

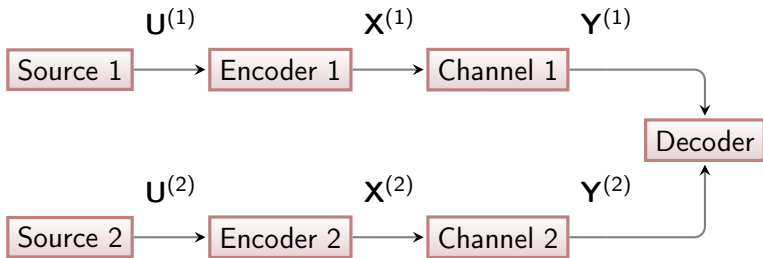
# Can Iterative Decoding for Erasure Correlated Sources be Universal?

Arvind Yedla

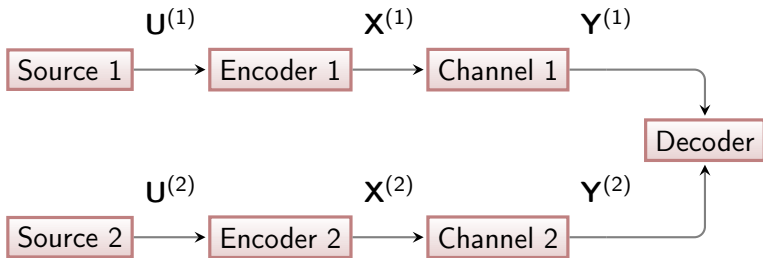
(Joint work with Henry Pfister & Krishna Narayanan)

Texas A&M University

## Channel Model

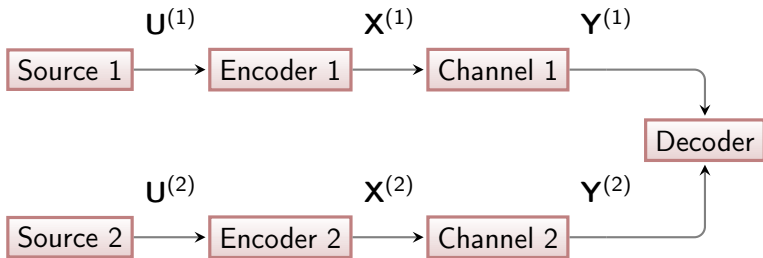


## Channel Model



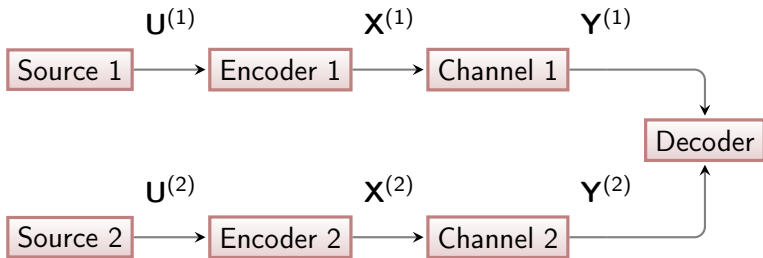
- Correlated Discrete Memoryless Sources
  - Known correlation model
  - Source rates  $H(\mathbf{U}^{(1)})$  and  $H(\mathbf{U}^{(2)})$

## Channel Model



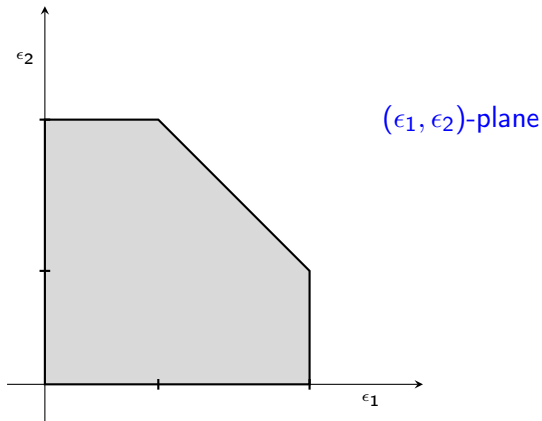
- Correlated Discrete Memoryless Sources
  - Known correlation model
  - Source rates  $H(\mathbf{U}^{(1)})$  and  $H(\mathbf{U}^{(2)})$
- No collaboration between the encoders
  - Independent Encoders
  - Rate  $R_i = (\text{no. of input bits})/(\text{no. of output bits})$

## Channel Model



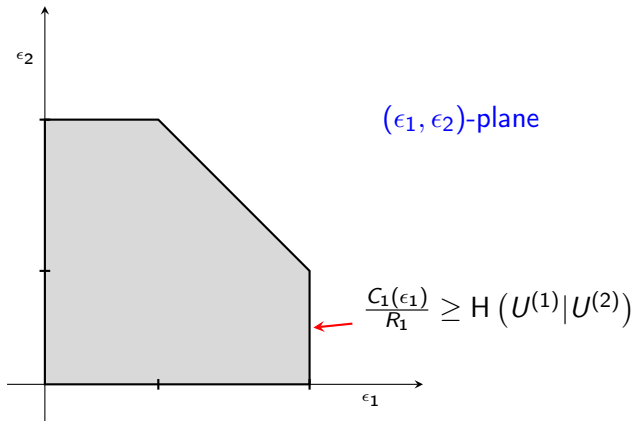
- Correlated Discrete Memoryless Sources
  - Known correlation model
  - Source rates  $H(\mathbf{U}^{(1)})$  and  $H(\mathbf{U}^{(2)})$
- No collaboration between the encoders
  - Independent Encoders
  - Rate  $R_i = (\text{no. of input bits}) / (\text{no. of output bits})$
- Two Independent Discrete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

# Slepian-Wolf Conditions



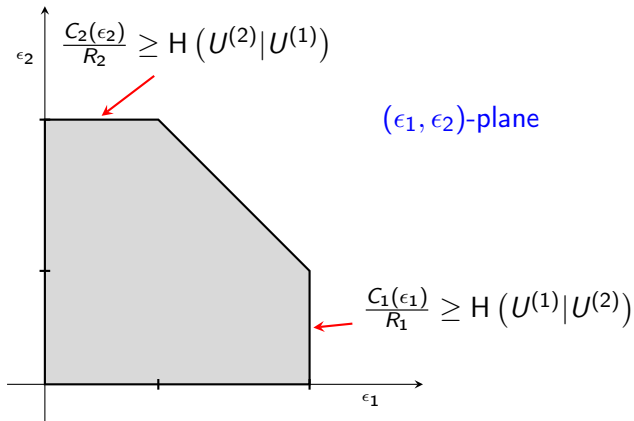
Slepian-Wolf region for a given rate pair  $(R_1, R_2)$

# Slepian-Wolf Conditions



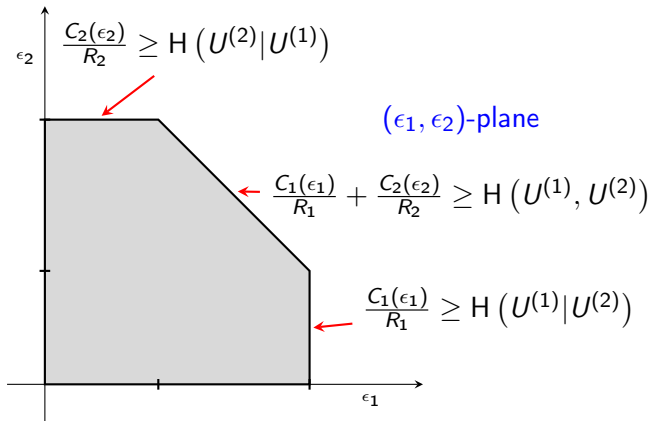
Slepian-Wolf region for a given rate pair  $(R_1, R_2)$

## Slepian-Wolf Conditions



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

## Slepian-Wolf Conditions



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

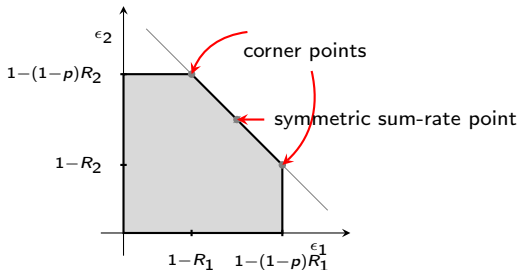
- $Z_i$  is a sequence of i.i.d. Bernoulli- $p$  random variables
- Source correlation defined by

$$\left( U_i^{(1)}, U_i^{(2)} \right) = \begin{cases} \text{i.i.d. Bernoulli } \frac{1}{2} \text{ r.v.s, if } Z_i = 0 \\ \text{same Bernoulli } \frac{1}{2} \text{ r.v. } U_i, \text{ if } Z_i = 1 \end{cases}$$

- Erasure channels with erasure rates  $\epsilon_1$  and  $\epsilon_2$
- Slepian-Wolf conditions simplify to

$$\begin{aligned} (1 - \epsilon_1) &\geq (1 - p)R_1 \\ (1 - \epsilon_2) &\geq (1 - p)R_2 \\ \frac{(1 - \epsilon_1)}{R_1} + \frac{(1 - \epsilon_2)}{R_2} &\geq 2 - p, \end{aligned}$$

# Capacity Region



- 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 ARR

## Erasures Correlation at the Decoder

- Decoder KNOWS the source correlation exactly
- The source correlation can be viewed as parity check constraints

$$Z_{\mathcal{P}'}\mathbf{X}^{(1)} = Z_{\mathcal{P}'}\mathbf{X}^{(2)}$$

at the decoder

- Notation
  - $\mathbf{Z}_k$  is a vector of i.i.d. Bernoulli-p random variables
  - $\mathcal{P}$  is index set corresponding to non-zero locations of  $\mathbf{Z}_k$
  - $Z = \text{diag}([\mathbf{Z}_k, \mathbf{0}])$
  - $Z_{\mathcal{P}'}$  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  $\mathbf{Z}_4 = [1 \ 0 \ 0 \ 1]^T$ , 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  $\mathbf{X}^{(1)}$  and  $\mathbf{X}^{(2)}$  are given by

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

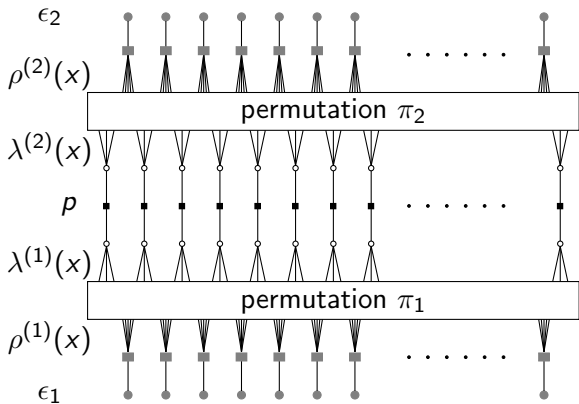
- Source bits are punctured before transmission
- Governing equations at the individual decoders are given by

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

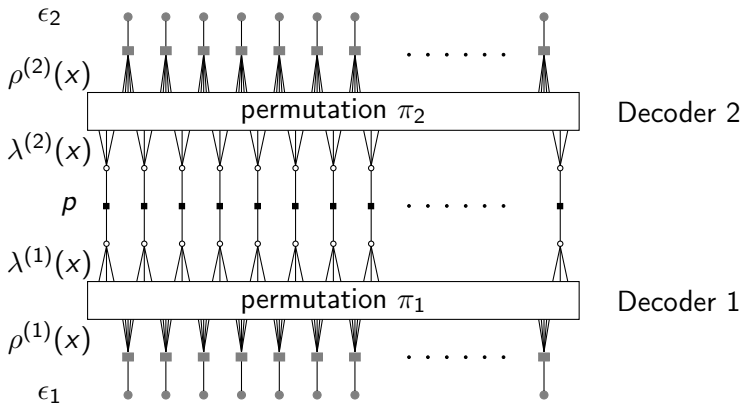
- Define  $H^{(i)} = \begin{bmatrix} G^{(i)T} & I_n \end{bmatrix}$  for  $i = 1, 2$
- The joint decoder solves the equation  $H\mathbf{X} = \mathbf{0}$ , where

$$H = \begin{bmatrix} H^{(1)} & 0 \\ 0 & H^{(2)} \\ Z_{\mathcal{P}'_1} & Z_{\mathcal{P}'_2} \end{bmatrix}$$

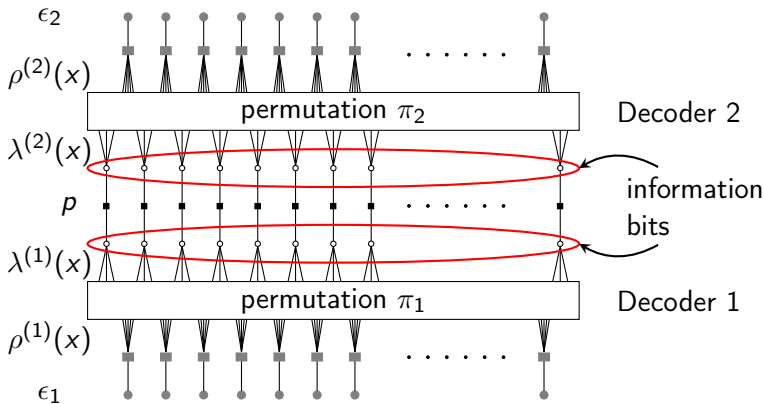
# Tanner Graph (LDGM)



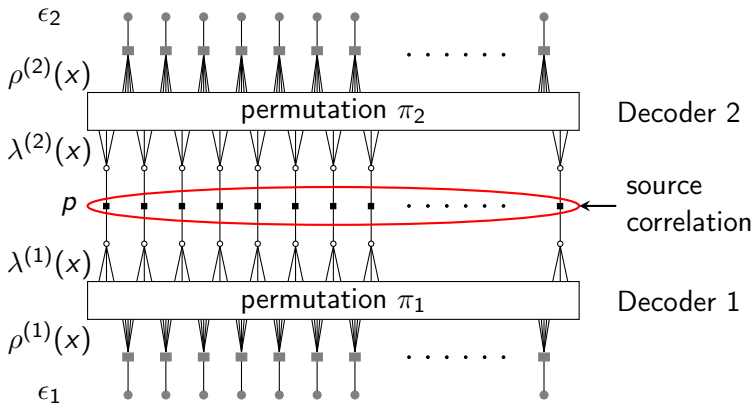
# Tanner Graph (LDGM)



# Tanner Graph (LDGM)



# Tanner Graph (LDGM)



## Code Design

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

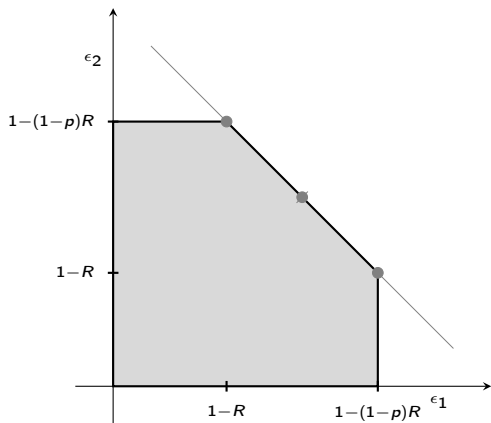
$$\begin{aligned}x_{i+1} &= 1 - (1 - \epsilon_1)\rho(1 - ((1 - p) + p\lambda(y_i))\lambda(x_i)) \\y_{i+1} &= 1 - (1 - \epsilon_2)\rho(1 - ((1 - p) + p\lambda(x_i))\lambda(y_i)),\end{aligned}$$

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

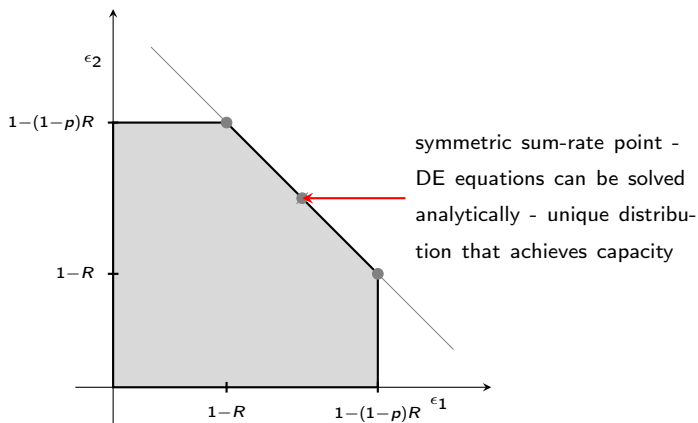
$$\sum_{1 \leq i \leq N} \rho_i \cdot \left(1 - ((1 - p) + pe^{\alpha(1-\epsilon)(x-1)})e^{\alpha(1-\epsilon)(x-1)}\right)^{i-1} < x$$

to maximize  $\rho_i/i$

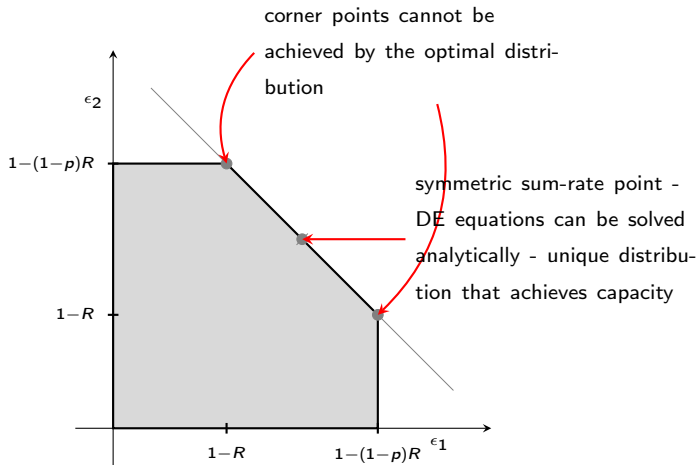
# Analysis of LT Codes



# Analysis of LT Codes



# Analysis of LT Codes



## LDPC Codes

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

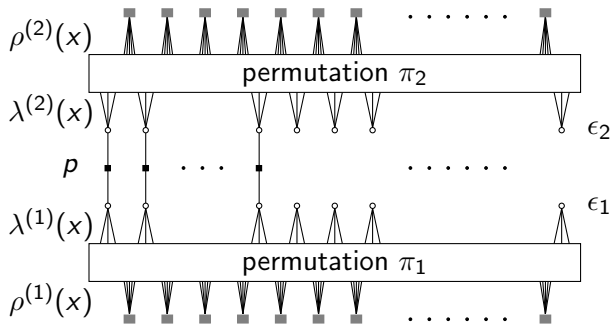
$$H^{(i)}\mathbf{X}^{(i)} = \mathbf{0}, \text{ for } i = 1, 2$$

- The joint decoder solves the equation  $H\mathbf{X} = \mathbf{0}$ , where

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

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

# Tanner Graph (LDPC)



## Code Design

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

$$x_{i+1} = \left[ \gamma((1 - \rho) + \rho L(1 - \rho(1 - y_i))) + (1 - \gamma)\epsilon_1 \right] \cdot \lambda(1 - \rho(1 - x_i))$$
$$y_{i+1} = \left[ \gamma((1 - \rho) + \rho L(1 - \rho(1 - x_i))) + (1 - \gamma)\epsilon_2 \right] \cdot \lambda(1 - \rho(1 - y_i))$$

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  $\lambda_i$

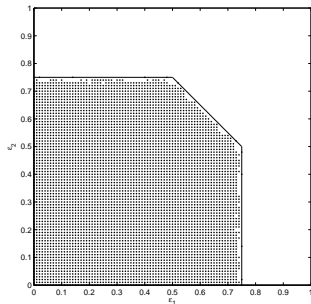
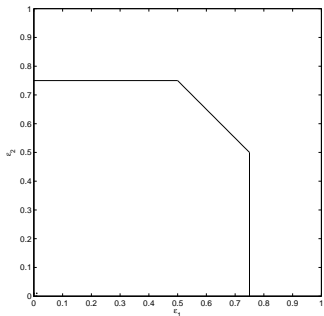
# Joint Decoding

- 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

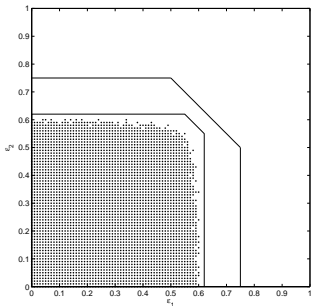
- LT Code I was optimized for the symmetric sum-rate point using linear programming
- LT code II was designed by adding constraints corresponding to the channel conditions at a corner point of the capacity region
- Results shown in  $(\epsilon_1, \epsilon_2)$ -plane for the rate pair  $(1/2, 1/2)$

# Performance of a Random Code

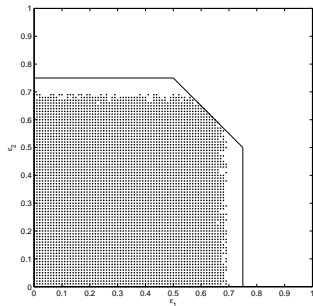


(a) ARR for a Random Code under Iterative Decoding (b) ARR for a Random Code under ML Decoding

# Performance of LT Code I

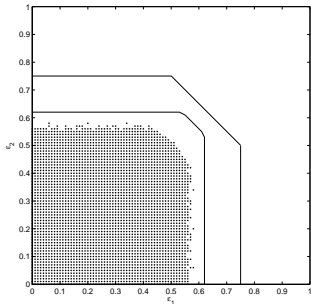


(c) ARR for LT Code I under Iterative Decoding

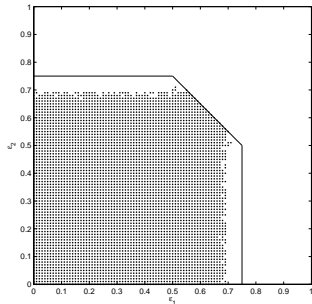


(d) ARR for LT Code I under ML Decoding

# Performance of LT Code II

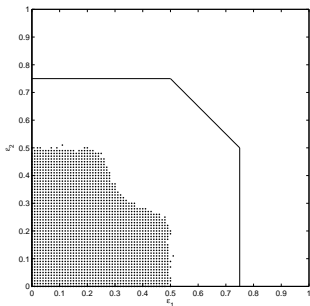


(e) ARR for LT Code II under Iterative Decoding

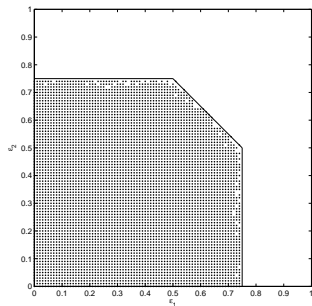


(f) ARR for LT Code II under ML Decoding

# 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!