Spatially-Coupled Codes for Side-Information Problems

Santhosh Kumar
Avinash Vem
Krishna Narayanan
Henry Pfister

June 30, 2014

Lossy Source Coding Problem

$X^n = (X_1, \cdots, X_n)$, $X_i \sim \text{Bernoulli}(\frac{1}{2})$

Binary code $\mathcal{C} = (n, k)$, rate $R = k/n$

Lossy Source Coding Problem

$$X^n = (X_1, \cdots, X_n),\ X_i \sim \text{Bernoulli}(\tfrac{1}{2})$$

Binary code $\mathcal{C} = (n, k)$, rate $R = k/n$

Lossy Source Coding

- Compress X^n to $\hat{X}^n \in \mathcal{C}$
- Min. Hamming distortion

$$D = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left|X_i - \hat{X}_i\right|$$

Lossy Source Coding Problem

$$X^n = (X_1, \cdots, X_n),\ X_i \sim \text{Bernoulli}(\tfrac{1}{2})$$

$$\text{Binary code } \mathcal{C} = (n, k),\ \text{rate } R = k/n$$

Lossy Source Coding

- Compress X^n to $\hat{X}^n \in \mathcal{C}$
- Min. Hamming distortion

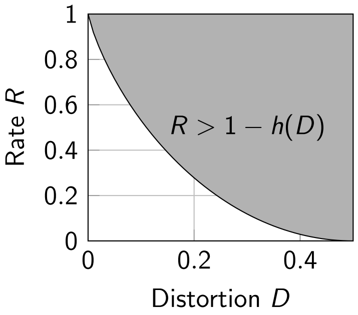
$$D = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} \left| X_i - \hat{X}_i \right|$$

- Rate-Distortion theory:

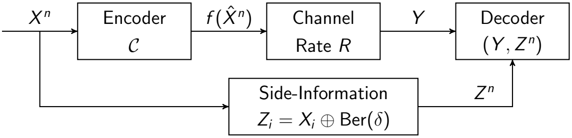
$$R > 1 - h(D)$$

- $h(\cdot)$ is binary entropy function

$$h(D) = -D \log_2 D - (1-D) \log_2 (1-D)$$



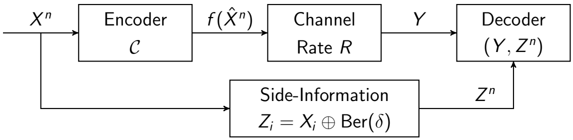
Side-Information Problems: Wyner-Ziv



Wyner-Ziv Formulation

- Side-information Z^n about X^n
- Decoder additionally has Z^n
- Say $Z_i = X_i \oplus \text{Ber}(\delta)$

Side-Information Problems: Wyner-Ziv

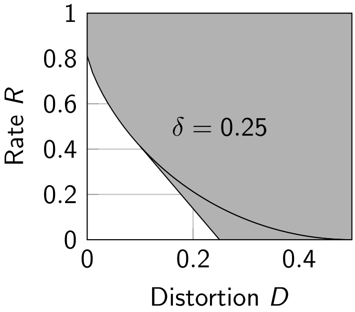


Wyner-Ziv Formulation

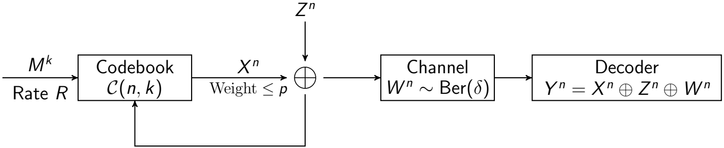
- Side-information Z^n about X^n
- Decoder additionally has Z^n
- Say $Z_i = X_i \oplus \text{Ber}(\delta)$
- Wyner-Ziv theory:

$$R > l.c.e\{h(D * \delta) - h(D), (\delta, 0)\}$$

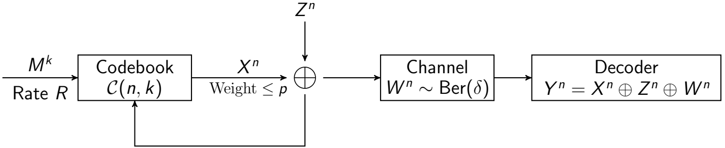
- $D * \delta = D(1 - \delta) + \delta(1 - D)$



Side-Information Problems: Gelfand-Pinsker



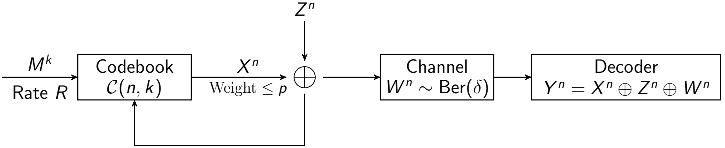
Side-Information Problems: Gelfand-Pinsker



Gelfand-Pinsker Formulation

- Message M^k encoded to $X^n \in \mathcal{C}$ with $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] \le p \le \frac{1}{2}$
- Side-information Z^n is available only at the encoder

Side-Information Problems: Gelfand-Pinsker

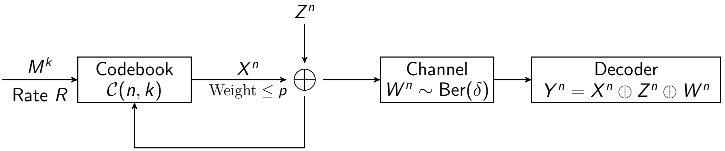


Gelfand-Pinsker Formulation

- Message M^k encoded to $X^n \in \mathcal{C}$ with $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] \leq p \leq \frac{1}{2}$
- Side-information Z^n is available only at the encoder
- The output at the decoder is

$$Y^n = X^n \oplus Z^n \oplus W^n, \quad \{W_i\} \sim \text{Ber}(\delta)$$

Side-Information Problems: Gelfand-Pinsker



Gelfand-Pinsker Formulation

- Message M^k encoded to $X^n \in \mathcal{C}$ with $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] \leq p \leq \frac{1}{2}$
- Side-information Z^n is available only at the encoder
- The output at the decoder is

$$Y^n = X^n \oplus Z^n \oplus W^n, \quad \{W_i\} \sim \mathrm{Ber}(\delta)$$

- Capacity region by Gelfand-Pinsker:

$$R < h(p) - h(\delta)$$

Main Result

Objective

- Construct low-complexity coding schemes that achieve the complete rate regions of Wyner-Ziv and Gelfand-Pinsker
  - Low-complexity encoding and decoding

Main Result

Objective

- Construct low-complexity coding schemes that achieve the complete rate regions of Wyner-Ziv and Gelfand-Pinsker
  - Low-complexity encoding and decoding

Idea

- Wainwright et al. used compound LDGM/LDPC codes with optimal encoding/decoding
- Message-passing algorithms have non-negligible gap

Main Result

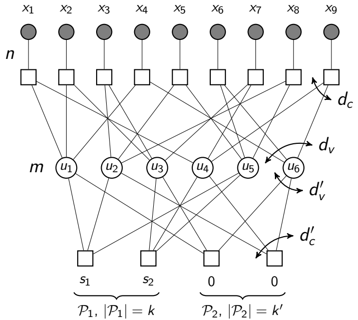
Objective

- Construct low-complexity coding schemes that achieve the complete rate regions of Wyner-Ziv and Gelfand-Pinsker
  - Low-complexity encoding and decoding

Idea

- Wainwright et al. used compound LDGM/LDPC codes with optimal encoding/decoding
- Message-passing algorithms have non-negligible gap
- Remedy via Spatial-Coupling
  - Channel coding in coupled compound codes (Kasai et al.)
  - Lossy source coding with spatially-coupled LDGM (Aref et al.)
  - Encoding with compound codes has additional challenges

Compound LDGM/LDPC Codes



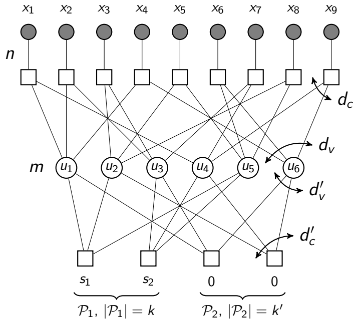
- Codebook $\mathcal{C}(n, m-k-k')$
- Message constraints

  $$u_1 \oplus u_2 \oplus u_5 = s_1, \quad u_1 \oplus u_3 \oplus u_6 = 0$$

- Codeword $(x_1, \cdots, x_9)$:

  $$x_1 = u_1 \oplus u_4, \qquad x_2 = \cdots$$

Compound LDGM/LDPC Codes



- Codebook $\mathcal{C}(n, m-k-k')$
- Message constraints

$$u_1 \oplus u_2 \oplus u_5 = s_1, \quad u_1 \oplus u_3 \oplus u_6 = 0$$

- Codeword $(x_1, \cdots, x_9)$:

$$x_1 = u_1 \oplus u_4, \qquad x_2 = \cdots$$

Key Properties

- Compound code is
  - a good source code under optimal encoding
  - a good channel code under optimal decoding

Good Code

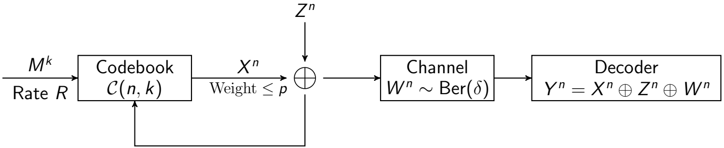
"Good" source code

- Rate of the code is $R = 1 - h(D) + \varepsilon$
- When this code is used to optimally encode $\mathrm{Ber}(\frac{1}{2})$
- The average Hamming distortion is at most D

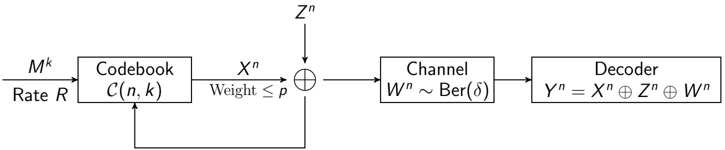
"Good" channel code

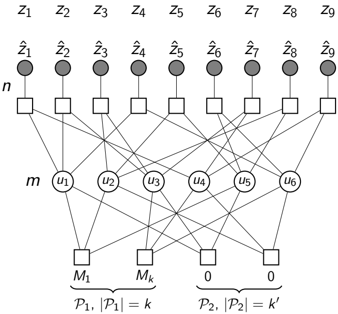
- Rate of the code is $R = 1 - h(\delta) - \varepsilon$
- When this code is used for channel coding on $\mathrm{BSC}(\delta)$
- Message est. under optimal decoding with error at most ε

Coding Scheme: Gelfand-Pinsker



Coding Scheme: Gelfand-Pinsker

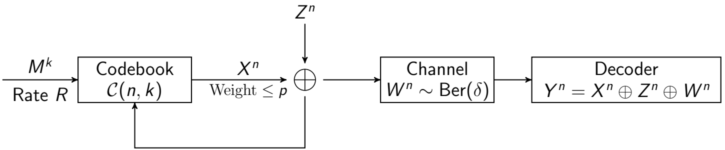


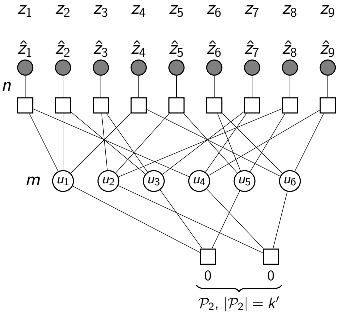


$\frac{m-k-k'}{n} \approx 1 - h(p) + \varepsilon \qquad \frac{m-k'}{n} \approx 1 - h(\delta) + \varepsilon$

- With message M^k, encode Z^n to $\hat{Z}^n$ (Distortion $\approx p$)
- Transmit $X^n = Z^n \oplus \hat{Z}^n$

Coding Scheme: Gelfand-Pinsker





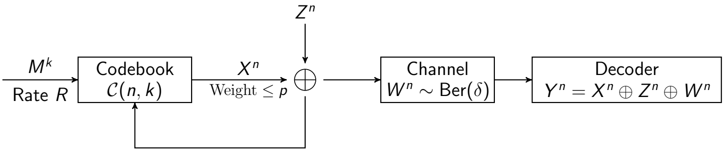
$$\frac{m-k'}{n} \approx 1 - h(\delta) + \varepsilon$$

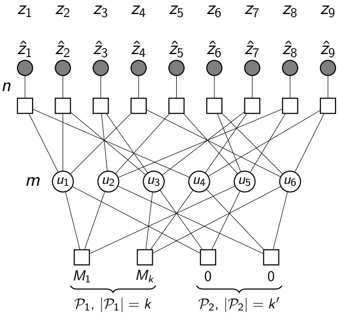
- With message M^k, encode Z^n to $\hat{Z}^n$ (Distortion $\approx p$)
- Transmit $X^n = Z^n \oplus \hat{Z}^n$
- Decoder has

$$Y^n = X^n \oplus Z^n \oplus W^n$$
$$= \hat{Z}^n \oplus W^n$$

- Decode $\hat{Z}^n$ and compute M^k

Coding Scheme: Gelfand-Pinsker





$\frac{m-k-k'}{n} \approx 1 - h(p) + \varepsilon \qquad \frac{m-k'}{n} \approx 1 - h(\delta) + \varepsilon$

- With message M^k, encode Z^n to $\hat{Z}^n$ (Distortion $\approx p$)
- Transmit $X^n = Z^n \oplus \hat{Z}^n$
- Decoder has

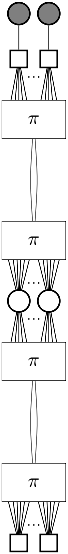
$$\begin{aligned} Y^n &= X^n \oplus Z^n \oplus W^n \\ &= \hat{Z}^n \oplus W^n \end{aligned}$$

- Decode $\hat{Z}^n$ and compute M^k
- $R = \frac{k}{n} \approx h(p) - h(\delta)$

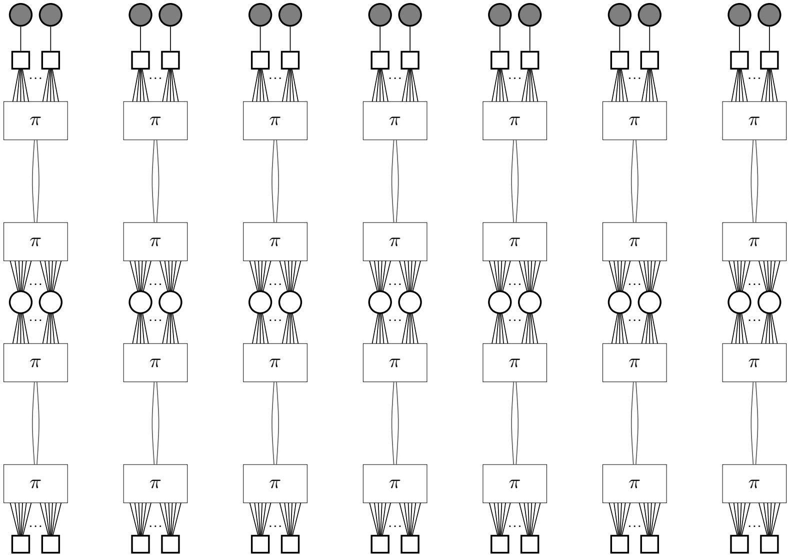
Remarks

- Need codes that are simultaneously good for channel and source coding
- Use message-passing algorithms instead of optimal
- Use spatial-coupling for goodness of codes under message-passing

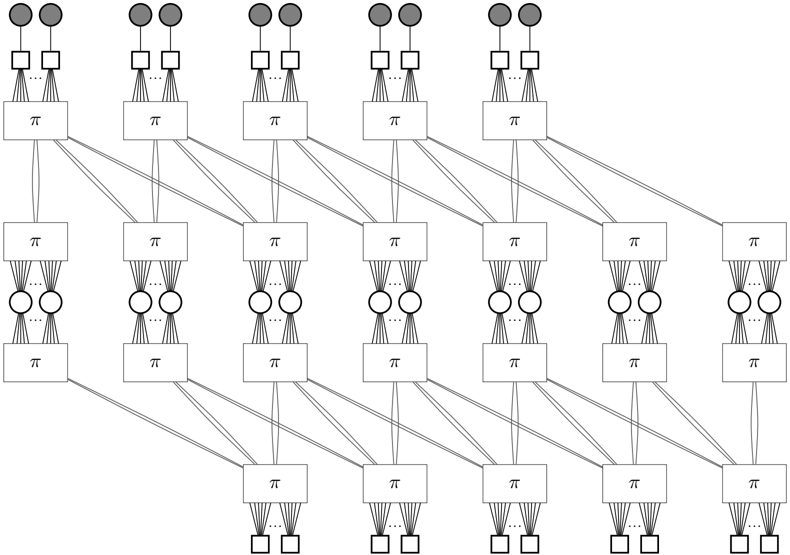
Spatially-Coupled Compound LDGM/LDPC Codes



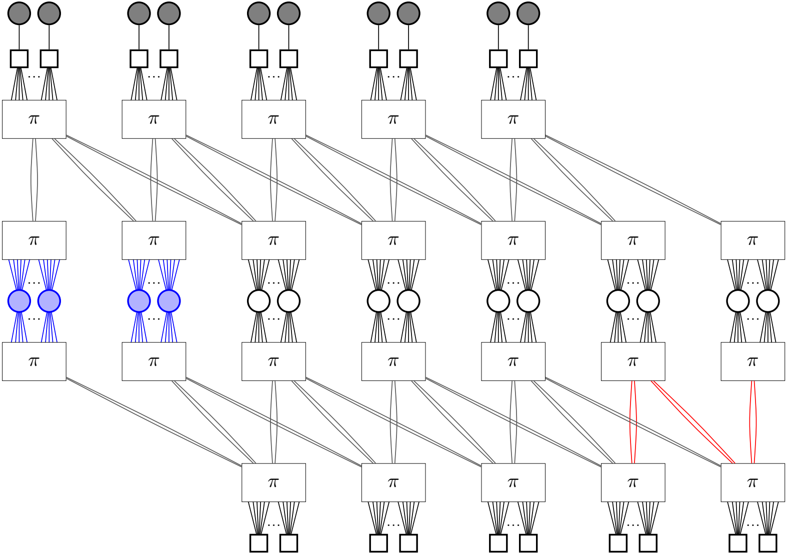
Spatially-Coupled Compound LDGM/LDPC Codes



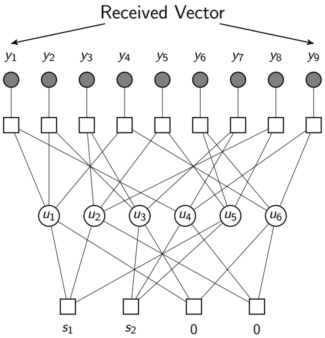
Spatially-Coupled Compound LDGM/LDPC Codes

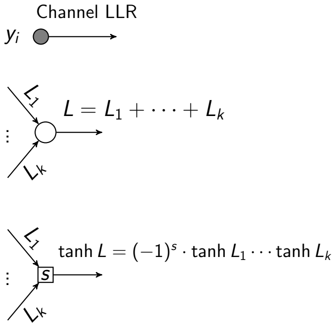


Spatially-Coupled Compound LDGM/LDPC Codes



Decoding in Spatially-Coupled Compound Codes

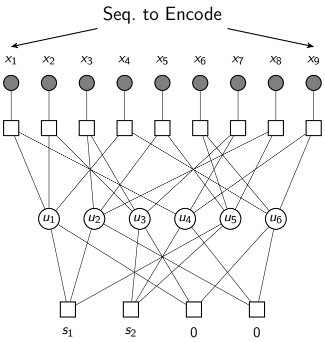


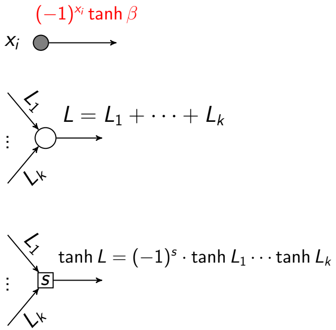


Remarks

- Standard message-passing algorithm
- Threshold saturation proven for SC compound codes on BEC
- Empirically observed for BMS channels

Encoding in Spatially-Coupled Compound Codes





Remarks

- Inverse temperature parameter β
- Message-passing rules are the same
- However, a crucial decimation step is needed

Encoding in SC Compound Codes: BPGD Algorithm

while There are active LDPC bit-nodes **do**
  for $t = 1$ to T **do**
    Run the BP equations
  end for
  Evaluate LLRs m_i for each LDPC bit-node
  Choose max. of $|m_i|$ in left-most w active sections
  if $|m_{i^*}| = 0$ **then**
    Set u_{i^*} to 0 or 1 uniformly randomly
  else
    Set u_{i^*} to 0 or 1 with prob. $\frac{1+\tanh m_{i^*}}{2}$ or $\frac{1-\tanh m_{i^*}}{2}$
  end if
  Decimate (remove) LDPC bit-node i^* and update parities
end while
If $\{u_i\}$ fail to satisfy LDPC checks, then **re-encode**

Encoding in SC Compound Codes: Remarks

- Randomization in setting u_{i^*} is crucial
- BPGD applied to uncoupled code always failed
- Spatially-coupled structure is crucial for successful encoding
  - In addition, distortion is close to optimal thresholds
  - Does not encode if decimated from both left and right
  - Does not encode if both left and right boundary is set to 0

Encoding in SC Compound Codes: Numerical Example

Block length (n)	4-cycles	Attempts $1/2/3/4/\geq 5$
9000	yes	5/3/5/2/35
9000	no	21/12/5/3/9
27000	no	35/15/0/0/0
45000	no	40/9/0/0/1
63000	no	44/6/0/0/0
81000	no	50/0/0/0/0

Remarks

- \# Attempts to encode 50 seq. in $(6,3)$ LDGM / $(3,6)$ LDPC
- $L = 20,\ w = 4,\ \beta = 0.65,\ T = 10$
- Removing 4-cycles dramatically improves success
- How much do 6-cycles matter?

Numerical Results: Wyner-Ziv

LDGM (d_v, d_c)	LDPC (d'_v, d'_c)	(L, w)	(D_*, δ_*)	(D, δ)
$(6, 3)$	(3,6)	(20,4)	(0.111,0.134)	(0.1174,0.122)
$(8, 4)$	(3,6)	(20,4)	(0.111,0.134)	(0.1149,0.120)
$(10, 5)$	(3,6)	(20,4)	(0.111,0.134)	(0.1139,0.122)

Remarks

- D_* and δ_* are calculated based on the rate of the respective code:

$$D_* = h^{-1}(1 - R1) \qquad \delta_* = h^{-1}(1 - R2)$$

- $n \approx 140000$, $\beta = 1.04$, $T = 10$

Numerical Results: Gelfand-Pinsker

LDGM (d_v, d_c)	LDPC (d'_v, d'_c)	(L, w)	(p_*, δ_*)	(p, δ)
$(6, 3)$	(3,6)	(20,4)	(0.215,0.157)	(0.2200,0.152)
$(8, 4)$	(3,6)	(20,4)	(0.215,0.157)	(0.2230,0.151)
$(10, 5)$	(3,6)	(20,4)	(0.215,0.157)	(0.2200,0.151)

Remarks

- p_* and δ_* are calculated based on the rate of the respective code:

$$p_* = h^{-1}(1 - R1) \qquad \delta_* = h^{-1}(1 - R2)$$

- $n \approx 140000,\ \beta = 0.65,\ T = 10$

Concluding Remarks

Conclusion

- Spatially-coupled codes achieve the rate regions of Wyner-Ziv and Gelfand-Pinsker problems
- Coupling structure is also crucial
  - to achieve optimum thresholds
  - for encoding to succeed with decimation

Open Questions

- Effect of degree profiles, short-cycles on encoding success
- Precise trade-offs with polar codes