ECEN 655: Advanced Channel Coding

Course Introduction

Henry D. Pfister

Department of Electrical and Computer Engineering
Texas A&M University

Outline

A Punctured History of ECC: In the Beginning

- 1945: Hamming experiments with ECC for computers
- 1948: Claude Shannon publishes treatise on information theory
- 1949: Golay publishes the binary Golay code
- 1950: Hamming publishes seminal Hamming parity-check code
- 1955: Elias publishes first paper on convolutional codes
- 1960: Reed-Solomon code paper and Peterson BCH decoding paper
- 1960: Gallager introduces low-density parity-check (LDPC) codes and iterative decoding in his PhD thesis

A Punctured History of ECC: Big Breakthroughs

- 1993: Berrou et al. revolutionize coding with turbo codes
- 1995: MacKay and Neal rediscover LDPC codes
- 1997: LMSSS approach capacity with irregular LDPC codes
- 2000: DVB-RCS and 3GPP standards are first to include turbo codes
- 2001: Zigangirov et al. introduce LDPC convolutional codes
- 2001: RU introduce density evolution to optimize LDPC codes
- 2002: Luby's rateless fountain codes achieve capacity on the BEC
- 2004: DVB-S2 standard is the first to include LDPC codes
- 200x: Optimized LDPC codes solve most coding problems

A Punctured History of ECC: Recent Advances

- 2009: Arikan's polar codes deterministically achieve capacity
- 2011: Kudekar, Richardson, and Urbanke discover threshold saturation and show LDPC convolutional codes achieve capacity
- 2012: Variants of LDPC convolutional codes (e.g., braided codes) used in practice for optical communication

Note: Many important advances (especially in algebraic coding) are neglected due to our focus on modern capacity approaching codes.

Applications of Error-Correcting Codes

- Many devices make use of error-correcting codes:
  - Compact Discs and DVDs
  - Cell Phones
  - Hard Disk Drives
  - The Internet
  - Flash Memory and RAM in your computer
  - Microprocessor Bus Connections
  - DNA Microarrays

Main Elements of this Course

Inference Problems:

- point-to-point communication
- multiple-access communication
- best assignment w/constraints
- rate distortion

Coding Schemes:

- product codes
- turbo codes
- LDPC codes
- polar codes

Mathematical Tools:

- factor graphs: marginalization and message passing
- probability: martingales, concentration, and Markov fields
- combinatorics: generating functions and duality

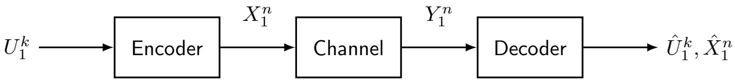
Goals of this Course (1)

1. Define maximum-likelihood (ML) decoding, maximum-a-posteriori (MAP) decoding, and a-posteriori-probability (APP) processing for inference problems.
2. Understand random codes (graphs) and code (graph) ensembles.
3. Find the factor graph for an inference problem and approximate its marginalization via message-passing on that factor graph.
4. Understand connections between: (i) the Gibb's free energy and APP processing, (ii) the Bethe free energy and sum-product processing.
5. Analyze the performance of message-passing decoding on the binary erasure channel (BEC) for both code ensembles and individual codes.

Goals of this Course (2)

6. Explain the gap between message-passing and MAP decoding and use EXIT functions to derive bounds on that gap on the BEC.
7. Analyze the performance of message-passing decoding for code ensembles on binary memoryless symmetric (BMS) channels.
8. Compute average weight enumerator and spectral shape of standard ensembles and use them to bound ML decoding thresholds.
9. Identify communications and signal processing problems where message-passing can be used implement detection and estimation.

Standard Point-to-Point Communication



- Encoder: Maps data $U_1^k \in \{0,1\}^k$ to codeword $X_1^n \in \mathcal{C} \subset \mathcal{X}^n$
- Channel: Randomly maps $X_1^n \in \mathcal{X}^n$ to $Y_1^n \in \mathcal{Y}^n$ (i.i.d. $\sim p(y|x)$)
- Decoder: Estimates information sequence $\hat{U}_1^k$ and codeword $\hat{X}_1^n$

- Information Theory: Shannon's Channel Coding Theorem
  - An *information rate* $R = k/n$ (bits/channel use) is achievable iff $R < I(X;Y)$, where $I(X;Y)$ is the *mutual information*
  - *Capacity* C is the maximum of $I(X;Y)$ over the input dist. $p(x)$
  - Proof based on using a random code from a suitable ensemble

Binary Memoryless Symmetric (BMS) Channels

For BMS Channels, capacity is achieved by:

- Uniform input distribution $(\Pr(X=0)=\Pr(X=1)=1/2)$
- Uniform random codes with maximum-likelihood (ML) decoding

Consider a BSC(p): (i.e., binary symmetric channel with error rate p)

- Capacity is $C = 1 - h(p)$, where $h(p) \triangleq p \log \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$
- Hamming ball: $B(y_1^n, m) = \{z_1^n \in \{0,1\}^n \mid d_H(y_1^n, z_1^n) \le m\}$
- For any $\epsilon > 0$, we find $\delta_n = \Pr\left(X_1^n \notin B(Y_1^n, (p+\epsilon)n)\right) \to 0$ by LLN
  - pn errors expected and prob. of $> (p+\epsilon)n$ errors vanishes as $n \to \infty$

Random Coding for the BSC (1)

Consider a random code where:

- For $i = 0, 1, \ldots, 2^k - 1$, ith codeword $X_1^n(i)$ is i.i.d. Bernoulli($\frac{1}{2}$)
- Codeword $X_1^n(j)$ is transmitted
- Decoder lists all codewords in ball $B(Y_1^n, (p + \epsilon)n)$ around received
    - Returns a codeword if exactly one codeword in ball
    - Declares failure otherwise
- A union bound on $P_e(j)$, the decoder error probability, gives

$$P_e(j) \leq \delta_n + \sum_{i=0, i \neq j}^{2^k} \Pr\left(X_1^n(i) \in B(Y_1^n, (p + \epsilon)n)\right)$$

Random Coding for the BSC (2)

- Assume $p+\epsilon \le \frac{1}{2}$. Since $X_1^n(i)$ is independent of Y_1^n for $i \neq j$,

$$\begin{aligned}\Pr\left(X_1^n(i) \in B(Y_1^n,(p+\epsilon)n)\right) &\le \frac{1}{2^n}\left|B(Y_1^n,(p+\epsilon)n)\right| \\ &= \frac{1}{2^n}\sum_{i=0}^{\lfloor (p+\epsilon)n\rfloor}\binom{n}{i} \le 2^{n[h(p+\epsilon)-1]}\end{aligned}$$

Random Coding for the BSC (2)

- Assume $p + \epsilon \leq \frac{1}{2}$. Since $X_1^n(i)$ is independent of Y_1^n for $i \neq j$,

$$\Pr\left(X_1^n(i) \in B(Y_1^n, (p+\epsilon)n)\right) \leq \frac{1}{2^n} \left|B(Y_1^n, (p+\epsilon)n)\right| = \frac{1}{2^n} \sum_{i=0}^{\lfloor (p+\epsilon)n \rfloor} \binom{n}{i} \leq 2^{n[h(p+\epsilon)-1]}$$

- If $R < C = 1 - h(p)$, then $R + h(p+\epsilon) - 1 < 0$ for some $\epsilon > 0$ and

$$P_e(j) \leq \delta_n + \sum_{i=0, i\neq j}^{2^k} \Pr\left(X_1^n(i) \in B(Y_1^n, (p+\epsilon)n)\right)$$

$$\leq \delta_n + 2^k 2^{n[h(p+\epsilon)-1]} = \delta_n + 2^{n \overbrace{[R + h(p+\epsilon) - 1]}^{<0}} \to 0$$

Random Linear Codes

- Also holds for codes defined by random $k \times n$ generator matrix
  - Only problem is that codewords are no longer independent r.v.
    - Given 2 codewords, distribution on a 3rd changes due to linearity
  - Need only to argue more carefully that, for $i \neq j$,

$$\Pr\left(X_1^n(i) \in B(Y_1^n, (p+\epsilon)n)\right) \leq \frac{1}{2^n} \left|B(Y_1^n, (p+\epsilon)n)\right|$$

- A few facts about random linear codes:
  - For fixed code, symmetry implies error prob. independent of j
  - Since $\mathbf{0} \in \mathcal{C}$, we order codewords so $X_1^n(0) = \mathbf{0}$ and choose $j = 0$
  - Then, $\Pr(X_1^n(i) = x_1^n \mid X_1^n(0) = \mathbf{0}) = \Pr(X_1^n(i) = x_1^n)$ and:
    - For $i \neq 0$, $X_1^n(i)$ is independent of both $X_1^n(0)$ and Y_1^n
    - $\Pr\left(X_1^n(i) \in B(Y_1^n, (p+\epsilon)n)\right) \leq \frac{1}{2^n}\left|B(Y_1^n, (p+\epsilon)n)\right|$

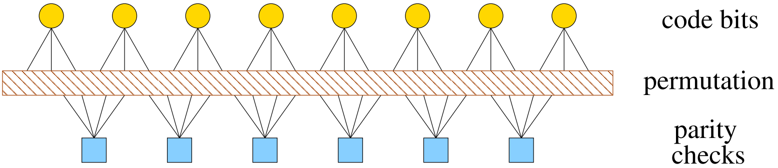
The 50 Year Challenge

- For general linear codes: storage and encoding is tractable
  - Requires nk bits for storage and nk boolean operations
- NP decision problems require "Yes" be verified in polynomial time
  - "Is there a codeword z_1^n s.t. $d_H(z_1^n, y_1^n) \leq e$?" is NP-complete!
- For the generator matrix G, ML decoding is the inference problem

$$\hat{x}_1^n = \arg \max_{x_1^n \in \{\mathbf{u}G \mid \mathbf{u} \in \{0,1\}^k\}} \Pr\left(Y_1^n = y_1^n \mid X_1^n = x_1^n\right)$$

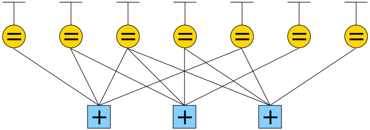
- Q: Is there a code structure that makes decoding tractable?

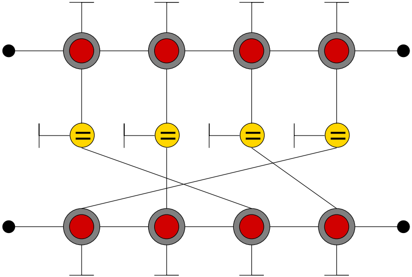
Low-Density Parity-Check (LDPC) Codes



- Linear codes defined by $xH^T = 0$ for all c.w. $x \in \mathcal{C}$
  - H is an $r \times n$ sparse parity-check matrix for the code
  - Ensembles defined by bit/check degrees and rand. perm.
- Bipartite Tanner graph
  - Bit (check) nodes associated with columns (rows) of H
  - Each check is attached to all bits that must satisfy the check

Sparse Graph Codes





- Codeword constraints defined via sparse factor graph
  - factor nodes define the constraints
  - variable nodes define the variables
  - half-edges represent observations (or degree-1 factor nodes)
- Three typical constraints
  - Equality (=): Edges are bits that must have the same value
  - Parity (+): Edges are bits that must sum to zero (mod 2)
  - Trellis: Bit edges must be compatible with state edges

Some Recent History

- Turbo Codes
  - Introduced in 1993 by Berrou, Glavieux, and Thitimajshima
  - Revolutionized coding theory with performance
  - McEliece et al.: turbo decoding = belief propagation (1998)

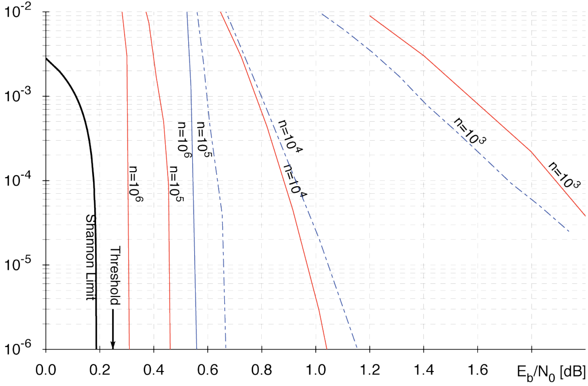
Some Recent History

- Turno Codes
    - Introduced in 1993 by Berrou, Glavieux, and Thitimajshima
    - Revolutionized coding theory with performance
    - McEliece et al.: turbo decoding = belief propagation (1998)
- Low Density Parity Check (LDPC) Codes
    - Introduced in 1960 by Gallager and then forgotten
    - Re-discovered by MacKay in 1995
    - Irregular LDPC achieves capacity on BEC (1997)
    - Density evolution for AWGN: 0.0045 dB from cap. (2001)

Some Recent History

- Turbo Codes
  - Introduced in 1993 by Berrou, Glavieux, and Thitimajshima
  - Revolutionized coding theory with performance
  - McEliece et al.: turbo decoding = belief propagation (1998)
- Low Density Parity Check (LDPC) Codes
  - Introduced in 1960 by Gallager and then forgotten
  - Re-discovered by MacKay in 1995
  - Irregular LDPC achieves capacity on BEC (1997)
  - Density evolution for AWGN: 0.0045 dB from cap. (2001)
- Sparse Graph Codes
  - Natural generalization that encompasses many code families
  - Low-complexity iterative decoding has outstanding performance

Turbo vs. LDPC Performance



- BER: Standard Turbo (blue) vs Irregular LDPC (red)
- From "The Capacity of LDPC Codes Under Message Passing Decoding", Richardson & Urbanke, Trans. IT 2001