On the Proof of Minimum r-Wise Distance for Linear Codes

Hsuan-Yin Lin

8 June 2018

Abstract

In [1], the *r-wise Hamming distance* for an arbitrary code (linear or nonlinear), is presented and can be seen as a generalized notion of the pairwise Hamming distance. Recently, in [2, Proposition 6] it was claimed that if we restrict the interested codes to be linear codes, the minimum r-wise Hamming distance is equivalent to the well-known *sth generalized Hamming weight* with $s = \lceil \log_2(r) \rceil$ [3], but without proof. In this report, we provide a detailed proof of it.

1 Preliminaries

1.1 Notation and Definitions

We first review a conventional notation in coding theory. A general codebook $\mathscr{C}^{(\mathsf{M},n)}$ with M codewords and with a blocklength n is usually written as an $\mathsf{M} \times n$ *codebook matrix* as below.

$$\mathscr{C}^{(\mathsf{M},n)} = \begin{pmatrix} — \ \boldsymbol{x}_1 \ — \\ \vdots \\ — \ \boldsymbol{x}_{\mathsf{M}} \ — \end{pmatrix} = \begin{pmatrix} | & | & & | \\ \boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_n \\ | & | & & | \end{pmatrix}, \tag{1}$$

where $\boldsymbol{x}_i$, $i \in \{1, 2, \ldots, \mathsf{M}\}$ are row-vectors and $\boldsymbol{c}_j^{(\mathsf{M})} = (c_{j,1} \cdots c_{j,\mathsf{M}})^{\mathsf{T}}$, $j \in \{1, \ldots, n\}$ is a column-vector. Note that the M rows correspond to the M codewords.

Next, we introduce the concept of *affinely independent (a.i.)* vectors, which will be used in the sequel.

Definition 1 (Affinely Independence)**.** A set of vectors $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$ is said to be *affinely independent (a.i.)* if the set $\{\boldsymbol{x}_1 - \boldsymbol{x}_0, \boldsymbol{x}_2 - \boldsymbol{x}_0, \ldots, \boldsymbol{x}_s - \boldsymbol{x}_0\}$ is *linearly independent (l.i.)*. The *affine subspace* generated by a.i. vectors $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$, denoted by $\mathrm{aff}(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)$, is the set of all *affine combinations* of a.i. vectors $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$, which can be represented as

$$\mathrm{aff}(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s) \triangleq \left\{ \boldsymbol{x} \colon \boldsymbol{x} = \sum_{i=0}^{s} \lambda_i \boldsymbol{x}_i \text{ with } \sum_{i=0}^{s} \lambda_i = 1 \right\}.$$

Note that since any permutation of a set of vectors $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$ should result in only one affine subspace, and hence to verify the affinely independence, one can check the linearly independence of $\{\boldsymbol{x}_0 - \boldsymbol{x}_i, \boldsymbol{x}_1 - \boldsymbol{x}_i, \ldots, \boldsymbol{x}_s - \boldsymbol{x}_i\}$ for any $i \in \{0, \ldots, s\}$.

We also remark that it can be shown that for a set of a.i. vectors $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_s\}$, any $\boldsymbol{x} \in \mathrm{aff}(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_s)$ has a unique representation

$$\boldsymbol{x} = \sum_{i=0}^{s} \lambda_i \boldsymbol{x}_i, \quad \sum_{i=0}^{s} \lambda_i = 1.$$

1.2 Weight/Distance Functions

In this subsection, we review several distance function definitions of an arbitrary code $\mathscr{C}^{(\mathsf{M},n)}$ in the literature. For the sake of simplicity, we sometimes will write $\mathscr{C}_s^{\mathrm{aff}} = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$ as a set of $(s+1)$ a.i. codewords, and $\mathscr{C}_r = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\}$ simply denotes a set of r codewords.

Definition 2 (Generalized Hamming Weight [3])**.** For a length-n binary vector $\boldsymbol{x} = (x_1 \ \ldots \ x_n)$, the support of $\boldsymbol{x}$ is the set of nonzero coordinates of $\boldsymbol{x}$, *i.e.*, $\chi(\boldsymbol{x}) \triangleq \{j \in \{1, \ldots, n\} \colon x_j \neq 0\}$. Moreover, the support of a set of vectors $\mathscr{C}_r$ is defined as follows.

$$\chi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r) \triangleq \bigcup_{i=1}^{r} \chi(\boldsymbol{x}_i).$$

Denote by $|\mathcal{I}|$ the cardinality of a set $\mathcal{I}$. The generalized Hamming weight $w(\mathscr{C}_r)$ of a code is defined as the cardinality of $\chi(\mathscr{C}_r)$, *i.e.*, $w(\mathscr{C}_r) \triangleq \big|\chi(\mathscr{C}_r)\big|$.

Definition 3 (Generalized r-wise Hamming Distance)**.** The generalized r-wise Hamming distance $d(\mathscr{C}_r)$ of a set of r length-n codewords $\mathscr{C}_r$ written as (1) is defined as

$$d(\mathscr{C}_r) = n - \big|\{j \in \{1, \ldots, n\} \colon c_{j,1} = c_{j,2} = \cdots = c_{j,r}\}\big|.$$

Remark 4.

1. The r-wise Hamming distance $d_{i_1 i_2 \cdots, i_r}(\mathscr{C}^{(\mathsf{M},n)})$ defined in [1, Def. 31] is equal to $d(\mathscr{C}_r)$ for $\mathscr{C}_r = \{\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \ldots, \boldsymbol{x}_{i_r}\}$, where $\{i_1, \ldots, i_r\} \subseteq \{1, \ldots, \mathsf{M}\}$.

2. For a given $\mathscr{C}_r$, $d(\mathscr{C}_r)$ is not necessarily equal to $w(\mathscr{C}_r)$, *e.g.*, $d(\{(0\,1\,0\,1), (1\,1\,0\,0)\}) = 2 < w(\{0\,1\,0\,1, 1\,1\,0\,0\}) = |\{1, 2, 4\}| = 3$.

3. For any set of vectors $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$,

$$d(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s) = d(\boldsymbol{0}, \boldsymbol{x}_1 - \boldsymbol{x}_0, \ldots, \boldsymbol{x}_s - \boldsymbol{x}_0) = w(\boldsymbol{x}_1 - \boldsymbol{x}_0, \ldots, \boldsymbol{x}_s - \boldsymbol{x}_0).$$

4. If a set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s\}$ is l.i. for $s \geq 1$, we must have

$$\begin{aligned} d(\boldsymbol{0}, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s) &= d(\mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s)) \\ &= w(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s) \text{ and} \\ w(\mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s)) &\leq w(\mathrm{span}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{s+1})), \end{aligned}$$

   where $\mathrm{span}(\cdot)$ represents the vector subspace spanned by a set of l.i. vectors.

△

1.3 Minimum Generalized Hamming Distance/Weight

After defined the weight functions properly, we now begin to define the minimum metric that we would like to compare.

Definition 5 (*s*th Generalized Hamming Weight [3])**.** The *sth generalized Hamming weight* is defined as the smallest support cardinality among all possible s-dimensional subcodes of $\mathscr{C}$. *I.e.*,

$$w_s(\mathscr{C}) \triangleq \min\{w(\mathscr{D}_s)\colon \mathscr{D}_s \subseteq \mathscr{C}, \mathscr{D}_s \text{ is linear and } \operatorname{rank}(\mathscr{D}_s) = s\}.$$

Definition 6 (Minimum r-Wise Hamming Distance)**.** The minimum r-Wise Hamming distance $d(\mathscr{C}_r)$ of a code $\mathscr{C}$ is the minimum of all possible r-wise Hamming distances $d(\mathscr{C}_r)$ for $\mathscr{C}_r \subseteq \mathscr{C}$. We denoted it by $d_r(\mathscr{C}) \triangleq \min\{d(\mathscr{C}_r)\colon \mathscr{C}_r \subseteq \mathscr{C}\}$.

Definition 7 (Generalized s-Distance [4])**.** The *generalized s-distance*, denoted by $d_s^{\text{aff}}(\mathscr{C})$, is the minimum cardinality of the support of $(s+1)$ a.i. codewords in $\mathscr{C}$, which is defined as $d_s^{\text{aff}}(\mathscr{C}) \triangleq \min\{d(\mathscr{C}_s)\colon \mathscr{C}_s^{\text{aff}} \subseteq \mathscr{C}\}$.

2 Main Theorem

In order to complete the proof of [2, Proposition 6], some lemmas have to be introduced first.

Lemma 8. *Let $\mathbb{F}_2^n$ denote the vector space of all length n binary vectors over the finite field $\mathbb{F}_2 \triangleq \{0,1\}$. Then any subset $\mathcal{A} \subseteq \mathbb{F}_2^n$ with $|\mathcal{A}| \geq 2^{s-1}+1$ must contain at least $(s+1)$ a.i. vectors.*

Proof. If the claim is not true, say there are only s a.i. vectors in $\mathcal{A}$, denote those a.i. vectors by $\mathcal{S} = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{s-1}\}$. Hence, by Definition 1, in $\mathbb{F}_2^n$ there are at most 2^{s-1} vectors that are affinely dependent to $\mathcal{S}$. Thus, $\operatorname{aff}(\mathcal{S})$ can be seen as the affine hull of $\mathcal{A}$,[1] and hence $\mathcal{A} \subseteq \operatorname{aff}(\mathcal{S})$. We then have $|\mathcal{A}| \leq 2^{s-1}$, which leads to a contradiction. □

Lemma 9. *Let $\mathcal{S} = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\} \subseteq \mathbb{F}_2^n$ be a.i. If $\boldsymbol{x} \in \operatorname{aff}(\mathcal{S})$, then we have*

$$d(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s) = d(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_s, \boldsymbol{x}).$$

Proof. By definition, we have $\boldsymbol{x} = \sum_{i=0}^{s} \lambda_i \boldsymbol{x}_i$ with $\sum_{i=1}^{s} \lambda_i = 1$. Therefore, if $c_{j,0} = c_{j,1} = \cdots = c_{j,s} = b \in \{0,1\}$, the jth coordinate of $\boldsymbol{x}$, denoted by $(\boldsymbol{x})_j$, becomes

$$(\boldsymbol{x})_j = \sum_{i=0}^{s} \lambda_i (c_{j,i}) = c_{j,i} \sum_{i=0}^{s} \lambda_i = b \cdot 1 = b.$$

This completes the proof (see Definition 3). □

We further introduce a basic lemma in linear vector spaces.

Lemma 10. *Let $\mathcal{V}$ be a finite-dimensional vector space with $\dim(\mathcal{V}) = v$, and $\mathcal{U} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_u\}$ be any subset of u l.i. vectors in $\mathcal{V}$, then there exists $\{\boldsymbol{x}_{u+1}, \ldots, \boldsymbol{x}_v\}$ such that $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_u, \boldsymbol{x}_{u+1}, \ldots, \boldsymbol{x}_v\}$ is l.i., and $\mathcal{V} = \operatorname{span}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_u, \boldsymbol{x}_{u+1}, \ldots, \boldsymbol{x}_v)$. In other words, $\mathcal{U}$ is a sub-basis of V.*

[1] The affine hull of a set is the smallest affine set containing the set.

Now, we are ready to accomplish the main goal.

Theorem 11 ([2, Proposition 6]). *For a given binary k-dimensional linear code $\mathscr{C}_{\mathrm{lin}}$ with length n, we have $d_s^{\mathrm{aff}}(\mathscr{C}_{\mathrm{lin}}) = w_s(\mathscr{C}_{\mathrm{lin}})$. Moreover, the minimum r-wise Hamming distance d_r for $2^{s-1}+1 \le r \le 2^s$ is equal to the generalized s-distance d_s^{aff} with $1 \le s \le k$. In other words,*

$$d_r(\mathscr{C}_{\mathrm{lin}}) = d_s^{\mathrm{aff}}(\mathscr{C}_{\mathrm{lin}}) \qquad \text{for } s = \lceil \log_2 r \rceil.$$

Proof. First, it is noted that every r codewords of a linear code $\mathscr{C}_{\mathrm{lin}}$ with its generator matrix G can be written as

$$\mathscr{C}_r = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r\} = \mathsf{U}_{r\times k}\mathsf{G}_{k\times n}.$$

Hence, we have

$$\begin{aligned}
d_r\big(\mathscr{C}_{\mathrm{lin}}\big) &= \min\big\{d(\mathscr{C}_r)\colon \mathscr{C}_r = \mathsf{U}_{r\times k}\mathsf{G}_{k\times n}\big\} \\
&\le \min\{d(\mathscr{C}_r)\colon \mathscr{C}_r = \mathsf{U}_{r\times k}\mathsf{G}_{k\times n} \text{ and } \operatorname{rank}(\mathscr{C}_r) = s\} \\
&= \min\{d(\operatorname{span}(\boldsymbol{x}_1-\boldsymbol{x}_0,\ldots,\boldsymbol{x}_s-\boldsymbol{x}_0))\colon \text{l.i. } \{\boldsymbol{x}_1-\boldsymbol{x}_0,\ldots,\boldsymbol{x}_s-\boldsymbol{x}_0\} \subseteq \mathscr{C}_{\mathrm{lin}}\} \quad (2)\\
&= \min\{d(\boldsymbol{0},\boldsymbol{x}_1-\boldsymbol{x}_0,\ldots,\boldsymbol{x}_s-\boldsymbol{x}_s)\colon \text{l.i. } \{\boldsymbol{x}_1-\boldsymbol{x}_0,\ldots,\boldsymbol{x}_s-\boldsymbol{x}_0\} \subseteq \mathscr{C}_{\mathrm{lin}}\} \quad (3)\\
&= \min\{d(\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_s)\colon \text{a.i. } \{\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_s\} \subseteq \mathscr{C}_{\mathrm{lin}}\} \quad (4)\\
&= d_s^{\mathrm{aff}}(\mathscr{C}_{\mathrm{lin}}) \\
&= \min\{w(\boldsymbol{x}_1-\boldsymbol{x}_0,\ldots,\boldsymbol{x}_s-\boldsymbol{x}_0)\colon \text{l.i. } \{\boldsymbol{x}_1-\boldsymbol{x}_0,\ldots,\boldsymbol{x}_s-\boldsymbol{x}_0\} \subseteq \mathscr{C}_{\mathrm{lin}}\} \\
&= w_s(\mathscr{C}_{\mathrm{lin}}) \quad (5)
\end{aligned}$$

where (2) holds since $\operatorname{rank}(\mathscr{C}_r) = s$, (3) follows from the 4th remark of Remark 4, and (4) holds from the linearity of the code.

Conversely, since

$$\begin{aligned}
d_r(\mathscr{C}_{\mathrm{lin}}) &= \min\big\{d(\mathscr{C}_r)\colon \mathscr{C}_r = \mathsf{U}_{r\times k}\mathsf{G}_{k\times n}\big\} \\
&= \min\big\{d(\mathscr{C}_r)\colon \text{ a.i. } \{\boldsymbol{x}_{i_0},\boldsymbol{x}_{i_1},\ldots,\boldsymbol{x}_{i_t}\} \subseteq \mathscr{C}_r = \mathsf{U}_{r\times k}\mathsf{G}_{k\times n} \text{ for } t \ge s\big\} \quad (6)\\
&= \min\big\{d(\boldsymbol{x}_{i_0},\boldsymbol{x}_{i_1},\ldots,\boldsymbol{x}_{i_t})\colon \\
&\qquad \text{a.i. } \{\boldsymbol{x}_{i_0},\boldsymbol{x}_{i_1},\ldots,\boldsymbol{x}_{i_t}\} \subseteq \mathscr{C}_r = \mathsf{U}_{r\times k}\mathsf{G}_{k\times n} \text{ for } t \ge s\big\} \quad (7)\\
&= \min\big\{w(\boldsymbol{x}_{i_1}-\boldsymbol{x}_{i_0},\ldots,\boldsymbol{x}_{i_t}-\boldsymbol{x}_{i_0})\colon \\
&\qquad \text{l.i. } \{\boldsymbol{x}_{i_1}-\boldsymbol{x}_{i_0},\ldots,\boldsymbol{x}_{i_t}-\boldsymbol{x}_{i_0}\} \subseteq \mathscr{C}_r \text{ for } t \ge s\big\} \quad (8)\\
&= \min\big\{w(\boldsymbol{x}_{j_1}-\boldsymbol{x}_{j_0},\ldots,\boldsymbol{x}_{j_s}-\boldsymbol{x}_{j_0})\colon \text{ l.i. } \{\boldsymbol{x}_{j_1}-\boldsymbol{x}_{j_0},\ldots,\boldsymbol{x}_{j_s}-\boldsymbol{x}_{j_0}\} \subseteq \mathscr{C}_r\big\} \quad (9)\\
&= \min\big\{d(\boldsymbol{x}_{j_0},\boldsymbol{x}_{j_1},\ldots,\boldsymbol{x}_{j_s})\colon \text{a.i. } \{\boldsymbol{x}_{j_0},\boldsymbol{x}_{j_1}\ldots,\boldsymbol{x}_{j_s}\} \subseteq \mathscr{C}_r\big\} \quad (10)\\
&\ge d_s^{\mathrm{aff}}(\mathscr{C}_{\mathrm{lin}})
\end{aligned}$$

where (6) follows from Lemma 8; (7) holds since any $\mathscr{C}_r$ in $\mathscr{C}_{\mathrm{lin}}$ is a subset of an affine subspace and Lemma 9 claims that in order to compute the generalized Hamming distance of an affine subspace, it is sufficient to evaluate the generalized Hamming distance of its a.i. vectors of the affine subspace; (9) follows from Lemma 10 and the 4th remark of Remark 4; finally, (8) and (10) hold because of the 3rd remark of Remark 4.

Therefore, we have

$$\begin{aligned}
d_s^{\mathrm{aff}}(\mathscr{C}_{\mathrm{lin}}) = \min\big\{d(\mathscr{C}_s^{\mathrm{aff}})\colon \mathscr{C}_s^{\mathrm{aff}} \subseteq \mathscr{C}_{\mathrm{lin}}\big\} &\le \min\big\{d(\mathscr{C}_s^{\mathrm{aff}})\colon \mathscr{C}_s^{\mathrm{aff}} \subseteq \mathscr{C}_r \subseteq \mathscr{C}_{\mathrm{lin}}\big\} \\
&= d_r(\mathscr{C}_{\mathrm{lin}}). \quad (11)
\end{aligned}$$

The proof is completed by combing the inequalities (5) and (11). □

References

[1] Hsuan-Yin Lin, Stefan M. Moser, and Po-Ning Chen, "Weak flip codes and their optimality on the binary erasure channel," 2018, to appear in *IEEE Transactions on Information Theory*. Available: https://arxiv.org/abs/1711.03310

[2] Christine Bachoc and Gilles Zémor, "Bounds for binary codes relative to pseudo-distances of k points," *Advances in Mathematics of Communications*, vol. 4, no. 4, pp. 547–565, 2010.

[3] Victor K. Wei, "Generalized Hamming weights for linear codes," *IEEE Transactions on Information Theory*, vol. 37, no. 5, pp. 1412–1418, September 1991.

[4] Gérard Cohen, Simon Litsyn, and Gilles Zémor, "Upper bounds on generalized distances," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 2090–2092, November 1994.