Universal Decoding for Asynchronous Slepian–Wolf Encoding

Neri Merhav

The Andrew & Erna Viterbi Faculty of Electrical Engineering
Technion – Israel Institute of Technology
Haifa 3200003, Israel

ITW 2020, April 2021
Asynchronous Slepian–Wolf (ASW) Encoding

ASW encoding – unknown relative delay: $X_i$ – correlated to $Y_{i+d}$.

Reasoning:

♣ No information exchange between the encoders.
♣ Different processing delays at the two encoders.
♣ Sampling clocks at the two locations are not synchronized.
♣ Relative delays between measurements due to different distances.
Earlier Work

♠ Willems (’88): delay unknown to encoders, known to decoder.
♠ Rimoldi & Urbanke (’77); Sun et al. (’10): source splitting; decoder waits.
♠ Matsuta & Uyematsu (’20): worst–case approach:

The joint distribution, $P_{XY}$, is only known to be in $S$.

Relative delay, $d$, is only known to be between $-\Delta n$ and $\Delta n$, $\Delta \in [0, 1)$.

Their main result:

\[
R_X \geq \sup_{P_{XY} \in S} [H(X|Y) + \Delta \cdot I(X; Y)] \\
R_Y \geq \sup_{P_{XY} \in S} [H(Y|X) + \Delta \cdot I(X; Y)] \\
R_X + R_Y \geq \sup_{P_{XY} \in S} [H(X, Y) + \Delta \cdot I(X; Y)]
\]
Earlier Work (Cont’d)

The worst case approach is pessimistic:

For example, if \( \mathcal{S} \) = the entire simplex of distributions over \( \mathcal{X} \times \mathcal{Y} \),

\[
R_x \geq \log |\mathcal{X}|
\]
\[
R_y \geq \log |\mathcal{Y}|
\]
\[
R_x + R_y \geq \log |\mathcal{X}| + \log |\mathcal{Y}|
\]

which is an uninteresting triviality.

Their analysis of the probability of error is also for the worst source in \( \mathcal{S} \).

Q1: But what if \( \{P_{XY}\} \) is not the worst source in \( \mathcal{S} \)?
Q2: Can one devise a universal decoder for unknown \( P_{XY} \) and \( d \)?
Earlier Work on Universal Decoding

Universal channel decoding:
◇ Goppa ('75) - the MMI decoder achieves capacity.
◇ Csiszár & Körner ('81) – MMI decoder achieves $E_r(R)$.
◇ Csiszár ('82); Ziv ('85); Feder & Lapidoth ('98); Feder & Merhav ('02).
◇ ...

Universal S–W decoding:
◇ Csiszár & Körner ('81): minimum entropy decoder.
◇ Oohama & Han ('94); Draper ('04); Chen et al. ('08); Merhav ('16),...
Main Contributions in This Work

♣ Proposing a universal decoder for unknown source and delay.

♣ Same random coding exponent as the MAP decoder.

♣ The lower bound is based on a lower bound to $\Pr\{\bigcup_{i,j} A_i \cap B_j\}$.

♣ Deriving a Lagrange–dual formula of the error exponent.

♣ Outlining a possible extension to sources with memory.
Problem Formulation

\[ \{(X_i, Y_{i+d})\} \text{ are i.i.d. pairs of finite–alphabet RVs } \sim P_{XY}. \]

\[ P_d(x, y) = \prod_{i=1}^{d} P_Y(y_i) \cdot \prod_{i=n-d+1}^{n} P_X(x_i) \cdot \prod_{i=1}^{n-d} P_{XY}(x_i, y_{i+d}). \]

Encoder \( X \): \( f(x) \sim \text{unif}\{1, 2, \ldots, 2^{nR_x}\}. \)

Encoder \( Y \): \( g(y) \sim \text{unif}\{1, 2, \ldots, 2^{nR_y}\}. \)

Neither \( P_{XY} \) nor \( d = \delta n \) are known.

MAP decoder: \( (\hat{x}, \hat{y}) = \arg\max_{x', y'} P_d(x', y') \)

General metric decoder: \( (\hat{x}, \hat{y}) = \arg\max_{x', y'} q(x', y') \).

Error probability: \( P_e = \Pr\{(\hat{x}, \hat{y}) \neq (x, y)\}. \)

Error exponent: \( E(R_x, R_y) = \lim_{n \to \infty} -\frac{\log P_e}{n} \).
The Proposed Universal Decoding Metric

For $0 \leq k \leq n$, define

$$u_k(x, y) = k \hat{H}(y^k_1) + (n - k) \hat{H}(x^{n-k}_1, y_{k+1}^n) + k \hat{H}(x^n_{n-k+1}),$$

$$v_k(x, y) = (n - k) \hat{H}(x^{n-k}_1 | y_{k+1}^n) + k \hat{H}(x^n_{n-k+1}),$$

$$w_k(x, y) = k \hat{H}(y^k_1) + (n - k) \hat{H}(y_{k+1}^n | x^{n-k}_1),$$

$$q_k(x, y) = \max\{u_k(x, y) - n(R_x + R_y), v_k(x, y) - nR_x, w_k(x, y) - nR_y\},$$

and finally, the universal decoding metric, $q$, is defined as

$$q(x, y) = \min_{0 \leq k \leq n} q_k(x, y).$$
Main Theorem

\[ E_{\text{univdec}}(R_x, R_y) = E_{\text{MAP}}(R_x, R_y) = \min\{E_{x|y}(R_x), E_{y|x}(R_y), E_{xy}(R_x, R_y)\}, \]

where

\[ E_{x|y}(R_x) = \max_{0 \leq \rho \leq 1} \rho \cdot \left[R_x - \delta H_{1/(1+\rho)}(X) - (1 - \delta) H_{1/(1+\rho)}(X|Y)\right] \]

\[ E_{y|x}(R_y) = \max_{0 \leq \rho \leq 1} \rho \cdot \left[R_y - \delta H_{1/(1+\rho)}(Y) - (1 - \delta) H_{1/(1+\rho)}(Y|X)\right] \]

\[ E_{xy}(R_x, R_y) = \max_{0 \leq \rho \leq 1} \rho \cdot \left[R_x + R_y - \delta H_{1/(1+\rho)}(X) - \delta H_{1/(1+\rho)}(Y) - (1 - \delta) H_{1/(1+\rho)}(X, Y)\right]. \]
Three types of errors: $X$ only, $Y$ only, and both $X$ and $Y$.
The decoding metric has three components correspondingly.
Uncertainty in $d$ - treated differently than that of $P_{XY}$.
Unconditional Rényi entropies correspond to independent segments.
Achievable rate region for error exponent $E$:

$$\mathcal{R}(E) = \{(R_x, R_y) : R_x \geq R_x(E), R_y \geq R_y(E), R_x + R_y \geq R_{xy}(E)\},$$

$$R_x(E) = \inf_{s \geq 1} \left[ sE + \frac{\delta H_s(X)}{1+s} + (1 - \delta) \frac{H_s(X|Y)}{1+s}\right]$$

and similar expressions for $R_y(E)$ and $R_{xy}(E)$.
Extension to Markov: replace empirical entropies by LZ metrics.
Conclusion

♠ Replacing the worst–case approach by an “adaptive” approach.
♠ Achieves not only the rate region, but also the error exponent.
♠ Non-trivial universality in $d$.
♠ Universal metric with three components, one for every type of error.
♠ Characterization of rate region for a given error exponent, $E$.
♠ Lower bound on $P[\cup_{i,j} A_i \cap B_j]$: could be useful elsewhere too.