A Generalization of the Rate-Distortion Function for Wyner-Ziv Coding of Noisy Sources in the Quadratic-Gaussian Case

David Rebollo-Monedero and Bernd Girod
Main Results

- Source Data $X_n$
- Reconstruction $\hat{X}_n$

WZ Encoder $\rightarrow$ WZ Decoder

$\hat{X}_n = \mu(y) + N \sim \mathcal{N}(0, \sigma^2)$

- **Direct Wyner-Ziv (WZ) coding**
  - Quadratic distortion
  - No rate loss due to unavailability of side information at encoder
  - Arbitrarily distributed side information
  - Closed-form expression for RD function

- **Extension to indirect WZ coding**
Outline

- RD function for WZ coding in the quadratic-Gaussian case
  - Previous work
  - Our generalization
  - Direct and indirect cases

- Experimental results
  - Simple Gaussian mixture
  - Information-theoretic vs. operational RD performance
Direct WZ Coding

Source Data \( \mathcal{R}_{X|Y}^{WZ}(\mathcal{D}) \geq \mathcal{R}_{X|Y}(\mathcal{D}) \) Reconstruction

Encoder \( X \xrightarrow{n} \hat{X} \xrightarrow{n} \) Decoder

- Distortion \( \mathcal{D} = \frac{1}{n} \sum_{i=1}^{n} E[d(X_i, \hat{X}_i)] \)
- Information-theoretic RD function \( \mathcal{R}(\mathcal{D}) \)
Previous Work on Direct WZ Coding

Source Data

\[ X_n \]

WZ Encoder

\[ \hat{X}_n \]

Reconstruction

\[ Y \]

Side Information

\[ N \sim \mathcal{N}(0, \sigma^2) \]

- Quadratic-Gaussian case [Wyner, Ziv, 76-78]
- Extension [Pradhan, Chou, Ramchandran, 03]
  - \( d(x, \hat{x}) = (x - \hat{x})^2 \)
  - \( Y \) arbitrarily distributed
  - \( \mathcal{R}_{X|Y}^{WZ}(\mathcal{D}) = \mathcal{R}_{X|Y}(\mathcal{D}) = \frac{1}{2} \log^+ \frac{\sigma^2}{\mathcal{D}} \)
Main Result on Direct WZ Coding

Source Data \( X_n \) \( \rightarrow \) WZ Encoder \( \rightarrow \) WZ Decoder \( \rightarrow \) Reconstruction \( \hat{X}_n \)

- \( N \sim \mathcal{N}(0, \sigma^2) \)

- Slightly relaxes conditions
  - \( d(x, \hat{x}) = (x - \hat{x})^2 \)
  - \( Y \) arbitrarily distributed
  - \( R_{X\mid Y}^{\text{WZ}}(D) = R_{X\mid Y}(D) = \frac{1}{2} \log^+ \frac{\sigma^2}{D} \)

- Direct proof, not using duality arguments
Indirect WZ Coding

Source Data $X^n$, $p(z^n | x^n, y^n)$, $Z^n$ → Encoder → Decoder → $\hat{X}^n$

Indirect Observation $Y^n$, $\mathcal{R}_{XZ|Y}^N(\mathcal{D}) \geq \mathcal{R}_{XZ|Y}^N(\mathcal{D})$

Reconstruction

Side Information $Y^n$
Previous Work on Indirect WZ Coding

- **Quadratic-Gaussian case** [Yamamoto, Itoh, 80]
  - $d(x, \hat{x}) = (x - \hat{x})^2$
  - $X, Y, Z$ jointly Gaussian
  - $R_{XZ|Y}^{NWZ}(D) = R_{X|Y}^N(D) = \frac{1}{2} \log \left( \frac{\sigma_Z^2 - \sigma_{X|Y}^2}{D - \sigma_x^2} \right)$

- **Similar, independent work** [Flynn, Gray, 87], [Draper, 02]
Main Result on Indirect WZ Coding

Source Data \( X \) \( \rightarrow \) \( Z \) \( \rightarrow \) NWZ Encoder \( \rightarrow \) NWZ Decoder \( \hat{X} \)

Indirect Observation \( \mu(y) \) \( \hat{X} \)

Reconstruction \( f(y) \) \( Y \) Side Information

\[ d(x, \hat{x}) = (x - \hat{x})^2 \]

- \( Y \) arbitrarily distributed
- \[ R_{XZ|Y}^{NWZ}(D) = R_{XZ|Y}^N(D) = \frac{1}{2} \log^+ \frac{\alpha^2 \sigma^2}{\mathbb{D} - \mathbb{E} \text{Var}[X|Y,Z]} \]
Intuitive Example

Source Data \( X \sim \mathcal{N}(0, 1) \) → Indirect Observation \( Z \)

\[
X \sim \mathcal{N}(0, 1)
\]

\[
\hat{X} \sim \mathcal{N}(0, 1)
\]

\[
Y = \frac{1}{2} - \frac{1}{2}, \text{ w.p. } 1/2
\]

\[
\hat{Y} = \frac{1}{2} - \frac{1}{2}, \text{ w.p. } 1/2
\]

\[
\hat{X} = \hat{Z} - Y
\]

\[
\mathcal{D} = \frac{1}{n} \mathbb{E} \left\| X^n - \hat{X}^n \right\|^2
\]

\[
\mathcal{D}_{NWZ} = 2^{-2\mathcal{R}} = \mathcal{D}_{NWZ}(\mathcal{R})
\]

\[
\mathcal{D} = 1.96 \cdot 2^{-2\mathcal{R}}
\]
Lloyd algorithm for noisy WZ quantization [Rebollo-Monedero, Girod, 05]

High-rate noisy WZ quantization [Rebollo-Monedero, Rane, Girod, 04]
- Lattice quantizer and Slepian-Wolf codec asymptotically optimal
- $D(R) \approx 2\pi e M_n 2^{-2R}$ ($M_n$ normalized moment of inertia)
RD Comparison

Distortion-Rate Function $D_{X|W|Z|Y}(R)$

Slope $\approx 6.02 \text{ dB/bit}$

- Red: Noisy WZ Quantizers ($n=1,2,3$)
- Dashed blue line: High-Rate Approximation ($n=1,2,3$)

$1/D \text{ [dB]}$

$R \text{ [bit]}$
Conclusions

- No rate loss incurred in WZ coding
  - Jointly Gaussian statistics are *not* necessary
  - Relaxed conditions in both the direct and indirect case
  - Side information arbitrarily distributed

- In the indirect case, condition on data similar to additive separability in high-rate noisy WZ quantization

- All conditions determined by conditional joint distribution of (data, observation) given side information, arbitrarily distributed

- Paper at [www.stanford.edu/~drebollo/publications.htm](http://www.stanford.edu/~drebollo/publications.htm)
A Generalization of the Rate-Distortion Function for Wyner-Ziv Coding of Noisy Sources in the Quadratic-Gaussian Case

David Rebollo-Monedero and Bernd Girod

Information Systems Lab.
Electrical Eng. Dept.
Stanford University
Outline

- Introduction and previous work

- Theoretic analysis
  - Indirect Wyner-Ziv coding
  - Generalization of the RD function in the quadratic-Gaussian case

- Example with experimental results
  - Simple Gaussian mixture
  - Information-theoretic vs. operational RD performance
Noisy Source Coding with Decoder Side Information

- Statistical dependence is known
- Side information *not* available at the encoder
- Is there a loss in rate-distortion (RD) performance?
Indirect Wyner-Ziv (WZ) Source Coding Problem

- In the quadratic-Gaussian case
  - No loss in RD performance
  - Closed expression for RD function
- We generalize the conditions in the current literature for
  - Direct case (source data observed)
  - Indirect case (noisy observation)
- Proofs do not use duality arguments
Some Fundamental Results on WZ Coding Theory

Directly observed data
- One-letter characterization and quadratic-Gaussian case [Wyner, Ziv, 76-78]
- Case in which source data is sum of arbitrarily distributed side information and independent Gaussian noise [Pradhan, Chou, Ramchandran, 03]
- Side-information-dependent distortion functions [Csiszár, Körner, 81], [Linder, Zamir, Zeger, 00]

Indirectly observed data
- One-letter characterization and quadratic-Gaussian case [Yamamoto, Itoh, 80], [Flynn, Gray, 87], [Draper, 02]
- Modified distortion functions to reduce indirect WZ problems to direct WZ problems [Witsenhausen, 80]
Definitions

- $X, Y, Z$ r.v. defined on a common probability space, taking values in alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$
- $((X_i, Y_i, Z_i))_{i \in \mathbb{Z}^+}$ i.i.d. drawings of $(X, Y, Z)$
- $\mathcal{X}^\ast$ measurable space. A distortion function is a measurable function $d : \mathcal{X}^\ast \times \mathcal{X}^\ast \to [0, \infty)$
- $n, M \in \mathbb{Z}^+$. A code consists of two measurable mappings
  - Encoder $q : \mathcal{Z}^n \to \{1, \ldots, M\}$
  - Decoder $\hat{x}^n : \{1, \ldots, M\} \times \mathcal{Y}^n \to \mathcal{X}^\ast^n$
The indirect WZ RD function is defined as
\[
R_{XZ|Y}^{NWZ}(D) = \inf \{ R | (R, D) \text{ achievable} \}
\]

<table>
<thead>
<tr>
<th>RATE-DISTORTION FUNCTION</th>
<th>( Y ) Also Available at Encoder (Conditional Coding)</th>
<th>( Y ) Available at Decoder Only (WZ Coding)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Case ( Z = X )</td>
<td>( R_{XZ</td>
<td>Y}(D) )</td>
</tr>
<tr>
<td>Indirect, Noisy Case</td>
<td>( R_{XZ</td>
<td>Y}^{N}(D) )</td>
</tr>
</tbody>
</table>

When does \( R_{XZ|Y}^{WZ}(D) = R_{XZ|Y}(D) \), or more generally, \( R_{XZ|Y}^{NWZ}(D) = R_{XZ|Y}^{N}(D) \), hold? Closed formula?
Generalization of the WZ RD Function in the Direct, Quadratic-Gaussian Case

- Direct case $Z = X$
- Theorem: $\mathcal{X} = \mathcal{X}^* = \mathbb{R}$, $d(x, \hat{x}) = (x - \hat{x})^2$, $\mu : \mathcal{Y} \rightarrow \mathbb{R}$ measurable, and $\sigma^2 > 0$. Suppose that $X \overset{d}{=} \mu(Y) + N$, with $N \sim \mathcal{N}(0, \sigma^2)$, independent from $Y$. Then, for all $D > 0$,
  \[ R_{X|Y}^{WZ}(D) = R_{X|Y}(D) = \frac{1}{2} \log^+ \frac{\sigma^2}{D} \]
- $\mathcal{Y}$ is arbitrary

\[ N \sim \mathcal{N}(0, \sigma^2) \]

This slightly relaxes the hypotheses in the literature

- Direct proof, not using duality arguments
Generalization of the WZ RD Function in the Indirect, Quadratic-Gaussian Case

Theorem: $\mathcal{X} = \hat{\mathcal{X}} = \mathcal{Z} = \mathbb{R}$, $d(x, \hat{x}) = (x - \hat{x})^2$, $\mu : \mathcal{Y} \to \mathbb{R}$ measurable, and $\sigma^2 > 0$. $D_\infty = E_{YZ} \text{Var}[X|Y, Z]$. Suppose

- $Z \overset{a.s.}{=} \mu(Y) + N$, $N \sim N(0, \sigma^2)$, independent from $Y$
- $X \overset{a.s.}{=} f(Y) + \alpha Z + N'$, where $\alpha \in \mathbb{R}$, $f : \mathcal{Y} \to \mathbb{R}$ measurable, $N'$ r.v. satisfying $E[N'|y, z] = 0$ for a.e. $y \in \mathcal{Y}$, $z \in \mathbb{R}$

Then, for all $D > D_\infty$,

$$R_{XZ|Y}^{NWZ}(D) = R_{XZ|Y}^N(D) = \frac{1}{2} \log^+ \frac{\alpha^2 \sigma^2}{D - D_\infty}$$
Modified Distortion Functions

- Proposition: Let $d : \mathcal{X} \times \hat{\mathcal{X}} \times \mathcal{Y} \times \mathcal{Z} \rightarrow [0, \infty)$ be measurable. Define $\tilde{d}(z, \hat{x}, y) = E[d(X, \hat{x}, y, z) | y, z]$. Let $Q$ be a r.v. in some alphabet $\mathcal{Q}$. Assume that $(X, Y) \leftrightarrow Z \leftrightarrow Q$ or $X \leftrightarrow (Y, Z) \leftrightarrow Q$, and that there exists a measurable function $\hat{x} : \mathcal{Q} \times \mathcal{Y} \rightarrow \hat{\mathcal{X}}$ such that $\hat{X} = \hat{x}(Q, Y)$. Then, $E\tilde{d}(X, \hat{X}, Y, Z) = E\tilde{d}(Z, \hat{X}, Y)$.

- Based on [Witsenhausen, 80]
- Can be used to reduce indirect WZ problem to direct WZ problem
- $Z$ becomes direct source data, replacing $X$
- Modified distortion function is side-information-dependent
- Markov conditions correspond to WZ and conditional cases, respectively
Overview of the Proofs

Indirect RD Function

Modified Distortion Function
\[ \tilde{d}(z, \hat{x}, y) = E[d(X, \hat{x}, y, z) \mid y, z] \]

Side-Information-Dependent Distortion Functions of the Form
\[ d(x, \hat{x}, y) = (\alpha x - \hat{x} + f(y))^2 \]

Direct RD Function

Conditional Gaussian Maximizes Conditional Differential Entropy Subject to Unconditional Power Constraint

Modification of Steps to Prove Shannon Lower Bound

Test Channel Technique To Find Non-Distributed RD Function in the Quadratic-Gaussian Case
Intuitive Example

- $Y$ discrete random state, uniformly distributed on $\mathcal{Y} = \{-1, 1\}$
- $Z = X + Y$, $X \sim \mathcal{N}(0, 1)$ independent from $Y$
- $D = \frac{1}{n} \mathbb{E} \|X^n - \hat{X}^n\|^2$
- RD function
  - If $Y$ were available at the encoder, $X = Z - Y$ could be encoded, and $D_{XZ|Y}(R) = 2^{-2R} = D_{NWX|Y}(R)$
  - If $Z$ were encoded and decoded ignoring $Y$, and we defined $\hat{X} = \hat{Z} - Y$, then $D(R) \simeq 1.96 2^{-2R}$
Operational RD Performance

- Operational coding assumes $\mathcal{R} = \frac{1}{n} H(Q|Y^n)$.

- Since $E[X|y, z] = z - y$ additively separable, according to the high-rate noisy WZ quantization theory in [Rebollo-Monedero, Rane, Girod, 04], for each $n$:
  - Lattice quantizer followed by SW coder asymptotically optimal
  - $D(\mathcal{R}) \simeq 2\pi e M_n \sigma^2 2^{-2 \mathcal{R}}$ ($M_n$ normalized moment of inertia)

- Lloyd algorithm extended for noisy WZ quantization design in [Rebollo-Monedero, Girod, 05]
RD Comparison

- Slope $\approx 6.02 \text{ dB/bit}$
- Distortion gap w.r.t. RD function $2\pi e M_n \approx 1.53, 1.37, 1.28 \text{ dB}$
Conclusions

- Conditions to ensure that no rate loss incurred in WZ coding
  - Jointly Gaussian statistics are *not* necessary
  - Relaxed conditions in both the direct and indirect case
  - Side information arbitrarily distributed, discrete or continuous
- In the indirect case, condition on the data similar to additive separability condition in high-rate noisy WZ quantization
- All conditions determined by conditional joint distribution of (data, observation) given side information for arbitrarily distributed side information