On More General Distributions of Random Binning for Slepian–Wolf Encoding

Neri Merhav

The Viterbi Faculty of Electrical and Computer Engineering

Technion – Israel Institute of Technology

Haifa 3200003, Israel

ISIT 2021, Melbourne, Victoria, Australia, July 2021

Background and Motivation

- Separate lossless comp. + joint decoding of corr. sources – revisited.
- Unlike in other code ensembles, S–W binning dist. is always uniform.
- Variable–rate S-W (VRSW) ensembles improve, but still – uniform.
- We address the question: why is that always the case?
- Partial answer: the ensemble is in the “compressed domain”.
- Satisfactory answer in terms of achievable rates.
- Not for trade-offs between err. probability and excess-length prob.

Model Setting

We consider a more general random binning as follows:
Given the two source vectors, $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$, randomly select resp. 'bins' $\boldsymbol{u} = (u_1, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, \ldots, v_n)$ using conditional distributions – random binning channels (RBCs):

$$A(\boldsymbol{u}|\boldsymbol{x}) \text{ and } B(\boldsymbol{v}|\boldsymbol{y})$$

and finally compress $\boldsymbol{u}$ and $\boldsymbol{v}$ (separately) to their entropies.

Decoder recovers $(\boldsymbol{u}, \boldsymbol{v})$; outputs the most likely $(\boldsymbol{x}, \boldsymbol{y})$ with bins $(\boldsymbol{u}, \boldsymbol{v})$.
We assume

$$A(\boldsymbol{u}|\boldsymbol{x}) \doteq \exp\{-nF(\hat{P}_{\boldsymbol{u}\boldsymbol{x}})\}; \quad B(\boldsymbol{v}|\boldsymbol{y}) \doteq \exp\{-nG(\hat{P}_{\boldsymbol{v}\boldsymbol{y}})\}.$$

Analyzable using the MoT and still rather general.

Discussion

We allow, not only non–uniform distributions, but also dependence on the source vectors.

Example:

Let $A(\boldsymbol{u}|\boldsymbol{x}) \propto \mathcal{I}\{d_H(\boldsymbol{x}, \boldsymbol{u}) \leq 1\}$.

Suppose $\boldsymbol{y}$ has already been decoded, and we now decode $\boldsymbol{x}$.

The decoder knows that $\boldsymbol{x}$ must satisfy $d_H(\boldsymbol{x}, \boldsymbol{u}) \leq 1$.

In other words, $\boldsymbol{u}$ serves as side info, in addition to $\boldsymbol{y}$.

Q: Can dependence between $\boldsymbol{u}$ and $\boldsymbol{x}$ can help?

Discussion

We allow, not only non–uniform distributions, but also dependence on the source vectors.

Example:

Let $A(\boldsymbol{u}|\boldsymbol{x}) \propto \mathcal{I}\{d_H(\boldsymbol{x},\boldsymbol{u}) \leq 1\}$.

Suppose $\boldsymbol{y}$ has already been decoded, and we now decode $\boldsymbol{x}$.

The decoder knows that $\boldsymbol{x}$ must satisfy $d_H(\boldsymbol{x},\boldsymbol{u}) \leq 1$.

In other words, $\boldsymbol{u}$ serves as side info, in addition to $\boldsymbol{y}$.

Q: Can dependence between $\boldsymbol{u}$ and $\boldsymbol{x}$ can help?

A: Unfortunately, no!

Discussion

We allow, not only non–uniform distributions, but also
dependence on the source vectors.

Example:

Let $A(\boldsymbol{u}|\boldsymbol{x}) \propto \mathcal{I}\{d_H(\boldsymbol{x},\boldsymbol{u}) \leq 1\}$.

Suppose $\boldsymbol{y}$ has already been decoded, and we now decode $\boldsymbol{x}$.

The decoder knows that $\boldsymbol{x}$ must satisfy $d_H(\boldsymbol{x},\boldsymbol{u}) \leq 1$.

In other words, $\boldsymbol{u}$ serves as side info, in addition to $\boldsymbol{y}$.

Q: Can dependence between $\boldsymbol{u}$ and $\boldsymbol{x}$ can help?

A: Unfortunately, no!

Additional motivation:

Robustness to unavailability of the other source to the decoder:

Trading off error prob., excess–length prob., and distortion.

Formulation

Let $A(\boldsymbol{u}|\boldsymbol{x}) \doteq \exp\{-nF(\hat{P}_{\boldsymbol{u}\boldsymbol{x}})\}$ and $B(\boldsymbol{v}|\boldsymbol{y}) \doteq \exp\{-nG(\hat{P}_{\boldsymbol{v}\boldsymbol{y}})\}$ be used for randomly drawing bins for every $\boldsymbol{x}$ and $\boldsymbol{y}$.

Denote $\boldsymbol{u} = f(\boldsymbol{x})$ and $\boldsymbol{v} = g(\boldsymbol{y})$.

Consider the ML decoder

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) = h[\boldsymbol{u}, \boldsymbol{v}] = \text{arg} \max_{\{(\boldsymbol{x},\boldsymbol{y}):\ f(\boldsymbol{x})=\boldsymbol{u},\ g(\boldsymbol{y})=\boldsymbol{v}\}} P(\boldsymbol{x}, \boldsymbol{y}).$$

Let $P_{\text{err}}(F, G) = \text{Pr}\{h[f(\boldsymbol{X}), g(\boldsymbol{Y})] \neq (\boldsymbol{X}, \boldsymbol{Y})\}$ and define the error exponent as:

$$\mathbf{E}_{\text{err}}(F, G) = \lim_{n\to\infty} \left[-\frac{\log P_{\text{err}}(F, G)}{n}\right].$$

Formulation (Cont'd)

We assume that $\boldsymbol{u}$ and $\boldsymbol{v}$ are compressed to their empirical entropies.

The excess code–length probability is

$$P_{\text{ecl}}(F,G) = \text{Pr}\{H(\hat{P}\boldsymbol{u}) \geq \tilde{R}_X,\ H(\hat{P}\boldsymbol{v}) \geq \tilde{R}_Y\},$$

and the excess code–length exponent is defined as

$$\mathbf{E}_{\text{ecl}}(F,G) = \lim_{n\to\infty}\left[-\frac{\log P_{\text{ecl}}(F,G)}{n}\right].$$

Every (F,G) incudes a point $(\mathbf{E}_{\text{err}}(F,G), \mathbf{E}_{\text{ecl}}(F,G))$ in the plane.

We are interested in the optimal tradeoff between them, e.g.,

$$E_{\text{err}}(E_0) = \max_{\{(F,G):\ \mathbf{E}_{\text{ecl}}(F,G)\geq E_0\}} \mathbf{E}_{\text{err}}(F,G).$$

Optimal RBCs

Since

$$\sum_{\boldsymbol{u}' \in \mathcal{T}(\boldsymbol{u}|\boldsymbol{x}} A(\boldsymbol{u}'|\boldsymbol{x}) = \sum_{\boldsymbol{u}' \in \mathcal{T}(\boldsymbol{u}|\boldsymbol{x})} \exp\{-nF(\hat{P}_{\boldsymbol{u}\boldsymbol{x}})\} \leq 1$$

it follows that

$$F(\hat{P}_{\boldsymbol{u}\boldsymbol{x}}) \geq \hat{H}_{\boldsymbol{u}\boldsymbol{x}}(U|X) \quad \forall\, \hat{P}_{\boldsymbol{u}\boldsymbol{x}},$$

with equality for at least one $\hat{P}_{\boldsymbol{u}|\boldsymbol{x}}$, for every $\hat{P}_{\boldsymbol{x}}$.

Optimal RBCs (Cont'd)

Since both exponents are "monotonically increasing with F",

$$F^*(\hat{P}_{\boldsymbol{ux}}) = \begin{cases} \hat{H}_{\boldsymbol{ux}}(U|X) & \text{for one } \hat{P}_{\boldsymbol{u}|\boldsymbol{x}} = Q_{U|X} \\ \infty & \text{elsewhere} \end{cases}$$

Similar statements apply to B and G.

In other words,

$$A^*(\boldsymbol{u}|\boldsymbol{x}) = \begin{cases} \frac{1}{|\mathcal{T}(Q_{U|X}|\boldsymbol{x})|} & \boldsymbol{u} \in \mathcal{T}(Q_{U|X}|\boldsymbol{x}) \\ 0 & \text{elsewhere} \end{cases}$$

$$B^*(\boldsymbol{v}|\boldsymbol{y}) = \begin{cases} \frac{1}{|\mathcal{T}(Q_{V|Y}|\boldsymbol{y})|} & \boldsymbol{u} \in \mathcal{T}(Q_{V|Y}|\boldsymbol{y}) \\ 0 & \text{elsewhere} \end{cases}$$

Error Exponents for Given $Q_{U|X}$ and $Q_{V|Y}$

$$\mathbf{E}_{\mathsf{err}}(Q_{U|X}, Q_{V|Y}) = \min\{\mathbf{E}_1(Q^*_{U|X}), \mathbf{E}_2(Q^*_{V|Y}), \mathbf{E}_3(Q^*_{U|X}, Q^*_{V|Y})\}$$

$$\begin{aligned}
\mathbf{E}_1(Q_{U|X}) &= \min_{Q_{UXY}} \{D(Q_{XY}\|P_{XY}) + H_Q(U|X) - \\
&\quad H_Q(U|X,Y) + [H_Q(U|X) - H_Q(X|Y,U)]_+\}, \\
\mathbf{E}_2(Q_{V|Y}) &= \min_{Q_{VXY}} \{D(Q_{XY}\|P_{XY}) + H_Q(V|Y) - \\
&\quad H_Q(V|X,Y) + [H_Q(V|Y) - H_Q(Y|X,V)]_+\}, \\
\mathbf{E}_3(Q_{U|X}, Q_{V|Y}) &= \min_{Q_{UVXY}} \{D(Q_{XY}\|P_{XY}) + \\
&\quad H_Q(U|X) + H_Q(V|Y) - H_Q(U,V|X,Y) + \\
&\quad [H_Q(U|X) + H_Q(V|Y) - H_Q(X,Y|U,V)]_+\}.
\end{aligned}$$

Dependencies seem to have a mixed impact on the error exponent...

Main Result

For a given Q_X (resp. Q_Y) and any associated conditional distribution, $Q_{U|X}$ (resp. $Q_{V|Y}$), let Q_U (resp. Q_V) be the induced marginal. Then,

$$
\begin{aligned}
\mathbf{E}_{\mathsf{ecl}}(Q_U, Q_V) &= \mathbf{E}_{\mathsf{ecl}}(Q_{U|X}, Q_{V|Y}), \\
\mathbf{E}_{\mathsf{err}}(Q_U, Q_V) &\geq \mathbf{E}_{\mathsf{err}}(Q_{U|X}, Q_{V|Y}), \\
\mathbf{E}_{\mathsf{err}}(Q_U, Q_V) &= \min\{E_1(Q_U), E_2(Q_V), E_3(Q_U, Q_V)\} \\
\mathbf{E}_1(Q_U) &= \min_{Q_{XY}} \{D(Q_{XY} \| P_{XY}) + \\
&\quad [H_Q(U) - H_Q(X|Y)]_+\} \\
\mathbf{E}_2(Q_V) &= \min_{Q_{XY}} \{D(Q_{XY} \| P_{XY}) + \\
&\quad [H_Q(V) - H_Q(Y|X)]_+\} \\
\mathbf{E}_3(Q_U, Q_V) &= \min_{Q_{XY}} \{D(Q_{XY} \| P_{XY}) + \\
&\quad [H_Q(U) + H_Q(V) - H_Q(X, Y)]_+\},
\end{aligned}
$$

where $H_Q(U)$ and $H_Q(V)$ denote the entropies of Q_U and Q_V.

Trading off with Distortion

It makes sense to create dependencies, $Q_{U|X}$ and $Q_{V|Y}$, if we wish to maintain distortion constraints, e.g.,

$$\max_{Q_{U|X}, Q_{V|Y}} \mathbf{E}_{\mathsf{err}}(Q_{U|X}, Q_{V|Y})$$

subject to the constraints:

$$\begin{aligned} \mathbf{E}_{\mathsf{ecl}}(Q_{U|X}, Q_{V|Y}) &\geq E_0 \\ \sum_{u,x} Q_{UX}(u,x) d_X(u,x) &\leq D_X \\ \sum_{v,y} Q_{VY}(v,y) d_Y(v,y) &\leq D_Y \end{aligned}$$

Limiting the distortion compromises the tradeoff.