The Generalized Stochastic Likelihood Decoder: Random Coding and Expurgated Bounds

Neri Merhav

Department of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 3200004, Israel

ISIT 2016, July 10–15, 2016.

The Likelihood Decoder

The likelihood decoder is a stochastic decoder that randomly selects an estimated message $\boldsymbol{x}_m$ by sampling the underlying posterior:

$$P(m|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{x}_m)}{\sum_{m'=1}^{M} P(\boldsymbol{y}|\boldsymbol{x}_{m'})}.$$

Motivation: Lends itself to easier analysis than the ordinary ML decoder.

Earlier work:

- Yassaee, Aref and Gohari (2013): network information theory.
- Song, Cuff and Poor (2014): source coding – likelihood encoder.
- Scarlett, Martinéz and Fábregas (2015): mismatched likelihood decoder. Matched case: optimal random coding error exponent.

Contributions

- A more general stochastic decoder: $P(m|\boldsymbol{y}) \propto \exp\{ng(\hat{P}_{\boldsymbol{x}_m\boldsymbol{y}})\}$.
- Tight error exponent in a single analysis.
- Extension to joint source–channel coding with side information.
- Expurgated exponent – at least as tight as the classical one.

Setup

- A DMC $\{W(y|x)\}$: $W(\boldsymbol{y}|\boldsymbol{x}) = \prod_{t=1}^{n} W(y_i|x_i)$.
- Random CCC $\mathcal{C} = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{M-1}\}$, $M = e^{nR}$. $\boldsymbol{X}_m \sim \{\mathcal{T}(Q_X)\}$.
- Generalized likelihood decoder (GLD):

$$P(m|\boldsymbol{y}) = \frac{\exp\{ng(\hat{P}_{\boldsymbol{x}_m \boldsymbol{y}})\}}{\sum_{m'=0}^{M-1} \exp\{ng(\hat{P}_{\boldsymbol{x}_{m'} \boldsymbol{y}})\}}.$$

Relevant choices of g:

- Ordinary LD: $g(\hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}) = \sum_{x,y} \hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}(x,y) \log W(y|x)$.
- With “temperature”: $g(\hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}) = \beta \sum_{x,y} \hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}(x,y) \log W(y|x)$.
- Mismatched LD: $g(\hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}) = \beta \sum_{x,y} \hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}(x,y) \log \tilde{W}(y|x)$.
- MMI LD: $g(\hat{P}_{\boldsymbol{x}_m \boldsymbol{y}}) = \beta I(\hat{P}_{\boldsymbol{x}_m \boldsymbol{y}})$.

Basic Result

Theorem: For two joint distributions, Q and Q', both on $\mathcal{X} \times \mathcal{Y}$, let

$$E_1(Q, Q', R) = [I(Q') - R + [g(Q) - g(Q')]_+]_+$$

Next define

$$E_2(Q, R) = \min_{\{Q' \colon Q'_X = Q_X,\ Q'_Y = Q_Y\}} E_1(Q, Q', R).$$

Then, the random coding error exponent of the GLD is given by

$$E(R) = \min_{Q} [D(Q \| Q_X \times W) + E_2(Q, R)].$$

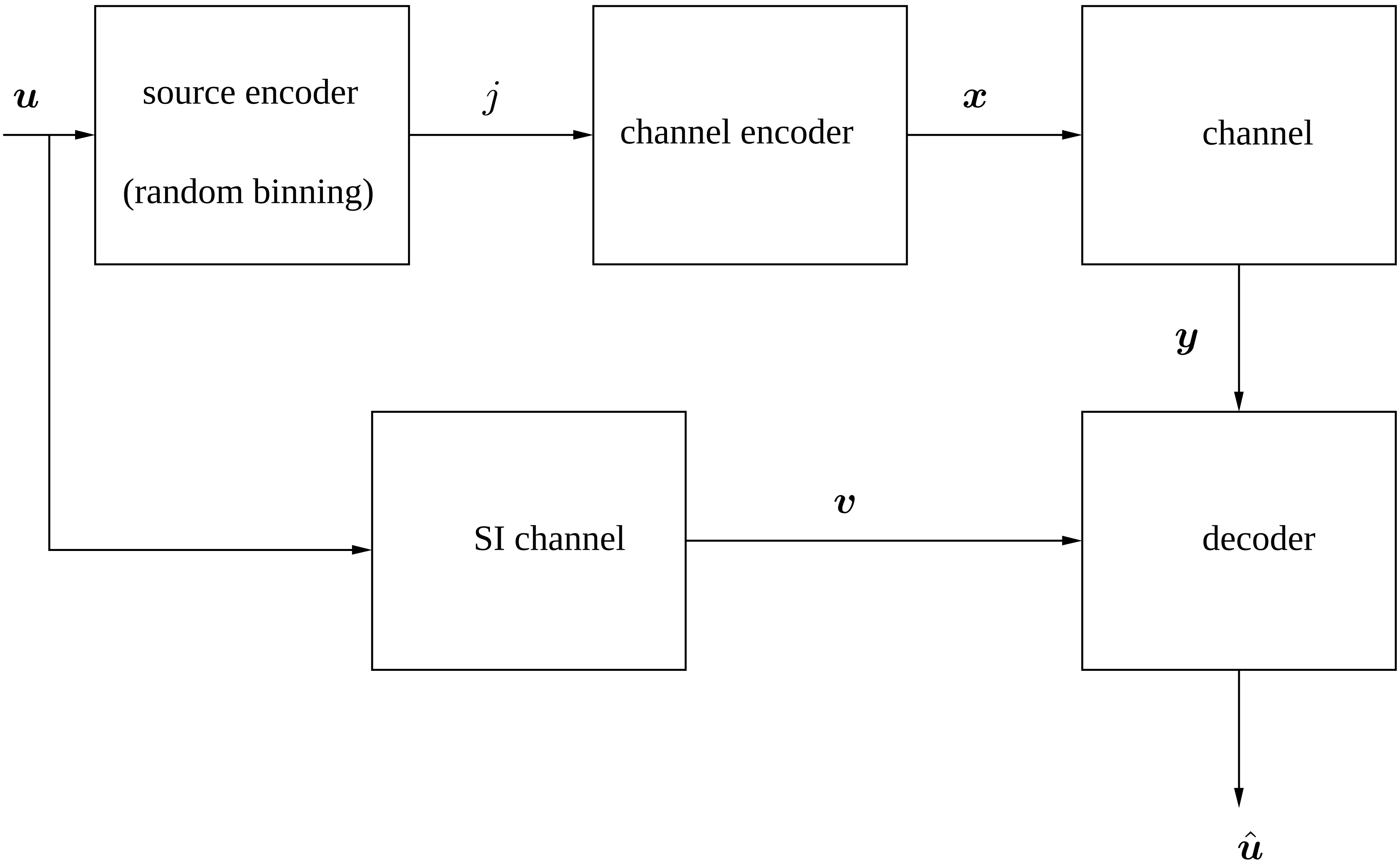
Comments

- The result by Scarlett *et al.* is obtained as a special case.
- Optimal exponent for $g(Q) = \beta \sum_{x,y} Q(x,y) \ln W(y|x)$, $\beta \geq 1$.
- Optimal exponent for $g(Q) = \beta I(Q)$, $\beta \geq 1$.
- $E(R) > 0$ for all $R \leq \min\{I(Q) : \ g(Q) \geq g(Q_X \times W)\}$.
- Single analysis, as opposed to separate upper and lower bounds.

Main Steps in the Derivation

$$\begin{aligned}
&\bar{P}_{\mathsf{e}}(\boldsymbol{x}_0,\boldsymbol{y}) \\
=\;& \mathbf{E}\left\{\frac{\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{x}_0\boldsymbol{y}})\}+\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}}\right\} \\
=\;& \int_0^1 \mathsf{Pr}\left\{\frac{\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{x}_0\boldsymbol{y}})\}+\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}}\geq t\right\}\mathsf{d}t \\
=\;& n\cdot\int_0^\infty e^{-n\theta}\mathsf{Pr}\left\{\frac{\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{x}_0\boldsymbol{y}})\}+\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}}\geq e^{-n\theta}\right\}\mathsf{d}\theta \\
\doteq\;& \int_0^\infty e^{-n\theta}\mathsf{Pr}\left\{\sum_{m=1}^{M-1}\exp\{ng(\hat{P}_{\boldsymbol{X}_m\boldsymbol{y}})\}\geq\exp\{n[g(\hat{P}_{\boldsymbol{x}_0\boldsymbol{y}})-\theta]\}\right\}\mathsf{d}\theta \\
\doteq\;& \int_0^\infty e^{-n\theta}\mathsf{Pr}\left\{\sum_{Q'}N_{\boldsymbol{y}}(Q')e^{ng(Q')}\geq\exp\{n[g(\hat{P}_{\boldsymbol{x}_0\boldsymbol{y}})-\theta]\}\right\}\mathsf{d}\theta
\end{aligned}$$

Joint/Separate Source–Channel Coding with SI



Motivation

Many customary models are coveed as special cases:

- Joint source–channel with/out SI (very large R).
- Pure Slepian–Wolf source coding (clean channel).
- Pure channel coding (uniform binary source, very large R).
- Systematic channel coding (SI channel = main channel).

Joint Source–Channel Likelihood Decoder

Randomly select the estimated source $\hat{\boldsymbol{u}}$ according to

$$P[\hat{\boldsymbol{u}} = \boldsymbol{u}|\boldsymbol{v}, \boldsymbol{y}] = \frac{P(\boldsymbol{u}, \boldsymbol{v})W(\boldsymbol{y}|\boldsymbol{x}[\boldsymbol{u}])}{\sum_{\boldsymbol{u}'} P(\boldsymbol{u}', \boldsymbol{v})W(\boldsymbol{y}|\boldsymbol{x}[\boldsymbol{u}'])}.$$

For a generalized version, consider

$$P[\hat{u} = \boldsymbol{u}|\boldsymbol{v}, \boldsymbol{y}] = \frac{\exp\{n[f(\hat{P}_{\boldsymbol{u}\boldsymbol{v}}) + g(\hat{P}_{\boldsymbol{x}[\boldsymbol{u}]\boldsymbol{y}})]\}}{\sum_{\boldsymbol{u}'} \exp\{n[f(\hat{P}_{\boldsymbol{u}'\boldsymbol{v}}) + g(\hat{P}_{\boldsymbol{x}[\boldsymbol{u}']\boldsymbol{y}})]\}},$$

for some given functions f and g.

Random Coding Exponent

$$h(Q_{UV}, Q_{XY}) = f(Q_{UV}) + g(Q_{XY}),$$

$$E_1(R, Q_{UV}) = \min_{Q_{U'V}} [[f(Q_{UV}) - f(Q_{U'V})]_+ + R - H(U'|V)]_+,$$

$$E_2(R) = \min_{Q_{UV}} \{D(Q_{UV} \| P_{UV}) + E_1(R, Q_{UV})\}$$

$$\begin{aligned} E_3(Q_{UV}, Q_{XY}, Q_{U'V}, Q_{X'Y}) \quad &= \quad [[h(Q_{UV}, Q_{XY}) - h(Q_{U'V}, Q_{X'Y})]_+ \\ & \quad\quad + I(X'; Y) - H(U'|V)]_+, \end{aligned}$$

$$E_4(Q_{UV}, Q_{XY}) = \min_{Q_{U'V}, Q_{X'Y}} E_3(Q_{UV}, Q_{XY}, Q_{U'V}, Q_{X'Y}).$$

$$E_5 = \min_{Q_{UV}, Q_{XY}} [D(Q_{UV} \| P_{UV}) + D(Q_{Y|X} \| W | Q_X) + E_4(Q_{UV}, Q_{XY})].$$

$$E(R) = \min\{E_2(R), E_5\}.$$

Expurgated Bound

Consider the expression of the conditional error probability:

$$P_{\mathsf{e}|m}(\mathcal{C}_n) = \sum_{m' \neq m} \sum_{\boldsymbol{y}} W(\boldsymbol{y}|\boldsymbol{x}_m) \cdot \frac{\exp\{ng(\hat{P}_{\boldsymbol{x}_{m'}\boldsymbol{y}})\}}{\exp\{ng(\hat{P}_{\boldsymbol{x}_m\boldsymbol{y}})\} + \underbrace{\sum_{m' \neq m} \exp\{ng(\hat{P}_{\boldsymbol{x}_{m'}\boldsymbol{y}})\}}_{Z_m(\boldsymbol{y})}}.$$

We show that for the vast majority of codes

$$Z_m(\boldsymbol{y}) \geq \exp\{n\alpha(R-\epsilon, \hat{P}_{\boldsymbol{y}})\} \qquad \forall m, \boldsymbol{y}$$

where

$$\alpha(R, Q_Y) = \sup_{\{Q_{X|Y}:\ I(Q_{XY}) \leq R\}} [g(Q_{XY}) - I(Q_{XY})] + R.$$

Expurgated Bound (Cont'd)

Defining

$$\Gamma(Q_{XX'}, R) = \inf_{Q_{Y|XX'}} \left\{ \mathbf{E}_Q \log[1/W(Y|X)] - H(Y|X, X') + [\max\{g(Q_{XY}), \alpha(R, Q_Y)\} - g(Q_{X'Y})]_+ \right\}$$

we have

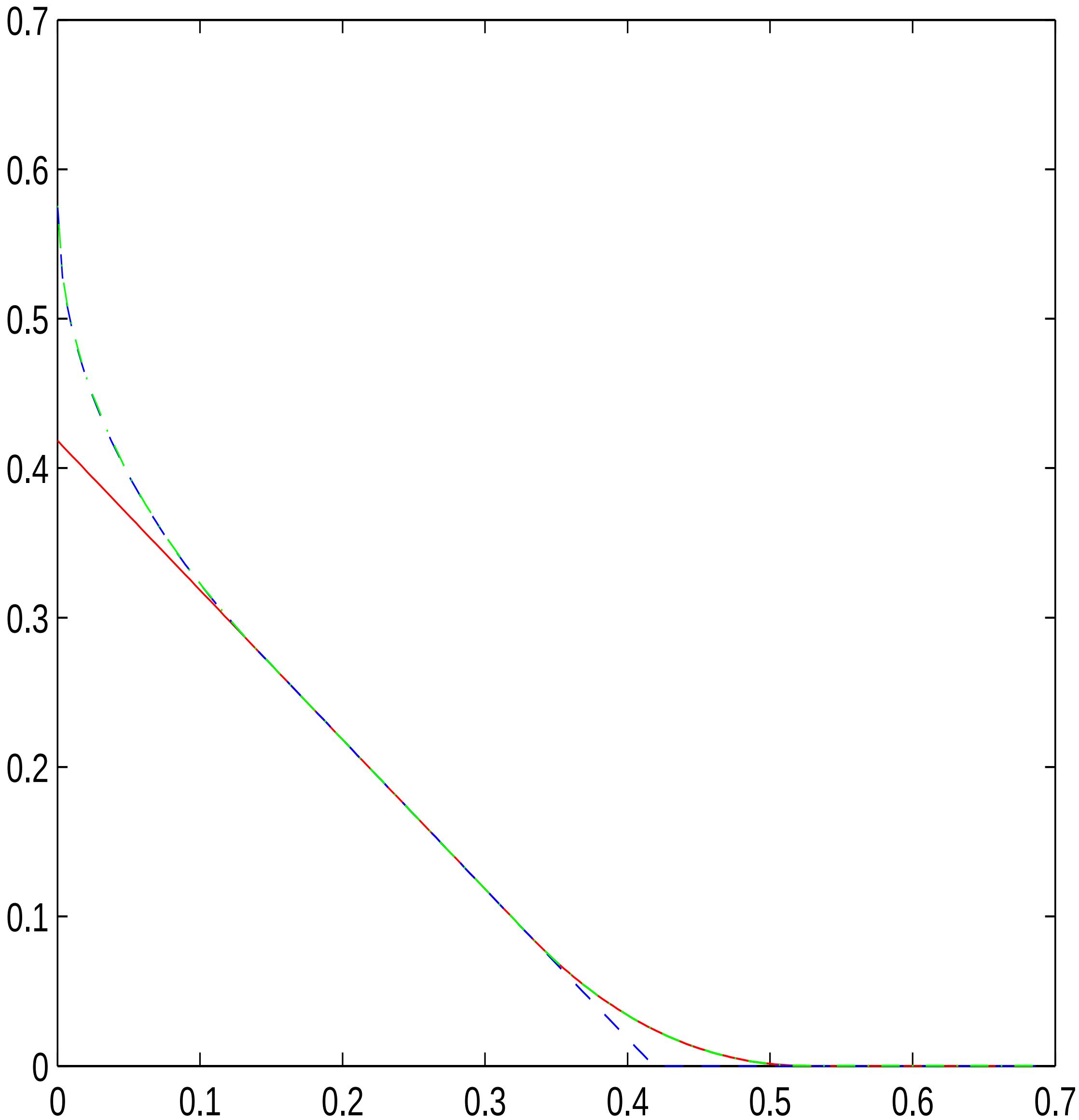
$$E_{\text{ex}}^{\text{gld}}(R, Q_X) = \inf_{\{Q_{XX'}:\ I_Q(X;X') \le R,\ Q_{X'} = Q_X\}} [\Gamma(Q_{XX'}, R) + I_Q(X;X')] - R.$$

We prove that for the ordinary LD, this is never worse than the classical expurgated bound (Csiszár–Körner–Marton, 1977).

Example – Z–Channel

Let $Q_X(0) = Q_X(1) = 1/2$ and consider the Z–channel

$$W(y|x) = \begin{cases} 0.9 & x = y = 0 \\ 0.1 & x = 0,\ y = 1 \\ 0 & x = 1,\ y = 0 \\ 1 & x = y = 1 \end{cases}$$



random coding, classical expurgated bound, and new expurgated bound.