Spatially Coupled LDPC Codes

Kenta Kasai

Tokyo Institute of Technology

30 Aug, 2013
We already have very good codes.

Efficiently-decodable asymptotically capacity-approaching codes

- Irregular LDPC Codes
  Achievability: proof is limited to BEC but likely achieve any memoryless channels.
  Drawback: non-universal and high error floors

- Polar Codes
  Achievability: proved for any memoryless channels
  Drawback: poor finite length performance
Optimal Finite-Length Error-Rate Approximation

\[
\log M^*(n, P_B) = nC(\epsilon) - \sqrt{nV(\epsilon)Q^{-1}(P_B)} + O(\log n)
\]

\[
P_B = Q\left(\frac{\sqrt{n(C(\epsilon) - R^*)}}{\sqrt{V(\epsilon)}}\right) (1 + o(1))
\]

\[
R^* := \frac{1}{n} \log M^*(n, P_B)
\]

\(M^*(n, P_B)\) is the maximum number of codewords of length-\(n\) codes that achieve the block error rate \(P_B\)

\(\epsilon\) : channel parameter

\(C(\epsilon)\) : channel capacity

\(V(\epsilon)\) : channel dispersion, e.g.,
\[
V(\epsilon) = \epsilon(1 - \epsilon) + \log^2\left(\frac{1-\epsilon}{\epsilon}\right) \text{ for } \text{BSC}(\epsilon)
\]
\[
V(\epsilon) = \epsilon(1 - \epsilon) \text{ for } \text{BEC}(\epsilon).
\]
Example: Fix BEC($\epsilon=0.5$) and BSC($\epsilon=0.11$), $C(\epsilon) = 0.5$ and $P_B = 10^{-3}$.

Who wants to fill the gap 1% away from capacity? $n = 10^5$ is sufficient.
Example: State-of-the-art finite-length codes. Fix $R^*$ and $P_B$ over BEC($\epsilon$)

Finite Polar Code
Infinite (2,4)-NBLDPCC GF($2^8$)
Finite (2,4)-NBLDPCC GF($2^8$)
Infinite (3,6)-BLDPCC
Finite (3,6)-BLDPCC
Finite Optimal Code
Uncoupled LDPC Code

Sparse parity-check matrix (column-weight 2, row-weight 4)

\[
H = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0
\end{pmatrix}
\]

, \ C = \{x \in \mathbb{F}_2^n \mid Hx^T = 0\}

Sparse graph representation (variable-degree 2, check-degree 4)
Efficient locally-optimal decoding uncoupled LDPC code

\[ x_5 + x_7 + x_{10} + x_{12} = 0 \quad (\star) \]

Suppose we observed \( x_5 =?, x_7 = 1, x_{10} = 1, x_{12} = 0 \) through the BEC(\( \epsilon \)), then from (\( \star \)) we know \( x_5 = 1 \).

- Iterate this until decoding is stuck.
- Locally-optimal decoding is efficient with computational cost \( O(n) \).
Locally-optimal decoder performance analysis of \((d_l, d_r)-codes over BEC(\epsilon)\)

Erasure probability of check node of degree \(d_r\)

\[ q(t) = 1 - (1 - p^{(t-1)})^{d_r-1} \]

Erasure probability of variable node of degree \(d_l\)

\[ p(t) = \epsilon q(t)^{d_l-1} \]

Putting together

\[ p(t) = \epsilon (1 - (1 - p^{(t-1)})^{d_r-1})^{d_l-1} \]
\[ = f(g(p^{(t-1)}); \epsilon) \]

How good the \((d_l, d_r)\)-code is measured by threshold

\[ \epsilon_{BP}(d_l, d_r) = \sup \{\epsilon \mid \lim_{t \to \infty} p(t) = 0\} \]
Globally optimal decoder performance analysis

\[ \hat{x}_i^{\text{MAP}}(y) = \arg\max_{x_i \in \mathbb{F}_2} P(x_i | y) \]

\[ \epsilon^{\text{MAP}} = \sup \left\{ \epsilon \mid \frac{1}{n} \sum_{i=1}^{n} \Pr(X_i \neq \hat{X}_i^{\text{MAP}}(Y)) = 0 \right\} \]

Globally optimal decoding is computationally complex with cost \( O(2^{nR}) \).
Graphical interpretation of $\epsilon^{BP}(d_l, d_r)$ via potential $U(x; \epsilon)$

Potential function

$$U(x; \epsilon) = xg(x) - G(x) - F(g(x); \epsilon)$$

$$\epsilon^{BP}(d_l, d_r) = \sup\{\epsilon | U'(x; \epsilon) > 0 \forall x(0, 1]\}$$

$$\epsilon^{POT}(d_l, d_r) := \sup\{\epsilon | \min_{x \in [0,1]} U(x; \epsilon) > 0\} \overset{\text{BEC}}{=} \epsilon^{MAP}(d_l, d_r)$$

Example: $(d_l = 3, d_r = 6)$ uncoupled code.

Reprinted from [Yedla et al.]
How to couple $L$ uncoupled $(d_l, d_r)$ codes to make a $(d_l, d_r, L)$ coupled code

1. Copy $L$ $(d_l, d_r)$ uncoupled codes

   \[ 1 \quad 2 \quad 3 \quad \ldots \quad L \]

2. Re-connect edges randomly with $w$ neighboring codes
Locally-optimal decoder performance analysis of \((d_l, d_r, L)\)-coupled codes over BEC(\(\epsilon\))

\(x_i\): check node erasure probability of \(i\)-th uncoupled code. DEMO.

\[ x_i^{(t+1)} = \frac{1}{w} \sum_{k=0}^{w-1} f \left( \frac{1}{w} \sum_{j=0}^{w-1} g(x_{i+j-k}^{(t)}; \epsilon_i-k) \right) \]

vector form:

\[ x^{(t+1)} = A^T f(Ag(x^{(t)}); \epsilon) \quad (A \text{ is some upper-triangle matrix}) \]
$(d_l, d_r, L, w)$ coupled codes achieve $\epsilon^{\text{POT}}(d_l, d_r)$

Definition: potential function of $(d_l, d_r, L, w)$ coupled codes

$$U(x; \epsilon) = g(x)^T x - G(x) - F(Ag(x); \epsilon)$$

Theorem [Yedla et al., 2012]
An $x \in [0, 1]^L$ is a stationary point of $U(x; \epsilon)$ if and only if $x$ is a fixed point of BP iteration $x = A^T f(Ag(x); \epsilon)$. For $\epsilon < \epsilon^{\text{POT}}(d_l, d_r)$, if $w$ is large enough, the only stationary point of $U(x; \epsilon)$ is 0.
History of Coupled LDPC Codes

- Felström and Zigangirov constructed coupled codes as LDPC convolutional codes and found they were better than uncoupled codes.

- Lentmaier found \((4, 8, L)\) coupled codes had low error rate at \(\epsilon\) above the BP threshold \((4, 8)\) uncoupled codes.

- Sridharan et al. found the BP threshold of coupled codes are very close to the MAP threshold of the uncoupled codes.
Kudekar et al. proved
\[
\lim_{w \to \infty} \epsilon^{BP}(d_l, d_r, L) = \epsilon^{MAP}(d_l, d_r).
\]

Yedla et al. proved
\[
\lim_{w \to \infty} \epsilon^{BP}(d_l, d_r, L, w) = \epsilon^{MAP}(d_l, d_r) = \epsilon^{POT}(d_l, d_r).
\]

Kudekar et al. reported empirical evidence that this phenomenon also occurs for MBIOS channels other than BEC.

Kudekar et al. reported empirical evidence that this phenomenon occurs for channels with memory and multi-access channels.
Practical Definition of Coupled Codes via Protograph

**Protograph**
A protograph is a bipartite graph defined by a set of variable nodes $V = \{v_1, \ldots, v_n\}$, a set of check nodes $C = \{c_1, \ldots, c_m\}$, and a set of edges $E$.

**Base matrix**
A protograph is defined by a base matrix $B \in \mathbb{F}_2^{m \times n}$.

$$
\begin{array}{c}
v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\
c_1 & 1 & 1 & 1 & 0 & 0 & 0 \\
c_2 & 0 & 1 & 1 & 1 & 0 & 0 \\
c_3 & 1 & 0 & 0 & 0 & 1 & 1 \\
c_4 & 0 & 0 & 0 & 1 & 1 & 1 \\
\end{array}
$$
A protograph code is defined by a parity-check matrix $H \in \mathbb{F}_2^{mM \times nM}$.

$$C := \{x \in \mathbb{F}_2^{nM} | Hx^T = 0\}$$

A parity-check matrix $H$ is derived from a base matrix $B$.

1 $\rightarrow$ $M \times M$ size random binary permutation matrix.

0 $\rightarrow$ $M \times M$ size zero matrix.

The parameter $M$ is referred to as lifting number.
\((d_l, d_r, L)\) coupled codes

- The base matrix \(B(d_l, d_r, L)\) is a \((L + d_l - 1) \times kL\) binary band matrix of band size \(d_r \times d_l\).

\[
B(d_l = 4, d_r = 12, L = 7) = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}.
\]
Efficient Termination of Spatially-Coupled Codes

K. Tazoe, K. Kasai and K. Sakaniwa, Tokyo Institute of Technology, Japan

The termination is equivalent to solve this equation.

Spatially-Coupled LDPC codes: (4,12,9) codes (left) and modified (4,12,9) codes (right) BEC(ε=0.3)

Denote a codeword by \((x_1, \ldots, x_{21})\). The parity-check equations are written by

\[
\sum_{j=1}^{d_i} P_{i,j} x_{i:k-d_i+j} = 0 \quad \text{(for } i = 0, \ldots, 10) \tag{1} \]

\(P_{i,j}'s\) are \(M \times M\) random permutation matrices. \(M \approx 1000\)

Rate of \((d_l, d_r)\) uncoupled codes \(R(d_l, d_r) = \frac{k - 1}{k}, k := \frac{d_r}{d_l}\)

Rate of \((d_l, d_r, L)\) coupled code \(R(d_l, d_r, L) = \frac{k - 1}{k} - \frac{d_l - 1}{kL}\)
Encoding of Coupled Codes

Ex. parity-check matrix for \((d_1 = 4, d_r = 12, L = 7)\)

\[
\begin{bmatrix}
  P_{1,10} & P_{1,11} & P_{1,12} \\
  P_{2,7} & P_{2,8} & P_{2,9} & P_{2,10} & P_{2,11} & P_{2,12} \\
  P_{3,4} & P_{3,5} & P_{3,6} & P_{3,7} & P_{3,8} & P_{3,9} & P_{3,10} & P_{3,11} & P_{3,12} \\
  P_{4,1} & P_{4,2} & P_{4,3} & P_{4,4} & P_{4,5} & P_{4,6} & P_{4,7} & P_{4,8} & P_{4,9} & P_{4,10} & P_{4,11} & P_{4,12} \\
  P_{5,1} & P_{5,2} & P_{5,3} & P_{5,4} & P_{5,5} & P_{5,6} & P_{5,7} & P_{5,8} & P_{5,9} & P_{5,10} & P_{5,11} & P_{5,12} \\
  P_{6,1} & P_{6,2} & P_{6,3} & P_{6,4} & P_{6,5} & P_{6,6} & P_{6,7} & P_{6,8} & P_{6,9} & P_{6,10} & P_{6,11} & P_{6,12} \\
  P_{7,1} & P_{7,2} & P_{7,3} & P_{7,4} & P_{7,5} & P_{7,6} & P_{7,7} & P_{7,8} & P_{7,9} & P_{7,10} & P_{7,11} & P_{7,12} \\
  P_{8,1} & P_{8,2} & P_{8,3} & P_{8,4} & P_{8,5} & P_{8,6} & P_{8,7} & P_{8,8} & P_{8,9} & P_{8,10} & P_{8,11} & P_{8,12} \\
  P_{9,1} & P_{9,2} & P_{9,3} & P_{9,4} & P_{9,5} & P_{9,6} & P_{9,7} & P_{9,8} & P_{9,9} & P_{9,10} & P_{9,11} & P_{9,12} \\
  P_{10,1} & P_{10,2} & P_{10,3} & P_{10,4} & P_{10,5} & P_{10,6} & P_{10,7} & P_{10,8} & P_{10,9} & P_{10,10} & P_{10,11} & P_{10,12}
\end{bmatrix}
\]

Denote a codeword by \((x_1, \ldots, x_{21})\). The parity-check equations are written by

\[
\sum_{j=1}^{d_r} P_{i,j} x_{i-k-d_r+j} = 0 \quad \text{(for } i = 0, \ldots, 10\text{)}
\]

The computational cost of the sequential encoding is \(O(LM)\). While the computational cost of the termination is \(O(M^2)\).
Termination

Termination calculates values of parity bits $x_{17}, x_{18}, x_{19}, x_{20}, x_{21}$ so that constraints $c_6, c_7, c_8, c_9, c_{10}$ are satisfied. The termination is equivalent to solve this equation.

\[
\begin{pmatrix}
P_{6,11} & P_{6,12} \\
P_{7,8} & P_{7,9} & P_{7,10} & P_{7,11} & P_{7,12} \\
P_{8,5} & P_{8,6} & P_{8,7} & P_{8,8} & P_{8,9} \\
P_{9,2} & P_{9,3} & P_{9,4} & P_{9,5} & P_{9,6} \\
P_{10,1} & P_{10,2} & P_{10,3}
\end{pmatrix}
\begin{pmatrix}
x_{17} \\
x_{18} \\
x_{19} \\
x_{20} \\
x_{21}
\end{pmatrix}
= \begin{pmatrix}
s_6 \\
s_7 \\
s_8 \\
s_9 \\
s_{10}
\end{pmatrix}
\]

To solve this equation involves multiplication of $5M \times 5M$ size not sparse matrix, so one needs $O(M^2)$ computational cost to finish the termination.

$s_i$ denote syndrome of $c_i$ which contains information of previously decided bits.
Modified \((d_1, d_r, L)\) codes

- Modified \((d_1, d_r, L)\) codes is protograph codes.
- The protograph of modified \((d_1, d_r, L)\) codes is obtained by removing \(d_1 - 2\) check nodes of the protograph of \((d_1, d_r, L)\) codes.
- The design rate is \(\frac{KL - (L + 1)}{KL} = \frac{k - 1}{k} - 1\)

Parity-check matrix of modified \((4,12,7)\) codes

\[
\begin{bmatrix}
P_{1,10} & P_{1,11} & P_{1,12} \\
P_{2,7} & P_{2,8} & P_{2,9} & P_{2,10} & P_{2,11} & P_{2,12} \\
P_{3,4} & P_{3,5} & P_{3,6} & P_{3,7} & P_{3,8} & P_{3,9} & P_{3,10} & P_{3,11} & P_{3,12} \\
P_{4,1} & P_{4,2} & P_{4,3} & P_{4,4} & P_{4,5} & P_{4,6} & P_{4,7} & P_{4,8} & P_{4,9} & P_{4,10} & P_{4,11} & P_{4,12} \\
P_{5,1} & P_{5,2} & P_{5,3} & P_{5,4} & P_{5,5} & P_{5,6} & P_{5,7} & P_{5,8} & P_{5,9} & P_{5,10} & P_{5,11} & P_{5,12} \\
P_{6,1} & P_{6,2} & P_{6,3} & P_{6,4} & P_{6,5} & P_{6,6} & P_{6,7} & P_{6,8} & P_{6,9} & P_{6,10} & P_{6,11} & P_{6,12} \\
P_{7,1} & P_{7,2} & P_{7,3} & P_{7,4} & P_{7,5} & P_{7,6} & P_{7,7} & P_{7,8} & P_{7,9} & P_{7,10} & P_{7,11} & P_{7,12} \\
P_{8,1} & P_{8,2} & P_{8,3} & P_{8,4} & P_{8,5} & P_{8,6} & P_{8,7} & P_{8,8} & P_{8,9}
\end{bmatrix}
\]

The modified coupled codes still have the same BP threshold as the unmodified coupled codes.
**Theorem** [Iyengar et al., 2012]
Consider windowed decoding of the \((d_l, d_r, w, L)\) spatially coupled ensemble over the binary erasure channel. Then for a target erasure rate \(\delta < \delta_*\), there exists a positive integer \(W_{\text{min}}(\delta)\) such that when the window size \(W \geq W_{\text{min}}(\delta)\) the WD threshold satisfies

\[
\epsilon^{\text{WD}}(d_l, d_r, w, W, \delta) \geq \left(1 - \frac{d_l d_r}{2} \delta^{\frac{d_l-2}{d_l-1}}\right) \\
\times \left(\epsilon^{\text{BP}}(d_l, d_r, L = \infty, w) - e^{-\frac{1}{B} \left(\frac{W}{w-1} - A \ln \ln \frac{D}{\delta} - C\right)}\right).
\]

Here \(A, B, C, D\) and \(\delta_*\) are strictly positive constants that depend only on the ensemble parameters \(d_l, d_r\) and \(w\).

This theorem says that the WD thresholds approach the BP threshold \(\epsilon^{\text{BP}}(d_l, d_r, L = \infty, w)\) at least exponentially fast in the ratio of the size of the window \(W\) to the coupling length \(w\) for a fixed target erasure probability \(\delta < \delta_*\).
Multi Dimensional Extension

Background

- Spatially-Coupled (SC) codes are constructed by coupling many regular low-density parity-check codes in a chain.
- Strongness: Spatial coupling improves BP threshold.
- Drawback: The decoding chain of SC codes aborts when facing burst erasure.

Goal

In this paper, we introduce multi-dimensional (MD) SC codes to circumvent the burst erasures.
Definition: 1D-SC \((d_l, d_r, L, w, \mathcal{Z})\) Codes

- Consider \(L\) sections on 1D discrete torus \(\mathbb{Z}_L\).
- Put \(M\) bit nodes of degree \(d_l\) and \(\frac{d_l}{d_r} M\) check nodes of degree \(d_r\) at each section \(i \in \{0, \ldots, L - 1\}\).
- Connect edges randomly between bit nodes at section \(i\) and check nodes at section \(i + j\) \((j \in \{0, \ldots, w - 1\})\) with \(Md_l/w\) edges.
- Shorten all the bit nodes at section \(i \in \mathcal{Z} = \{0, \ldots, w - 1\}\)

Figure: \((d_l = 3, d_r = 6, L = 5, w = 2, \mathcal{Z} = \{0, 1\})\) code
We Introduce Simple Representation of Connection in Order to Emphasize Symmetry of Decoding Messages

In the right figure, nodes are connected if the corresponding sections can be reached within 2 edges in the left figure.
Definition: \((d_l, d_r, L, w, \mathcal{Z})\) 2D-SC Code

- Consider \(L^2\) sections on 2D-torus \(\mathbb{Z}_L^2\).
- Put \(M\) bit nodes of degree \(d_l\) and \(\frac{d_l}{d_r} M\) check nodes of degree \(d_r\) at each section \(i \in \{0, \ldots, L - 1\}^2\).
- Connect randomly bit nodes at section \(i\) and check nodes at section \(i + j\) (\(j \in \{0, \ldots, w - 1\}^2\)) with \(Md_l/w^2\) edges.
- Shorten all the bit nodes at section \(i \in \mathcal{Z} \subset \mathbb{Z}_L^2\).

Example

![Diagram for w = 2](image1)

![Diagram for w = 3](image2)

Figure: \(w = 2\)

Figure: \(w = 3\)
Density Evolution Predicts BP Erasure Probability $p_{i_\ell}^{(\ell)}$ at round $\ell$ at section $i$

\[ p_{i_\ell}^{(0)} = \begin{cases} 
0 & (i \in \mathcal{Z}) \\
\epsilon & (i \notin \mathcal{Z}) \\
\epsilon + \alpha & \text{(burst erasures occur at section } i) 
\end{cases} \]

\[ p_{i_\ell}^{(\ell)} = p_{i_\ell}^{(0)} \left( \sum_{i} \frac{1}{w^D} \left( 1 - \left( 1 - \sum_{k} \frac{1}{w^D} p_{i+j-k}^{(\ell-1)} \right)^{d_r-1} \right)^{d_l-1} \right) \]

We are interested in BP threshold $\epsilon^*$ such that

BP threshold: $\epsilon^* = \sup \left\{ \epsilon > 0 \mid \lim_{\ell \to \infty} \mathbb{P}_b^{(\ell)} = 0 \right\}$

BP erasure probability: $\mathbb{P}_b^{(\ell)} = \frac{1}{L^D - \# \mathcal{Z}} \sum_{i \in \mathbb{Z}_L^D} p_{i_\ell}^{d_l}$
2D-SC Codes with Line Segment Zero Domain are Equivalent to 1D-SC Codes

Proposition:

\[
\epsilon_{2D}^*(d_l, d_r, L, w, Z_{2D}) = \epsilon_{1D}^*(d_l, d_r, L, w, Z_{1D})
\]

\[
Z_{2D} = \{(i, j) \mid i \in \{0, 1, \ldots, w - 1\}, j \in [L]\}
\]

\[
Z_{1D} = \{0, 1, \ldots, w - 1\}
\]

Example

- Left figure shows initial constellation \(p^{(0)}_{(i_1, i_2)}\) of 2D-SC codes.
- There are \(L^2 = 201^2\) sections
- Black line at middle represents zero domain \(Z\)
- Red sections are set with \(\epsilon = 0.480\)
- Demonstrate the density evolution of (3,6) 2D-SC codes
Burst Error is a Problem for SC Codes

- SC codes have a nice structure for windowed encoding and decoding.
- Interleaving over whole coded bits kills the structure.
- Burst errors stops the decoding chain.
- Increasing coupling number $L$ does not help. Increasing randomized window size $w$ does help. But large $w$ is not preferable.

- We assume interleaver in each section.
  Then burst erasures can be viewed as output from $\text{BEC}(\epsilon + \alpha)$.
- We assume burst erasures occur randomly at some $O(1)$ sections.
1D-SC decoding is a chain reaction like dominoes. If one of dominoes is too heavy to be toppled, the decoding chain reaction aborts.

Example: \((d_l = 3, d_r = 6, L = 101, w = 4, \mathcal{Z} = \{-1, 0, +1\})\)

\[\text{BEC}(\epsilon = 0.48) + \text{burst erasures BEC}(\epsilon + \alpha = 0.52)\]
2D-SC Codes are More Robust to Burst Erasures than 1D-SC Codes

Theorem: Necessity Condition of Decoding Burst Erasure
The MD \((d_l, d_r, L, \omega, \mathcal{Z})\) code of dimension \(D\) cannot recover any burst section erasure if \(\epsilon + \alpha > \epsilon_{BP}(d_l, d_r)w^D\), where \(\epsilon_{BP}(d_l, d_r)\) is the BP threshold of uncoupled \((d_l, d_r)\) codes.
Square Zero Domain Mitigates Rate-loss

- 2D-SC codes have $L^2$ sections. But 2D-SC codes with line segment zero domain have rate-loss $O(1/L)$ which is worse than 1D-SC codes.
- Rate-loss of 2D-SC codes with square zero domain with size $z$ turns out $O(1/L^2)$

- Demonstrate density evolution
- Square zero domain size $z$ has to be set large to grow.
- Zero domain surface needs to be flat to grow in the discrete world.
- Small size zero domain degrades BP threshold.
Square Zero Domain Mitigates Rate-loss
Conclusion

- We proposed multi-dimensional SC codes.
- 2D-SC codes achieve the same BP threshold as 1D-SC codes when line-segment zero domains are used as the shortened domain.
- 2D-SC codes are more robust against burst section erasures than 1D-SC codes.
Good points of spatially-coupled LDPC codes

Minimum distance and water-fall tradeoff
Irregular LDPC codes often suffer from the dilemma between minimum
distance and water-fall performance. Spatially-coupled LDPC codes have,
at the same time, large minimum distance and capacity-achieving
water-fall performance.

Universality
Sason et al. proved that, in the limit of large column-weight and
row-weight, regular LDPC codes achieve the capacity of any MBIOS
channels under MAP decoding.

\[
\lim_{d_l, d_r \to \infty} \epsilon^{\text{MAP}}(d_l, d_r) = \epsilon^{\text{Shannon}}(d_l, d_r)
\]

Hence, it is expected that coupled codes universally achieve the capacity.
Drawback of coupled codes

**Drawback**

Infinite column-weight and row-weight are required to strictly achieve the capacity, which leads to infinite computation per iteration.

**Bad news**

- Sason and Urbanke showed the codes achieving $1 - \epsilon$ of capacity under MAP decoding have average degree at least $O(\ln \frac{1}{\epsilon})$.
- This result is valid only for the codes without punctured bits.
Capacity-Achieving Codes with Bounded Degree

- MacKay-Neal (MN) Codes
  - Murayama et al. Tanaka and Saad observed by *replica method* that MN codes were capacity-achieving under MAP decoding with bounded degree for BSC and AWGNC.

- Hsu-Anastasopoulos (HA) Codes
  - Hsu and Anastasopoulos proved HA codes achieve the capacity of any MBIOS channels under MAP decoding with bounded degree.
Coupled MN codes achieve Shannon limit

For any $d_l, d_r, d_g \geq 2$, we observed the BP thresholds of spatially coupled $(d_l, d_r, d_g)$-MN codes are very close to the Shannon limit for the BEC. Obata et al. proved that

$$\epsilon_{\text{POT}}^{\text{MN}}(d_l = 2, d_r, d_g = 2) = \epsilon_{\text{Shannon}}^{\text{MN}}(d_l = 2, d_r, d_g = 2).$$

Protograph of spatially coupled (8,4,5)-MN code
Coupled HA codes achieve Shannon limit

For any $d_l, d_r, d_g \geq 2$, we observed the BP thresholds of spatially coupled $(d_l, d_r, d_g)$-HA codes are very close to the Shannon limit for the BEC. Obata et al. proved that

$$\epsilon_{\text{POT}}^{\text{HA}}(d_l, d_r = 2, d_g = 2) = \epsilon_{\text{Shannon}}^{\text{HA}}(d_l, d_r = 2, d_g = 2).$$

Protograph of spatially coupled $(4,8,5)$-HA code
Uncoupled Codes and How to Couple Them

\((d_l, d_r)\)-regular codes

\(x^{d_r} \quad \cdots \quad x^{d_l} \)

Density evolution updates scalars \(x^{(\ell)} \in \mathbb{R}, \ (\ell = 0, 1, \ldots)\)

\((d_l, d_r, L, w)\)-coupled codes

\(x^{d_r} \quad \cdots \quad x^{d_l} \)

Density evolution updates vectors \(x^{(\ell)} \in \mathbb{R}^L, \ (\ell = 0, 1, \ldots)\)
Coupled Hsu-Anastasopoulos (HA) Codes Achieve Capacity with BP Decoding with Bounded Degree

- Achieve the capacity of BEC with **bounded** degree (proved).
- Numerical results show **universality** over BMS channels, e.g., BAWGNC [Mitchell et al., 12].
Block Codes and Rateless Codes

Block codes
Block codes are designed to generate $n$ fixed-length coded bits from $k$ information bits so that the coding rate $R = \frac{k}{n}$ approaches the capacity $C$ with vanishing decoding error probability.

Rateless Codes
Rateless codes are designed to generate limitless sequence of coded bits from $k$ information bits so that receivers can recover the $k$ information bits from arbitrary $(1 + \alpha)k/C =: n$ received bits with vanishing overhead $\alpha$ and with vanishing decoding error probability.

Two preceding rateless codes
1. Raptor Codes
2. Coupled Rateless LDGM Codes
Raptor Codes Have Unbounded Degree
Scheme [Shokrollahi, 03]

- High-rate regular LDPC **pre-code**
- Each output bit is the sum of $d$ randomly selected pre-coded bits w.p. $\Omega_d$
- Optimized irregular output degree distribution $\Omega(x) = \sum_d \Omega_d x^d$

What is strong
Raptor codes **achieve capacity** of BEC($\epsilon$).

Drawback
- Raptor codes need to have **unbounded maximum degree** $d$ of $\Omega(x)$ to achieve capacity, which leads to a computation complexity problem both at encoders and decoders.
- Being optimized for only BEC, raptor codes are **not universal** for BMSC.
Coupled Rateless LDGM Codes Have Non-zero Error Probability

Scheme [Aref and Urbanke, 11]
- Coupled rateless LDGM code
- No precode

What is strong
- Achieves area threshold of the underlying code
- Coupled rateless LDGM codes are likely universal

Drawback
- Degree 0 information bit nodes exist, which bounds decoding error probability away from 0.
We Propose Using Coupled HA Codes for Rateless Coding

Do the codes achieve capacity of BEC or achieve zero-overhead? If so how small the degree $d_l, d_r, d_g$ can be?
Density Evolution Predicts Error Probability

\( i \): section and \( \ell \): BP iteration round

\( p_i^{(\ell)} \): erasure probability of BP messages sent from bit nodes to check nodes

\( s_i^{(\ell)} \): erasure probability of BP messages sent from bit nodes to channel nodes

\[
\begin{align*}
p_i^{(0)} &= s_i^{(0)} = \begin{cases} 0 & i \not\in [0, L - 1] \\ 1 & i \in [0, L - 1] \end{cases} \\
\Lambda(x) &= \exp[-l_{\text{avg}}(1 - x)], \\
p_i^{(\ell+1)} &= \left( \frac{1}{w} \sum_{j=0}^{w-1} (1 - (1 - \frac{1}{w} \sum_{k=0}^{w-1} p_i^{(\ell)} d_{r-1}))^{d_l-1} \right) \cdot \Lambda\left( \frac{1}{w} \sum_{j=0}^{w-1} (1 - (1 - \epsilon)(1 - \frac{1}{w} \sum_{k=0}^{w-1} s_i^{(\ell)} d_{g-1})^{d_l-1}) \right) \\
s_i^{(\ell+1)} &= \left( \frac{1}{w} \sum_{j=0}^{w-1} (1 - (1 - \frac{1}{w} \sum_{k=0}^{w-1} p_i^{(\ell)} d_{r-1}))^{d_l} \right) \cdot \Lambda\left( \frac{1}{w} \sum_{j=0}^{w-1} (1 - (1 - \epsilon)(1 - \frac{1}{w} \sum_{k=0}^{w-1} s_i^{(\ell)} d_{g-1})^{d_l-1}) \right)
\end{align*}
\]
We are interested in minimum overhead $\alpha$ such that decoding erasure probability goes to 0.

From density evolution, overhead threshold $\alpha_L^*$ is calculated as follows

$$\alpha_L^* := \inf \{ \alpha > 0 \mid P_b^{(\infty)}(L) = 0 \} ,$$

$$P_b^{(\ell)} := \frac{1}{L} \sum_{i=0}^{L-1} p_i^{(\ell)} ,$$

$$l_{avg} = \frac{d_g}{1 - \epsilon} \frac{LR_{pre}(L)(1 + \alpha)}{L + w - 1}$$

From the rateloss effect, we need large $L \to \infty$ to have $\alpha_L^* = 0$

- Do the proposed codes achieve zero overhead threshold
  $$\lim_{L \to \infty} \alpha_L^* = 0$$

- If so how small the degree $d_l, d_r, d_g$ can be?
Observed proposed codes achieve zero overhead threshold with some parameters $d_l, d_r, d_g$. How small can they be?

- We numerically observed that the proposed codes achieve zero overhead threshold $\lim_{L \to \infty} \alpha_L^* = 0$ if $d_l \geq 3$, $d_g \geq 3$ and $d_r \geq 4$.
- When $d_l = 2$, we numerically observed that the proposed codes achieve zero overhead threshold with some parameter tuples $(d_l = 2, d_r, d_g)$.
Necessary Condition of Capacity-Achieving Coupled Precoded Regular Rateless Codes

Theorem:
For capacity-achieving \((d_l = 2, d_r, d_g, L, w)\) precoded rateless codes have to satisfy

\[
d_g \geq \frac{d_r \ln(d_r - 1)}{d_r - 2}.
\]

This implies \(d_r > 2\) or \(d_g > 2\) is necessary.

Outline of Proof: Taylor expansion around 0 gives

\[
p^{(\ell+1)} = P_L p^{(\ell)} + o(p^{(\ell)}),
\]

where \(P_L\) is an \(L \times L\) band matrix as follows

\[
(P_L)_{i,j} = \begin{cases} 
\frac{w - |i-j|}{w^2} (d_r - 1) \lambda(\epsilon) & (|i-j| \leq w) \\
0 & (|i-j| > w)
\end{cases}
\]

Evaluate the spectrum radius \(\rho(P_L)\) of \(P_L\) gives a necessary condition.

\[
1 > \rho(P_L) = \max_{x \in \mathbb{R}^L: x \neq 0} \frac{x^\top P_L x}{x^\top x} \geq \frac{1^\top P_L 1}{1^\top 1}
\]

\[
L \to \infty \quad (d_r - 1) e^{-(\overline{\text{avg}})^* L (1-\epsilon)}.
\]
Conclusion

Achievement

▶ Empirically observed that precoded rateless codes achieve the capacity over the BEC.
▶ Derive a necessary condition of capacity-achieving codes.

Ongoing work

▶ Prove proposed codes achieve capacity of BEC via potential function. (completed)
▶ An extension to BMS channels and a proof for universal capacity-achievability.