The Exponential Distribution in Rate Distortion Theory: The Case of Compression with Independent Encodings

Uri Erez[1] Jan Østergaard[2] Ram Zamir[1]

[1]Tel Aviv University

[2]Aalborg University

IEEE Data Compression Conference, March 2020

Outline

- Motivations
- Independent encodings — a powerful source coding paradigm (as we demonstrate)
- The one-sided exponential source
- Jointly optimal decoding – select the maximum
- Asymptotical optimal encoding at low rates
- Example
- Conclusions and future work

Motivations

- To combat packet erasures on erroneous packet-switched networks, one may introduce channel coding such as forward error correction (FEC) codes.
- FEC codes suffers from the cliff-wall effect: If too few packets are received, no recovery is possible.
- An alternative is to use joint source-channel coding techniques such as multiple-description (MD) coding.
- In MD coding, a graceful degradation in the case of packet loss is possible.
- Any single description yields a desired performance.
- If more descriptions are received, the quality increases, thereby avoiding the cliff-wall effect.
- Unfortunately, it is difficult to design good MD codes for general sources and distortion measures.

Independent Encodings

- Independent Encodings are a special case of multiple-description (MD) coding.
- With *Independent Encodings*, k descriptions are constructed so they are invidually rate-distortion optimal.
- We do not control the performance achieved using any sub set of descriptions.
- This is the case of no excess marginal rates in MD coding.
- We define $Y_1, \ldots, Y_k$, to be *independent* encodings of the common source X, if and only if, they are conditionally mutually independent given X

The Quadratic Gaussian Case

- Independent encodings for the Gaussian case and under MSE is well-understood [Østergaard & Zamir, ISIT'11].
- In this case the optimal decoder is the best linear estimator.
- There exists simple closed-form expressions, which imply that a low resolution (rate), independent encoding is jointly rate-distortion optimal for any number of encodings.
  [Østergaard & Zamir, ISIT'11].

- The question we address in this paper, is whether we can find another nice source and distortion measure, which leads to simple analytic expressions, and which is rate-distortion optimal at low resolution.

The one-sided exponential distribution

- The one-sided exponential distribution is given by

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0,$$

  with mean $\mathbb{E}[X] = \lambda^{-1}$.
- Let x be the source and y the reproduction of the coder
- The one-sided absolute error criterion is given by:

$$d(x, y) = \begin{cases} x - y, & \text{if } x \geq y, \\ \infty, & \text{if } x < y. \end{cases}$$

- The rate-distortion function for the exponential source with one-sided error criterion is given by [Verdu'96]:

$$R(D) = \begin{cases} -\log(\lambda D), & 0 \leq D \leq \lambda^{-1}, \\ 0, & D > \lambda^{-1}. \end{cases} \tag{1}$$

Backward test channel

- Let $Z + Y = X$ denote the *backward* test channel whose optimal conditional "output" distribution is given by [Verdu'96]:

$$f_{X|Y}(x|y) = \frac{1}{D} e^{-\frac{(x-y)}{D}}, \quad x \geq y \geq 0.$$

- The channel noise, Z, is one-sided exponential distributed with parameter $1/D$.
- Z and Y are here mutually independent.
- Let the k independent encodings be given as:

$$Y_i + Z_i = X, \quad i = 1, \ldots, k,$$

  where Z_i, Z_j are mutually independent for $i \neq j$.
- The following Markov chains apply: $Y_i - X - Y_j, \forall i, j$.

Select-Max Estimator — Optimal Decoder

- Given k independent encodings, we define the *select-max* estimator of X as:

$$\hat{Y} = \max\{Y_1, \ldots, Y_k\}$$

- **Lemma 1**

  Given k independent encodings $Y_1, \ldots, Y_k$ of the one-sided exponential source, the select-max estimator $\hat{Y} = \max\{Y_1, \ldots, Y_k\}$ is an optimal estimator under the one-sided absolute distortion measure.

Total Distortion using Optimal Decoding

- **Lemma 2**

  Let $Y_1, \ldots, Y_k$, be independent encodings of the one-sided exponential source, each resulting in distortion D.

  The total expected distortion $\bar{D}$ due to using the select-max estimator on the k encodings $Y_1, \ldots, Y_k$ is given by:

$$\bar{D} = \frac{D}{k - (k-1)\lambda D}.$$

Distortion Rate Function

- Let the distortion be given by $D = \frac{1-\epsilon}{\lambda}$, for any $1 > \epsilon > 0$.
- Then the rate for a single description is given by the RDF:

$$R(D) = -\log_2(D\lambda) = -\log_2(1-\epsilon).$$

- The DRF is given by:

$$D(R) = \frac{1}{\lambda}2^{-R}.$$

- Inserting the rate of k times the single-description rate into the DRF leads to a distortion of:

$$D(kR) = \frac{1}{\lambda}2^{k\log_2(1-\epsilon)} = \frac{1}{\lambda}(1-\epsilon)^k.$$

Operational Distortion-Rate Function

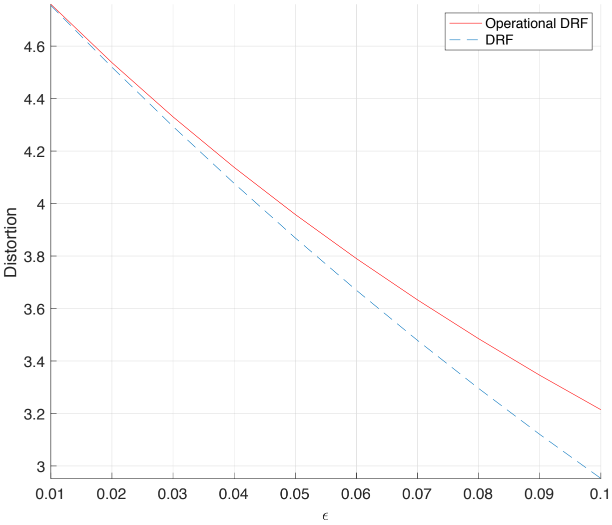
- The operational DRF using k independent encodings and the select-max estimator yields:

$$\bar{D} = \frac{1-\epsilon}{\lambda(1+\epsilon(k-1))}.$$

- It is easy to show that the ratio of the derivatives of the DRF and the ODRF wrt. ϵ tends to 1 as $\epsilon \to 0$, i.e.:

$$\lim_{\epsilon \to 0} \frac{\frac{\partial}{\partial \epsilon} \frac{1}{\lambda}(1-\epsilon)^k}{\frac{\partial}{\partial \epsilon} \frac{1-\epsilon}{\lambda(1+\epsilon(k-1))}} = \lim_{\epsilon \to 0} \frac{-k(1-\epsilon)^{k-1}}{-k(1+\epsilon(k-1))^{-2}} = 1,$$

Example



- $D = (1 - \epsilon)\lambda^{-1}, \quad 0 < \epsilon \leq 1$
- $\lambda = 0.2, \lambda^{-1} = 5$, and $k = 5$ parallel channels.

Conclusions

- We introduced the concept of *independent encodings*, which is a special case of multiple-description coding.
- Each description is rate-distortion optimal by itself but they are not necessarily jointly rate-distortion optimal.
- For the case of exponential sources and one-sided error distortion criteria, we derived the optimal decoder (select-max) given k independent encodings.
- At small rates and under the one-sided error criterion, it is nearly rate-distortion optimal, to encode the one-sided exponential source into k independent encodings and then reconstruct using the the maximum of the received descriptions.

Future work: Multi-round independent encodings

- Assume we encode the source into k independent encodings to be transmitted over a network with feedback.
- We then receive feedback upon which independent encodings are received.
- We can now estimate the source from the received descriptions, and encode the residual into k new independent encodings.
- Thus, we combine MD coding and successive refinement.
- Each round uses a very small sum-rate and is therefore nearly rate-distortion optimal.
- Any desired quality can be achieved in a finite number of rounds.