# Erasure/List Exponents for Slepian–Wolf Decoding

Neri Merhav

Department of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 32000, Israel

ITW 2013, Sevilla, Spain, September 2013.

# Related Work on Error Exponents for S–W Decoding

- Gallager (1976, unpublished): same technique as in channel decoding.
- Csiszár, Körner & Marton (1977,80): universal achievability.
- Csiszár and Körner (1981): linear codes; expurgated exponents.
- Csiszár (1982) + Oohama & Han (1994): coded side information.
- Kelly & Wagner (2011): improvements at high rates.

# In This Work

- Generalized decoding for the S-W problem: erasure and list decoding.
- Draper & Martinian (2007): List decoding – fixed & deterministic list.
- We analyze tradeoffs between exponents similarly as in Forney (1968).
- Erasure option: no decoding when the confidence level is low.
- List option: tentative candidates – final decision after further processing.

# In This Work (Cont'd)

- We analyze error exponents using two methods:
  - The Gallager/Forney method.
  - Distance enumeration – inspired by the random energy model (REM).
- Second method:
  - Always at least as tight.
  - May be better by an arbitrarily large factor.
  - Sometimes simpler and easier to calculate.
- Variable–rate coding: improves the exponents.

# Some Definitions

Let $(\boldsymbol{X}, \boldsymbol{Y}) \sim \prod_{i=1}^{n} P(x_i, y_i)$.

$\boldsymbol{x}$ – source to be encoded.

$\boldsymbol{y}$ – side info @ decoder.

Encoder: $f : \mathcal{X}^n \to \{0, 1, \ldots, M-1\}$, $M = e^{nR}$.

$$z = f(\boldsymbol{x}).$$

Random binning: For every $\boldsymbol{x} \in \mathcal{X}^n$, $z$ is selected independently at random from $\{0, 1, \ldots, M-1\}$.

# Some Definitions (Cont'd)

Erasure/list decoder: Given $\boldsymbol{y} \in \mathcal{Y}^n$ and $z$, calculate for all $\hat{\boldsymbol{x}} \in f^{-1}(z)$:

$$\frac{P(\hat{\boldsymbol{x}}, \boldsymbol{y})}{\sum_{\boldsymbol{x}' \in f^{-1}(z) \setminus \{\hat{\boldsymbol{x}}\}} P(\boldsymbol{x}', \boldsymbol{y})}.$$

If $\geq e^{nT}$, $\hat{\boldsymbol{x}}$ is a candidate.

- If there are no candidates – an erasure is declared.
- If there is exactly one candidate – ordinary decoding: $\hat{\boldsymbol{x}} =$candidate.
- If there is more than one candidate – a list is of all candidates is created.

Define $\mathcal{E}_1$ as the event where the real $\boldsymbol{x}$ is not a candidate.
Let $E_1(R,T) =$ exponent of Pr$\{\mathcal{E}_1\}$. The other exponent

$$E_2(R,T) = \begin{cases} \text{decoding error exp} & \text{erasure mode} \\ \text{expected list size exp} & \text{list mode} \end{cases} = E_1(R,T) + T.$$

# Common Starting Point

$$
\begin{aligned}
\mathrm{Pr}\{\mathcal{E}_1\} &= \sum_{\boldsymbol{x},\boldsymbol{y}} P(\boldsymbol{x},\boldsymbol{y})\mathcal{I}\left\{\frac{e^{nT}\sum_{\boldsymbol{x}'\neq\boldsymbol{x}} P(\boldsymbol{x}',\boldsymbol{y})\mathcal{I}[f(\boldsymbol{x}')=f(\boldsymbol{x})]}{P(\boldsymbol{x},\boldsymbol{y})} > 1\right\} \\
&\leq \sum_{\boldsymbol{x},\boldsymbol{y}} P(\boldsymbol{x},\boldsymbol{y})\left[\frac{e^{nT}\sum_{\boldsymbol{x}'\neq\boldsymbol{x}} P(\boldsymbol{x}',\boldsymbol{y})\mathcal{I}[f(\boldsymbol{x}')=f(\boldsymbol{x})]}{P(\boldsymbol{x},\boldsymbol{y})}\right]^s \\
&= e^{nsT}\sum_{\boldsymbol{x},\boldsymbol{y}} P^{1-s}(\boldsymbol{x},\boldsymbol{y})\left[\sum_{\boldsymbol{x}'\neq\boldsymbol{x}} P(\boldsymbol{x}',\boldsymbol{y})\mathcal{I}[f(\boldsymbol{x}')=f(\boldsymbol{x})]\right]^s.
\end{aligned}
$$

# The Gallager/Forney Approach

Use

$$\Pr\{\mathcal{E}_1\} \leq e^{nsT} \sum_{\boldsymbol{x},\boldsymbol{y}} P^{1-s}(\boldsymbol{x},\boldsymbol{y}) \left( \left[ \sum_{\boldsymbol{x}' \neq \boldsymbol{x}} P(\boldsymbol{x}',\boldsymbol{y}) \mathcal{I}[f(\boldsymbol{x}') = f(\boldsymbol{x})] \right]^{s/\rho} \right)^{\rho} \qquad \rho \geq s$$

$$\leq e^{nsT} \sum_{\boldsymbol{x},\boldsymbol{y}} P^{1-s}(\boldsymbol{x},\boldsymbol{y}) \left( \sum_{\boldsymbol{x}' \neq \boldsymbol{x}} P^{s/\rho}(\boldsymbol{x}',\boldsymbol{y}) \mathcal{I}[f(\boldsymbol{x}') = f(\boldsymbol{x})] \right)^{\rho}.$$

and then, for the ensemble average, use Jensen's inequality with the limitation $\rho \leq 1$.

Two potential points of losing exponential tightness:

- The inequality $(\sum_i a_i)^r \leq \sum_i a_i^r$, $0 \leq r \leq 1$.
- Jensen's inequality.

# The Resulting Error Exponent

$$\overline{\mathsf{Pr}\{\mathcal{E}_1\}} \leq e^{-nE_1(R,T)},$$

where

$$E_1(R,T) = \sup_{0 \leq s \leq \rho \leq 1} [E_0(\rho, s) + \rho R - sT],$$

with

$$E_0(\rho, s) = -\ln \left[ \sum_{y \in \mathcal{Y}} P(y) \sum_{x \in \mathcal{X}} P^{1-s}(x|y) \left( \sum_{x' \in \mathcal{X}} P^{s/\rho}(x'|y) \right)^{\rho} \right].$$

# Extension to Variable–Rate Coding

Instead of fixed rate $R$, let the rate be $R(\boldsymbol{x})$, a function that depends on $\boldsymbol{x}$ only via the type class (header + random binning in each type), e.g., $R(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} r(x_i)$. Then, the above extends to:

$$\tilde{E}_1(R,T) = \sup_{0\le s\le\rho\le 1}\ \sup_{\{\boldsymbol{r}:\ \mathbf{E}\{r(X)\}\le R,\ r(x)>0\ \forall\ x\in\mathcal{X}\}} [\tilde{E}_0(\rho,s) - sT],$$

where

$$\tilde{E}_0(\rho,s) = -\ln\left[\sum_{y\in\mathcal{Y}} P(y)\sum_{x\in\mathcal{X}} P^{1-s}(x|y)\left(\sum_{x'\in\mathcal{X}} P^{s/\rho}(x'|y)e^{-r(x')}\right)^{\rho}\right].$$

Closed–form optimization of $\{r(x)\}$ s.t. $\mathbf{E}\{r(X)\} \le R$ is easy at least for $\rho = 1$ and the improvement in the exponent can be assessed.

# Type Class Enumeration Method

Back to fixed–rate, instead of the above, use:

$$\mathbf{E}\left\{\left[\sum_{\boldsymbol{x}'\neq\boldsymbol{x}} P(\boldsymbol{x}'|\boldsymbol{y})\mathcal{I}[f(\boldsymbol{x}')=f(\boldsymbol{x})]\right]^{s}\right\} = \mathbf{E}\left[\sum_{\mathcal{T}(\boldsymbol{x}'|\boldsymbol{y})} P(\boldsymbol{x}'|\boldsymbol{y})N(\boldsymbol{x}'|\boldsymbol{x},\boldsymbol{y})\right]^{s}$$

$$\doteq \sum_{\mathcal{T}(\boldsymbol{x}'|\boldsymbol{y})} P^{s}(\boldsymbol{x}'|\boldsymbol{y})\mathbf{E}\{N^{s}(\boldsymbol{x}'|\boldsymbol{x},\boldsymbol{y})\}$$

where $N(\boldsymbol{x}'|\boldsymbol{x},\boldsymbol{y})$ is the type class enumerator:

$$N(\boldsymbol{x}'|\boldsymbol{x},\boldsymbol{y}) = \left|T(\boldsymbol{x}'|\boldsymbol{y})\bigcap f^{-1}[f(\boldsymbol{x})]\right|.$$

$\mathbf{E}\{N^{s}(\boldsymbol{x}'|\boldsymbol{x},\boldsymbol{y})\}$ can be assessed using simple large–deviations considerations.

# The Binary Case

Let $X$ and $Y$ be BSS's connected via a BSC with crossover probability $p$. Then,

$$\overline{\text{Pr}\{\mathcal{E}_1\}} \leq e^{-nE_1'(R,T)},$$

where

$$E_1'(R,T) = \sup_{s \geq 0} E_1'(R,T,s)$$

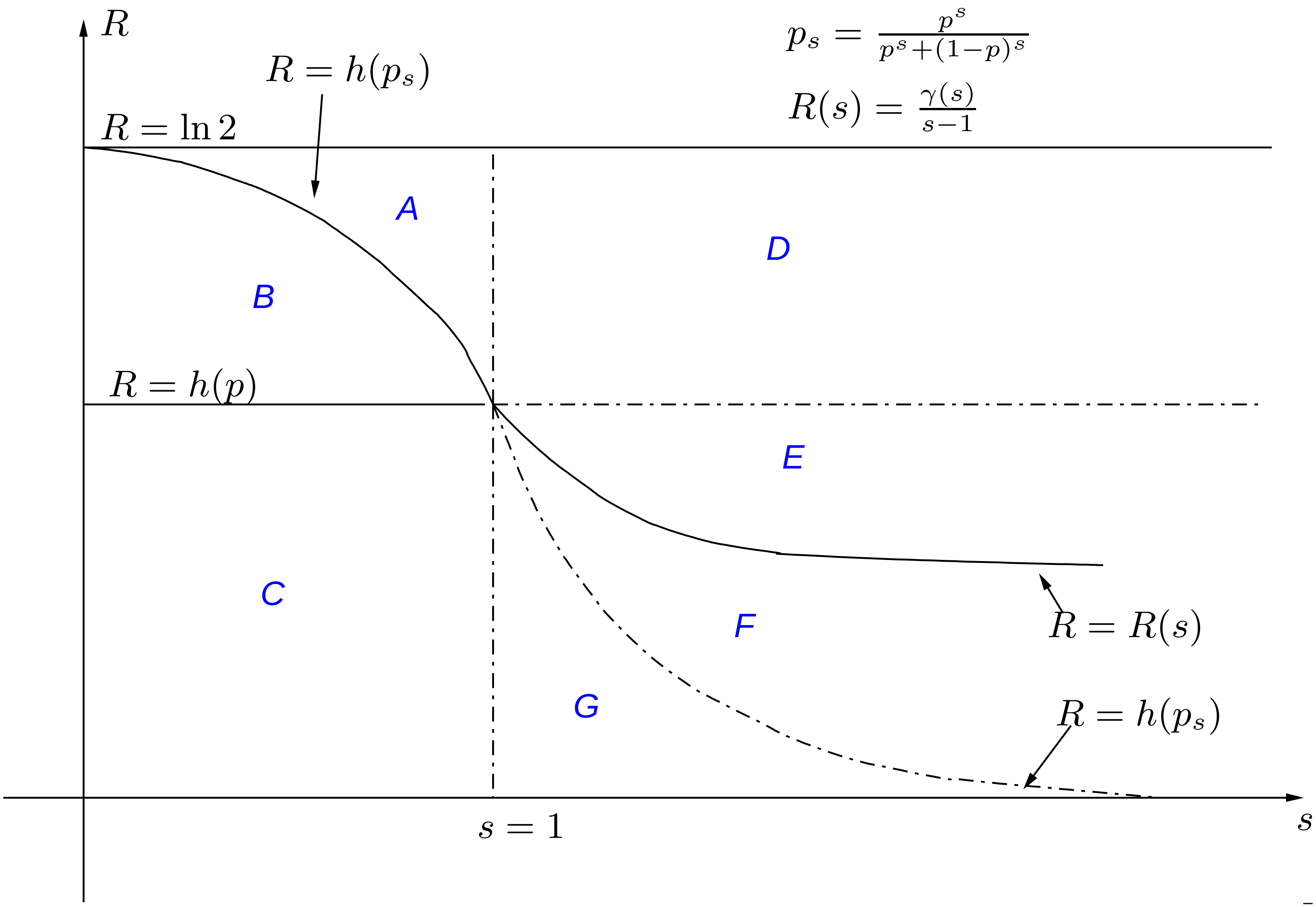
$$E_1'(R,T,s) = \begin{cases} s(R-T) + \gamma(s) & (s,R) \in C \cup F \cup G \\ s[R - T + D(h^{-1}(R) \| p)] + \gamma(s) & (s,R) \in B \\ R - sT + \gamma(s) + \gamma(1-s) & (s,R) \in A \cup D \cup E \end{cases}$$

$$\gamma(s) = -\ln[p^{1-s} + (1-p)^{1-s}]$$

and where the sets A–G are defined in the following figure.

The analysis can be extended to general DMS's.

# Phase Diagram for $E_1'(R, T, s)$



# Comparison Between $E_1(R,T)$ and $E_1'(R,T)$

$E_1'(R,T) \geq E_1(R,T)$ always.

For some regions in the plane $R$—$T$, $E_1'(R,T)$ may be larger than $E_1(R,T)$ by an arbitrarily large factor!

1. For $R > h(p)$ and $T < \ln \frac{p}{1-p}$:

$$E_1(R,T) \leq R + |T| < \infty; \qquad E_1'(R,T) = \infty.$$

2. Consider the case of very weakly correlated sources, i.e., $p = \frac{1}{2} - \epsilon$, $|\epsilon| \ll 1$. For $R \in [h(p), \ln 2]$ and $T = -\tau\epsilon^2$ with $\tau > 4$:

$$E_1(R,T) \leq (\tau + 2)\epsilon^2, \quad E_1'(R,T) \geq \left[\frac{\tau(\tau+8)}{16} - 1\right]\epsilon^2.$$

Both examples work thanks to the fact that $s$ take arbitrarily large values, not just in $[0,1]$.

# Summary

- Trade–offs between random coding exponents for erasure/list decoding.
- The type–class enum. method is never worse and sometimes a lot better.
- Optimization range of $s$ is unlimited.
- Only one parameter to optimize, rather than two.
- Variable–rate encoding can be handled also.
- Extendable to the case where both $X$ and $Y$ are encoded (separately).