Advanced Topics in Information Theory

Lecture Notes

Stefan M. Moser

© Copyright Stefan M. Moser

Signal and Information Processing Lab
ETH Zürich
Zurich, Switzerland

Department of Electrical and Computer Engineering
National Chiao Tung University (NCTU)
Hsinchu, Taiwan

You are welcome to use these lecture notes for yourself, for teaching, or for any other noncommercial purpose. If you use extracts from these lecture notes, please make sure that their origin is shown. The author assumes no liability or responsibility for any errors or omissions.

Version 2.10.
Compiled on 12 May 2017.
For the latest version see http://moser-isi.ethz.ch/scripts.html
\[ \tilde{Q} \]

\[ Q^* \]

\[ \mathcal{F} \]

> 90 degrees
independent description
49 points

dependent description
45 points
<table>
<thead>
<tr>
<th>$\hat{x}_1$</th>
<th>$\hat{x}_2$</th>
<th>$\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>0.3578</td>
<td>1.3288</td>
<td>0.8433</td>
</tr>
<tr>
<td>0.3973</td>
<td>1.4011</td>
<td>0.8992</td>
</tr>
<tr>
<td>0.4202</td>
<td>1.4450</td>
<td>0.9326</td>
</tr>
<tr>
<td>0.4336</td>
<td>1.4714</td>
<td>0.9525</td>
</tr>
<tr>
<td>0.4414</td>
<td>1.4872</td>
<td>0.9643</td>
</tr>
<tr>
<td>0.4461</td>
<td>1.4966</td>
<td>0.9714</td>
</tr>
<tr>
<td>0.4488</td>
<td>1.5022</td>
<td>0.9755</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>·</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4528</td>
<td>1.5104</td>
<td>0.9816</td>
</tr>
</tbody>
</table>
information rate distortion function (with discontinuity)

\[ R^I(D) \]

\[ R^I(D + \epsilon) \]
The figure illustrates the tradeoff between rate and distortion in signal processing. The line represents the sets of points \((\lambda \mathbb{E}[d(X, \hat{X})], I(X; \hat{X}))\) for varying values of \(\lambda\). The point on the line corresponds to the optimal distortion-rate function, which is the point on the line that achieves the minimum rate for a given distortion. The line has a slope of \(-\lambda\), indicating the tradeoff between rate and distortion. The axes represent distortion on the horizontal axis and rate on the vertical axis.
\[ \lambda E_{q^*}[d(X, \hat{X})] = \lambda D \]

\[ I_{q^*}(X; \hat{X}) = R(D) \]

\[ R_0(q^*, \lambda) \]

\[ R_0(q^*, \lambda) \]

\[ R(\cdot) \]

achievable by \( q^* \)
\[ R_0(q^*, \lambda) \]
\[ R_0(q', \lambda) \]

achievable by \( q^* \)

achievable by \( q' \)
here a discontinuity is possible

line below $R(D)$
slope discontinuities

two different tangents with slopes $-\lambda_1$ and $-\lambda_2$
the sources \( \tilde{Q} \) that do not work because for the given \( R \) and \( D \):

\[
R(\tilde{Q}, D) > R
\]

\[
\inf \mathcal{D}(\tilde{Q} \parallel Q) = \mathcal{D}^*\]
| $Q_{\hat{X}^{(1)},{\hat{X}^{(2)}}|X(\cdot,\cdot|0)}$ | $\hat{X}^{(2)}$ | $Q_{\hat{X}^{(1)}|X(\cdot|0)}$ |
|----------------|--------------|----------------|
| $\hat{X}^{(1)}$ | | |
| 0 | 0, 1 | 0, 1 |
| 1 | $3 - 2\sqrt{2}$, $\sqrt{2} - 1$ | 2 - $\sqrt{2}$ |
| $Q_{\hat{X}^{(2)}|X(\cdot|0)}$ | | |
| 0 | 2 - $\sqrt{2}$, $\sqrt{2} - 1$ | |

| $Q_{\hat{X}^{(1)},{\hat{X}^{(2)}}|X(\cdot,\cdot|1)}$ | $\hat{X}^{(2)}$ | $Q_{\hat{X}^{(1)}|X(\cdot|1)}$ |
|----------------|--------------|----------------|
| $\hat{X}^{(1)}$ | | |
| 0 | 0, 0 | 0 |
| 1 | 0, 1 | 1 |
| $Q_{\hat{X}^{(2)}|X(\cdot|1)}$ | | |
| 0 | 0, 1 | |

<table>
<thead>
<tr>
<th>( Q_{\hat{X}^{(1)}, \hat{X}^{(2)} (\cdot, \cdot)} )</th>
<th>( \hat{X}^{(2)} )</th>
<th>( Q_{\hat{X}^{(1)} (\cdot)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{X}^{(1)} )</td>
<td>0</td>
<td>( \frac{3}{2} - \sqrt{2} )</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>( \frac{\sqrt{2}}{2} - \frac{1}{2} )</td>
</tr>
<tr>
<td>( Q_{\hat{X}^{(2)} (\cdot)} )</td>
<td>( 1 - \frac{\sqrt{2}}{2} )</td>
<td>( \frac{\sqrt{2}}{2} )</td>
</tr>
</tbody>
</table>
codeword 1

$e^{nR'}$

bin 1

bin 2

bin 3

bin $(e^{nR}-1)$

bin $e^{nR}$

separate compression and decompression

joint encoding

<table>
<thead>
<tr>
<th></th>
<th>$Q_{X,Y}(\cdot, \cdot)$</th>
<th>Taichung $Y$</th>
<th>Hsinchu total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hsinchu $X$</td>
<td>rain 0.445</td>
<td>rain 0.445</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>sun 0.055</td>
<td>sun 0.055</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Taichung total 0.5</td>
<td>Taichung total 0.5</td>
<td></td>
</tr>
</tbody>
</table>
convex hull of $C^a \cup C^b$
inactive constraint

inactive constraint
\( R^{(2)} \)

\[ C \left( \frac{E^{(2)}}{\sigma^2} \right) \]

\( C \left( \frac{E^{(2)}}{E^{(1)} + \sigma^2} \right) \)

\( R^{(1)} + R^{(2)} = C \left( \frac{E^{(1)} + E^{(2)}}{\sigma^2} \right) \)

\[ \alpha = \frac{E^{(1)}}{E^{(1)} + E^{(2)}} \]

TDMA for \( \alpha \) from 0 to 1

\[ \alpha = 0 \]

\[ \alpha = 1 \]
\[
R^{(1)} = -R^{(0)} + I(U; Y^{(2)}) + I(X; Y^{(1)}|U)
\]

\[ R^{(2)} = -R^{(1)} + I(U^{(1)}; Y^{(1)}) + I(U^{(2)}; Y^{(2)}) - I(U^{(1)}; U^{(2)}) \]

\[ R^{(1)} = I(U^{(1)}; Y^{(1)}) \]
\( f: \mathcal{U}^{(1)} \times \mathcal{U}^{(2)} \rightarrow \mathcal{X} \)
\[
\frac{1}{2} \log \left( 1 + \frac{E}{\sigma_{(2)}^2} \right)
\]

\[
\alpha = 0
\]

\[
\alpha = 1
\]

Stefan M. Moser, Advanced Topics in Information Theory, Version 2.10.
\[ Q_{Y^{(2)}, Y^{(3)} | X^{(1)}, X^{(2)}} \]

Terminal 4 \rightarrow \hat{M} \rightarrow \text{Terminal 1} \rightarrow M

\text{Channel} \rightarrow S_2 \rightarrow X^{(2)} \rightarrow Y^{(2)} \rightarrow \text{Terminal 2}

\text{Channel} \rightarrow S_1 \rightarrow Y^{(4)} \rightarrow \text{Terminal 4}

\text{Channel} \rightarrow S_3 \rightarrow X^{(1)} \rightarrow \text{Terminal 1}

\text{Channel} \rightarrow S_4 \rightarrow X^{(3)} \rightarrow \text{Terminal 3}

\[ \hat{M}^{(1)} \quad \text{Dest. 1} \quad \text{Dec. } \psi^{(1)} \quad Y^{(1)} \quad \text{Channel} \quad Q_{Y^{(1)},Y^{(2)}}^n |X^{(1)},X^{(2)} \quad X^{(1)} \quad \text{Enc. } \phi^{(1)} \quad M^{(1)} \quad \text{Uniform Source 1} \]

\[ \hat{M}^{(2)} \quad \text{Dest. 2} \quad \text{Dec. } \psi^{(2)} \quad Y^{(2)} \quad \text{Channel} \quad Q_{Y^{(1)},Y^{(2)}}^n |X^{(1)},X^{(2)} \quad X^{(2)} \quad \text{Enc. } \phi^{(2)} \quad M^{(2)} \quad \text{Uniform Source 2} \]
\[ a_{12} = 0.15, \ a_{21} = 0.05 \]

\[ a_{12} = 0.35, \ a_{21} = 0.25 \]

\[ a_{12} = 0.55, \ a_{21} = 0.45 \]

\[ a_{12} = 0.85, \ a_{21} = 0.75 \]

\[ a_{12} = 1.15, \ a_{21} = 1.15 \]

\[ a_{12} = 2.15, \ a_{21} = 2.15 \]
\( R^{(2)} \leq 3.26 \) bits

\( R^{(1)} + 2R^{(2)} \leq 7.09 \) bits

\( R^{(1)} + R^{(2)} \leq 4.19 \) bits

\( 2R^{(1)} + R^{(2)} \leq 7.09 \) bits

\( R^{(1)} \leq 3.26 \) bits

\( a_{12} = a_{21} = 0.1 \)
The diagram illustrates the relationship between $d_{\text{sym}}^*$ and $a$, with axes labeled as follows:

- Vertical axis: $d_{\text{sym}}^*$
- Horizontal axis: $a$

The x-axis is divided into intervals: $\frac{1}{2}$, $\frac{2}{3}$, 1, $\frac{3}{2}$, 2, with corresponding labels for the y-axis:

- weak
- medium
- strong
- very strong

The graph shows a decreasing trend from $\frac{3}{2}$ to $\frac{2}{3}$, followed by an increasing trend to 1, and finally a horizontal line at 1 for $a > 1$. The diagram is from Stefan M. Moser, *Advanced Topics in Information Theory*, Version 2.10.