Lower Bounds on the Minimum Pseudo-Weight of Linear Codes

Pascal O. Vontobel and Ralf Koetter*

(Submitted to ISIT 2004)

November 30, 2003

Abstract

We discuss two techniques for obtaining lower bounds on the (AWGN channel) pseudo-weight of binary linear codes. Whereas the first bound is based on the largest and second-largest eigenvalues of a matrix associated with the parity-check matrix of a code, the second bound is given by the solution to a linear program.

Keywords: Linear codes, parity-check matrix, pseudo-weight, eigenvalues, linear program.

1 Introduction

In order to obtain bounds on the maximum-likelihood decoding performance of a linear code, one needs to know the minimum Hamming weight of the code and the multiplicity of the minimum Hamming weight codewords (or even better, the whole Hamming weight spectrum of the code.) As argued in [1], if one wants to assess the performance under iterative message-passing decoding, one needs to study the pseudo-weight of pseudo-codewords, i.e. one needs to find the minimum pseudo-weight (or even better, the whole pseudo-weight spectrum). As observed in [1], the pseudo-codewords characterized by the fundamental polytope/cone, give an astonishingly accurate picture.[1] Given a code, note that the minimum Hamming weight is a function of the code whereas the minimum pseudo-weight is a function of the parity-check matrix describing the code.

Finding the minimum Hamming weight of a linear code is known to be a hard problem.[2] Obtaining the minimum AWGN channel pseudo-weight seems not to be an easy task either; therefore, one has to find techniques that yield upper and lower bounds on the minimum pseudo-weight. Related problems include finding the minimum stopping set size[3] and finding the minimum fractional/max-fractional distance.[4]

In [1] we discussed two ways of obtaining upper bounds on the minimum (AWGN channel) pseudo-weight: one of them was based on searching for low-weight pseudo-codewords in the fundamental cone, the other was based on the so-called canonical completion. In this paper we now introduce two techniques for obtaining lower bounds: the first one is a purely algebraic eigenvalue-based bound (see Sec. 3), whereas the second is a linear-programming-based bound (see Sec. 4).

*Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1308 West Main Street, Urbana, IL 61801, USA, `vontobel@ifp.uiuc.edu`, `koetter@uiuc.edu`. The first author was supported by NSF Grants CCR 99-84515 and CCR-0105719.

[1] One more reason to study the minimum pseudo-weight of pseudo-codewords is that for (certain relaxations) of the linear programming decoder by Feldman and Karger [2] the characterization by the fundamental polytope/cone is actually exact.

[2] The papers [3, 4] discuss this issue; it seems though that for codes like turbo and LDPC codes, the problem might not be as hard as the general case, see e.g. [5, 6].

[3] Its hardness for general codes was established in [7] by modifying the proof of [3].

[4] The fractional and the max-fractional distance [2] are lower bounds on the binary symmetric channel pseudo-weight [8]; [2] gives lower bounds on the max-fractional distance in terms of the girth and it is also shown that the fractional distance can be computed efficiently.

2 Definitions

All vectors will be row vectors. The L_1 and the L_2 norm of a vector $\mathbf{x} \in \mathbb{R}^n$ of length n are $||\mathbf{x}||_1 \triangleq \sum_{i=1}^{n} |x_i|$ and $||\mathbf{x}||_2 \triangleq \sqrt{\sum_{i=1}^{n} |x_i|^2}$, respectively. All codes will be binary linear codes; $w_{\mathrm{H}}^{\min}(\mathsf{C})$ denotes the minimum Hamming weight of a linear code C.

Definition 1 *Let $\mathbf{H}$ be the parity-check matrix of a binary linear code C. We let $\mathcal{V} \triangleq \mathcal{V}(\mathbf{H})$ be the set of column indices of $\mathbf{H}$ and $\mathcal{R} \triangleq \mathcal{R}(\mathbf{H})$ be the set of row indices of $\mathbf{H}$, respectively. For each $r \in \mathcal{R}$, we let $\mathcal{V}_r \triangleq \mathcal{V}_r(\mathbf{H}) \triangleq \{v \in \mathcal{V} \mid [\mathbf{H}]_{r,v} = 1\}$. Furthermore, for any $r \in \mathcal{R}$ and any vector $\mathbf{x}$ of length $|\mathcal{V}|$, we let $\mathbf{x}_{\mathcal{V}_r}$ be the vector that has only the entries of $\mathbf{x}$ whose indices are in $\mathcal{V}_r$. We call C a (j,k)-regular code if the uniform column weight of $\mathbf{H}$ is j and the uniform row weight of $\mathbf{H}$ is k.*

Therefore, a vector $\mathbf{x} \in \mathbb{F}_2^n$ is a codeword of C if and only if the modulo-2 sum of the entries of $\mathbf{x}_{\mathcal{V}_r}$ equals zero for each $r \in \mathcal{R}$. Throughout the paper we will consider a code of length n, i.e. $|\mathcal{V}| = n$.

Definition 2 (Positive orthant) *Let $\mathcal{O}^{(n)}$ be the positive orthant of the n-dimensional real space, i.e., $\mathcal{O}^{(n)} \triangleq \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0 \text{ for all } i = 1, \ldots, n\}$. Moreover, we let $\dot{\mathcal{O}}^{(n)}$ be the punctured positive orthant, i.e., $\dot{\mathcal{O}}^{(n)} \triangleq \mathcal{O}^{(n)} \setminus \{\mathbf{0}\}$.*

Definition 3 (Additive white Gaussian noise channel (AWGNC) pseudo-weight [9, 8]) *Let $\mathbf{x} \in \dot{\mathcal{O}}^{(n)}$. The AWGNC pseudo-weight $w_{\mathrm{p}}(\mathbf{x})$ of $\mathbf{x}$ is given by*

$$w_{\mathrm{p}}(\mathbf{x}) \triangleq w_{\mathrm{p}}^{\mathrm{AWGNC}}(\mathbf{x}) \triangleq \frac{||\mathbf{x}||_1^2}{||\mathbf{x}||_2^2}.$$

Remark 4 *The pseudo-weight as defined in Def. 3 is invariant under scaling by a positive scalar, i.e. $w_{\mathrm{p}}(\alpha \cdot \mathbf{x}) = w_{\mathrm{p}}(\mathbf{x})$ for any $\alpha > 0$ and any $\mathbf{x} \in \dot{\mathcal{O}}^{(n)}$.*

Definition 5 (Fundamental Polytope/Cone) *The fundamental polytope $\mathcal{P}(\mathbf{H})$ and the fundamental cone $\mathcal{K}(\mathbf{H})$of a linear code C with parity-check matrix $\mathbf{H}$ were introduced in [1]. Moreover, $\dot{\mathcal{P}}(\mathbf{H}) \triangleq \mathcal{P}(\mathbf{H}) \setminus \{\mathbf{0}\}$ and $\dot{\mathcal{K}}(\mathbf{H}) \triangleq \mathcal{K}(\mathbf{H}) \setminus \{\mathbf{0}\}$ will denote the punctured fundamental polytope and cone, respectively.*

Reformulating the mathematical definition given in [1], one obtains the following simple characterization of the fundamental cone.

Theorem 6 *Let $\mathbf{H}$ be a parity-check matrix of a code C of length n. A necessary and sufficient condition for a vector $\mathbf{x} \in \mathcal{O}^{(n)}$ to be in the fundamental cone $\mathcal{K}(\mathbf{H})$ is that for each $r \in \mathcal{R}$ and for each $v \in \mathcal{V}_r$ we must have*

$$\sum_{v' \in \mathcal{V}_r \setminus \{v\}} x_{v'} \geq x_v.$$

All these inequalities can be expressed as $\mathbf{K}\mathbf{x}^{\mathsf{T}} \geq \mathbf{0}$ for some $\mathbf{K} \triangleq \mathbf{K}(\mathbf{H})$.

Proof: Omitted. □

Definition 7 (Minimum Pseudo-Weight) *For a given parity-check matrix $\mathbf{H}$ of a code C, the minimum AWGNC pseudo-weight is defined to be*

$$w_{\mathrm{p}}^{\min}(\mathbf{H}) \triangleq \min_{\mathbf{x} \in \dot{\mathcal{P}}(\mathbf{H})} w_{\mathrm{p}}(\mathbf{x}) \overset{(*)}{=} \min_{\mathbf{x} \in \dot{\mathcal{K}}(\mathbf{H})} w_{\mathrm{p}}(\mathbf{x}),$$

where equality $(*)$ follows from the scaling-invariance of $w_{\mathrm{p}}(\,\cdot\,)$ and the properties of $\dot{\mathcal{P}}(\mathbf{H})$ and $\dot{\mathcal{K}}(\mathbf{H})$ [1].

Note that for any parity-check matrix $\mathbf{H}$ of a binary linear code C we have $w_{\mathrm{p}}^{\min}(\mathbf{H}) \leq w_{\mathrm{H}}^{\min}(\mathsf{C})$.

3 An Eigenvalue-Based Lower Bound on the Minimum Pseudo-Weight

The following lemma will prove useful for our eigenvalue-based lower bound.

Lemma 8 *Let $\mathbf{x} \in \mathcal{K}(\mathbf{H})$ be a vector in the fundamental cone of $\mathbf{H}$. Then, for any $r \in \mathcal{R}$ we have*

$$\left(\sum_{v \in \mathcal{V}_r} x_v\right)^2 \geq 2 \cdot \left(\sum_{v \in \mathcal{V}_r} x_v^2\right).$$

Proof: For any $r \in \mathcal{R}$ we get

$$\left(\sum_{v \in \mathcal{V}_r} x_v\right)^2 = \left(\sum_{v \in \mathcal{V}_r} x_v\right) \cdot \left(\sum_{v' \in \mathcal{V}_r} x_{v'}\right)$$
$$= \left(\sum_{v \in \mathcal{V}_r} x_v \underbrace{\left(\sum_{v' \in \mathcal{V}_r} x_{v'}\right)}_{\overset{(*)}{\geq} 2x_v}\right) \geq 2\left(\sum_{v \in \mathcal{V}_r} x_v^2\right),$$

where (∗) follows from Th. 6 □

Theorem 9 *Let C be a (j,k)-regular code of length n defined by the parity-check matrix $\mathbf{H}$ and let the corresponding Tanner graph have one component. Let $\mathbf{L} \overset{\triangle}{=} \mathbf{H}^\mathsf{T}\mathbf{H}$ and let μ_1 and μ_2 be the largest and second-largest eigenvalue, respectively, of $\mathbf{L}$. Then the minimum Hamming weight and the minimum AWGNC pseudo-weight are lower bounded by*

$$w_\mathrm{H}^\mathrm{min}(\mathsf{C}) \geq w_\mathrm{p}^\mathrm{min}(\mathbf{H}) \geq n \cdot \frac{2j-\mu_2}{\mu_1-\mu_2}.$$

Remark 10 *Interestingly, the lower bound given in Th. 9 is the same as the bit-oriented lower bound given by Tanner [10] for the minimum Hamming weight of a binary code.*

Proof: The proof is very much inspired by the proof of the bit-oriented lower bound given by Tanner [10], although some of the equalities and inequalities hold because of other (more general) reasons. Let $\mathbf{1}$ be a vector of length n containing only ones. Then, the pseudo-weight of any pseudo-codeword $\mathbf{x} \in \dot{\mathcal{K}}(\mathbf{H})$ can be rewritten as

$$w_\mathrm{p}(\mathbf{x}) = \frac{||\mathbf{x}||_1^2}{||\mathbf{x}||_2^2} = \frac{(\mathbf{x}\cdot\mathbf{1}^\mathsf{T})^2}{||\mathbf{x}||_2^2}. \qquad (1)$$

The crucial idea is to define $\mathbf{y} \overset{\triangle}{=} \mathbf{x}\cdot\mathbf{H}^\mathsf{T}$, which is a vector of length $|\mathcal{R}|$, and to try to get a lower and an upper bound on $||\mathbf{y}||_2^2$.

First, let us derive a lower bound on $||\mathbf{y}||_2^2$.

$$\begin{aligned} ||\mathbf{y}||_2^2 &= \sum_{r\in\mathcal{R}} y_i^2 = \sum_{r\in\mathcal{R}} \left(\sum_{v\in\mathcal{V}_r} x_v\right)^2 \overset{(*)}{\geq} \sum_{r\in\mathcal{R}} 2\left(\sum_{v\in\mathcal{V}_r} x_v^2\right) \\ &= 2\sum_{r\in\mathcal{R}}\sum_{v\in\mathcal{V}_r} x_v^2 \overset{(**)}{=} 2j\cdot||\mathbf{x}||_2^2, \qquad (2) \end{aligned}$$

where (∗) follows from Lemma 8 and (∗∗) from the fact that in the double sum $\sum_{r\in\mathcal{R}}\sum_{v\in\mathcal{V}_r}$ every term x_v^2, $v\in\mathcal{V}$, appears exactly j times.

Secondly, let us derive an upper bound on $||\mathbf{y}||_2^2$. To this end, let us assume that $\mathbf{L}$ has s distinct eigenvalues $\mu_1 > \mu_2 > \cdots > \mu_s$. Let $\mathbf{z}^{(v)}$ be the projection of $\mathbf{x}$ onto the v-th eigenspace, $v\in\{1,\ldots,s\}$. (Because $\mathbf{L}$ is a symmetric matrix, the algebraic and the geometric multiplicities are equal for each eigenvalue and all eigenspaces are orthogonal on each other.) It can easily be checked that $\mathbf{1}$ is a left eigenvector of $\mathbf{L}$ with eigenvalue $\mu_1 = j\cdot k$ whose multiplicity is 1. (Multiplicity 1 follows from the assumption that the Tanner graph has one component.) Therefore, the projection of $\mathbf{x}$ onto the first eigenspace is $\mathbf{z}^{(1)} = (\mathbf{x}\cdot\mathbf{1}^\mathsf{T})/(\mathbf{1}\cdot\mathbf{1}^\mathsf{T})\cdot\mathbf{1} = (1/n)\cdot(\mathbf{x}\cdot\mathbf{1}^\mathsf{T})\cdot\mathbf{1}$, whose squared L_2-norm equals

$$||\mathbf{z}^{(1)}||_2^2 = \frac{1}{n^2}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2\cdot||\mathbf{1}||_2^2 = \frac{1}{n}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2. \qquad (3)$$

We also have $||\mathbf{x}||_2^2 = \sum_{v=1}^s ||\mathbf{z}^{(v)}||_2^2$, from which

$$\sum_{v=2}^s ||\mathbf{z}^{(v)}||_2^2 = ||\mathbf{x}||_2^2 - ||\mathbf{z}^{(1)}||_2^2 = ||\mathbf{x}||_2^2 - \frac{1}{n}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2 \qquad (4)$$

follows. Using this partial results, we can now try to upper bound $||\mathbf{y}||_2^2$. We get

$$\begin{aligned} ||\mathbf{y}||_2^2 &= ||\mathbf{x}\cdot\mathbf{H}^\mathsf{T}||_2^2 = \mathbf{x}\cdot\mathbf{H}^\mathsf{T}\cdot\mathbf{H}\cdot\mathbf{x}^\mathsf{T} = \mathbf{x}\cdot\mathbf{L}\cdot\mathbf{x}^\mathsf{T} \\ &= \left(\sum_{\ell=1}^s \mathbf{z}^{(\ell)}\right)\cdot\mathbf{L}\cdot\left(\sum_{\ell'=1}^s \mathbf{z}^{(\ell')\mathsf{T}}\right) \\ &= \sum_{\ell=1}^s\sum_{\ell'=1}^s \mu_\ell\cdot\mathbf{z}^{(\ell)}\cdot\mathbf{z}^{(\ell')\mathsf{T}} = \sum_{\ell=1}^s \mu_\ell\cdot||\mathbf{z}^{(\ell)}||_2^2 \\ &= \mu_1\cdot||\mathbf{z}^{(1)}||_2^2 + \left(\sum_{\ell=2}^s \mu_\ell\cdot||\mathbf{z}^{(\ell)}||_2^2\right) \\ &\overset{(*)}{\leq} \mu_1\cdot||\mathbf{z}^{(1)}||_2^2 + \mu_2\cdot\left(\sum_{\ell=2}^s ||\mathbf{z}^{(\ell)}||_2^2\right) \\ &\overset{(**)}{=} \mu_1\cdot\frac{1}{n}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2 + \\ &\qquad \mu_2\cdot\left(||\mathbf{x}||_2^2 - \frac{1}{n}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2\right) \\ &= (\mu_1-\mu_2)\cdot\frac{1}{n}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2 + \mu_2\cdot||\mathbf{x}||_2^2, \qquad (5) \end{aligned}$$

where (∗) follows from $\mu_\ell < \mu_2$ for $\ell\in\{3,\ldots,s\}$ (equality can happen if $s=2$ or if $\mathbf{x}$ lies in the subspace spanned by the first two eigenspaces) and (∗∗) from (3) and (4). Combining (2) and (5) we obtain $(\mu_1-\mu_2)\cdot\frac{1}{n}\cdot\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2 + \mu_2\cdot||\mathbf{x}||_2^2 \geq ||\mathbf{y}||_2^2 \geq 2j\cdot||\mathbf{x}||_2^2$. Because $\mu_1 > \mu_2$, we have $\mu_1-\mu_2 > 0$, which allows us to formulate

$$\frac{\left(\mathbf{x}\cdot\mathbf{1}^\mathsf{T}\right)^2}{||\mathbf{x}||_2^2} \geq n\cdot\frac{2j-\mu_2}{\mu_1-\mu_2},$$

This, combined with (1), leads to the desired result in the theorem. □

Remark 11 *All known cases, where Tanner's bit-oriented lower bound on the minimum Hamming weight gives non-trivial results, give now also non-trivial results on the minimum pseudo-weight. Codes from partial geometries [11], which include projective planes [12] (finite generalized triangles) and finite generalized quadrangles [13], belong to this set. Note that for codes from projective planes and for some codes from generalized quadrangles the above lower bound on the minimum pseudo-weight matches the minimum Hamming weight, therefore for these codes the minimum pseudo-weight equals the minimum Hamming weight.*

Theorem 12 *Consider a binary code of length n whose automorphism group is two-transitive on the bits and whose dual code has minimum Hamming weight $w_{\mathrm{H}}^{\mathrm{min}\perp}(\mathsf{C})$. Let $\mathbf{H}$ be the matrix consisting of all vectors in the dual code whose Hamming weight equals $w_{\mathrm{H}}^{\mathrm{min}\perp}(\mathsf{C})$. Then,*

$$w_{\mathrm{H}}^{\mathrm{min}}(\mathsf{C}) \geq w_{\mathrm{p}}^{\mathrm{min}}(\mathbf{H}) \geq \frac{n-1}{w_{\mathrm{H}}^{\mathrm{min}\perp}(\mathsf{C})-1}+1.$$

(We assume that the above parity-check matrix $\mathbf{H}$ spans indeed the whole dual code; if not, then the lower bound is for an even larger code.)

Proof: In App. E of [14] the above lower bound for the minimum Hamming weight was derived from Tanner's bit-oriented bound on the minimum Hamming weight. But because Tanner's bit-oriented lower bound on the minimum Hamming weight and the lower bound in Th. 9 give the same value for the given parity-check matrix $\mathbf{H}$, the theorem follows. □

4 An LP-Based Lower Bound on the Minimum Pseudo-Weight

Our linear programming lower bound on the minimum pseudo-weight was originally very much inspired by the linear programming lower bound on the minimum Hamming weight as presented by Tanner [10]. But finally, its form is quite different. Actually, the present form reminds much more of the "Lift and Project" technique in Sec. 5.4.2 of [2] which was used to obtain a modification of the linear programming decoder. But the approach in [2] is used to constrain the fundamental polytope whereas we are interested in relaxing the fundamental cone. Note moreover that in [10] and in [2] an important ingredient is the relation $x_i = x_i^2$ (which holds because the components of the vector $\mathbf{x}$ were desired to be 0 or 1), but this does not hold anymore for components of pseudo-codewords.

The lower bounds on the minimum pseudo-weight that will be presented in this section are based on the following lemma, which can be considered as a form of relaxed optimization. This relaxation makes sense especially in the cases where the new optimization problem is simpler and can be solved efficiently.

Lemma 13 *Let $\mathcal{S}$ and $\mathcal{S}'$ be two sets, let f be a function with domain $\mathcal{S}$, and let f' be a function with domain $\mathcal{S}'$. If for each $\mathbf{x} \in \mathcal{S}$ there exists at least one $\mathbf{x}' \in \mathcal{S}'$ such that $f(\mathbf{x}) \leq f'(\mathbf{x}')$, then*

$$\max_{\mathbf{x}\in\mathcal{S}} f(\mathbf{x}) \leq \max_{\mathbf{x}'\in\mathcal{S}'} f'(\mathbf{x}').$$

Proof: Let $\mathbf{x}^* \in \mathcal{S}$ be a vector that achieves the maximum in $\max_{\mathbf{x}\in\mathcal{S}} f(\mathbf{x})$. Because for each $\mathbf{x} \in \mathcal{S}$ there exists at least one $\mathbf{x}' \in \mathcal{S}'$ such that $f(\mathbf{x}) \leq f'(\mathbf{x}')$, there must exist a $\mathbf{x}'^* \in \mathcal{S}'$ such that $f(\mathbf{x}^*) \leq f'(\mathbf{x}'^*)$. Therefore,

$$\max_{\mathbf{x}\in\mathcal{S}} f(\mathbf{x}) = f(\mathbf{x}^*) \leq f'(\mathbf{x}'^*) \leq \max_{\mathbf{x}'\in\mathcal{S}'} f'(\mathbf{x}'),$$

which proves the statement in the lemma. □

Definition 14 *Let C be a code of length n with parity-check matrix $\mathbf{H}$ and let $\mathbf{K} \triangleq \mathbf{K}(\mathbf{H})$ be as given in Th. 6. We introduce the sets*

$$\begin{aligned}\mathcal{K} \triangleq \mathcal{K}(\mathbf{H}) &\triangleq \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{K}\mathbf{x}^\mathsf{T} \geq \mathbf{0}^\mathsf{T} \text{ and } \mathbf{x} \geq \mathbf{0}\right\},\\ \mathcal{K}_1 \triangleq \mathcal{K}_1(\mathbf{H}) &\\ &\triangleq \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{K}\mathbf{x}^\mathsf{T} \geq \mathbf{0}^\mathsf{T}, \mathbf{x} \geq \mathbf{0}, ||\mathbf{x}||_1 = 1\right\}\\ &= \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{K}\mathbf{x}^\mathsf{T} \geq \mathbf{0}^\mathsf{T}, \mathbf{x} \geq \mathbf{0}, \mathbf{x}\cdot\mathbf{1}^\mathsf{T} = 1\right\}.\end{aligned}$$

Theorem 15 *Let C be a code of length n with parity-check matrix $\mathbf{H}$, let $\mathcal{V} \triangleq \mathcal{V}(\mathbf{H})$ as in Def. 1, and let $\mathbf{K} \triangleq \mathbf{K}(\mathbf{H})$ as in Th. 6. Let the entries of a vector $\mathbf{y} \in \mathbb{R}^{(\mathcal{V}^2)}$ be indexed by $(v,w) \in \mathcal{V}^2$. Furthermore, for $v \in \mathcal{V}$ we let $\mathbf{y}_{(v,:)}$ be the sub-vector (of length $|\mathcal{V}|$) of $\mathbf{y}$ consisting of all entries with index (v,w), $w \in \mathcal{V}$,* *and for $w \in \mathcal{V}$ we let $\mathbf{y}_{(:,w)}$ be the sub-vector (of length $|\mathcal{V}|$) of $\mathbf{y}$ consisting of all entries with index (v, w), $v \in \mathcal{V}$. Then, the minimum Hamming weight and the minimum pseudo-weight $w_{\mathrm{p}}^{\min}(\mathbf{H})$ can be lower bounded by*

$$w_{\mathrm{H}}^{\min}(\mathsf{C}) \geq w_{\mathrm{p}}^{\min}(\mathbf{H}) \geq \frac{1}{\max_{\mathbf{y}\in\mathcal{K}_1'} f'(\mathbf{y})}, \qquad (6)$$

where

$$f'(\mathbf{y}) \overset{\triangle}{=} \sum_{v\in\mathcal{V}} y_{(v,v)}$$

and

$$\mathcal{K}_1' \overset{\triangle}{=} \left\{ \mathbf{y} \in \mathbb{R}^{(\mathcal{V}^2)} \;\middle|\; \begin{array}{l} \mathbf{y} \geq \mathbf{0}, \mathbf{y}\cdot\mathbf{1}^{\mathsf{T}} = 1, \\ \mathbf{K}\mathbf{y}_{(v,:)}^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}} \textit{ for all } v \in \mathcal{V}, \\ \mathbf{K}\mathbf{y}_{(:,w)}^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}} \textit{ for all } w \in \mathcal{V} \end{array} \right\}.$$

Note that the maximization problem in the denominator on the right-hand side of (6) is a linear optimization program.

Proof: Using the sets defined in Def. 14 we have

$$\frac{1}{w_{\mathrm{p}}^{\min}(\mathbf{H})} = \max_{\mathbf{x}\in\mathcal{K}} \frac{||\mathbf{x}||_2^2}{||\mathbf{x}||_1^2} = \max_{\mathbf{x}\in\mathcal{K}_1} ||\mathbf{x}||_2^2 \overset{(*)}{\leq} \max_{\mathbf{y}\in\mathcal{K}_1'} f'(\mathbf{y}),$$

where in $(*)$ we have used Lemma 13. In order to show that this relaxation is indeed valid, we have to show that for each $\mathbf{x} \in \mathcal{K}_1$ there exists a $\mathbf{y} \in \mathcal{K}_1'$ such that $||\mathbf{x}||_2^2 \leq f'(\mathbf{y})$. So, choose a vector $\mathbf{x} \in \mathcal{K}_1$. Then, let $\mathbf{y}$ have entries $y_{(v,w)} \overset{\triangle}{=} x_v \cdot x_w$. This, inter alia, implies that $\mathbf{y}_{(v,:)} = x_v \cdot \mathbf{x}$ for each $v \in \mathcal{V}$ and that $\mathbf{y}_{(:,w)} = x_w \cdot \mathbf{x}$ for each $w \in \mathcal{V}$. Let us first show that $\mathbf{y} \in \mathcal{K}_1'$.

- By assumption, $\mathbf{x} \geq \mathbf{0}$. Therefore, for each $(v, w) \in \mathcal{V}^2$ we have $y_{v,w} = x_v \cdot x_w \geq 0 \cdot 0 = 0$ and so $\mathbf{y} \geq \mathbf{0}$.
- By assumption, $\mathbf{x} \cdot \mathbf{1}^{\mathsf{T}} = 1$, i.e. $\sum_{v\in\mathcal{V}} x_v = 1$. Therefore, $\mathbf{y} \cdot \mathbf{1}^{\mathsf{T}} = \sum_{v\in\mathcal{V}} \sum_{w\in\mathcal{V}} y_{(v,w)} = \sum_{v\in\mathcal{V}} \sum_{w\in\mathcal{V}} x_v x_w = \big(\sum_{v\in\mathcal{V}} x_v\big) \cdot \big(\sum_{w\in\mathcal{V}} x_v\big) = 1 \cdot 1 = 1$.
- Let $v \in \mathcal{V}$. By assumption, $x_v \geq 0$ and $\mathbf{K}\mathbf{x}^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}}$. It follows that also the scaled vector (scaled by a non-negative value) $\mathbf{y}_{(v,:)} = x_v \cdot \mathbf{x}$ fulfills $\mathbf{K}\mathbf{y}_{(v,:)}^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}}$.
- Let $w \in \mathcal{V}$. By assumption, $x_w \geq 0$ and $\mathbf{K}\mathbf{x}^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}}$. It follows that also the scaled vector (scaled by a non-negative value) $\mathbf{y}_{(:,w)} = x_w \cdot \mathbf{x}$ fulfills $\mathbf{K}\mathbf{y}_{(:,w)}^{\mathsf{T}} \geq \mathbf{0}^{\mathsf{T}}$.

Finally, $f'(\mathbf{y}) = \sum_{v\in\mathcal{V}} y_{(v,v)} = \sum_{v\in\mathcal{V}} x_v^2 = ||\mathbf{x}||_2^2$ and so certainly $||\mathbf{x}||_2^2 \leq f'(\mathbf{y})$. □

There are many extensions/modifications to this technique. We briefly mention some of them.

- An alternative formulation of Th. 15 is as follows. Instead of $\mathbf{y} \in \mathbb{R}^{(\mathcal{V}^2)}$, we consider the matrix $\mathbf{Y} \in \mathbb{R}^{|\mathcal{V}|\times|\mathcal{V}|}$. The function $f'(\mathbf{y})$ then becomes the trace of $\mathbf{Y}$, etc.
- Instead of the ansatz $y_{(v,w)} = x_v \cdot x_w$ based on quadratic terms, one can also use the ansatz $y_{(v,w,u)} = x_v \cdot x_w \cdot x_u$ based on cubic terms. The vector $\mathbf{y}$ can then be represented by a cube where in all three directions the content must be in the fundamental cone. The cost function has the form $\frac{1}{3}\big(\sum_{v,w} y_{(v,v,w)} + \sum_{v,w} y_{(v,w,v)} + \sum_{v,w} y_{(w,v,v)}\big)$. This procedure can be extended to a quartic term approach, a quintic term approach etc. Obviously the complexity grows.
- Modifying the proof of Th. 15 appropriately, one can set $y_{(w,v)} \overset{\triangle}{=} y_{(v,w)}$ for all $v, w \in \mathcal{V}$; this is based on the observation that $x_i \cdot x_j = x_j \cdot x_i$ for a pseudo-codeword $\mathbf{x}$. Additionally, if the parity-check matrix has some symmetries, this can be used to reduce the complexity of the linear program by a factor proportional to the size of the symmetry group.
- An approach to improve the linear programming bound is to assume that x_v is the largest component. Then $y_{w,v} \geq y_{w,v'}$ and $y_{v,w} \geq y_{v',w}$ for all w and all v'. Executing the linear programming bound for all possible $v \in \mathcal{V}$ and taking the least lower bound gives also a lower bound on the minimum pseudo-weight. But note that this improvement is not compatible with using symmetries of the parity-check matrix. The only symmetry of $\mathbf{y}$ that can be used is $y_{(v,w)} = y_{(w,v)}$ for all $v, w \in \mathcal{V}$.
- The approach in Th. 15 can be generalized to get lower bounds on the minimum pseudo-weight of codes described by factor graphs with state

nodes, e.g. tail-biting trellises. For tail-biting trellises one can also formulate a fundamental polytope/cone.[5]

- To all the linear programs formulated above one can formulate a dual linear program, see e.g. [15]. For the above linear programs, the cost function of the dual linear program turns out to have actually a rather simple form. This allows one to use simple optimization heuristics as e.g. gradient-based methods.

  Whereas only the optimal point of the primal linear program leads to a true lower bound on the minimum pseudo-weight, any feasible point of the dual linear program is actually a lower bound on the minimum pseudo-weight. Therefore, we do not need a guarantee that an optimization algorithm of the dual linear program really achieves the optimum.

Note that whereas the eigenvalue-based technique (Th. 9) gives nontrivial results only for certain code families, this second technique (Th. 15) gives nontrivial results for any code, but is computationally also more demanding.

We have some preliminary numerical results using this technique and its extensions. For codes from projective planes the lower bound equals the minimum Hamming weight (as was the case for Th. 9, see Rem. 11). For the $[155, 64, 20]$ code by Tanner [16] (for which an upper bound on the minimum pseudo-weight is 16.4) we obtained 9.3 (by the quadratic approach) and 10.8 (a feasible point from the dual linear program of the quadratic approach with the fourth extension/modification mentioned above).

References

[1] R. Koetter and P. O. Vontobel, "Graph-covers and iterative decoding of finite-length codes," in *Proc. 3rd Intern. Conf. on Turbo Codes and Related Topics*, (Brest, France), pp. 75–82, Sept. 1–5 2003.

[2] J. Feldman, *Decoding Error-Correcting Codes via Linear Programming.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 2003.

[3] E. Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. on Inform. Theory*, vol. IT–24, no. 3, pp. 384–386, 1978.

[4] A. Vardy, "The intractability of computing the minimum distance of a code," *IEEE Trans. on Inform. Theory*, vol. IT–43, no. 6, pp. 1757–1766, 1997.

[5] C. Berrou and S. Vaton, "Computing the minimum distances of linear codes by the error impulse method," in *Proc. IEEE Intern. Symp. on Inform. Theory*, (Lausanne, Switzerland), p. 5, June 30–July 5 2002.

[6] X.-Y. Hu, "On the computation of minimum distance of low-density parity-check codes," *preprint*, 2003.

[7] H. Pishro-Nik and F. Fekri, "On LDPC codes over the erasure channel," in *Proc. 41st Allerton Conf. on Communications, Control, and Computing*, (Allerton House, Monticello, Illinois, USA), October 1–3 2003.

[8] G. D. Forney, Jr., R. Koetter, F. R. Kschischang, and A. Reznik, "On the effective weights of pseudocodewords for codes defined on graphs with cycles," in *Codes, Systems, and Graphical Models (Minneapolis, MN, 1999)* (B. Marcus and J. Rosenthal, eds.), vol. 123 of *IMA Vol. Math. Appl.*, pp. 101–112, Springer Verlag, New York, Inc., 2001.

[9] N. Wiberg, *Codes and Decoding on General Graphs.* PhD thesis, Linköping University, Sweden, 1996.

[10] R. M. Tanner, "Minimum-distance bounds by graph analysis," *IEEE Trans. on Inform. Theory*, vol. IT–47, no. 2, pp. 808–821, 2001.

[11] S. J. Johnson and S. R. Weller, "Codes for iterative decoding from partial geometries," in *Proc. IEEE Intern. Symp. on Inform. Theory*, (Lausanne, Switzerland), p. 310, June 30–July 5 2002.

[12] R. Lucas, M. Fossorier, Y. Kou, and S. Lin, "Iterative decoding of one-step majority logic decodable codes based on belief propagation," *IEEE Trans. on Comm.*, vol. COMM-48, pp. 931–937, June 2000.

[13] P. O. Vontobel and R. M. Tanner, "Construction of codes based on finite generalized quadrangles for iterative decoding," in *Proc. IEEE Intern. Symp. on Inform. Theory*, (Washington, D.C., USA), p. 223, June 24–29 2001.

[14] P. O. Vontobel, *Algebraic Coding for Iterative Decoding.* PhD thesis, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 2003. Available under `http://www.ifp.uiuc.edu/~vontobel`.

[15] D. Bertsimas and J. N. Tsitsiklis, *Linear Optimization.* Belmont, MA: Athena Scientific, 1997.

[16] R. M. Tanner, D. Sridhara, and T. Fuja, "A class of group-structured LDPC codes," in *Proc. of ICSTA 2001*, (Ambleside, England), 2001.

[5] In the same way as certain LP relaxations for LDPC codes considered by Feldman [2] correspond to our fundamental polytope/cone, certain LP relaxations of tail-biting trellises considered in [2] correspond to the fundamental polytope/cone of tail-biting trellises.