
Understanding Pseudocodewords

Judy L. Walker

`jwalker@math.unl.edu`

University of Nebraska

Overview

- General Case: linear programming (LP) decoding
- Pseudocodewords in LP decoding
- LDPC codes
- Pseudocodewords in graph decoding of LDPC codes
- Pseudocodewords and the fundamental cone
- Algebraic description of pseudocodewords of cycle codes
- Back to the general LDPC case

Idea of LP Decoding

- Let $C \subseteq \mathbb{F}_2^n$ be determined by a parity check matrix H .
- H determines a *parity polytope* $\Omega(H)$.
- A received vector \tilde{y} determines a *cost function* γ .
- The goal of LP decoding is to minimize γ on $\Omega(H)$.

More precisely ...

The parity polytope

$\Omega(H)$ is the set of points $(x_1, \dots, x_n) \in \mathbb{R}^n$ satisfying:

- For each i , we have $0 \leq x_i \leq 1$.
- For each row \mathbf{r} of H and each subset V of

$$S(\mathbf{r}) := \{i \mid r_i = 1\}$$

such that $|V|$ is odd, we have

$$\sum_{i \in V} x_i \leq |V| - 1 + \sum_{j \in S(\mathbf{r}) \setminus V} x_j.$$

The cost function

Suppose $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_n)$ is received, and assume $\mathbf{y} = (y_1, \dots, y_n)$ was sent. Then γ is the *log-likelihood ratio*:

$$\gamma_i = \ln \left(\frac{\Pr(\tilde{y}_i | y_i = 0)}{\Pr(\tilde{y}_i | y_i = 1)} \right)$$

and we wish to minimize

$$\sum_{i=1}^n \gamma_i x_i$$

over all $\mathbf{x} = (x_1, \dots, x_n) \in \Omega(H)$.

Properties of $\Omega(H)$

Feldman showed:

- $\Omega(H)$ is *proper* for C , i.e.,

$$\Omega(H) \cap \{0, 1\}^n = C.$$

Hence, if LP decoding outputs a vector that *looks like* a codeword, then it *is* a codeword and it is the *ML* codeword.

- $\Omega(H)$ is *C -symmetric*.

Hence, the probability of error in LP decoding is independent of what codeword was transmitted, and so we may assume **0** was transmitted.

Observations

- The output of the LP decoder is the point $\mathbf{x} = (x_1, \dots, x_n) \in \Omega(H)$ where the minimum value of the cost function $\sum \gamma_i x_i$ occurs.
- This point \mathbf{x} will be a vertex of $\Omega(H)$.
- The only way \mathbf{x} can be a codeword is if $\mathbf{x} = \mathbf{0}$.
- Otherwise, the coordinates of \mathbf{x} are rational and so there is some positive integer β such that $\beta\mathbf{x}$ is a vector of nonnegative integers.
- Since \mathbf{x} is a solution to the LP, the cost of \mathbf{x} is nonpositive. Hence the cost of $\beta\mathbf{x}$ is nonpositive.

Pseudocodewords (LP)

Definition: A *pseudocodeword* is an integer vector of the form $\beta \mathbf{x}$ where β is a positive integer and \mathbf{x} is a vertex of $\Omega(H)$.

Theorem: (Feldman) Assuming $\mathbf{0}$ was transmitted, the LP decoder fails *if and only if* there is a pseudocodeword having nonpositive cost.

The Fundamental Cone

Since we may assume that $\mathbf{0}$ was transmitted, we may remove constraints defining $\Omega(H)$ not optimized at $\mathbf{0}$. Recall:

$\Omega(H)$ is the set of points $(x_1, \dots, x_n) \in \mathbb{R}^n$ satisfying:

- For each i , we have $0 \leq x_i \leq 1$.
- For each row \mathbf{r} of H and each subset V of

$$S(\mathbf{r}) := \{i \mid r_i = 1\}$$

such that $|V|$ is odd, we have

$$\sum_{i \in V} x_i \leq |V| - 1 + \sum_{j \in S(\mathbf{r}) \setminus V} x_j.$$

So we have $x_i \geq 0$ for each i and those constraints with $|V| = 1$.

The Fundamental Cone (con't)

$\mathcal{K}(H)$ is the set of points $(x_1, \dots, x_n) \in \mathbb{R}^n$ satisfying:

- For each i , we have $x_i \geq 0$.
- For each row \mathbf{r} of H and each $i \in S(\mathbf{r})$, we have

$$x_i \leq \sum_{j \in S(\mathbf{r}) \setminus \{i\}} x_j.$$

This is a cone in \mathbb{R}^n called the *fundamental cone*.

Goal: Understand the points, especially the (integer multiples of) vertex points, in $\Omega(H)$ or $\mathcal{K}(H)$.

Tanner Graph

Let $H = (h_{ji})$ be an $r \times n$ matrix of 0's and 1's. To H , we associate the bipartite *Tanner graph* T :

- left vertices \leftrightarrow columns of H
 n “bit nodes” X_1, \dots, X_n
- right vertices \leftrightarrow rows of H
 r “check nodes” f_1, \dots, f_r
- edges \leftrightarrow 1's in H
 $\{X_i, f_j\}$ is an edge $\iff h_{ji} = 1$
 \iff the i^{th} bit is involved in the j^{th} check

Remark

Note that the Tanner graph T records the matrix H , and hence the code C , *graphically*:

- a binary assignment of the bit nodes (c_1, \dots, c_n) is a codeword in C

if and only if

- the binary sum of the values at the neighbors of each check node is zero.

LDPC Codes

If H (and hence T) is *sparse*, then C is called a *low density parity check (LDPC)* code.

LDPC codes come equipped with an *iterative message-passing decoding algorithm* which is *extremely efficient* and corrects, with high probability, many *more error patterns* than guaranteed by the minimum distance.

Graph Decoding

Received word \implies assignment of 0 or 1
(plus reliability) at each bit node.

Then:

- Bit nodes broadcast to check nodes.
- Check nodes make estimates based on what they receive.
- Check nodes broadcast back to bit nodes.
- Bit nodes make estimates based on what they receive.
- Repeat.

Local Issues

The decoding algorithm acts *locally*: at each stage, decisions are made at each vertex, based only on information coming from neighbors of that vertex.

Algorithm's strength: Speed

Algorithm's weakness: Non-optimality

The algorithm *cannot distinguish* between the original Tanner graph and any finite, unramified cover of the Tanner graph!

Codewords in covers

Every codeword in *every* code \tilde{C} corresponding to *every* cover \tilde{T} of the Tanner graph is competing with the codewords in C to be the best explanation of the received word.

Liftings: If $(c_1, \dots, c_n) \in C$, then

$$(c_{(1,1)} : \dots : c_{(1,M)}, \dots, c_{(n,1)} : \dots : c_{(n,M)})$$

is in every \tilde{C} .

Others: There are others too.

Pseudocodewords (LDPC)

Definition: Let

$$\tilde{\mathbf{c}} = (c_{(1,1)} : \dots : c_{(1,M)}, \dots, c_{(n,1)} : \dots : c_{(n,M)})$$

be a codeword in the code corresponding to some finite cover of T . The *pseudocodeword* associated to $\tilde{\mathbf{c}}$ is the vector

$$\mathbf{p} = \mathbf{p}(\tilde{\mathbf{c}}) = (p_1, \dots, p_n) \in \mathbb{N}^n$$

with

$$p_i = \#\{k \mid c_{(i,k)} \neq 0\}.$$

Goal: Characterize the pseudocodewords.

Pseudocodewords and the Fundamental Cone

Theorem: (Koetter/Li/Vontobel/Walker)

Let $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{N}^n$. The following are equivalent:

- \mathbf{p} is a pseudocodeword.
- $\mathbf{p} \in \mathcal{K}(H)$ and $H\mathbf{p}^t = \mathbf{0} \in \mathbb{F}_2^r$.

Theorem: (Koetter/Li/Vontobel/Walker)

The rays through the pseudocodewords are dense in the fundamental cone.

More precisely: for every $\mathbf{v} \in \mathcal{K}(H)$ and every $\varepsilon > 0$, there is a pseudocodeword \mathbf{p} such that $\|\alpha\mathbf{p} - \mathbf{v}\| < \varepsilon$ for some $\alpha > 0$.

Cycle Codes

Suppose T is a Tanner graph and every bit node of T has degree 2. Then we can associate a *normal graph* N to T :

Vertices of N : Check nodes of T

Edges of N : Neighborhoods of bit nodes of T , i.e.,

$\{f, f'\}$ is an edge in N



There is a bit node x of T with $\partial(x) = \{f, f'\}$.

Cycle Codes (con't)

In this situation, the LDPC code with Tanner graph T is exactly the binary linear span of the *characteristic vectors* of the *simple cycles* of N .

The pseudocodewords are the closed paths on N which become simple cycles on some cover of N .

Using work of Wiberg, Feldman observes that if C is a cycle code, then LP decoding and graph (sum-product) decoding have identical performance.

Edge Zeta Function (Hashimoto)

Definition: Let N be a graph with edges e_1, \dots, e_n . To the edge e_i , assign the variable u_i . The *monomial* $g(E)$ of the path $E = (e_{i_1}, \dots, e_{i_k})$ is

$$g(E) = u_{i_1} \dots u_{i_k}.$$

The *edge zeta function* of N is

$$\zeta_N(u_1, \dots, u_n) = \prod_{[E]} (1 - g(E))^{-1},$$

where $[E]$ runs through all *equivalence classes* of *reduced, primitive* cycles in N .

Edge Zeta Function (con't)

The edge zeta function is a rational function:

Theorem: (Bass; Stark/Terras)

$$\zeta_N(u_1, \dots, u_n) = \frac{1}{\det(I - UM)} = \frac{1}{\det(I - MU)}$$

where

- I is a $2n \times 2n$ identity matrix
- U is a diagonal matrix with entries $u_1, \dots, u_n, u_1, \dots, u_n$
- M is a *directed edge matrix* for N .

Pseudocodewords and the Zeta Function

Theorem: (Koetter/Li/Vontobel/Walker)

Let C be a cycle code of length n with normal graph N .
Then

$$(p_1, \dots, p_n)$$

is a pseudocodeword *if and only if*

$$u_1^{p_1} \dots u_n^{p_n}$$

occurs in the power series expansion of $\zeta_N(u_1, \dots, u_n)$
with nonzero coefficient.

Question: Can we find an *algebraic description* for the pseudocodewords of a *general* code described by a parity check matrix?

Bit-even Tanner Graphs

Observation: Let $C = \ker(H_0)$ where H_0 is $r \times n$ and let T_0 be the corresponding Tanner graph.

Let H be the $2r \times n$ matrix formed by duplicating every row of H_0 and let T be the corresponding Tanner graph.

Then:

- H and T still define C
- H and H_0 have the same fundamental cone.
Same pseudocodewords!
- T is *bit-even*

Proposition: (Koetter/Li/Vontobel/Walker)

Let T be a bit-even Tanner graph with corresponding code C . Let \hat{C} be the *cycle code* with normal graph T .

Then C is a *punctured subcode* of \hat{C} .

Describing C in terms of \hat{C}

Proposition: (Koetter/Li/Vontobel/Walker)

Let T be a bit-even Tanner graph for a code C . Then the *codewords* in C correspond to disjoint unions of edge-simple cycles on T such that at each bit node x of T , either *all or none* of the edges incident to x occur.

Proposition: (Koetter/Li/Vontobel/Walker)

Let T be a bit-even Tanner graph for a code C . Then the *pseudocodewords* correspond to disjoint unions of reduced cycles on T in which all edges incident to any given bit node occur the *same number* of times.

General LDPC Codes and Zeta Functions

Theorem: (Koetter/Li/Vontobel/Walker)

The vector

$$(p_1, \dots, p_n) \in \mathbb{N}^n$$

is a pseudocodeword for C if and only if

$$\prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq d_i}} u_{(i,j)}^{p_i}$$

appears in the power series expansion of ζ_T .

Summary

- The notion of *pseudocodewords* arises in LP decoding and in graph decoding.
- For LP decoding, pseudocodewords describe *exactly* when decoding failure occurs.
- For graph decoding, this is less precise (except in the case of *cycle codes*).
- Pseudocodewords are described *analytically* via the *fundamental cone*.
- In the case of cycle codes, pseudocodewords are described *algebraically* via the *edge zeta function* of the normal graph.

Summary (con't)

- For more general LDPC codes, pseudocodewords are described *algebraically* via the edge zeta function of the *Tanner graph* (but this is less than satisfactory).
 - Lots of edges.
 - Not all monomials correspond to pseudocodewords.
 - No proof (yet) that the “punctured” zeta function is rational.