}{q-1}} = \frac{2}{9}.$$

- Compute the log-Sobolev constant of W_δ:

$$\alpha(W_\delta) = \frac{(q-2)\delta}{(q-1)\log(q-1)} = \frac{1}{9\log(2)} \approx 0.1603\,.$$

- Use domination of Dirichlet forms to get the LSI:

$$D\left(f^2\mathbf{u} \,||\, \mathbf{u}\right) \leq \frac{1}{0.1603}\,\mathcal{E}_V(f, f)\,.$$

Conclusion: Deriving Log-Sobolev Inequalities

Consider the irreducible channel:

$$V = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 1/3 & 5/12 & 1/4 \end{bmatrix}$$

with $q = 3$, stationary pmf $\mathbf{u} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, and minimum entry $\nu = \frac{1}{6}$.

- Generate 3-ary symmetric channel $W_\delta \in \mathbb{R}_{\mathsf{sto}}^{3\times 3}$ such that $W_\delta \succeq_{\mathsf{ln}} V$:

$$\delta = \frac{\nu}{1 - (q-1)\nu + \frac{\nu}{q-1}} = \frac{2}{9}.$$

- Compute the log-Sobolev constant of W_δ:

$$\alpha(W_\delta) = \frac{(q-2)\delta}{(q-1)\log(q-1)} = \frac{1}{9\log(2)} \approx 0.1603\,.$$

- Use domination of Dirichlet forms to get the LSI:

$$D\left(f^2\mathbf{u}\,||\,\mathbf{u}\right) \leq \frac{1}{0.3570}\,\mathcal{E}_V(f,f) \leq \frac{1}{0.1603}\,\mathcal{E}_V(f,f)\,.$$

Thank You!