Minimum Expected Distortion in Gaussian Source Coding with Uncertain Side Information

Chris T. K. Ng¹, Chao Tian²
Andrea J. Goldsmith¹, Shlomo Shamai (Shitz)³

¹Stanford University, ²EPFL, ³Technion

ITW September 2007, Lake Tahoe, CA
Introduction

- Lossy data compression:
  - Side information at the decoder can help reduce the distortion.
- Side information may be acquired over unreliable wireless channels.
  - E.g., wireless sensor network: a helper sensor transmits correlated measures over the wireless channel to be used as side information at the decoder.
- We consider the Gaussian source coding problem where:
  - There is a fading analog side-information channel.
  - Decoder knows the realization of the side-information channel.
  - Encoder knows only the fading distribution.
- Want to allocation the Heegard-Berger source coding rate among the fading states:
  - To minimize the expected distortion.
Source Coding with Uncertain Analog Side Information

- Combine digital transmission with analog transmission.
  - The analogy side information is correlated with the digital description.
  - The digital transmitter does not know the fading realization of the analog channel.
Motivation

- Wireless sensor networks.
  - Compression of sensor measurements.
  - Side information is transmitted by a cooperative sensor over a fading wireless channel.

- Hybrid digital-analog communications.

- Systematic lossy source-coding over a fading channel without CSI at the transmitter.
  - Upgrading of legacy communications systems
  - A digital channel is added to augment an existing analog channel.
  - The analog reception is used as side information.
Related Works

- Without side information:
  - Standard rate-distortion function.
- Side-information channel has no fading [Wyner-Ziv ’78].
- Side-information is also available at the encoder [Gray ’73, Zamir ’96].
- Combination of decode and encode side information [Fleming, Effros ’06].
- Systematic lossy source channel coding [Shamai, Verdu, Zamir ’98].
- Source coding with degraded side information at the decoder [Heegard, Berger ’85, Kaspi ’94].
- Successive refinement/scalable source coding with degraded side information [Steinberg, Merhav ’04, Tian, Diggavi ‘07].
System Model – Fading Analog Side Information Channel

Encoder: rate constraint $R_X$.

Decoder: side information over uncoded analog channel subject to slow fading.

Decode knows fading realization but encoder knows only distribution.

Motivation: E.g., wireless sensor network.

Minimize expected squared error distortion:

- Optimally allocate rate among the layers ($M$ fading states) in the Heegard-Berger rate-distortion function.

Source

$X \sim \mathcal{N}(0, \sigma_X^2)$

Encoding

$R_X$

Decoder

$S \sim f_S(s)$

Reconstruction

$\hat{X}$

Side Information

$Y' = \sqrt{S}X + Z$

$Z \sim \mathcal{N}(0, 1)$
Heegard-Berger Rate-Distortion Function

- Rate required to simultaneously satisfy a set of distortion constraints.
  - For a set of degraded side-information random variables:
    \[ X \leftrightarrow Y_M \leftrightarrow Y_{M-1} \leftrightarrow \cdots \leftrightarrow Y_1 \]

- Heegard-Berger rate distortion function:
  \[ R_{HB}(D) = \min_{W_1^M \in P(D)} \sum_{i=1}^{M} I(X;W_i|Y_i,W_{i-1}^i) \]

  - Minimization over \( W_1^M \) such that:
    \[ W_1^M \leftrightarrow X \leftrightarrow Y_M \leftrightarrow Y_{M-1} \leftrightarrow \cdots \leftrightarrow Y_1 \]
    \[ \mathbb{E}[d_i(X,\hat{X}_i)] \leq D_i, \quad i = 1,\ldots,M. \]

- Side information can be statistically degraded.
- Squared error distortion measures:
  \[ d_i(X,\hat{X}_i) = (X - \hat{X}_i)^2 \]
Gaussian Source under Squared Error Distortion

- Two discrete fading states: $S = \{s_1, s_2\}$.

\[ R_{HB}(D_1, D_2) = \min_{W_1, W_2 \in P(D_1, D_2)} \{I(X; W_1|Y_1) + I(X; W_2|Y_2, W_1)\} \]

- For Gaussian source:

\[ R_{HB}(D_1, D_2) = -\frac{1}{2} \log(s_1 + \sigma_x^{-2}) \]
\[ + \min_{W_1} \left\{ -\frac{1}{2} \log \left( 1 + (s_2 - s_1) \text{VAR}[X|Y_1, W_1] \right) \right. \]
\[ + \min_{W_2} \left\{ -\frac{1}{2} \log(\text{VAR}[X|Y_2, W_1, W_2]) \right\} \].

- Pick $W_1, W_2$ such that the distortion constraints are satisfied.

\[ \max_{W_1} \text{VAR}[X|Y_1, W_1] = \min(\text{VAR}[X|Y_1], D_1) \]
\[ \max_{W_2} \text{VAR}[X|Y_2, W_1, W_2] = \min(\text{VAR}[X|Y_2, W_1], D_2) \]
Optimal Distortion Trade-off and Rate Allocation

- Need to find optimal operating point on the distortion region:
  - For a given fading distribution.
  - To minimize expected distortion.
Two Discrete Fading States - Optimal Rate Allocation

- Rate allocated to the better fading state, $R_2^*$, is not monotonic.
  - When $s_2$ is large, $R_2^*$ declines as $E[D]^*$ is dominated by worse fading state.
  - Optimal rate heavily skewed toward to worse fading state; $R_2^* > 0$ only when $p_2$ is large.
Multiple Discrete Fading States

- The Heegard-Berger rate-distortion function extends directly to the case when the side-information channel has multiple discrete fading states ($M > 2$).

$$R_{HB}(D) = -\frac{1}{2} \log(\sigma_X^{-2} + s_1) - \frac{1}{2} \log \tilde{D}_M - \frac{1}{2} \sum_{i=1}^{M-1} \log\left(1 + (s_{i+1} - s_i)\tilde{D}_i\right)$$

$$\tilde{D}_i \triangleq \min(D_i, (\tilde{D}_{i-1}^{-1} + s_i - s_{i-1})^{-1})$$

- For $M > 2$, we did not find closed-form expressions for the optimal rate allocation and the corresponding minimum expected distortion $E[D]^*$. However, $E[D]$ minimization can be formulated as a convex optimization problem.
Expected Distortion Minimization

- Convex optimization problem:
  - Can be computed efficiently.

\[
\begin{align*}
\text{minimize} & \quad \mathbf{p}^T \mathbf{D} \\
\text{subject to} & \quad 0 \leq D_i, \quad 0 \leq \tilde{D}_i, \quad i = 1, \ldots, M \\
& \quad - \frac{1}{2} \log(\sigma_X^{-2} + s_1) - \frac{1}{2} \log \tilde{D}_M \\
& \quad - \frac{1}{2} \sum_{i=1}^{M-1} \log(1 + (s_{i+1} - s_i)\tilde{D}_i) \leq R_X \\
& \quad \tilde{D}_1 \leq (\sigma_X^{-2} + s_1)^{-1} \\
& \quad \tilde{D}_i \leq (\tilde{D}_{i-1}^{-1} + s_i - s_{i-1})^{-1}, \quad i = 2, \ldots, M \\
& \quad \tilde{D}_i \leq D_i, \quad i = 1, \ldots, M.
\end{align*}
\]
Optimal rate allocation concentrates at lowest layer: $R^*_1 = R_X$.

When the source coding rate is large:
- Uncertain side information is almost no more useful than no side information.
- Reduce analog channel uncertainty.
Continuous Fading Distribution

- Consider the expected-distortion-rate function:

\[
E[D] = p^T D = p_1 D_1 + p_2 D_2 + \cdots + p_M D_M
\]

\[
D_i = \left[\frac{1}{(((\sigma_X^{-2} + s_1)2^{2R_1} + s_2 - s_1)2^{2R_2} + \cdots)2^{2R_M}}\right]^{-1}
\]

- We consider the limiting process as \(s_{i+1} - s_i\) tends to 0:

\[
E[D] = \int_0^\infty f_S(s)u'(s)(\sigma_X^{-2} + u(s))^{-1} ds
\]

\[
u(s) \triangleq \int_0^s 2^{-2} \int_0^t R(r) dr dt
\]

- Boundary conditions: \(u'(0) = 1, u'(\infty) = 2^{-2R_x}\).
Optimal Continuous Rate Distribution

- Necessary condition for optimality as given by the Euler-Lagrange equation:

\[ \frac{d}{ds} \left( \frac{\partial F}{\partial u'} \right) - \frac{\partial F}{\partial u} = 0 \]

\[ F(s, u, u') \triangleq f_S(s)u'(s) \left( \sigma_X^{-2} + u(s) \right)^{-1} \]

- The condition evaluates to:

\[ \frac{f_S'(s)}{\sigma_X^{-2} + u(s)} = 0 \]

- We suppose in general for the given fading distribution, \( f''_S(s) \neq 0 \); then no \( u(s) \) satisfies the Euler-Lagrange condition.
  - We conjecture that the optimal rate allocation is discrete even when the fading distribution is continuous and smooth.
Conservative Rate Allocation

- Under Rayleigh fading, assume the rate allocation concentrates at the worse fading state \((s = 0)\), i.e., \(R(s) = R_x \delta(s)\).

- Under such conservative rate allocation, the expected distortion is:

\[
E[D] = \int_{0}^{\infty} \frac{f_s(s)}{\sigma_X^{-2}2^{2R_x} + s} ds
\]

\[
= -\bar{S}^{-1} \exp(\bar{S}^{-1} \sigma_X^{-2}2^{2R_x}) \text{Ei}(-\bar{S}^{-1} \sigma_X^{-2}2^{2R_x})
\]

- When \(R_x\) is large:

\[
E[D] \approx \int_{0}^{K} \frac{f_s(s)}{\sigma_X^{-2}2^{2R_x} + s} ds \approx \int_{0}^{K} \frac{f_s(s)}{\sigma_X^{-2}2^{2R_x}} ds
\]

\[
\approx \sigma_X^22^{-2R_x} = D_{\text{No-SI}}.
\]

- Uncertain side information is almost no more useful than no side information.
Conclusions

- We considered the problem of Gaussian source coding under squared error distortion:
  - With uncertain side information at the decoder.
- When the side-information channel has two discrete fading states:
  - Derived the optimal rate allocation between the fading states, and the corresponding minimum expected expected distortion.
- Optimal rate allocation is conservative:
  - Rate is allocated to the better fading state only if it is highly probable.
  - The risk of the better fading state not being realized dominates the expected distortion.
Conclusions – continued

- When the side-information channel has multiple discrete fading states:
  - The expected distortion minimization can be formulated as a convex optimization problem.
  - Can be computed efficiently.

- Discretized Rayleigh fading:
  - Optimal rate allocation concentrates at the worst fading state ($s = 0$).
  - Uncertain side information is negligibly more useful than no side information when the source coding rate is large.

- Continuous fading distribution:
  - We conjecture that the optimal rate allocation is discrete even when the fading distribution is continuous and smooth.