Statistical Physics of Random Binning

Neri Merhav

Department of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 3200004, Israel

ISIT 2015, Hong Kong, June 2015.

Some Background

- Slepian & Wolf (‘73) – (almost) lossless compression with SI @ decoder.
- Gallager (‘76) – random coding error exponents.
- Csiszár, Körner & Marton (‘77,‘80) – universal achievability.
- Csiszár, Körner (‘81) – same with linear codes + expurgated bounds.
- Csiszár (’82) + Oohama & Han (‘94) – coded SI.
- Kelly & Wagner (‘11) – improvements for high rates.

This Work

- Exponential error bounds for random binning.
- A statistical–mechanical perspective – finite temperature decoding.
- Phase diagram in the rate vs. temperature plane.
- Similarities and differences relative to channel coding.
- Exact random coding error exponent – phase transitions.
- Extensions: mismatch, universality, variable rates, joint coding.

Problem Setup and Preliminaries

- $\{(X_i, Y_i)\}_{i=1}^N$ – N independent copies of $(X, Y) \sim P(x, y)$.
- $\boldsymbol{X} = (X_1, \ldots, X_N)$ – source to be compressed.
- $\boldsymbol{Y} = (Y_1, \ldots, Y_N)$ – side info @ decoder.
- Random binning – e^{NR} bins.
- Encoder: $u = f(\boldsymbol{x})$, $f : \mathcal{X}^N \to \{1, \ldots, e^{NR}\}$.
- Inverse image (= bin) of $\boldsymbol{x}$: $f^{-1}(u) = \{\boldsymbol{x} : f(\boldsymbol{x}) = u\}$.

Block–level MAP decoder:

$$\hat{\boldsymbol{x}} = \text{argmax}_{\boldsymbol{x} \in f^{-1}(u)} P(\boldsymbol{x}|\boldsymbol{y}) = \text{argmax}_{\boldsymbol{x} \in f^{-1}(u)} P(\boldsymbol{x}, \boldsymbol{y}).$$

Symbol–level MAP decoder:

$$\hat{x}_i = \text{argmax}_{x \in \mathcal{X}} P(x_i = x, \boldsymbol{y}) = \text{argmax}_{x \in \mathcal{X}} \sum_{\boldsymbol{x}:\ x_i = x} P(\boldsymbol{x}, \boldsymbol{y}), \qquad i = 1, 2, \ldots, N.$$

Finite–Temperature Decoding (Rujmasán, ‘93)

$$\hat{x}_i = \text{argmax}_{x \in \mathcal{X}} \sum_{\boldsymbol{x}:\ x_i = x} P^{\beta}(\boldsymbol{x}, \boldsymbol{y}), \qquad \beta > 0.$$

Motivation:

- Common framework for both SL and BL MAP ($\beta = 1$, $\beta \to \infty$, resp.).
- Mismatch due to uncertainty (e.g., double BSS):
  - $\beta < 1$ – pessimistic decoder
  - $\beta > 1$ – optimistic decoder
- Gallager–style bounds include probabilities raised to some power.

The Finite–Temperature Posterior

Define

$$P_\beta(\boldsymbol{x}|\boldsymbol{y},u) = \begin{cases} \frac{P^\beta(\boldsymbol{x},\boldsymbol{y})}{\sum_{\boldsymbol{x}'\in f^{-1}(u)} P^\beta(\boldsymbol{x}',\boldsymbol{y})} & \boldsymbol{x}\in f^{-1}(u) \\ 0 & \text{elsewhere} \end{cases}$$

or, the Boltzmann distribution:

$$P_\beta(\boldsymbol{x}|\boldsymbol{y},u) = \begin{cases} \frac{\exp\{-\beta\mathcal{E}(\boldsymbol{x},\boldsymbol{y})\}}{\sum_{\boldsymbol{x}'\in f^{-1}(u)} \exp\{-\beta\mathcal{E}(\boldsymbol{x}',\boldsymbol{y})\}} & \boldsymbol{x}\in f^{-1}(u) \\ 0 & \text{elsewhere} \end{cases}$$

with energy function: $\mathcal{E}(\boldsymbol{x},\boldsymbol{y}) \stackrel{\triangle}{=} -\ln P(\boldsymbol{x},\boldsymbol{y})$ and partition function $Z(\beta|\boldsymbol{y},u)$.

We first study the phase diagram of the corresponding “physical system”.

- Ordinary random coding $\leftrightarrow$ Random Energy Model (REM).
- Random binning $\leftrightarrow$ Random Dilution Model (RDM).

Some Quick Background on the REM and RDM

In channel coding, the analogous posterior is

$$P_\beta(\boldsymbol{x}|\boldsymbol{y}) = \begin{cases} \frac{P^\beta(\boldsymbol{y}|\boldsymbol{x})}{\sum_{\boldsymbol{x}'\in\mathcal{C}} P^\beta(\boldsymbol{y}|\boldsymbol{x}')} & \boldsymbol{x}\in\mathcal{C} \\ 0 & \text{elsewhere} \end{cases}$$

which is the Boltzmann distribution with $\mathcal{E}(\boldsymbol{x},\boldsymbol{y}) \stackrel{\triangle}{=} -\ln P(\boldsymbol{y}|\boldsymbol{x})$.

In random coding the $\boldsymbol{x}$'s are drawn independently at random:

- $\mathcal{E}(\boldsymbol{X},\boldsymbol{y})$ are independent given $\boldsymbol{y}$.
- Analogous to the REM, which has random i.i.d. energy levels.
- The REM (and hence also $P_\beta(\cdot|\boldsymbol{y})$) undergoes a ϕ–transition:
  - Below a critical temperature ($\beta > \beta_c$) – zero–entropy –glassy phase.
  - Above critical temperature – positive entropy – paramagnetic phase.

The RDM

Consider $Z(\beta) = \sum_{\boldsymbol{x}} e^{-\beta \mathcal{E}(\boldsymbol{x})}$ $\quad \beta = \frac{1}{kT}$ inverse temperature

The randomly diluted version is

$$Z_{\mathsf{D}}(\beta) = \sum_{\boldsymbol{x}} I(\boldsymbol{x}) e^{-\beta \mathcal{E}(\boldsymbol{x})} = \sum_{\boldsymbol{x}} e^{-\beta[\mathcal{E}(\boldsymbol{x}) + \Psi(\boldsymbol{x})]}$$

where $\{I(\boldsymbol{x})\}$ are i.i.d. Bernoulli RV's with

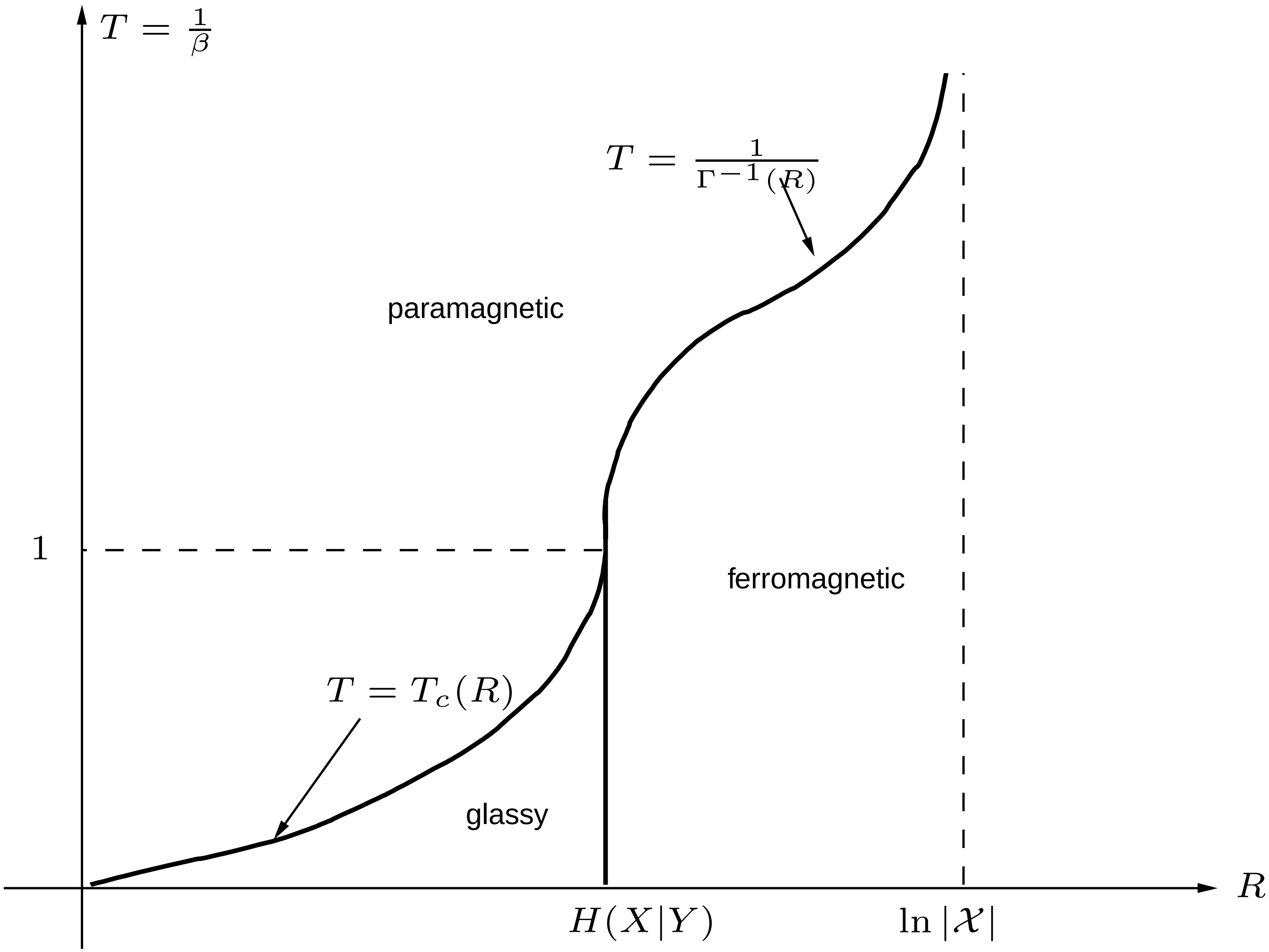
$$\Pr\{I(\boldsymbol{x}) = 1\} = \Pr\{\Psi(\boldsymbol{x}) = 0\} = 1 - \Pr\{I(\boldsymbol{x}) = 0\} = 1 - \Pr\{\Psi(\boldsymbol{x}) = \infty\} = e^{-NR}.$$

The RDM also has a glassy phase transition, similar to the REM.
Relevance to random binning:

$$Z(\beta|\boldsymbol{y}, u) = \sum_{\boldsymbol{x}} I[\boldsymbol{x} \in f^{-1}(u)] \cdot P^{\beta}(\boldsymbol{x}, \boldsymbol{y}).$$

However, for the correct $\boldsymbol{x}$, $I[\boldsymbol{x} \in f^{-1}(u)] \equiv 1$.

Phase Diagram of $Z(\beta|\boldsymbol{y}, u)$



$$\Gamma(\beta) = \beta H(X,Y) + \sum_y P(y) \ln \left[\sum_x P^\beta(x,y) \right]$$

$$\beta_c(R) = s'[s^{-1}(R)] \quad s(\epsilon) = \max\{H_Q(X|Y) : \ \mathbf{E}_Q \ln[1/P(X,Y)] \le \epsilon\}.$$

Discussion

Phase diagram $\sim$ mirror image of channel coding (Mézard & Monatanari, ‘09).

Reason: SW coding at rate $R \leftrightarrow$ channel coding at rate $(H - R)$.

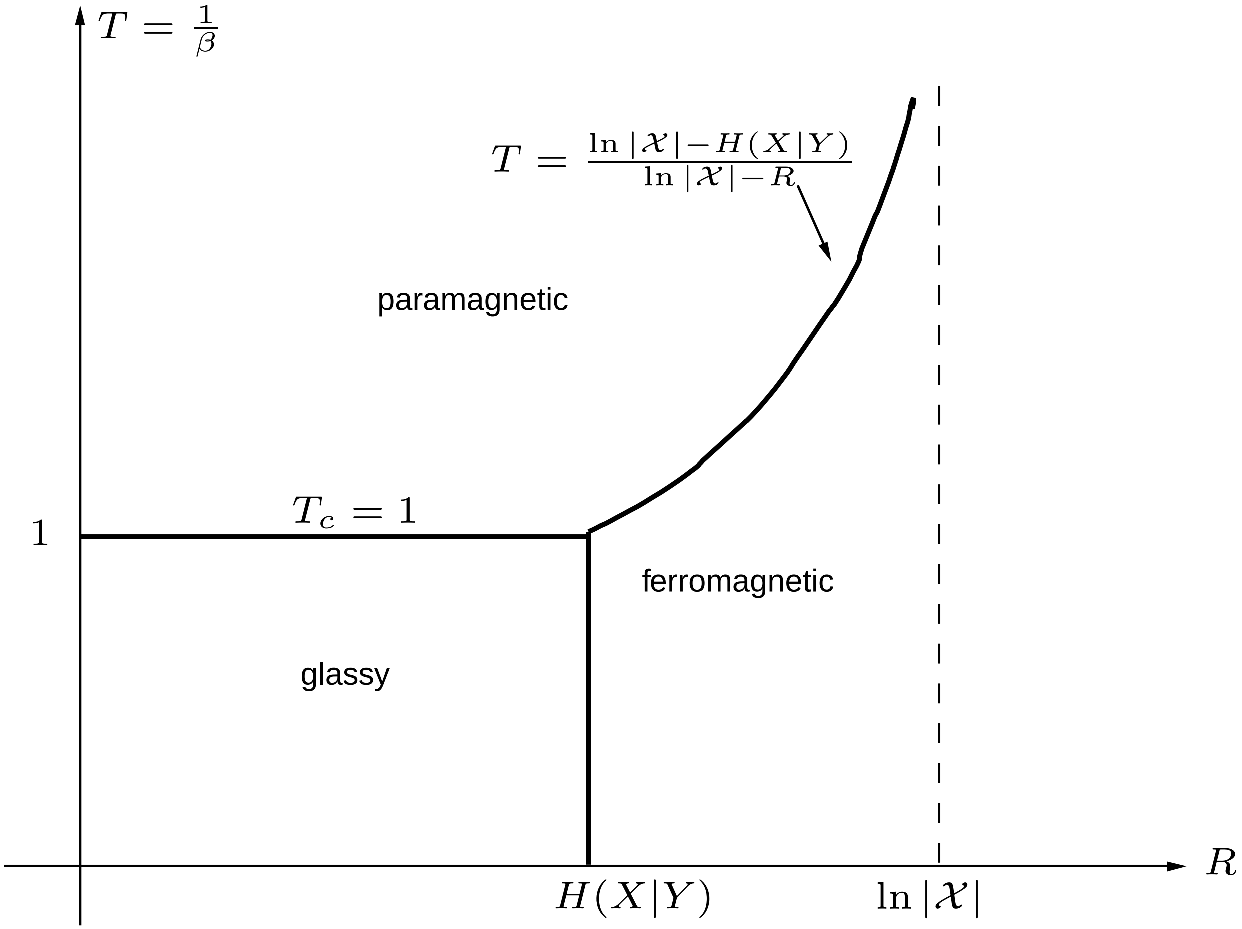
But there are a few non–trivial differences:

- Typical $|f^{-1}(u)|$ (RV) $\sim |\mathcal{X}|^N e^{-NR}$. Only $e^{N(H-R)}$ are in $\mathcal{T}(P)$.
  - Different from channel coding – fixed codebook size.
  - A–typical bin members may affect large deviations behavior.
- Prior – not necessarily uniform.
- Compositions of “codewords” are random.

Extensions and Variations

- Variable–rate SW coding (type–dependent rate).
- Mismatched decoding: decoding according to $\tilde{P}(\boldsymbol{x}|\boldsymbol{y})$.
- Universal decoding: minium empirical conditional entropy decoding.
- Full SW problem: separate codings & joint decoding of $(\boldsymbol{X}, \boldsymbol{Y})$.

Universal Decoding



Full SW Coding: Encoding and Decoding Both X and Y

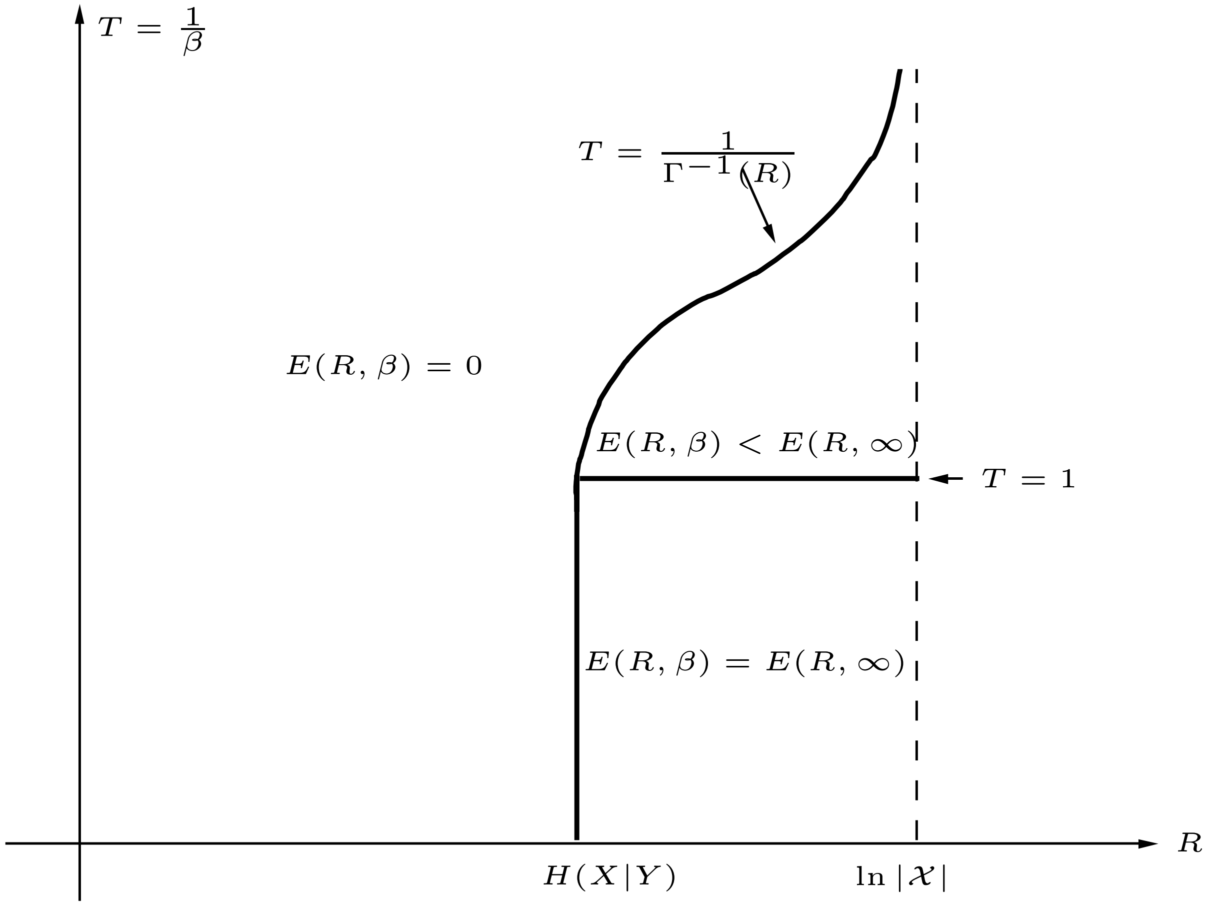
$$Z(\beta|u,v) = \sum_{\boldsymbol{x},\boldsymbol{y}} I[\boldsymbol{x} \in f^{-1}(u)] \cdot I[\boldsymbol{y} \in g^{-1}(u)] \cdot P^{\beta}(\boldsymbol{x},\boldsymbol{y}).$$

- 4 partial partition functions, corresponding to in/correct decoding of $\boldsymbol{X}$, $\boldsymbol{Y}$.
- Each partial function has 3 phases ...

For $\beta \leq 1$, reliable decoding occurs if:

$$\begin{aligned} R_X &> \beta H(X,Y) + \mathbf{E}\ln\left[\sum_x P^{\beta}(x,Y)\right] \\ R_Y &> \beta H(X,Y) + \mathbf{E}\ln\left[\sum_y P^{\beta}(X,y)\right] \\ R_X + R_Y &> \beta H(X,Y) + \ln\left[\sum_{x,y} P^{\beta}(x,y)\right] \end{aligned}$$

Exact Random Binning Error Exponent $E(R, \beta)$



The single–letter expression of $E(R, \beta)$ is available in the paper.