The Generalized Stochastic Likelihood Decoder: Random Coding and Expurgated Bounds

Neri Merhav

Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 3200004, Israel

The Likelihood Decoder

The likelihood decoder is a stochastic decoder that randomly selects an estimated message $x_m$ by sampling the underlying posterior:

$$P(m|y) = \frac{P(y|x_m)}{\sum_{m'=1}^{M} P(y|x_{m'})}.$$  

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

Earlier work:

- Yassaee, Aref and Gohari (2013): network information theory.
Contributions

- A more general stochastic decoder: \( P(m|y) \propto \exp\{ng(\hat{P}_{x_m}y)\} \).
- 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 \(\{W(y|x)\}: W(y|x) = \prod_{i=1}^{n} W(y_i|x_i)\).
- Random CCC \(C = \{x_0, x_1, \ldots, x_{M-1}\}, M = e^{nR}. X_m \sim \{T(Q_X)\}\).
- Generalized likelihood decoder (GLD):

\[
P(m|y) = \frac{\exp\{ng(\hat{P}_{x_m}y)\}}{\sum_{m'=0}^{M-1} \exp\{ng(\hat{P}_{x_{m'}}y)\}}.
\]

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

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

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

$$E(R) = \min_Q [D(Q\|Q_X \times W) + E_2(Q, R)].$$
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 
\[ g(Q) = \beta I(Q), \beta \geq 1. \]

\[ E(R) > 0 \text{ for all } R \leq \min\{I(Q) : g(Q) \geq g(Q_X \times W)\}. \]

Single analysis, as opposed to separate upper and lower bounds.
Main Steps in the Derivation

\[ \bar{P}_e(x_0, y) \]

\[ = \mathbb{E} \left\{ \frac{\sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\}}{\exp\{ng(\hat{P}_{x_0} y)\} + \sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\}} \right\} \]

\[ = \int_0^1 \Pr \left\{ \frac{\sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\}}{\exp\{ng(\hat{P}_{x_0} y)\} + \sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\}} \geq t \right\} dt \]

\[ = n \cdot \int_0^\infty e^{-n\theta} \Pr \left\{ \frac{\sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\}}{\exp\{ng(\hat{P}_{x_0} y)\} + \sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\}} \geq e^{-n\theta} \right\} d\theta \]

\[ = \int_0^\infty e^{-n\theta} \Pr \left\{ \sum_{m=1}^{M-1} \exp\{ng(\hat{P}_{X_m} y)\} \geq \exp\{n[g(\hat{P}_{x_0} y) - \theta]\} \right\} d\theta \]

\[ = \int_0^\infty e^{-n\theta} \Pr \left\{ \sum_{Q'} N_{y(Q')} e^{ng(Q')} \geq \exp\{n[g(\hat{P}_{x_0} y) - \theta]\} \right\} d\theta \]
Joint/Separate Source–Channel Coding with SI

- $u$: Source encoder (random binning)
- $j$: Channel encoder
- $x$: Channel
- $y$: SI channel
- $v$: Decoder
- $\hat{u}$: Decoder output
Many customary models are covered 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 $\hat{u}$ according to

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

For a generalized version, consider

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

for some given functions $f$ and $g.$
Random Coding Exponent

\[ h(Q_{UV}, Q_{XY}) = f(Q_{UV}) + g(Q_{XY}), \]

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

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

\[ E_3(Q_{UV}, Q_{XY}, Q_{U'V}, Q_{X'Y}) = \left[ [h(Q_{UV}, Q_{XY}) - h(Q_{U'V}, Q_{X'Y})]_+ + I(X'|Y) - H(U'|V) \right]_+, \]

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

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

\[ E(R) = \min \{ E_2(R), E_5 \}. \]
Consider the expression of the conditional error probability:

\[
P_{e|m}(C_n) = \sum_{m' \neq m} \sum_y W(y|x_m) \cdot \frac{\exp\{ng(\hat{P}_{x_m}y)\}}{\exp\{ng(\hat{P}_{x_m}y)\} + \sum_{m' \neq m} \exp\{ng(\hat{P}_{x_m}y)\}}.
\]

We show that for the vast majority of codes

\[
Z_m(y) \geq \exp\{n\alpha(R - \epsilon, \hat{P}_y)\} \quad \forall m, y
\]

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

$$\Gamma(Q_{XX'}, R) = \inf_{Q_{YY|XX'}} \left\{ \mathbf{E}_Q \log [1/W(Y|X)] - H(Y|X, X') + \right.$$ 

$$\left[ \max \{g(Q_{XY}), \alpha(R, Q_Y)\} - g(Q_{X'Y}) \right]_{+} \right\}$$

we have

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

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.