A Brief Introduction to Spatially-Coupled Codes and Threshold Saturation

Henry D. Pfister

Based on joint work with Yung-Yih Jian, Santhosh Kumar, Krishna R. Narayanan, Phong S. Nguyen, and Arvind Yedla

Chinese University of Hong Kong
June 22nd, 2015
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
C.1. Coding for Discrete-Time Memoryless Channels

- Transition probability: \( P_{Y|X}(y|x) \) for \( x \in \mathcal{X} \) and \( y \in \mathcal{Y} \)
- Transmit a length-\( n \) codeword \( x \in C \subset \mathcal{X}^n \)

C.2. Shannon Capacity

- Random code of rate \( R \triangleq \frac{1}{n} \log_2 |C| \) (bits per channel use)
- As \( n \to \infty \), reliable transmission if \( R < C \triangleq \max_{p(x)} I(X; Y) \)

C.3. Example: the binary erasure channel BEC(\( \varepsilon \))

- Bits sent perfectly (with prob. \( 1 - \varepsilon \)) or erased (with prob. \( \varepsilon \))
- Capacity: \( C = 1 - \varepsilon = \text{fraction unerased bits} \)
- Roughly one info bit transmitted for each unerased reception
Low-Density Parity-Check (LDPC) Codes

- Linear codes with a sparse parity-check matrix $H$
  - Regular $(l, r)$: $H$ has $l$ ones per column and $r$ ones per row
  - Irregular: number of ones given by degree distribution $(\lambda, \rho)$
  - Introduced by Gallager in 1960; largely forgotten until 1995

- Tanner Graph
  - An edge connects check node $i$ to bit node $j$ if $H_{ij} = 1$
  - Naturally leads to message-passing iterative (MPI) decoding
Decoding LDPC Codes

- **Belief-Propagation (BP) Decoder**
  - Low-complexity message-passing decoder by Gallager
  - Probability estimates are passed along edges in the Tanner graph
  - Updates based on assuming *incoming estimates are independent*

- **Density Evolution (DE)**
  - Tracks *distribution of messages* during iterative decoding
  - BP noise threshold can be *computed via DE*
  - Long codes decode almost surely if DE predicts success

- **Maximum A Posteriori (MAP) Decoder**
  - Optimum decoder that chooses the *most likely codeword*
  - *Infeasible in practice* due to enormous number of codewords
  - MAP noise threshold can be bounded using EXIT curves
A Little History

Robert Gallager introduced LDPC codes in 1962 paper

Low-Density Parity-Check Codes

R. G. GALLAGER

Summary—A low-density parity-check code is a code specified by a parity-check matrix with the following properties: each column contains a small fixed number $j \geq 3$ of 1's and each row contains a small fixed number $k > j$ of 1's. The typical minimum distance of these codes increases linearly with block length for a fixed rate and fixed $j$. When used with maximum likelihood decoding on a sufficiently quiet binary-input symmetric channel, the typical probability of decoding error decreases exponentially with block length for a fixed rate and fixed $j$.

A simple but nonoptimum decoding scheme operating directly from the channel a posteriori probabilities is described. Both the equations. We call the set of digits contained in a parity-check equation a parity-check set. For example, the first parity-check set in Fig. 1 is the set of digits (1, 2, 3, 5).

The use of parity-check codes makes coding (as distinguished from decoding) relatively simple to implement. Also, as Elias [3] has shown, if a typical parity-check code of long block length is used on a binary symmetric channel, and if the code rate is between critical rate and channel capacity, then the probability of decoding error

Judea Pearl defined general belief-propagation in 1986 paper

Fusion, Propagation, and Structuring in Belief Networks

Judea Pearl
Cognitive Systems Laboratory, Computer Science Department, University of California, Los Angeles, CA 90024, U.S.A.

Recommended by Patrick Hayes

ABSTRACT

Belief networks are directed acyclic graphs in which the nodes represent propositions (or variables), the arcs signify direct dependencies between the linked propositions, and the strengths of these dependencies are quantified by conditional probabilities. A network of this sort can be used to represent the generic knowledge of a domain expert, and it turns into a computational architecture if the links are used not merely for storing factual knowledge but also for directing and activating the data flow in the computations which manipulate this knowledge.
Message-Passing Decoding for the BEC (1)

- Bit and check nodes define the set of valid codewords
  - **Circles** represent a single bit value shared by checks
  - **Squares** assert attached bits sum to 0 mod 2

- Iterative decoding on the binary erasure channel (BEC)
  - Estimates of bit values are passed along edges in phases
  - 1st phase: bits pass messages to adjacent checks
  - 2nd phase: checks pass messages to adjacent bits
  - Each output message depends on other input messages
  - Messages are always *either the correct value or an erasure*
Message-Passing Decoding for the BEC (2)

- Message passing rules for the BEC
  - Bits pass an erasure only if all other inputs are erased
  - Checks pass the correct value only if all other inputs are correct

![Diagram showing message passing rules for the BEC]
Message-Passing Decoding for the BEC (2)

- Message passing rules for the BEC
  - Bits pass an erasure only if all other inputs are erased
  - Checks pass the correct value only if all other inputs are correct

![Diagram of message passing rules for the BEC](image-url)
Message-Passing Decoding for the BEC (2)

- Message passing rules for the BEC
  - Bits pass an erasure only if all other inputs are erased
  - Checks pass the correct value only if all other inputs are correct

![Diagram of message-passing rules for the BEC](image)
Message-Passing Decoding for the BEC (2)

- Message passing rules for the BEC
  - Bits pass an erasure only if all other inputs are erased
  - Checks pass the correct value only if all other inputs are correct

- If input messages are independently erased with prob. $x$

\[
\begin{align*}
&\text{1} \\
\Rightarrow &\text{1} \\
\Rightarrow &\text{1}
\end{align*}
\]

\[
\begin{align*}
&\text{1} \\
\Rightarrow &\text{0} \\
\Rightarrow &\text{?}
\end{align*}
\]

\[
\begin{align*}
&\text{x} \\
\Rightarrow &\text{e} \\
\Rightarrow &\text{x}
\end{align*}
\]

\[
\begin{align*}
&\text{x} \\
\Rightarrow &\text{x}
\end{align*}
\]
Message-Passing Decoding for the BEC (2)

- Message passing rules for the BEC
  - Bits pass an erasure only if all other inputs are erased
  - Checks pass the correct value only if all other inputs are correct

```
1  ▶  1
?  ▶  ?
?
```

- If input messages are independently erased with prob. \( x \)

```
x  ▶  \( \varepsilon \)
x  ▶  \varepsilon x^3
x  ▶  \varepsilon x^3
```

```
x  ▶  ?
```

```
x  ▶  ?
```

A Brief Introduction to Spatially-Coupled Codes and Threshold Saturation
Message-Passing Decoding for the BEC (2)

- Message passing rules for the BEC
  - Bits pass an erasure only if all other inputs are erased
  - Checks pass the correct value only if all other inputs are correct

- If input messages are independently erased with prob. $x$

\[
\varepsilon \rightarrow \varepsilon x^3 \quad \text{and} \quad 1 - (1 - x)^3
\]
Computation Graph and Density Evolution

- Computation graph for a (3,4)-regular LDPC code
  - Illustrates decoding from the perspective of a single bit-node
  - For long random LDPC codes, the graph is typically a tree
  - Allows density evolution to track message erasure probability
  - On each level, messages are independently erased with fixed prob.

\[
\begin{align*}
\tilde{x}_3 &= \varepsilon y_2^3 \\
y_2 &= 1 - (1 - x_2)^3 \\
x_2 &= \varepsilon y_1^2 \\
y_1 &= 1 - (1 - x_1)^3 \\
x_1 &= \varepsilon
\end{align*}
\]
A Brief Introduction to Spatially-Coupled Codes and Threshold Saturation

**Computation Graph and Density Evolution**

- Computation graph for a (3,4)-regular LDPC code
  - Illustrates decoding from the perspective of a single bit-node
  - For long random LDPC codes, the graph is typically a tree
  - Allows density evolution to track message erasure probability
  - On each level, messages are independently erased with fixed prob.

\[
\begin{align*}
\tilde{x}_3 &= \varepsilon y_2^3 \\
y_2 &= 1 - (1 - x_2)^3 \\
x_2 &= \varepsilon y_1^2 \\
y_1 &= 1 - (1 - x_1)^3 \\
x_1 &= 0.600
\end{align*}
\]
Computation Graph and Density Evolution

- Computation graph for a (3,4)-regular LDPC code
  - Illustrates decoding from the perspective of a single bit-node
  - For long random LDPC codes, the graph is typically a tree
  - Allows density evolution to track message erasure probability
  - On each level, messages are independently erased with fixed prob.

\[ \tilde{x}_3 = \varepsilon y_2^2 \]
\[ y_2 = 1 - (1 - x_2)^3 \]
\[ x_2 = \varepsilon y_1^2 \]
\[ y_1 = 0.936 \]
\[ x_1 = 0.600 \]
Computation Graph and Density Evolution

- Computation graph for a (3,4)-regular LDPC code
  - Illustrates decoding from the perspective of a single bit-node
  - For long random LDPC codes, the graph is typically a tree
  - Allows density evolution to track message erasure probability
  - On each level, messages are independently erased with fixed prob.
Computation Graph and Density Evolution

- Computation graph for a (3,4)-regular LDPC code
  - Illustrates decoding from the perspective of a single bit-node
  - For long random LDPC codes, the graph is typically a tree
  - Allows density evolution to track message erasure probability
  - On each level, messages are independently erased with fixed prob.

\[ \tilde{x}_3 = \varepsilon y_3^3 \]
\[ y_2 = 0.894 \]
\[ x_2 = 0.526 \]
\[ y_1 = 0.936 \]
\[ x_1 = 0.600 \]
Computation Graph and Density Evolution

- Computation graph for a (3,4)-regular LDPC code
  - Illustrates decoding from the perspective of a single bit-node
  - For long random LDPC codes, the graph is typically a tree
  - Allows density evolution to track message erasure probability
  - On each level, messages are independently erased with fixed prob.

$x_1 = 0.600$
$y_1 = 0.936$
$x_2 = 0.526$
$y_2 = 0.894$
$\tilde{x}_3 = 0.429$
Density Evolution (DE) for LDPC Codes

Density evolution for a (3, 4)-regular LDPC code:

$$x_{\ell+1} = \varepsilon \left(1 - (1 - x_{\ell})^3\right)^2$$

Decoding Thresholds:

$$\varepsilon^{BP} \approx 0.647$$

$$\varepsilon^{MAP} \approx 0.746$$

$$\varepsilon^{Sh} = 0.750$$

- DE tracks bit-to-check msg erasure rate $$x_{\ell}$$ after $$\ell$$ iterations
Density Evolution (DE) for LDPC Codes

Density evolution for a $(3, 4)$-regular LDPC code:

$$x_{\ell+1} = \varepsilon \left(1 - (1 - x_{\ell})^3\right)^2$$

Decoding Thresholds:

$$\varepsilon^{BP} \approx 0.647$$
$$\varepsilon^{MAP} \approx 0.746$$
$$\varepsilon^{Sh} = 0.750$$

- DE tracks bit-to-check msg erasure rate $x_{\ell}$ after $\ell$ iterations
- $x_{\ell}$ decreases to a limit $x_\infty(\varepsilon)$ that depends on $\varepsilon$
Density Evolution (DE) for LDPC Codes

Density evolution for a \((3, 4)\)-regular LDPC code:

\[ x_{\ell+1} = \varepsilon \left( 1 - (1 - x_{\ell})^3 \right)^2 \]

Decoding Thresholds:

\[ \varepsilon^{\text{BP}} \approx 0.647 \]
\[ \varepsilon^{\text{MAP}} \approx 0.746 \]
\[ \varepsilon^{\text{Sh}} = 0.750 \]

- DE tracks bit-to-check msg erasure rate \(x_\ell\) after \(\ell\) iterations
- \(x_\ell\) decreases to a limit \(x_\infty(\varepsilon)\) that depends on \(\varepsilon\)
- As \(n \to \infty\), decoding succeeds if \(\varepsilon\) less than the BP noise threshold
  - \(\varepsilon^{\text{BP}} = \sup\{\varepsilon \in [0, 1] \mid x_\infty(\varepsilon) = 0\}\) (easily computed numerically)
EXtrinsic Information Transfer (EXIT) Curves

- (3,4)-regular LDPC code
  - Codeword $(X_1, \ldots, X_n)$
  - Received $(Y_1, \ldots, Y_n)$
EXtrinsic Information Transfer (EXIT) Curves

- (3,4)-regular LDPC code
  - Codeword \((X_1, \ldots, X_n)\)
  - Received \((Y_1, \ldots, Y_n)\)

- BP EXIT curve via DE
  - This code: \(h_{\text{BP}}(\varepsilon) = (x_\infty(\varepsilon))^3\)
  - 0 below BP threshold 0.647
EXtrinsic Information Transfer (EXIT) Curves

- (3,4)-regular LDPC code
  - Codeword \((X_1, \ldots, X_n)\)
  - Received \((Y_1, \ldots, Y_n)\)

- **BP EXIT curve** via DE
  - This code: \(h^{BP}(\varepsilon) = (x_{\infty}(\varepsilon))^3\)
  - 0 below BP threshold 0.647

- **MAP EXIT curve** is extrinsic entropy \(H(X_i|Y_{\sim i})\) vs. channel \(\varepsilon\)
  - 0 below MAP threshold 0.746
  - Area under curve equals rate \(R\)
  - Upper bounded by BP EXIT
EXtrinsic Information Transfer (EXIT) Curves

- (3,4)-regular LDPC code
  - Codeword \((X_1, \ldots, X_n)\)
  - Received \((Y_1, \ldots, Y_n)\)

- BP EXIT curve via DE
  - This code: \(h_{BP}(\varepsilon) = (x_\infty(\varepsilon))^3\)
  - 0 below BP threshold 0.647

- MAP EXIT curve is extrinsic entropy \(H(X_i | Y_{\sim i})\) vs. channel \(\varepsilon\)
  - 0 below MAP threshold 0.746
  - Area under curve equals rate \(R\)
  - Upper bounded by BP EXIT

- MAP threshold upper bound \(\varepsilon^{MAP}\)
  - \(\varepsilon\) s.t. area under BP EXIT equals \(R\)
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
Spatially-Coupled LDPC Codes: \((l, r, L, w)\) Ensemble

\[\cdots -L \cdots -2 \quad -1 \quad 0 \quad 1 \quad 2 \cdots L\]
Spatially-Coupled LDPC Codes: \((l, r, L, w)\) Ensemble

Historical Notes
- LDPC convolutional codes introduced by FZ in 1999
- Shown to have near optimal noise thresholds by LSZC in 2005
- \((l, r, L, w)\) ensemble proven to achieve capacity by KRU in 2011
Spatially-Coupled LDPC Codes: \((l, r, L, w)\) Ensemble

Historical Notes
- LDPC convolutional codes introduced by FZ in 1999
- Shown to have near optimal noise thresholds by LSZC in 2005
- \((l, r, L, w)\) ensemble proven to achieve capacity by KRU in 2011

\[
l = 3, \quad w = 3, \quad r = 4,\]

\[
\begin{align*}
\pi_{-L-2} & \quad \pi_{-L-1} & \quad \pi_{-L} & \quad \cdots & \quad \pi_{-2} & \quad \pi_{-1} & \quad \pi_0 & \quad \pi_1 & \quad \pi_2 & \quad \cdots & \quad \pi_L & \quad \pi_{L+1} & \quad \pi_{L+2} \\
\pi'_{-L} & \quad \pi'_{-2} & \quad \pi'_{-1} & \quad \cdots & \quad \pi'_0 & \quad \pi'_1 & \quad \pi'_2 & \quad \cdots & \quad \pi'_L & \quad \pi'_{L+1} & \quad \pi'_{L+2} \\
\end{align*}
\]
Spatially-Coupled LDPC Codes: \((l, r, L, w)\) Ensemble

Historical Notes

- LDPC convolutional codes introduced by FZ in 1999
- Shown to have near optimal noise thresholds by LSZC in 2005
- \((l, r, L, w)\) ensemble proven to achieve capacity by KRU in 2011
Iterative Decoding Threshold Analysis for LDPC Convolutional Codes

Michael Lentmaier, Member, IEEE, Arvind Sridharan, Member, IEEE, Daniel J. Costello, Jr., Life Fellow, IEEE, and Kamil Sh. Zigangirov, Fellow, IEEE
Threshold Saturation via Spatial Coupling: Why Convolutional LDPC Ensembles Perform So Well over the BEC

Shrinivas Kudekar, Member, IEEE, Thomas J. Richardson, Fellow, IEEE, and Rüdiger L. Urbanke
Density Evolution for the $((l, r, L, w)\text{-SC LDPC Ensemble}$

\[
x_{i}^{(l+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - (1 - x_{i+j-k}^{(\ell)})^{r-1} \right) \right)^{l-1} \mathbf{1}_{[-L, L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[(3, 4, 16, 3)\)-SC Ensemble with \(\varepsilon = 0.70\]

\[
x^{(\ell+1)}_i = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - (1 - x^{(\ell)}_{i+j-k})^{r-1} \right) \right)^l 1_{[-L, L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[
\begin{align*}
(3, 4, 16, 3)\text{-SC Ensemble with } \varepsilon &= 0.70 \\
\end{align*}
\]

\[
x^{(\ell+1)}_i = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( \left( 1 - x^{(\ell)}_{i+j-k} \right)^{r-1} \right) \right)^{l-1} 1_{[-L, L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[(3, 4, 16, 3)\)-SC Ensemble with \(\varepsilon = 0.70\]

\[
x_i^{(\ell+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - (1 - x_{i+j-k}^{(\ell)})^{r-1} \right) \right)^{l-1} \mathbf{1}_{[-L, L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[(3, 4, 16, 3)\)-SC Ensemble with \(\varepsilon = 0.70\)

\[
x^{(\ell+1)}_i = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left(1 - (1 - x^{(\ell)}_{i+j-k})^{r-1}\right)^{l-1} \right) \mathbf{1}_{[-L,L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[(3, 4, 16, 3)\)-SC Ensemble with \(\varepsilon = 0.70\]

\[
x_i^{(\ell+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - (1 - x_{i+j-k}^{(\ell)})^{r-1} \right) \right)^{l-1} \text{1}_{[-L,L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[(3, 4, 16, 3)\)-SC Ensemble with \(\varepsilon = 0.70\)

\[
x^{(\ell+1)}_i = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - (1 - x^{(\ell)}_{i+j-k})^{r-1} \right) \right)^{l-1} 1_{[-L, L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\((3, 4, 16, 3)\)-SC Ensemble with \(\varepsilon = 0.70\)

\[
x_i^{(\ell+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - \left( 1 - x_{i+j-k}^{(\ell)} \right)^{r-1} \right) \right)^{l-1} 1_{[-L, L+w-1]}(i-k)
\]
Density Evolution for the \((l, r, L, w)\)-SC LDPC Ensemble

\[
x_i^{(\ell+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \varepsilon \left( \frac{1}{w} \sum_{j=0}^{w-1} \left( 1 - \left( 1 - x_{i+j-k}^{(\ell)} \right)^{r-1} \right) \right)^{l-1} \mathbf{1}_{[-L, L+w-1]}(i-k)
\]
Properties of Threshold Saturation

<table>
<thead>
<tr>
<th>$l$</th>
<th>$r$</th>
<th>$\varepsilon_{BP}$</th>
<th>$\varepsilon_{MAP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>6</td>
<td>0.4294</td>
<td>0.4882</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>0.3834</td>
<td>0.4977</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0.3416</td>
<td>0.4995</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>0.3075</td>
<td>0.4999</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>0.2798</td>
<td>0.5000</td>
</tr>
</tbody>
</table>

- Spatial coupling achieves the MAP threshold as $w \to \infty$
  - BP threshold typically decreases after $l = 3$
  - MAP threshold is increasing in $l, r$ for fixed rate

- Benefits and Drawbacks
  - For fixed $L$, minimum distance grows linearly with block length
  - Rate loss of $O(w/L)$ is a big obstacle in practice
Threshold Saturation via Spatial Coupling

- **General Phenomenon** (observed by Kudekar, Richardson, Urbanke)
  - BP threshold of the spatially-coupled system converges to the **MAP threshold** of the uncoupled system
  - Can be proven rigorously in many cases!
Threshold Saturation via Spatial Coupling

- **General Phenomenon** (observed by Kudekar, Richardson, Urbanke)
  - **BP threshold** of the spatially-coupled system converges to the **MAP threshold** of the uncoupled system
  - Can be proven rigorously in many cases!

- Connection to statistical physics
  - Factor graph defines system of coupled particles
  - Valid sequences are **ordered crystalline structures**

---

http://www.youtube.com/watch?v=Xe8vJrIvDQM
Threshold Saturation via Spatial Coupling

- **General Phenomenon** (observed by Kudekar, Richardson, Urbanke)
  - BP threshold of the spatially-coupled system converges to the MAP threshold of the uncoupled system
  - Can be proven rigorously in many cases!

- Connection to statistical physics
  - Factor graph defines system of coupled particles
  - Valid sequences are ordered crystalline structures

- Between BP and MAP threshold, system acts as supercooled liquid
  - Correct answer (crystalline state) has minimum energy
  - Crystallization (i.e., decoding) does not occur without a seed
  - Ex.: ice melts at 0 °C but freezing w/o a seed requires −48.3 °C

http://www.youtube.com/watch?v=Xe8vJrlvDQM
Why is Spatial Coupling Interesting?

- Breakthroughs: first practical constructions of
  - universal codes for binary-input memoryless channels [KRU12]
  - information-theoretically optimal compressive sensing [DJM11]
  - universal codes for Slepian-Wolf and MAC problems [YJNP11]
  - codes \(\rightarrow\) capacity with iterative hard-decision decoding [JNP12]
  - codes \(\rightarrow\) rate-distortion limit with iterative decoding [AMUV12]
Why is Spatial Coupling Interesting?

- Breakthroughs: first practical constructions of
  - universal codes for binary-input memoryless channels [KRU12]
  - information-theoretically optimal compressive sensing [DJM11]
  - universal codes for Slepian-Wolf and MAC problems [YJNP11]
  - codes \(\rightarrow\) capacity with iterative hard-decision decoding [JNP12]
  - codes \(\rightarrow\) rate-distortion limit with iterative decoding [AMUV12]

- It allows rigorous proof in many cases
  - Original proofs [KRU11/12] quite specific to LDPC codes
  - Our proof is for increasing scalar/vector recursions [YJNP12/13]
Why is Spatial Coupling Interesting?

- Breakthroughs: first practical constructions of
  - universal codes for binary-input memoryless channels [KRU12]
  - information-theoretically optimal compressive sensing [DJM11]
  - universal codes for Slepian-Wolf and MAC problems [YJNP11]
  - codes $\rightarrow$ capacity with iterative hard-decision decoding [JNP12]
  - codes $\rightarrow$ rate-distortion limit with iterative decoding [AMUV12]

- It allows rigorous proof in many cases
  - Original proofs [KRU11/12] quite specific to LDPC codes
  - Our proof is for increasing scalar/vector recursions [YJNP12/13]

- Spatial coupling as a proof technique [GMU13]
  - For a large random factor graph, construct a coupled version
  - Use DE to analyze BP decoding of coupled system
  - Compare uncoupled MAP with coupled BP via interpolation
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
Universality over Unknown Parameters

The Achievable Channel Parameter Region (ACPR)

- For a sequence of coding schemes involving one or more parameters, the parameter region where decoding succeeds in the limit.
- In contrast, a capacity region is a rate region for fixed channels.

**Diagram**

- MAC-ACPR boundary for rate 1/2.
Universality over Unknown Parameters

- The Achievable Channel Parameter Region (ACPR)
  - For a sequence of coding schemes involving one or more parameters, the parameter region where decoding succeeds in the limit
  - In contrast, a capacity region is a rate region for fixed channels

- Properties
  - For fixed encoders, the ACPR depends on the decoders
  - For example, one has $\text{BP-ACPR} \subseteq \text{MAP-ACPR}$
  - Often, $\exists$ unique maximal ACPR given by information theory
Universality over Unknown Parameters

- The Achievable Channel Parameter Region (ACPR)
  - For a sequence of coding schemes involving one or more parameters, the parameter region where decoding succeeds in the limit
  - In contrast, a capacity region is a rate region for fixed channels

- Properties
  - For fixed encoders, the ACPR depends on the decoders
  - For example, one has BP-ACPR $\subseteq$ MAP-ACPR
  - Often, $\exists$ unique maximal ACPR given by information theory

- Universality
  - A sequence of encoding/decoding schemes is called universal if: its ACPR equals the optimal ACPR
  - Channel parameters are assumed unknown at the transmitter
  - At the receiver, the channel parameters are easily estimated
2-User Binary-Input Gaussian Multiple Access Channel

- Fixed noise variance
- Real channel gains $h_1$ and $h_2$ not known at transmitter
- Each code has rate $R$
- MAC-ACPR denotes the information-theoretic optimal region
A Little History: SC for Multiple-Access (MAC) Channels

- KK consider a binary-adder erasure channel (ISIT 2011)
  - SC exhibits threshold saturation for the joint decoder

- YNPN consider the Gaussian MAC (ISIT/Allerton 2011)
  - SC exhibits threshold saturation for the joint decoder
  - For channel gains $h_1, h_2$ unknown at transmitter, SC provides universality

- Others consider CDMA systems without coding
  - TTK show SC improves BP demod of standard CDMA
  - ST prove saturation for a SC protograph-style CDMA
Spatially-Coupled Factor Graph for Joint Decoder
Spatially-Coupled Factor Graph for Joint Decoder

$2L + 1$
Spatially-Coupled Factor Graph for Joint Decoder

\[ 2L + 1 \]
DE Performance of the Joint Decoder

\[ \alpha_2 = \left| h_2 \right|^2 \]

\[ \alpha_1 = \left| h_1 \right|^2 \]

MAC-ACPR boundary for rate 1/2
DE Performance of the Joint Decoder

![Graph showing DE Performance of the Joint Decoder]

- $\alpha_1 = |h_1|^2$
- $\alpha_2 = |h_2|^2$
- BP-ACPR, LDPC(3,6)
- MAC-ACPR boundary for rate $1/2$
DE Performance of the Joint Decoder

\[ \alpha_1 = |h_1|^2 \]

\[ \alpha_2 = |h_2|^2 \]

- BP-ACPR, LDPC(3, 6)
- BP-ACPR, SC(3, 6, 64, 5)
- MAC-ACPR boundary for rate 1/2

A Brief Introduction to Spatially-Coupled Codes and Threshold Saturation
DE Performance of the Joint Decoder

\[ \alpha_1 = |h_1|^2 \]

\[ \alpha_2 = |h_2|^2 \]

MAC-ACPR boundary for rate $1/2$

BP-ACPR, LDPC$(4, 8)$
DE Performance of the Joint Decoder

\[ \alpha_2 = |h_2|^2 \]

\[ \alpha_1 = |h_1|^2 \]

BP-ACPR, LDPC(4, 8)

BP-ACPR, SC(4, 8, 64, 5)

MAC-ACPR boundary for rate 1/2
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
An Abstract Approach to Spatial Coupling

Let $f : \mathcal{X} \to \mathcal{X}$ and $g : \mathcal{X} \to \mathcal{X}$ be strictly increasing $C^2$ functions on $\mathcal{X} = [0, 1]$ with $f(0) = g(0) = 0$. Then, the scalar recursion (from $x^{(0)} = 1$)

$$
\begin{align*}
y^{(\ell+1)} &= g \left( x^{(\ell)} \right) \\
x^{(\ell+1)} &= f \left( y^{(\ell+1)} \right)
\end{align*}
$$
Let $f : \mathcal{X} \to \mathcal{X}$ and $g : \mathcal{X} \to \mathcal{X}$ be strictly increasing $C^2$ functions on $\mathcal{X} = [0, 1]$ with $f(0) = g(0) = 0$. Then, the scalar recursion (from $x^{(0)} = 1$)

\begin{align*}
y^{(\ell+1)} &= g \left( x^{(\ell)} \right) = 1 - (1 - x)^3 \\
x^{(\ell+1)} &= f \left( y^{(\ell+1)} \right) = \varepsilon x^2
\end{align*}

Ex. (3,4) LDPC
An Abstract Approach to Spatial Coupling

Let $f: \mathcal{X} \to \mathcal{X}$ and $g: \mathcal{X} \to \mathcal{X}$ be strictly increasing $C^2$ functions on $\mathcal{X} = [0, 1]$ with $f(0) = g(0) = 0$. Then, the scalar recursion (from $x^{(0)} = 1$)

$$y^{(\ell+1)} = g\left(x^{(\ell)}\right) = 1 - (1 - x)^3$$

$$x^{(\ell+1)} = f\left(y^{(\ell+1)}\right) = \varepsilon x^2$$

- Characterizes the performance of a single system

Ex. (3,4) LDPC
An Abstract Approach to Spatial Coupling

Let $f: \mathcal{X} \to \mathcal{X}$ and $g: \mathcal{X} \to \mathcal{X}$ be strictly increasing $C^2$ functions on $\mathcal{X} = [0, 1]$ with $f(0) = g(0) = 0$. Then, the scalar recursion (from $x^{(0)} = 1$)

\[
y^{(\ell+1)} = g\left(x^{(\ell)}\right) = 1 - (1 - x)^3
\]
\[
x^{(\ell+1)} = f\left(y^{(\ell+1)}\right) = \varepsilon x^2
\]

- Characterizes the performance of a single system
- Monotonicity and continuity of $f, g$ imply convergence: $x^{(\ell)} \downarrow x^{(\infty)}$ and $y^{(\ell)} \downarrow y^{(\infty)}$ to a fixed point

Ex. (3,4) LDPC
The coupled recursion is given, for $i \in [N+w-1]$, by $x_i^{(0)} = 1$ and

$$y_i^{(\ell+1)} = g\left(x_i^{(\ell)}\right)$$

$$x_i^{(\ell+1)} = \sum_{j=1}^{N+w-1} A_{j,i} f\left(\sum_{k=1}^{N} A_{j,k} y_k^{(\ell+1)}\right)$$

$$A = \frac{1}{w} \begin{bmatrix}
1 & 1 & \cdots & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & \cdots & 1 & 0 & 0 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 0 & 1 & 1 & \cdots & 1
\end{bmatrix}$$
The coupled recursion is given, for \( i \in [N+w-1] \), by \( x_i^{(0)} = 1 \) and
\[
\begin{align*}
y_i^{(\ell+1)} &= g \left( x_i^{(\ell)} \right) \\
x_i^{(\ell+1)} &= \sum_{j=1}^{N+w-1} A_{j,i} f \left( \sum_{k=1}^{N} A_{j,k} y_k^{(\ell+1)} \right)
\end{align*}
\]

In vector notation, we have \( x^{(0)} = 1 \) and
\[
\begin{align*}
y^{(\ell+1)} &= g \left( x^{(\ell)} \right) \\
x^{(\ell+1)} &= A^\top f \left( A y^{(\ell+1)} \right)
\end{align*}
\]
The coupled recursion is given, for \( i \in [N + w - 1] \), by \( x_i^{(0)} = 1 \) and

\[
y_i^{(\ell+1)} = g \left( x_i^{(\ell)} \right)
\]

\[
x_i^{(\ell+1)} = \sum_{j=1}^{N+w-1} A_{j,i} f \left( \sum_{k=1}^{N} A_{j,k} y_k^{(\ell+1)} \right)
\]

In vector notation, we have \( x^{(0)} = 1 \) and

\[
y^{(\ell+1)} = g \left( x^{(\ell)} \right)
\]

\[
x^{(\ell+1)} = A^\top f \left( A y^{(\ell+1)} \right)
\]

- Monotonicity and continuity of \( f, g \) again imply convergence: \( x_i^{(\ell)} \downarrow x_i^{(\infty)} \) and \( y_i^{(\ell)} \downarrow y_i^{(\infty)} \) to a coupled fixed point
Example: DE for Spatially-Coupled LDPC Codes

$N$ bit-node and $M \triangleq N + w - 1$ check-node sections

- Erasure probability by section: $x, y \in [0, 1]^M$ and $v, z \in [0, 1]^N$
- Coupling matrix: $A_{j,i}$ is the fraction of $\pi_j$ edges attached to $\pi_i$
- Vector updates: $[f(y)]_j = f(y_j)$, $j \in [N]$ and $[g(x)]_i = g(x_i)$, $i \in [M]$
Single-System Potential and Convergence to Zero

Let the potential function $U_s(x)$ of the recursion be

$$U_s(x) = \int_0^x \left( z - f(g(z)) \right) \frac{g'(z)}{g'(z)} \, dz$$

where $f(g(z)) \leq z \geq 0$ if $f(g(z)) \leq z \geq 0$.

$$= xg(x) - G(x) - F(g(x)),$$

where $F(x) = \int_0^x f(z) \, dz$ and $G(x) = \int_0^x g(z) \, dz$.

Recursion $x^{(\ell)} \to 0$ if $f(g(z)) < z$ for $z \in (0, 1]$.

Coupled $x^{(\ell)} \to 0$ as $w \to \infty$ if $U_s(x) > 0$ for $x \in (0, 1]$.
Let the potential function $U_s: \mathcal{X} \rightarrow \mathbb{R}$ of the scalar recursion be

$$U_s(x) \triangleq \int_0^x (z - f(g(z)))g'(z)dz.$$
Let the potential function $U_s : \mathcal{X} \rightarrow \mathbb{R}$ of the scalar recursion be

$$U_s(x) \triangleq \int_0^x (z - f(g(z)))g'(z)dz.$$
Let the potential function $U_s : \mathcal{X} \rightarrow \mathbb{R}$ of the scalar recursion be

$$U_s(x) \triangleq \int_0^x (z - f(g(z))) g'(z) dz.$$
The Potential Function and Threshold Saturation

Let the potential function \( U_s : \mathcal{X} \rightarrow \mathbb{R} \) of the scalar recursion be

\[
U_s(x) \triangleq \int_0^x (z - f(g(z)))g'(z)dz.
\]
Let the potential function $U_s : \mathcal{X} \rightarrow \mathbb{R}$ of the scalar recursion be

$$U_s(x) \triangleq \int_0^x (z - f(g(z)))g'(z)dz.$$ 

**Threshold Saturation**

If $f(g(x)) < x$ for $x \in (0, \delta)$ and $\min_{x \in [\delta, 1]} U_s(x) > 0$, then $\exists w_0 < \infty$:

for $w > w_0$, only fixed point of coupled recursion is $x^{(\infty)} = 0$
Let the potential function \( U_s : \mathcal{X} \rightarrow \mathbb{R} \) of the scalar recursion be

\[
U_s(x) \triangleq \int_0^x (z - f(g(z)))g'(z)dz.
\]

**Threshold Saturation**

If \( f(g(x)) < x \) for \( x \in (0, \delta) \) and \( \min_{x \in [\delta, 1]} U_s(x) > 0 \), then \( \exists w_0 < \infty \):

for \( w > w_0 \), only fixed point of coupled recursion is \( x^{(\infty)} = 0 \)
Let the potential function $U_s : \mathcal{X} \rightarrow \mathbb{R}$ of the scalar recursion be

$$U_s(x) \triangleq \int_0^x (z - f(g(z)))g'(z)dz.$$ 

**Threshold Saturation**

If $f(g(x)) < x$ for $x \in (0, \delta)$ and $\min_{x \in [\delta, 1]} U_s(x) > 0$, then $\exists w_0 < \infty$:

for $w > w_0$, only fixed point of coupled recursion is $x(\infty) = 0$
A Little History: Threshold Saturation Proofs

- the BEC by KRU in 2010
  - Established many properties and tools used by later approaches
- the Curie-Weiss model in physics by HMU in 2010
- CDMA using a GA by TTK in 2011
- CDMA with outer code via GA by Truhachev in 2011
- compressive sensing using a GA by DJM in 2011
- regular codes on BMS channels by KRU in 2012
- increasing scalar and vector recursions by YJNP in 2012
- irregular LDPC codes on BMS channels by KYMP in 2012
- non-decreasing scalar recursions by KRU in 2012
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
Rate-Distortion, Wyner-Ziv, and Gelfand-Pinsker

- Rate Distortion (RD) Problem
  - What is the **minimum data rate** to transmit a source with average distortion less than $D$?
Rate-Distortion, Wyner-Ziv, and Gelfand-Pinsker

- **Rate Distortion (RD) Problem**
  - What is the minimum data rate to transmit a source with average distortion less than $D$?

- **Wyner-Ziv (WZ) Problem**
  - WZ extends RD to the case of side-information at the decoder
Rate-Distortion, Wyner-Ziv, and Gelfand-Pinsker

- Rate Distortion (RD) Problem
  - What is the minimum data rate to transmit a source with average distortion less than $D$?

- Wyner-Ziv (WZ) Problem
  - WZ extends RD to the case of side-information at the decoder

- Gelfand-Pinsker (GP) Problem
  - Channel coding with non-causal side-information at transmitter
Rate-Distortion, Wyner-Ziv, and Gelfand-Pinsker

- Rate Distortion (RD) Problem
  - What is the minimum data rate to transmit a source with average distortion less than $D$?

- Wyner-Ziv (WZ) Problem
  - WZ extends RD to the case of side-information at the decoder

- Gelfand-Pinsker (GP) Problem
  - Channel coding with non-causal side-information at transmitter

- WZ and GP problems arise naturally in network information theory
Belief-Propagation Guided Decimation (BPGD)

- RD-type problems are **challenging** for graph codes with BP decoding
  - They require quantization of an *arbitrary sequence* to a codebook
  - BP converges only if received sequence is “close” to a codeword
  - But, vanishing fraction of total space is “close” to codewords

▶ When the received vector is not “close” to a codeword
▶ BP decoder typically converges to a non-informative fixed point
▶ There are exponentially many codewords with low distortion
▶ But, the decoder just cannot pick one
▶ The bias of a bit is defined to be $|\ln \frac{P(X=0)}{P(X=1)}|$
▶ To force convergence, bits are sequentially “decimated”:
  1. The BP decoder is run for a fixed number of iterations
  2. A bit with large bias is sampled and “decimated”
Belief-Propagation Guided Decimation (BPGD)

- RD-type problems are challenging for graph codes with BP decoding
  - They require quantization of an arbitrary sequence to a codebook
  - BP converges only if received sequence is “close” to a codeword
  - But, vanishing fraction of total space is “close” to codewords

- When the received vector is not “close” to a codeword
  - BP decoder typically converges to a non-informative fixed point
  - There are exponentially many codewords with low distortion
  - But, the decoder just cannot pick one
  - The bias of a bit is defined to be \( |LLR| = \left| \ln \frac{P(X=0)}{P(X=1)} \right| \)
Belief-Propagation Guided Decimation (BPGD)

- RD-type problems are challenging for graph codes with BP decoding
  - They require quantization of an arbitrary sequence to a codebook
  - BP converges only if received sequence is “close” to a codeword
  - But, vanishing fraction of total space is “close” to codewords

- When the received vector is not “close” to a codeword
  - BP decoder typically converges to a non-informative fixed point
  - There are exponentially many codewords with low distortion
  - But, the decoder just cannot pick one

  The bias of a bit is defined to be $|LLR| = \left| \ln \frac{P(X=0)}{P(X=1)} \right|$

- To force convergence, bits are sequentially “decimated”:
  1. The BP decoder is run for a fixed number of iterations
  2. A bit with large bias is sampled and “decimated”
Once Again, Spatial-Coupling Comes to the Rescue

Rate Distortion
- SC low-density generator matrix (LDGM) codes can approach the RD limit with BPGD [AMUV12]

Wyner-Ziv and Gelfand-Pinsker
- SC compound LDGM/LDPC codes can approach the WZ/GP limits with BPGD [KVNP14]
Outline

Review of LDPC Codes and Density Evolution

Spatially-Coupled Graphical Models

Universality for Multiuser Scenarios

Abstract Formulation of Threshold Saturation

Wyner-Ziv and Gelfand-Pinsker

Conclusions
Summary and Open Problems

- **Spatial Coupling**
  - Powerful technique for designing and understanding factor graphs
  - Related to the statistical physics of supercooled liquids
  - General proof of threshold saturation for scalar recursions

- **Interesting Open Problems**
  - Clever constructions to reduce the rate-loss due to termination
  - Finding new problems where SC gives real benefits
  - Proving SC codes with decimation achieve the rate-distortion limit
Thanks for your attention