Exponential Error Bounds on Parameter Modulation–Estimation for DMC's

Neri Merhav

Department of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 32000, Israel

ISIT 2013, Istanbul, Turkey, July 2013.

Background and Motivation

A random parameter U is to be conveyed to a given destination via a DMC

$$p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{t=1}^{n} p(y_t|x_t).$$

For large n, how well can the decoder estimate U if we have the freedom to design the modulator $\boldsymbol{X} = f(U) = (f_1(U), \ldots, f_n(U))$?

- Discrete–time analogue of the “waveform communication” problem.
- Can be viewed both from the perspectives of IT and estimation theory.
  - IT: Joint source–channel coding with source block length = 1.
  - Estimation theory: Most bounds – derivable for a given modulator.
  - Exceptions: DPT bounds and Chazan–Ziv–Zakai bounds.

Background and Motivation (Cont'd)

- We use both techniques for upper and lower bounds on $\mathbf{E}|\hat{U} - U|^\rho$, $\rho \geq 0$.
- Modulator opt. $\to \mathbf{E}|\hat{U} - U|^\rho$ can be made exponentially small in n.
- Accordingly, we wish to find the fastest achievable exponential rate:

$$\inf \mathbf{E}|\hat{U} - U|^\rho \doteq e^{n\mathcal{E}(\rho)}.$$

- Bounds on $\mathcal{E}(\rho)$ – related to coding error exponents and capacity.
  - Small ρ – random coding exponent.
  - Large ρ – expurgated exponent.
  - Upper and lower bounds asymp. agree both for $\rho \to 0$ and $\rho \to \infty$.
  - Extension to parameter vectors: bounds coincide for large dimension.

Problem Formulation

- Let $U \sim$ Unif$[-1/2, +1/2]$ and let $p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{t=1}^{n} p(y_t|x_t)$ be given.
- Modulator: $\boldsymbol{X} = f_n(U)$.
- Estimator: $\hat{U} = g_n(\boldsymbol{Y})$.
- Upper limiting exponent:

$$\overline{\mathcal{E}}(\rho) \stackrel{\triangle}{=} \limsup_{n\to\infty} \left[-\frac{1}{n} \ln \left(\inf_{f_n, g_n} \mathbf{E}|\hat{U} - U|^{\rho} \right) \right].$$

- Lower limiting exponent:

$$\underline{\mathcal{E}}(\rho) \stackrel{\triangle}{=} \liminf_{n\to\infty} \left[-\frac{1}{n} \ln \left(\inf_{f_n, g_n} \mathbf{E}|\hat{U} - U|^{\rho} \right) \right].$$

- We seek upper bounds on $\overline{\mathcal{E}}(\rho)$ and lower bounds on $\underline{\mathcal{E}}(\rho)$.
- When do the upper bounds and the lower bounds come close?

Some Definitions

$$E_0(\rho, q) = -\ln\left(\sum_{y\in\mathcal{Y}}\left[\sum_{x\in\mathcal{X}} q(x)p(y|x)^{1/(1+\rho)}\right]^{1+\rho}\right), \quad \rho \geq 0.$$

$$E_0(\rho) = \max_q E_0(\rho, q), \qquad \bar{E}_0(\rho) = \mathsf{UCE}\{E_0(\rho)\}.$$

$$E_x(\varrho) = -\varrho \ln\left(\sum_{x,x'\in\mathcal{X}} q(x)q(x')\left[\sum_{y\in\mathcal{Y}}\sqrt{p(y|x)p(y|x')}\right]^{1/\varrho}\right).$$

$$E_{ex}(R) = \sup_{\varrho\geq 1}[E_x(\varrho) - \varrho R].$$

It is well known that

$$E_{ex}(0) = \sup_{\varrho\geq 1} E_x(\varrho) = -\sum_{x,x'\in\mathcal{X}} q(x)q(x')\ln\left[\sum_{y\in\mathcal{Y}}\sqrt{p(y|x)p(y|x')}\right].$$

Upper Bound

Theorem 1: For every $\rho \geq 0$,

$$\overline{\mathcal{E}}(\rho) \leq \overline{E}(\rho) \triangleq \begin{cases} \overline{E}_0(\rho) & \rho \leq \rho_0 \\ E_{ex}(0) & \rho > \rho_0 \end{cases}$$

where ρ_0 is the (unique) solution to the equation $\overline{E}_0(\rho) = E_{ex}(0)$.

Main idea of the proof: By Chebychev's inequality

$$\mathbf{E}|\hat{U} - U|^\rho \geq e^{-n\rho R}\text{Pr}\{|\hat{U} - U| \geq e^{-nR}\} \geq e^{-n\rho R} \cdot e^{-nE(R)},$$

where the 2nd inequality comes from the sub-optimality of a decoder based on $\hat{U}$. Now, use upper bounds on $E(R)$ and maximize over R.

Lower Bound

Let

$$R_- = \inf\{R : \ E_{ex}(R) \text{ is attained by } \rho = 1\}$$

$$R_+ = \sup\{R : \ E_r(R) \text{ is attained by } \rho = 1\}$$

$$\rho_+ = \frac{E_0(1) - R_+}{R_+}; \quad \rho_- = \frac{E_0(1) - R_-}{R_-}.$$

Theorem 2: For every $\rho \geq 0$,

$$\underline{\mathcal{E}}(\rho) \geq \underline{E}(\rho) \stackrel{\triangle}{=} \begin{cases} \sup_{0 \leq s \leq 1} \rho E_0(s)/(s+\rho) & \rho \leq \rho_+ \\ \rho E_0(1)/(1+\rho) = \rho E_x(1)/(1+\rho) & \rho_+ < \rho \leq \rho_- \\ \sup_{s \geq 1} \rho E_x(s)/(s+\rho) & \rho > \rho_- \end{cases}$$

The proof is by analyzing a scheme based on uniform quantization of U with spacing of e^{-nR} and assigning a (random) codeword to each lattice point. Here, $R = \underline{E}(\rho)/\rho$.

Discussion

- Both bounds are given in terms of $E_0(\cdot)$ for small ρ, and $E_x(\cdot)$ for large ρ.
- Both are subjected to phase transitions.
- For small ρ: $\overline{E}(\rho) \sim \underline{E}(\rho) \sim \rho C$.
- For large ρ: $\overline{E}(\rho) \sim \underline{E}(\rho) \sim E_{ex}(0)$.
- Two last points imply: “separation” theorem in these two extremes, although source block-length$= 1$!

Data processing Upper Bound

Another upper bound, based on generalized data processing inequalities:

$$\overline{\mathcal{E}}(\rho) \leq E_{DPT}(\rho),$$

where

$$E_{DPT}(\rho) = \inf_{k>1} \inf_{\alpha_1,\ldots,\alpha_k} \sup_q \frac{E(\alpha_1,\ldots,\alpha_k,q)}{\sum_{i=1}^k \zeta_\rho(\alpha_i)},$$

with

$$E(\alpha_1,\ldots,\alpha_k,q) = -\ln\left[\sum_{y\in\mathcal{Y}} \prod_{i=1}^k \left(\sum_{x_i\in\mathcal{X}} q(x_i)p(y|x_i)^{\alpha_i}\right)\right]$$

and

$$\zeta_\rho(\alpha) = \min\left\{\alpha, \frac{1-\alpha}{\rho}\right\}.$$

This bound is at least as tight as $\overline{E}(\rho)$, but we have not found (yet) an example that it is strictly so.

Example – Very Noisy Channel

The *very noisy channel* model is characterized by

$$p(y|x) = p(y)[1 + \epsilon(x, y)], \qquad |\epsilon(x, y)| \ll 1.$$

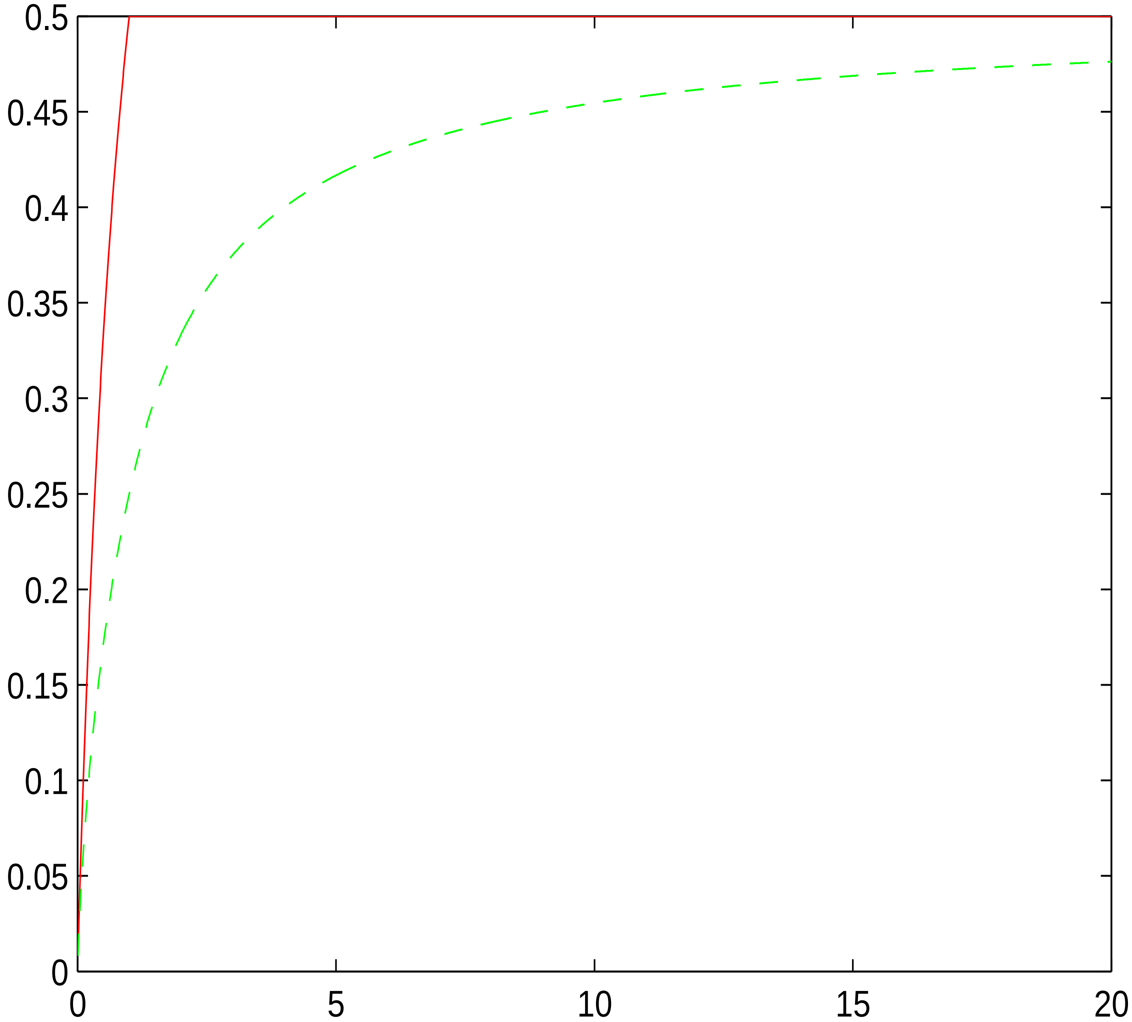
Here, we have:

$$\underline{E}(\rho) = \begin{cases} \frac{\rho}{(1+\sqrt{\rho})^2} \cdot C & \rho < 1 \\ \frac{\rho}{1+\rho} \cdot \frac{C}{2} & \rho \geq 1 \end{cases}$$

$$\overline{E}(\rho) = \begin{cases} \frac{\rho}{1+\rho} \cdot C & \rho \leq 1 \\ \frac{C}{2} & \rho > 1 \end{cases}$$

and

$$E_{DPT}(\rho) = C \cdot \inf_{k>1} \inf_{\alpha_1, \ldots, \alpha_k} \frac{1 - \sum_{i=1}^{k} \alpha_i^2}{\sum_{i=1}^{k} \zeta_\rho(\alpha_i)}.$$


0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20

Multidimensional Parameter

Let $\boldsymbol{U} = (U_1, \ldots, U_d) \sim$ Unif$[-1/2, +1/2]^d$. We are after bounds on

$$\sum_{i=1}^{d} e^{nr_i} \mathbf{E} \| \hat{U}_i - U_i |^{\rho}, \qquad \forall i \; r_i \geq 0, \quad \min_i r_i = 0.$$

The extension of the upper bound is

$$\bar{E}(\rho, d) = \begin{cases} E_0\left(\frac{\rho}{d}\right) - \frac{1}{d} \sum_{i=1}^{d} r_i & \rho/d \leq \rho_0 \\ E_{ex}(0) - \frac{\rho_0}{\rho} \sum_{i=1}^{d} r_i & \rho/d > \rho_0 \end{cases}$$

For $r_1 = \ldots = r_d = 0$, it is the same bound as before, except that ρ is replaced by ρ/d.

Thus, for d large, $\overline{E}(\rho/d) \sim \rho C/d$ coincides with the lower bound, achieved by forming a code on a lattice of close to $e^{nC/d}$ points in each dimension.

Summary

- We derived upper and lower bounds on exponential rates of $\mathbf{E}|\hat{U} - U|^\rho$, for every $\rho \geq 0$.
- The exponent bounds are related to random coding bounds and expurgated bounds.
- Asymptotic optimality of a simple separation–based scheme for both small and large ρ.
- We derived an alternative data processing upper bound.
- We extended the upper bound to the multidimensional case.
- Challenge: Close (or at least reduce) the gap for intermediate values of ρ.