On the Joint Decoding of LDPC Codes and Finite-State Channels via Linear Programming

Byung-Hak Kim
(joint with Henry D. Pfister)

Texas A&M University
College Station

International Symposium on Information Theory
Austin, TX
June 14, 2010

Motivation: Feldman's Open Question (2003)

"...In practice, channels are generally not memoryless due to physical effects in the communication channel.

Even coming up with a proper linear cost function for an LP to use in these channels is an interesting question.

The notions of pseudocodeword and fractional distance would also need to be reconsidered for this setting..."

- Feldman '03

Previous Approaches for Channels with Memory (1)

- Graph-Based Decoding on ISI by Taghavi and Siegel (ISIT07, arXiv07)
    - Linear programming decoding of channel and joint decoding with an LDPC code
    - Suboptimal binary representation of channel
- Interior Point Decoding for Linear Vector Channel by Wadayama (ISIT08, IT10)
    - Relaxed ML decoding based on quadratic programming
    - Limited to linear channel and no ML certificate property

Previous Approaches for Channels with Memory (2)

- Linear-Programming Receivers by Flanagan (Allerton08, arXiv09)
  - This work predates ours and is more general
  - But, we discovered it only just before submission
  - Our description, based on the trellis of the channel, is conceptually simpler, but mathematically identical to his “Efficient LP Relaxation”.

Main Differences

1. Our main application focuses on predicting joint-decoding performance at high SNR from simulations at low SNR
2. Our simulation results highlight the fact that of LP joint decoding is sometimes superior to joint iterative decoding (e.g., at high SNR)

Main Goal: Analysis of Joint Iterative Decoding

Memoryless Channel

- Iterative Decoding
- Linear Programming Decoding
  - ML Certificate
  - Pseudo-codewords
  - Fractional Distance
- Finite-Length Analysis

Finite-State Channel

- Joint Iterative Decoding
- Linear Programming Joint-Decoding?
  - ML Certificate?
  - Pseudo-codewords?
  - Generalized Distance?
- New Insight?

Outline

1. Introduction: LP Decoding and the Finite-State Channel
2. Main Result 1: LP Joint-Decoding and Pseudo-Codewords
3. Main Result 2: Decoder Performance and Error-Rate Estimation
4. Conclusion and Open Problems

LP Decoding for Memoryless Channels (1)

- ML decoding can be written as a linear program
- For a code $\mathcal{C}$, let $\mathrm{CH}(\mathcal{C})$ be the convex hull of all codewords
- A binary code $C_1, \ldots, C_n$ sent through channel, $Y_1, \ldots, Y_n$ received
- ML decoding: minimize negative log likelihood

$$\gamma_i \triangleq \log\left(\frac{\Pr(Y_i = y_i \,|\, C_i = 0)}{\Pr(Y_i = y_i \,|\, C_i = 1)}\right)$$

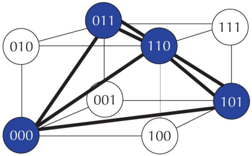
- ML decoding can be written as

$$\min \gamma^T \mathbf{x}$$
$$\mathbf{x} \in \mathrm{CH}(\mathcal{C})$$

LP Decoding for Memoryless Channels (2)

- Unfortunately, CH$(\mathcal{C})$ cannot be described efficiently
- However a way to approximate: Relax the polytope $\mathcal{P}(H)$
- How to relax? Intersection of local codeword polytopes defined by local parity constraints
- LP decoding can be written as

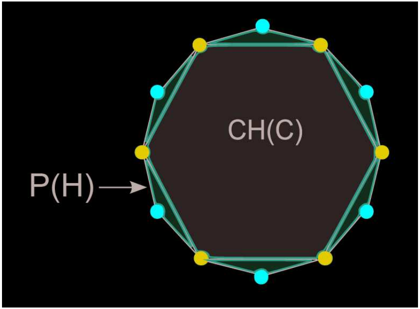
$$\min \gamma^T \mathbf{x}$$
$$\mathbf{x} \in \mathcal{P}(H)$$



- Integer solutions provide an ML certificate

Pseudo-codewords

- Possible outputs of LP decoder



1. Codeword with ML certificate (e.g., $[0 \;\; 1 \;\; 0 \;\; 1 \;\; 0 \;\; 1]$)
2. Fractional vertex (e.g., $[1 \;\; \frac{1}{2} \;\; 1 \;\; \frac{1}{2} \;\; 1 \;\; \frac{1}{2}]$)

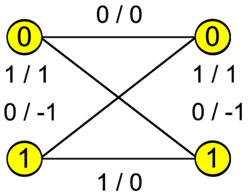
Finite-State Channel (FSC)

Qualitative Definition

- A FSC is a model for communication systems with memory
- Each output depends only on the current input and channel state instead of the entire past

FS Intersymbol Interference (FSISI) Channel

- A FSC whose next state is a deterministic function of the current state and input (Edges labelled by input/output)
- Dicode channel

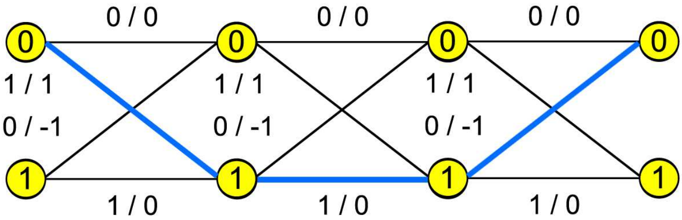


LP Joint-Decoding (1)

- Consider the ML joint-decoding of the optimum edge path
- For each edge e, let $g(e) \in \{0, 1\}$ be the indicator function of the event that the ML edge path traverses that edge

Example

Consider a codeword $[1\,1\,0]$, $\mathbf{g} = [0\,1\,0\,0;\ 0\,0\,0\,1;\ 0\,0\,1\,0]$



LP Joint-Decoding (2)

- ML joint-decoding: minimize negative log likelihood

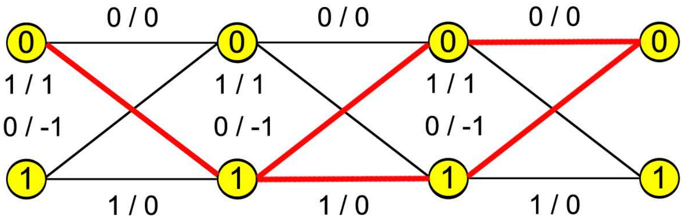
$$b(e) \triangleq -\log \Pr\left(\text{observation}|\text{edge traversed}\right) + \text{initial state term}$$

- Let $\mathcal{P}_{\mathcal{T}}(H)$ be the polytope of all valid trellis paths satisfying local codeword constraints on trellis by relaxing $g(e) \in [0,1]$
- The LP joint-decoding can be written as

$$\min \sum_{e \in \mathcal{E}} g(e) b(e)$$
$$\mathbf{g} \in \mathcal{P}_{\mathcal{T}}(H)$$

Joint-Decoding Pseudo-codewords

- Fractional outputs of LP joint-decoder



- Trellis-wise Pseudo-codeword: $\mathbf{g} = \left[0\,1\,0\,0;\; 0\,0\,\frac{1}{2}\,\frac{1}{2};\; \frac{1}{2}\,0\,\frac{1}{2}\,0\right]$
- Symbol-wise Pseudo-codeword: $\left[1,\, \frac{1}{2},\, \frac{1}{2}\right]$

ML Certificate Property

Theorem

The LP joint-decoder computes

$$\arg \min_{\mathbf{g} \in \mathcal{P}_{\mathcal{T}}(H)} \sum_{e \in \mathcal{E}} g(e) b(e)$$

and outputs a joint ML edge-path if $\mathbf{g}$ *is integral.*

Remark

- If the channel is a FSISI channel, the LP joint decoder outputs a joint ML codeword when $\mathbf{g}$ is integral
- Otherwise, integer valued solutions are not necessarily ML codeword (e.g., finite-state fading channel)

Performance Analysis

Theorem

Assume FSISI channel with AWGN, then, the pairwise error probability between a codeword $\mathbf{c}$ *and a pseudo-codeword* $\mathbf{p}$ *is*

$$Pr\left(\mathbf{c} \to \mathbf{p}\right) = Q\left(\frac{d_{gen}\left(\mathbf{c},\, \mathbf{p}\right)}{2\sigma}\right),$$

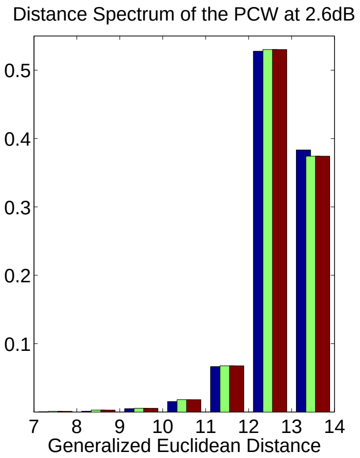
where $d_{gen}\left(\mathbf{c},\, \mathbf{p}\right)$ *is the Generalized Euclidean distance from* $\mathbf{c}$ *to* $\mathbf{p}$*.*

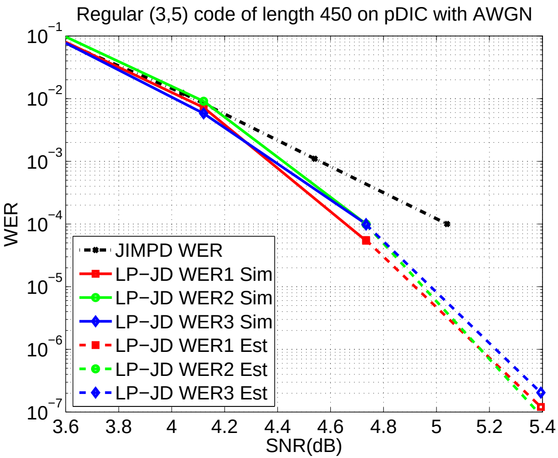
Union Bound on Word Error Rate

$$P_{w|\mathbf{c}} \leq \sum_{d_{gen} \in \mathcal{G}(\mathbf{c})} K_{d_{gen}}(\mathbf{c})\, Q\left(\frac{d_{gen}}{2\sigma}\right),$$

where: $K_{d_{gen}}(\mathbf{c})$ is the number of PCW $\mathbf{p}$ at distance d_{gen} from $\mathbf{c}$
$\mathcal{G}(\mathbf{c})$ is the set of $d_{gen}\left(\mathbf{c},\, \mathbf{p}\right)$

Error Rate Results and Predictions





Conclusion

Main Contributions

1. Comparison of linear-programing joint decoding with the joint iterative decoding for LDPC codes on FSCs
2. Simulation-based semi-analytic method for estimating the error rate of LDPC codes on FSISI channels at high SNR

Avenues for Future Research

1. Introduce a low-complexity Message-Passing Solver for the LP joint decoder
2. Derive the asymptotic noise threshold of the LP joint decoding using modified forms of density evolution

Thank you