Can Iterative Decoding for Erasure Correlated Sources be Universal?

Arvind Yedla
(Joint work with Henry Pfister & Krishna Narayanan)

Texas A&M University
Channel Model

Source 1 $\rightarrow$ Encoder 1 $\rightarrow$ Channel 1 $\rightarrow$ Decoder

Source 2 $\rightarrow$ Encoder 2 $\rightarrow$ Channel 2

$U^{(1)}$, $X^{(1)}$, $Y^{(1)}$, $U^{(2)}$, $X^{(2)}$, $Y^{(2)}$
Channel Model

- **Correlated Discrete Memoryless Sources**
  - Known correlation model
  - Source rates $H(U^{(1)})$ and $H(U^{(2)})$
Channel Model

- Correlated Discrete Memoryless Sources
  - Known correlation model
  - Source rates $H(U^{(1)})$ and $H(U^{(2)})$
- No collaboration between the encoders
  - Independent Encoders
  - Rate $R_i = (\text{no. of input bits})/(\text{no. of output bits})$
Channel Model

- Correlated Discrete Memoryless Sources
  - Known correlation model
  - Source rates $H(U^{(1)})$ and $H(U^{(2)})$
- No collaboration between the encoders
  - Independent Encoders
  - Rate $R_i = (\text{no. of input bits})/(\text{no. of output bits})$
- Two Independent Discrete Memoryless Channels
  - Parametrized by a single parameter $\epsilon$
  - Capacities $(C_1(\epsilon_1), C_2(\epsilon_2))$
Slepian-Wolf Conditions

Slepian-Wolf region for a given rate pair \((R_1, R_2)\)
Slepian-Wolf Conditions

\[
\frac{C_1(\epsilon_1)}{R_1} \geq H(U^{(1)}|U^{(2)})
\]

Slepian-Wolf region for a given rate pair \((R_1, R_2)\)
Slepian-Wolf Conditions

\[
\frac{C_2(\epsilon_2)}{R_2} \geq H(U^{(2)}|U^{(1)})
\]

\[
\frac{C_1(\epsilon_1)}{R_1} \geq H(U^{(1)}|U^{(2)})
\]

Slepian-Wolf region for a given rate pair \((R_1, R_2)\)
Slepian-Wolf Conditions

\[
\frac{C_2(\epsilon_2)}{R_2} \geq H(U^{(2)}|U^{(1)})
\]

\[
\frac{C_1(\epsilon_1)}{R_1} + \frac{C_2(\epsilon_2)}{R_2} \geq H(U^{(1)}, U^{(2)})
\]

\[
\frac{C_1(\epsilon_1)}{R_1} \geq H(U^{(1)}|U^{(2)})
\]

Slepian-Wolf region for a given rate pair \((R_1, R_2)\)
Achievable Region

Definition
For a given pair of encoding functions and a joint decoding algorithm, the achievable rate region (ARR) is defined as the set of all channel parameters \((\epsilon_1, \epsilon_2)\) for which the encoder/decoder combination achieves an arbitrarily low probability of error in the limit of infinite blocklengths.
Achievable Region

Definition
For a given pair of encoding functions and a joint decoding algorithm, the achievable rate region (ARR) is defined as the set of all channel parameters \((\epsilon_1, \epsilon_2)\) for which the encoder/decoder combination achieves an arbitrarily low probability of error in the limit of infinite blocklengths.

For the particular problem considered here, random coding with ML decoding achieves the entire Slepian-Wolf region and hence the capacity region is the Slepian-Wolf region.
Special Case: Erasure Correlation

- $Z_i$ is a sequence of i.i.d. Bernoulli-$p$ random variables
- Source correlation defined by
  \[ (U_i^{(1)}, U_i^{(2)}) = \begin{cases} 
  \text{i.i.d. Bernoulli } \frac{1}{2} \text{ r.v.s, if } Z_i = 0 \\
  \text{same Bernoulli } \frac{1}{2} \text{ r.v. } U_i, \text{ if } Z_i = 1
  \end{cases} \]
- Erasure channels with erasure rates $\epsilon_1$ and $\epsilon_2$
- Slepian-Wolf conditions simplify to
  \[
  (1 - \epsilon_1) \geq (1 - p)R_1 \\
  (1 - \epsilon_2) \geq (1 - p)R_2 \\
  \frac{(1 - \epsilon_1)}{R_1} + \frac{(1 - \epsilon_2)}{R_2} \geq 2 - p,
  \]
Can design graph based codes with near optimal performance under iterative decoding for a point on the sum-rate constraint.

- Requires a-priori knowledge of the channel parameters.
- Unrealistic for several practical situations.
Notions of Universality under Iterative Decoding

- Universality w.r.t. channel types
  - Different BMS channels with the same capacity
  - A single code that performs well over these channels
Notions of Universality under Iterative Decoding

• Universality w.r.t. channel types
  • Different BMS channels with the same capacity
  • A single code that performs well over these channels

• Universality w.r.t. channel parameters
  • Rateless codes perform well on one channel type with different capacities
  • LT codes are universal for the single user scenario (for the erasure channel)
  • Multi-user scenarios not studied
Notions of Universality under Iterative Decoding

- Universality w.r.t. channel types
  - Different BMS channels with the same capacity
  - A single code that performs well over these channels

- Universality w.r.t. channel parameters
  - Rateless codes perform well on one channel type with different capacities
  - LT codes are universal for the single user scenario (for the erasure channel)
  - Multi-user scenarios not studied

- Universality w.r.t. the correlation model
  - Encoder/Decoder do NOT depend on source statistics
  - Coding performance is asymptotically the same
  - Random coding is universal for the Slepian-Wolf problem
  - Can be shown to be the same as universality w.r.t. channel types for symmetric correlation
Universality

Definition
Any encoder/decoder pair which can communicate the sources over all channel parameters which satisfy the Slepian-Wolf conditions is called universal

- Random Codes under ML decoding are universal (for the erasure channel)
  - Impractical due to large complexity
- Time Sharing between corner points can achieve the entire Slepian-Wolf region
  - Can achieve corner points via peeling
  - Requires knowledge of the channel
Problem Statement

- Investigate the existence of graph based codes and iterative decoding algorithms that are *universal*
- Find encoder/decoder pairs that result in large ARRs
Erasure Correlation at the Decoder

- Decoder KNOWS the source correlation exactly
- The source correlation can be viewed as parity check constraints
  \[ Z_{\mathcal{P}^\prime}X^{(1)} = Z_{\mathcal{P}^\prime}X^{(2)} \]
  at the decoder
- Notation
  - \( Z_k \) is a vector of i.i.d. Bernoulli-\( p \) random variables
  - \( \mathcal{P} \) is index set corresponding to non-zero locations of \( Z_k \)
  - \( Z = \text{diag} ([Z_k, 0]) \)
  - \( Z_{\mathcal{P}^\prime} \) is the sub-matrix of \( Z \) restricted to the rows indexed by \( \mathcal{P} \)
Example: Correlation at the Decoder

If we choose to encode 4 source bits to 7 bits, and the source correlation is given by $Z_4 = [1 \ 0 \ 0 \ 0 \ 1]^T$, then the parity-check constraints can be written as

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
X_1^{(1)} \\
X_2^{(1)} \\
X_3^{(1)} \\
X_4^{(1)} \\
X_5^{(1)} \\
X_6^{(1)} \\
X_7^{(1)} \\
\end{bmatrix}
= 
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
X_1^{(2)} \\
X_2^{(2)} \\
X_3^{(2)} \\
X_4^{(2)} \\
X_5^{(2)} \\
X_6^{(2)} \\
X_7^{(2)} \\
\end{bmatrix}
$$
LDGM Codes

- Encoding using LDGM codes defined in terms of generator matrices $G^{(1)}$ and $G^{(2)}$
- The encoded sequences $X^{(1)}$ and $X^{(2)}$ are given by

$$X^{(i)} = \begin{bmatrix} U^{(i)} \\ G^{(i)^T} U^{(i)} \end{bmatrix}$$

- Source bits are punctured before transmission
- Governing equations at the individual decoders are given by

$$\begin{bmatrix} G^{(i)^T} & I_n \end{bmatrix} X^{(i)} = 0, \text{ for } i = 1, 2$$

- Define $H^{(i)} = \begin{bmatrix} G^{(i)^T} & I_n \end{bmatrix}$ for $i = 1, 2$
- The joint decoder solves the equation $HX = 0$, where

$$H = \begin{bmatrix} H^{(1)} & 0 \\ 0 & H^{(2)} \\ Z_{P'}, & Z_{P'} \end{bmatrix}$$
Tanner Graph (LDGM)

\[ \tau_1 \]

\[ \tau_2 \]

\[ \rho^{(2)}(x) \]

\[ \lambda^{(2)}(x) \]

\[ \rho^{(1)}(x) \]

\[ \lambda^{(1)}(x) \]

\[ \epsilon_{1} \]

\[ \epsilon_{2} \]

permutation \( \pi_1 \)

permutation \( \pi_2 \)
Tanner Graph (LDGM)

\[ \rho^{(1)}(x) \]
\[ \epsilon_1 \]

\[ \lambda^{(1)}(x) \]
\[ \rho^{(2)}(x) \]
\[ \epsilon_2 \]

permutation \( \pi_1 \)
permutation \( \pi_2 \)

Decoder 1
Decoder 2
Tanner Graph (LDGM)

Decoder 1

Decoder 2

information bits
Tanner Graph (LDGM)

Decoder 1

Decoder 2

source correlation

permutation $\pi_1$

permutation $\pi_2$

$\epsilon_1$

$\epsilon_2$

$\rho^{(2)}(x)$

$\rho^{(1)}(x)$

$\lambda^{(2)}(x)$

$\lambda^{(1)}(x)$
Code Design

• The density evolution equations in terms of the generator-node to variable-node messages can be written as follows

\[ x_{i+1} = 1 - (1 - \epsilon_1)\rho (1 - ((1 - p) + p\lambda(y_i)) \lambda(x_i)) \]
\[ y_{i+1} = 1 - (1 - \epsilon_2)\rho (1 - ((1 - p) + p\lambda(x_i)) \lambda(y_i)) , \]

• We use Linear Programming to design LT codes with the following constraints

\[ \sum_{1 \leq i \leq N} \rho_i \cdot \left(1 - ((1 - p) + pe^{\alpha(1-\epsilon)(x-1)}) e^{\alpha(1-\epsilon)(x-1)}\right)^{i-1} < x \]

to maximize \( \rho_i/i \)
Analysis of LT Codes

\[ \epsilon_1 = (1 - p)R \]

\[ 1 - (1 - p)R \]

\[ 1 - R \]

\[ 1 - R \]

\[ 1 - (1 - p)R \]

\[ \epsilon_2 \]
Analysis of LT Codes

\[ \epsilon_1 - (1-p)R \]

\[ \epsilon_2 \]

symmetric sum-rate point - DE equations can be solved analytically - unique distribution that achieves capacity
Analysis of LT Codes

corner points cannot be achieved by the optimal distribution

symmetric sum-rate point - DE equations can be solved analytically - unique distribution that achieves capacity
LDPC Codes

- Encoding using systematic LDPC codes defined in terms of parity-check matrices $H^{(1)}$ and $H^{(2)}$
- Denote the encoded sequences by $X^{(1)}$ and $X^{(2)}$
- Systematic bits are punctured before transmission
- Governing equations at the individual decoders are given by

$$H^{(i)}X^{(i)} = 0, \text{ for } i = 1, 2$$

- The joint decoder solves the equation $HX = 0$, where

$$H = \begin{bmatrix}
H^{(1)} & 0 \\
0 & H^{(2)} \\
Z_{P'} & Z_{P'}
\end{bmatrix}$$

is the stacked parity check matrix and $X = [X^{(1)}, X^{(2)}]$. 
Tanner Graph (LDPC)

\[ \rho^{(2)}(x) \quad \ldots \ldots \ldots \quad \epsilon_2 \]

\[ \lambda^{(2)}(x) \quad \ldots \quad \pi_2 \quad \epsilon_2 \]

\[ p \quad \ldots \quad \pi_1 \quad \epsilon_1 \]

\[ \lambda^{(1)}(x) \quad \ldots \quad \rho^{(1)}(x) \quad \ldots \ldots \ldots \]

\[ \rho^{(1)}(x) \quad \ldots \ldots \ldots \]
The density evolution equations in terms of the check-node to variable-node messages can be written as follows:

\[
\begin{align*}
    x_{i+1} &= \left[ \gamma((1 - p) + pL(1 - \rho(1 - y_i))) + (1 - \gamma)\epsilon_1 \right] \cdot \\
    & \quad \lambda(1 - \rho(1 - x_i)) \\
    y_{i+1} &= \left[ \gamma((1 - p) + pL(1 - \rho(1 - x_i))) + (1 - \gamma)\epsilon_2 \right] \cdot \\
    & \quad \lambda(1 - \rho(1 - y_i))
\end{align*}
\]

where \(\gamma\) is the fraction of bits that are punctured and \(L^{(i)}(x)\), for \(i = 1, 2\), are the degree distributions (from the node perspective) corresponding to the information bits.

We cannot use Linear Programming to optimize the left degree distributions for LDPC codes, as the constraints are non-linear in \(\lambda_i\).
Joint Decoding

- Maximum-likelihood Decoder
  - Equivalent to solving a system of linear equations
  - Let $\mathcal{E}_1$, $\mathcal{E}_2$ (and $\bar{\mathcal{E}}_1$, $\bar{\mathcal{E}}_2$) denote the index sets of erasures (and non-erasures) corresponding to the received vectors $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$
  - For the binary case, solve $H_{\mathcal{E}} \mathbf{Y}_{\mathcal{E}} = H_{\bar{\mathcal{E}}} \mathbf{Y}_{\bar{\mathcal{E}}}$, where $H$ is the stacked parity-check matrix and $H_{\mathcal{E}}$ is the sub-matrix corresponding to the columns index by $\mathcal{E}$ and $\mathbf{Y} = [\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}]$

- Iterative Decoder
  - Iterative decoding is performed by message passing on the Tanner graph corresponding to the stacked parity-check matrix
Simulation Results

- LT Code I was optimized for the symmetric sum-rate point using linear programming.
- LT code II was designed by adding constraints corresponding to the channel conditions at a corner point of the capacity region.
- Results shown in $(\epsilon_1, \epsilon_2)$-plane for the rate pair $(1/2, 1/2)$. 
Performance of a Random Code

(a) ARR for a Random Code under Iterative Decoding  
(b) ARR for a Random Code under ML Decoding
(c) ARR for LT Code I under Iterative Decoding  
(d) ARR for LT Code I under ML Decoding
Performance of LT Code II

(e) ARR for LT Code II under Iterative Decoding  
(f) ARR for LT Code II under ML Decoding
Performance of (4, 6) LDPC Codes

(g) ARR for a (4, 6) LDPC Code under Iterative Decoding

(h) ARR for a (4, 6) LDPC Code under ML Decoding
Main Results

- LT codes can achieve the symmetric sum-rate point using joint iterative decoding
- LT codes cannot be universal under joint iterative decoding
- Simulation results comparing the performance of different ensembles with ML decoding and iterative decoding
Conclusions and Future Work

- The (4, 6) LDPC code ensemble with a punctured systematic encoder is nearly universal with maximum likelihood decoding.
- Results indicate that the problem with universality lies more with the message passing decoder.
- Can consider enhancements to the iterative decoder to increase the achievable rate region.
- Consider the performance of optimized protograph ensembles under iterative decoding.
Thank You!