On the Multiple Description Coding Problem with One Semi-deterministic Distortion Measure

Lin Zhou (ECE, NUS)
Joint work with Prof. Mehul Motani (ECE, NUS)

Globecom, Singapore, Dec. 2017
First-Order Coding Region: DMSes

- A set of conditional distributions
  \[ \mathcal{P}(D_1, D_2) := \{ (P_{X_i|X_1}, P_{X_i|X_2}) : \mathbb{E}[d_i(X, \hat{X}_i)] \leq D_i, i \in \{1, 2\} \} \]

- A set of rate pairs: given \( (P_{X_i|X_1}, P_{X_i|X_2}) \),
  \[ \mathcal{R}(P_{X_i|X_1}, P_{X_i|X_2}) := \left\{ (R_1, R_2) : R_1 \geq I(X; \hat{X}_1), R_2 \geq H(Y), R_1 + R_2 \geq H(Y) + I(\hat{X}_1; Y) + I(X; \hat{X}_1 \hat{X}_2|Y) \right\} \]

Theorem 1 (Fu and Yeung, 2002)

For any DMS under bounded distortion measures, we have

\[ \mathcal{R}(D_1, D_2|P_X) = \bigcup_{(P_{X_i|X_1}, P_{X_i|X_2}) \in \mathcal{P}(P_X, D_1, D_2)} \mathcal{R}(P_{X_i|X_1}, P_{X_i|X_2}) \]

Definitions

- Separable distortion measures: \( d_i(X^n, \hat{X}_i^n) := \frac{1}{n} \sum_{i=1}^n d_i(X_i, \hat{X}_i) \)
  where \( d_i : X \times \hat{X} \rightarrow [0, \infty) \), \( i \in \{1, 2\} \).
- Joint excess-distortion and error probability
  \[ P_{e,n}(D_1, D_2) := \Pr \left\{ d_1(X^n, \hat{X}_1^n) > D_1 \text{ or } d_2(X^n, \hat{X}_2^n) > D_2 \text{ or } \hat{Y}^n \neq Y^n \right\} \]
- \( (R_1, R_2) \) is \( (D_1, D_2) \)-achievable if \( \exists \) a sequence of \( (n, M_1, M_2) \)-codes such that
  \[ \lim_{n \to \infty} \frac{1}{n} \log M_i \leq R_i, i \in \{1, 2\} \text{ and } \lim_{n \to \infty} P_{e,n} = 0. \]
- First-order coding region \( \mathcal{R}(D_1, D_2|P_X) \): closure of all \( (D_1, D_2) \)-achievable pairs.

Extreme Points of First-Order Coding Region

- Minimum rate for encoder \( f_1 \)
  \[ \min \left\{ R_1 : \exists R_2 \text{ s.t. } (R_1, R_2) \in \mathcal{R}(D_1, D_2|P_X) \right\} = \min_{P_{X|X}|\mathcal{X} \subseteq \mathcal{D}(X; \hat{X}_1)} I(X; \hat{X}_1) = R(P_X; D_1) \]
- Minimum rate for encoder \( f_2 \)
  \[ \min \{ R_2 : \exists R_1 \text{ s.t. } (R_1, R_2) \in \mathcal{R}(D_1, D_2|P_X) \} = H(Y) \]
- Minimum sum-rate with fixed rate \( R_1 \) for encoder \( f_1 \)
  \[ \min \{ R_1 + R_2 : (R_1, R_2) \in \mathcal{R}(D_1, D_2|P_X) \} = H(Y) + R(R_1, D_1, D_2|P_X), \]
  where
  \[ R(R_1, D_1, D_2|P_X) := \min_{(P_{X|X}|\mathcal{X} \subseteq \mathcal{D}(X; \hat{X}_1)), R_1 \geq I(X; \hat{X}_1)} I(X; \hat{X}_1) + I(X; \hat{X}_1 \hat{X}_2|Y). \]
Optimal Test Channels: Preliminaries

- Convex optimization problem: \( R(R_1, D_1, D_2|P_X) \)
  \[
  I(\hat{X}_2; Y) + I(X; \hat{X}_1 \hat{X}_2|Y) = I(XY; \hat{X}_2) + I(X; \hat{X}_2|Y \hat{X}_1) \\
  = I(X; \hat{X}_2) + I(X; \hat{X}_2|Y \hat{X}_1).
  \]

- Optimal solutions to the dual problem:
  \[
  s^* := -\frac{\partial R(R_1, D_1, D_2|P_X)}{\partial R} \bigg|_{R=R_1} , \\
  t_1^* := -\frac{\partial R(R_1, D_1, D_2|P_X)}{\partial D} \bigg|_{D=D_1} , \\
  t_2^* := -\frac{\partial R(R_1, D_1, D_2|P_X)}{\partial D} \bigg|_{D=D_2}.
  \]

- Definitions: given distributions \((Q_{\hat{X}_1}, Q_{\hat{X}_2}|Y \hat{X}_1)\) and \((x, y, \hat{x}_1)\),
  \[
  \beta_2(x, y, \hat{x}_1 | Q_{\hat{X}_1}, Q_{\hat{X}_2}|Y \hat{X}_1) := \left\{ \mathbb{E}_{Q_{\hat{X}_1}, Q_{\hat{X}_2}|Y \hat{X}_1} \left[ \exp(-t_2^* d_2(x, \hat{x}_2)) | Y = y, \hat{X}_1 = \hat{x}_1 \right] \right\}^{-1},
  \]
  \[
  \beta(x, y | Q_{\hat{X}_1}, Q_{\hat{X}_2}|Y \hat{X}_1) := \left\{ \mathbb{E}_{Q_{\hat{X}_1}} \left[ \exp \left( -t_1^* d_1(x, \hat{x}_1) + \log \beta_2(x, y, \hat{x}_1 | Q_{\hat{X}_1}, Q_{\hat{X}_2}|Y \hat{X}_1) \right) \right] \right\}^{-1}.
  \]

Optimal Test Channels: Properties

**Lemma 2**

A pair of test channels \((P^*_{\hat{X}_1|X \hat{x}_1}, P^*_{\hat{X}_2|X \hat{x}_1})\) achieves \( R(R_1, D_1, D_2|P_X) \) if and only if

- For all \((x, y, \hat{x}_1, \hat{x}_2)\) such that \( y = g(x) \),
  \[
  P^*_{\hat{X}_1|X \hat{x}_1}(\hat{x}_1|x) = P^*_{\hat{X}_1}(\hat{x}_1) \beta(x, y | P^*_{\hat{X}_1}, P^*_{\hat{X}_2}|Y \hat{X}_1) \\
  \times \exp \left( -t_1^* d_1(x, \hat{x}_1) + \log \beta_2(x, y, \hat{x}_1 | P^*_{\hat{X}_1}, P^*_{\hat{X}_2}|Y \hat{X}_1) \right),
  \]

- For all \((x, y, \hat{x}_1, \hat{x}_2)\) such that \( y = g(x) \) and \( P^*_{\hat{X}_1|X \hat{x}_1}(\hat{x}_1|x) > 0 \)
  \[
  P^*_{\hat{X}_2|X \hat{x}_1}(\hat{x}_2|x, \hat{x}_1) = P^*_{\hat{X}_2|Y \hat{X}_1}(\hat{x}_2|y, \hat{x}_1) \beta_2(x, y, \hat{x}_1 | P^*_{\hat{X}_1}, P^*_{\hat{X}_2}|Y \hat{X}_1) \\
  \times \exp \left( -t_2^* d_2(x, \hat{x}_2) \right).
  \]

Rate-distortion-tilted information density

- Definition: given optimal test channels \((P^*_{\hat{X}_1|X \hat{x}_1}, P^*_{\hat{X}_2|X \hat{x}_1})\)
  \[
  j(x, y | R_1, D_1, D_2, P_X) := (1 + s^*) \log \beta(x, y | P^*_{\hat{X}_1|X \hat{x}_1}, P^*_{\hat{X}_2|X \hat{x}_1}) - s^* R_1 - t_1^* D_1 - t_2^* D_2.
  \]

- Properties
  - Expected value
    \[
    R(R_1, D_1, D_2|P_X) = \mathbb{E}[j(X, Y | R_1, D_1, D_2, P_X)].
    \]
  - Expansion: for \((x, y, \hat{x}_1, \hat{x}_2)\) such that \( y = g(x) \) and \( P^*_{\hat{X}_1}(\hat{x}_1) P^*_{\hat{X}_2|Y \hat{X}_1}(\hat{x}_2|g(x), \hat{x}_1) > 0 \), we have
    \[
    j(x, y | R_1, D_1, D_2, P_X) = (1 + s^*) \log \frac{P^*_{\hat{X}_1|X \hat{x}_1}(\hat{x}_1|x)}{P^*_{\hat{X}_1}(\hat{x}_1)} + \log \frac{P^*_{\hat{X}_2|X \hat{x}_1}(\hat{x}_2|g(x), \hat{x}_1)}{P^*_{\hat{X}_2|Y \hat{X}_1}(\hat{x}_2|y, \hat{x}_1)} \\
    + t_1^* (d_1(x, \hat{x}_1) - D_1) + t_2^* (d_2(x, \hat{x}_2) - D_2).
    \]
Non-Asymptotic Converse Bound: Preliminaries

- Distortion-tilted information density\(^3\), i.e.,
  \[ j(x, D_1 | P_X) := -\log \left( \sum_{\hat{x}_1} P_{\hat{x}_1}^* \exp(-t^*(d_1(x, \hat{x}_1) - D_1)) \right), \]
where \(P_{\hat{x}_1}^*\) is induced by the source distribution \(P_X\) and the optimal test channel \(P_{\hat{x}_1}^*: X \rightarrow \hat{x}_1\) for the rate-distortion function \(R(P_X, D_1)\) and \(t^* = -\frac{\partial R(P_X, D_1)}{\partial D_1}|_{D_1=D}. \)
- Three sets: given any \(\gamma \geq 0\)
  \[ A_1 := \{(x, y) : j(x, D_1 | P_X) \geq \log M_1 + \gamma\}, \]
  \[ A_2 := \{(x, y) : -\log P_Y(y) \geq \log M_2 + \gamma\}, \]
  \[ A_3 := \{(x, y) : j(x, y | R_1, D_1, D_2, P_X) \geq \log M_1 M_2 + s^* \log M_1 + (1 + s^*) \gamma\}. \]


Non-Asymptotic Converse Bound

Lemma 3

Any \((1, M_1, M_2)\)-code satisfies that for any \(\gamma \geq 0\),
\[ P_{e,n}(D_1, D_2) \geq \Pr \{(X, Y) \in (A_1 \cup A_2 \cup A_3)\} - 4\exp(-\gamma). \]
- Proof invokes the properties of optimal test channels and arguments by Kostina and Verdú\(^4\);
- Hold for arbitrary source, discrete or continuous;
- Can specialize to the successive refinement problem\(^5\) and generalize our previous result\(^6\).

\(^4\)“A new converse in rate-distortion theory”, CISS, pp. 1–6, 2012

Second-Order Asymptotics for DMSes: Definition

- Achievable second-order pair: Given a rate-distortion tuple \((R_1, R_2, D_1, D_2)\) and any \(\epsilon \in [0, 1]\), \((L_1, L_2)\) is second-order achievable if \(\exists\) a sequence of \((n, M_1, M_2)\)-codes such that
  \[ \lim_{n \to \infty} \sup \frac{\log M_i - n R_i}{\sqrt{n}} \leq L_i, \ i = 1, 2, \]
  \[ \lim_{n \to \infty} P_e(n, D_1, D_2) \leq \epsilon. \]
- Second-order coding region:
  \[ \mathcal{L}(R_1, R_2, D_1, D_2, \epsilon) := \{(L_1, L_2) : (L_1, L_2) \text{ is second-order achievable given } (R_1, R_2, D_1, D_2, \epsilon)\}. \]
Under mild conditions, for any $\varepsilon \in [0, 1)$, we have

- **Case (i):** $R_1 = R(P_X, D_1)$ and $R_1 + R_2 > R(R_1, D_1, D_2 | P_X) + H(P_Y)$

  \[
  \mathcal{L}(R_1, R_2, D_1, D_2, \varepsilon) = \{(L_1, L_2) : L_1 \geq \sqrt{V(P_X, D_1)Q^{-1}(\varepsilon)}\}.
  \]

- **Case (ii):** $R_2 = R(P_X, D_1)$ and $R_1 + R_2 = R(R_1, D_1, D_2 | P_X) + H(P_Y)$

  \[
  \mathcal{L}(R_1, R_2, D_1, D_2, \varepsilon)
  = \{(L_1, L_2) : (1 + s^*)L_1 + L_2 \geq 1 - \varepsilon\}.
  \]

### Summary

- Revisit the Fu-Yeung problem, a special case of the multiple description coding where the El Gamal-Cover inner bound is tight.
- Our results: optimal test channels for the minimum sum-rate function, non-asymptotic converse bound, second-order asymptotics for DMSes.
- Large and moderate deviations available in our full version on arXiv:1708.05496.
- Derive tight non-asymptotic converse bound.
- Derive second-order coding region for Gaussian memoryless sources under quadratic distortion measures for the multiple description coding.

---