An Algebraic View to Gradient Descent Decoding

M. Borges Quintana and M.A. Borges Trenard
Facultad de Matemática y Computación
Universidad de Oriente
Santiago de Cuba, Cuba
mijail@mbq.uo.edu.cu  mborges@mabt.uo.edu.cu

I. Márquez-Corbella and E. Martínez-Moro
SINGACOM group
Universidad de Valladolid
Castilla, Spain
http://www.singacom.uva.es
imarquez@agt.uva.es edgar@maf.uva.es

Abstract—There are two gradient descent decoding procedures for binary codes proposed independently by Liebler and by Ashikhmin and Barg. Liebler in his paper [15] mentions that both algorithms have the same philosophy but in fact they are rather different. The purpose of this communication is to show that both algorithms can be seen as two ways of understanding the reduction process to be different (see section II for formal definitions of the algorithms). The main tool used for showing this is the Gröbner representation of the monoid associated to the linear code.

I. INTRODUCTION

From now on a code $C$ will be a binary linear code of length $n$ and dimension $k$, i.e. a $k$-dimensional linear subspace of $\mathbb{F}_2^n$ where $\mathbb{F}_2$ is the field of two elements. Let $d(\cdot, \cdot)$, $wt(\cdot)$ be the Hamming distance and the Hamming weight on $\mathbb{F}_2^n$ respectively. Let $d$ denote the minimal Hamming distance of the code $C$.

Given a code $C$ and let $r$ be a received word in $\mathbb{F}_2^n$ the complete decoding problem (CDP) addresses to determine a codeword $c \in C$ that is closest to $r$. The $t$-bounded distance decoding problem (t-BDP) is to determine a codeword $c \in C$ such that $d(r, c) \leq t$ if such codeword exists. If $t = \left\lfloor \frac{d-1}{2}\right\rfloor$ then the solution of the t-BDP is unique and if $t = \rho$ the covering radius then t-BDP is the same as CDP. Both problems are quite related to the coset weights problem (t-CWP) that can be stated as follows, given a binary $r \times n$ matrix and an $r$-dimensional vector $s$ and $t \in \mathbb{Z}_{\geq 0}$, does a binary vector $e \in \mathbb{F}_2^j$ exist such that $w(e) \leq t$ and $He = s$? All these problems have been shown to be NP-complete [2], [3] even if preprocessing is allowed [10].

Recently complete decoding and particularly gradient descent complete decoding have gain new interest related to the decoding of LDPC codes, in fact Liebler in [15] says that there is not a clear answer to the question of which parameters of a code could help to recognize and to implement a gradient descent decoding function for the code having the the coset leaders as output. Moreover in the same paper the author makes a distinction between two gradient descent decoding algorithms (GDDA) that we will denote by leader GDDA (l-GDDA) and test-set GDDA (ts-GDDA) that are claimed to be different (see section II for formal definitions of the algorithms).

The purpose of this work is to show that both algorithms can be seen as two ways of understanding the reduction process within algebraic monoid structure related to the code. For that aim the main tool used will be the Gröbner representation of the monoid associated to the linear code [4]. The structure of the paper will be as follows. Section II will show the two gradient descent decoding algorithms, Section III will give a brief review to the Gröbner representation of a code and its associate structures, Section IV will show the main result, i.e. how the two GDD algorithms can be seen as reduction associated to the Gröbner representation of the code.

II. GRADIENT DESCENT DECODING ALGORITHMS

In this section we will briefly describe two gradient descent decoding algorithms. The first one will be the leader GDDA and can be stated as follows. Let us denote by $y$ the coset in $\mathbb{F}_2^n / C$ containing $y$ and $wt(y)$ the weight of one of its leaders.

Algorithm 1. l-GDDA

Input: $r$ the received word.

Output: A codeword $c \in C$ that is closest to $r$

Repeat until $wt(r) = 0$

a) Compute $r'$ such that $wt(r - r') = 1$ and $wt(r) \geq wt(r')$

b) $r \leftarrow r'$

Return $c = r$.

Note that in each of the steps of the algorithm the vector $r$ changes between different cosets of $\mathbb{F}_2^n / C$ until it arrives to the $0$ coset, i.e. the code itself. This is essentially the same as syndrome decoding broken up in smaller steps. The paper [15] presents the first such construction method of a gradient function $\gamma : \mathbb{F}_2^n / C \rightarrow \mathbb{Z}$ such that is a strictly increasing function of $wt(r)$ for performing such a l-GDDA.

For understanding the next GDD algorithm we will need some knowledge of minimal (support) codewords. The support of a codeword $c \in C$ will be the set of its non-zero positions, i.e. supp$(c) = \{i \mid c_i \neq 0\}$.

Definition 2. A codeword $m$ in the code $C$ is said to be minimal if there is no other codeword $c \in C$ such that supp$(c) \subseteq$ supp$(m)$.

We will denote by $M_C$ the set of all the minimal codewords of $C$. The usual way of defining a steepest descent method in the Hamming space is to construct a test set $T \subseteq \mathbb{F}_2^n$. A test set is a set of codewords such that every word $y$ either lies in $V(0)$ (the Voronoy region of the all-zero vector) or there is a
t \in \mathcal{T} \text{ such that } \text{wt}(y - t) < \text{wt}(y). \text{ The gradient-like or test set GDDA is stated as follows (see \cite{2} for further details and correctness of the algorithm)}.

**Algorithm 3. ts-GDDA**

**Input:** r the received word.

**Output:** A codeword \( c \in \mathcal{C} \) is closest to r

\( c \leftarrow 0 \)

Repeat until no \( t \in \mathcal{T} \) is found such that

\( \text{wt}(r - t) < \text{wt}(r) \)

a) \( c \leftarrow c + t \)

b) \( r \leftarrow r - t \)

Return c.

It is pointed in \cite{2} that setting \( \mathcal{T} = \mathcal{M}_C \) in the previous Algorithm 3 the so call minimal vector algorithm performs complete minimum distance decoding.

### III. Gröbner Representation and Related Structures

In this section we will show some basic results on the Gröbner representation of a code \( C \). In fact it is related to the additive representation of the monoid \( \mathbb{F}_2^n/C \). We will try to keep this section as Gröbner basis technology-free as possible. For some references on Gröbner representations of codes and its implementations see \cite{4,5,6,7}. Let \( e_i \in \mathbb{F}_2^n \) be the vector with all its entries 0 but a 1 in the \( i \)-th position.

**Definition 4.** A Gröbner representation of \( \mathbb{F}_2^n/C \) is a pair \( N, \phi \) where \( N \) is a transversal of the cosets in \( \mathbb{F}_2^n/C \) (i.e. one element of each coset) such that \( 0 \in N \) and for each \( n \in N \setminus \{0\} \) there exists a \( e_i, i \in \{1, 2, \ldots, n\} \) such that \( n = n' + e_i \) with \( n' \in N \) and a mapping

\[ \phi : N \times \{e_1, e_2, \ldots, e_n\} \to N \]

such that the image of the pair \((n, e_i)\) is the element representing the coset that contains \( n + e_i \).

The word Gröbner is not casual as we will see it with the following construction. Let us consider the binomial ideal

\[ \mathcal{I}_C = \langle \{x^{aw_1} - x^{aw_2} | w_1 - w_2 \in C\} \rangle \subseteq \mathbb{K}[x_1, \ldots, x_n] \]

where the characteristic crossing function \( \triangledown : \mathbb{Z}_2^n \to \mathbb{Z}^n \) replaces the class of 0,1 by the same symbols regarded as integers; \( \mathbb{K} \) is an arbitrary field and if \( \triangledown w = (w_1, \ldots, w_n) \) then \( x^{aw} = \prod x_i^{a_i} \). If we consider a degree compatible ordering \( \prec \) we can compute a Gröbner basis \( \mathcal{G}_\prec \) w.r.t. the ideal \( \mathcal{I}_C \) the normal form of any monomial \( \prod x_i^{a_i} \) corresponds with the syndrome of the word \( \triangledown(w_1, \ldots, w_n) \) where the map \( \triangledown \) is reduction modulo 2. Thus we can take \( N \) in Definition 4 as the vectors \( \triangledown(w_1, \ldots, w_n) \) such that \( \prod x_i^{a_i} \) is a normal form w.r.t. \( \mathcal{G}_\prec \), i.e. the syndromes of the code. Note also that \( \phi \) is just given by the multiplication tables of the normal forms times the variables \( x_i \) in the ring \( \mathbb{K}[x_1, \ldots, x_n]/\mathcal{I}_C \). This is standar way of representing the quotient by an ideal \( \mathcal{I}_C \) using the FGLM algorithm (see \cite{17} chapter 29 for a complete reference on Gröbner basis topics). Moreover, the Gröbner representation of a code can be computed with a modification of the FGLM algorithm \cite{5}, one implementation in GAP \cite{11} of this algorithm can be found in \cite{7}.

The binomial ideal \( \mathcal{I}_C \) can be seen also as a kernel of a modular integer linear program problem stated as follows. Let \( H \in \mathbb{Z}^{m \times n} \) be a \( m \times n \) matrix such that \( \triangledown H \) is a parity check matrix of \( C \) and \( b \in \mathbb{Z}^m \).

\[ IP_H(b) = \begin{cases} \min \{(1, \ldots, 1) \cdot u^t\} & u \in \mathbb{Z}_{>0}^n \quad \text{H} \cdot u^t = b \mod 2. \end{cases} \]

Ikegami and Kaji \cite{12} studied the kernel of this problem related with the maximum likelihood decoding problem. It has been also studied in \cite{16} in order to describe the combinatorics of the minimal codewords of the code \( C \).

Associated to the Gröbner representation we can define the border of a code \cite{9} as follows

**Definition 5.** Let \( C \) be a code and \( H \) a parity check matrix of \( C \), let \((N, \phi)\) be a Gröbner representation of \( \mathbb{F}_2^n/C \). Then the border of the code \( C \) w.r.t. \((N, \phi)\) is the set

\[ B(C) = \{(n_i + e_i, n_2) | i \in \{1, \ldots, n\}, n_1 + e_i \neq n_2, n_1, n_2 \in N \text{ and } H \cdot (n_1 + e_i) = H \cdot n_2\}. \]

An important remark is that both components of an element in the set \( B(C) \) are in the same coset, i.e. their sum is in the code. We can also describe the border in as

\[ B(C) = \{(n + e_i, \phi(n, e_i)) | i \in \{1, \ldots, n\}, n \in N\}\setminus\{(x, x)\}. \]

The border of a code \( B(C) \) is associate to the border basis of the ideal \( \mathcal{I}_C \). The connection between the ideal \( \mathcal{G}_\prec \) comes from the well known fact that that every Gröbner basis with respect to a degree-compatible term ordering can be extended to a border basis (see \cite{14} p. 281ff) but not every border basis is an extension of a Gröbner basis. The preference of border bases over Gröbner bases in our case arises from the iterative generation of linear syzygies, inherent in the linear algebra algorithm used in \cite{8}, which allows for successively approximating the basis degree-by-degree, i.e. weight-by-weight.

### IV. Gradient Descent Decoding and Reduction

Given a code \( C \) and its corresponding Gröbner representation \((N, \phi)\) we can accomplish two types of reduction that we will see are associated to Algorithms 1, 3 above. Thus both algorithms obey to the same algebraic structure.

**A. Reduction by \( \phi \)**

We shall define the reduction of an element \( n \in N \) w.r.t. \( e_i \) as the element \( n' = \phi(n, e_i) \) and we will denote it by \( n \rightarrow n' \). For each \( y \in \mathbb{F}_2^n \), \( y = 0 + \sum y_i e_i \) for some \( i_j \in \{1, \ldots, n\} \), thus we can iterate a finite number of reductions to find the representative of the coset \( \triangledown \) containing \( y \). Note that in the case that we use \( \prec \) defined above the representatives of the classes correspond with coset leaders, we will consider that this is
the case from now on. This gives us the following gradient
descent decoding algorithm.

**Algorithm 6.** \((N, \phi)\)-reduction GDDA

**Input:** \(r\) the received word.
**Output:** A codeword \(c \in C\) that is closest to \(r\)

**Forward step**
\[ r = \sum_{j=1}^{\ell} e_{i_j}. \] Compute \(n \in N\) corresponding to the
coset \(T\), i.e.

a) \(n = 0\).

b) For \(j = 1, \ldots, s\) do
    \[ n \rightarrow_{i_j} n', \quad n \leftarrow n'. \]

**Backward step**

While \(n \neq 0\)

a) Compute \(r'\) such that \(r' = r + e_{i_j}\) and
    \[ \text{wt}(n) \geq \text{wt}(\phi(n, e_{i_j})) \]

b) \(r \leftarrow r', \quad n \leftarrow \phi(n, e_{i_j}).\)

**Return:** \(c = r\).

Note that the previous algorithm is somehow redundant,
since at the end of the forward step we end with the coset
leader \(n\) of the class \(\mathcal{T}\), thus we can decode without performing
the forward step. Anyway we have stated this way to see the resemblance with Algorithm 11. We can modify our algorithm
capturing the needed information of the Gröbner representation as follows.

**Definition 7.** Let \((N, \phi)\) Gröbner representation of \(\mathbb{F}_{2}^{n}/C\) and \(\{n_{i}\}_{i=1}^{2^{n-k}}\) an ordering on \(N\) with \(n_{1} = 0\). We will denote by \((N^{\ast}, \phi^{\ast})\) the following pair.

\[ N^{\ast} = \{(i, w_{i}) \in \mathbb{Z}_{\geq 0}^{2} \mid w_{i} = \text{wt}(n_{i}), i = 1, \ldots, 2^{n-k}\} \]

\[ \phi^{\ast} : N^{\ast} \times \{e_{1}, e_{2}, \ldots, e_{n}\} \rightarrow N^{\ast} \]

\[ (i, w_{i}) \phi^{\ast}((i, w_{i}), e_{j}) \] such that \(\phi^{\ast}((i, w_{i}), e_{j}) = (i_{j}, w_{i_{j}})\) if \(n_{i_{j}} = \phi(n_{i}, e_{j})\) and \(w_{i_{j}} = \text{wt}(n_{i_{j}})\).

In other words, we keep track only on the ordering of the normal forms representing each coset and the weight of one of its leaders. Note that \((N^{\ast}, \phi^{\ast})\) can be easily computed from a Gröbner representation \((N, \phi)\) w.r.t. a degree compatible ordering \(\prec\) for \(N^{\ast}\) is the normal forms are coset leaders. Moreover, the way of computing a Gröbner representation by FGLM techniques gives us an incremental construction of \(N^{\ast}\) ordered non-decreasingly on the second component (see 11 for further details), i.e.

\[ (i, w_{i}), (j, w_{j}) \in N^{\ast} \text{ and } i < j \Rightarrow w_{i} \leq w_{j}. \]

Now it is clear that we can decode using only \((N^{\ast}, \phi^{\ast})\), thus we can avoid storing the normal forms in the Gröbner representation.

**Algorithm 8.** \((N^{\ast}, \phi^{\ast})\)-reduction GDDA

**Input:** \(r\) the received word.

**Output:** A codeword \(c \in C\) that is closest to \(r\)

**Forward step**
\[ r = \sum_{j=1}^{\ell} e_{i_{j}}. \] Compute \(\ell\) \(e_{i_{j}}\) corresponding to the coset \(T\), i.e.

a) \(i = 1, w_{1} = 0\).

b) For \(j = 1, \ldots, s\) do
    \[ \phi^{\ast}((i, w_{i}), e_{i_{j}}) = (i', w'_{i_{j}}), \quad (i, w_{i}) \leftarrow (i', w'_{i_{j}}) \]

**Return** \(i = \ell\).

**Backward step**

While \(i \neq 1\)

a) Compute \(r'\) such that \(r' = r + e_{i_{j}}\) and
    \[ w_{i} \geq w'_{i_{j}} \]

where \(w'_{i_{j}}\) is the second component of \(\phi^{\ast}((i, w_{i}), e_{i_{j}})\).

b) \(r \leftarrow r', \quad (i, w_{i}) \leftarrow \phi^{\ast}((i, w_{i}), e_{i_{j}})\).

**Return:** \(c = r\).

As an intermediate result, from the forward step we already
know if the coset has a correctable leader if \(w_{t} \leq t = \lfloor \frac{d-1}{2} \rfloor\), in that case the backward step gives us a unique solution, if \(t > t\) then there could be multiple ways of doing the backtracking step depending on the number of leaders in the \(t\)-coset. Also this algorithm can be used to answer the \(t\)-CWP problem.

Note that the backward step is just the \(1\)-GDDA in Algorithm 11 As pointed by Liebler 15 in each step of the backtracking procedure we change of coset till we arrive to the \(\mathcal{U}\) coset.

**B. Border reduction**

Now taking into account the information on the border of the code \(B(C)\), we can make a similar reduction subtituting in each step the first component of an element of the border by the second one. More formally, let \((b_{1}, b_{2}) = b \in B(C)\), we define the head and the tail of \(b\) as

\[ \text{head}(b) = b_{1}, \quad \text{tail}(b) = b_{2} \in \mathbb{F}_{2}^{n}. \]

As pointed before \(\text{head}(b) + \text{tail}(b)\) is a codeword of \(C\) for all \(b \in B(C)\) and by its definition \(3\) it is clear that the information in the border allows complete decoding. The information in the border is somehow redundant, we can reduce the number of codeword in it needed for decoding.

**Definition 9.** A set \(R(C)\) is the reduced border for the code \(C\) with respect to the order \(\prec\) if \(R(C) \subseteq B(C)\) and it fulfills the following conditions:

1) For each pair \((n, e_{i})\) such that \(n + e_{i}\) is a head in \(B(C)\) there exists an element in \(R(C)\) such that its head is \(h\) where
    \[ \text{supp}(h) \subseteq \text{supp}(n + e_{i}). \]

2) Given two elements in \(R(C)\) and \(h_{1}, h_{2}\) their heads, then we have that
    \[ \text{supp}(h_{1}) \not\subseteq \text{supp}(h_{2}) \text{ and } \text{supp}(h_{2}) \not\subseteq \text{supp}(h_{1}). \]
Thus \( R(C) \) is the set with smallest cardinal that allows us a gradient-like test set decoding using reductions.

**Proposition 10.** Let us consider the set of codewords in \( C \) given by

\[
\text{Min}_{\text{red}}(C) = \{ \text{head}(b) + \text{tail}(b) \mid b \in R(C) \} \subseteq C.
\]

Then \( \text{Min}_{\text{red}}(C) \subseteq M_C \).

**Proof:** Let \( \text{head}(b) + \text{tail}(b) = c \) where \( b \in R(C) \) and suppose \( c \notin M_C \), then there exists \( c' \in C \) such that \( \text{supp}(c') \subseteq \text{supp}(c) \). Let \( c_1 \) be a vector such that \( \text{supp}(c_1) = \text{supp}(c) \cap \text{supp}(\text{head}(b)) \), thus \( c_2 = c - c_1 \) fulfills \( \text{supp}(c_2) \subseteq \text{supp}(\text{tail}(b)) \). Let \( m \) be the maximum between \( c_1 \) and \( c_2 \), therefore \( \text{supp}(m) \subseteq \text{supp}(c) \), and \( m \) is smaller than \( \text{head}(b) \) and \( \text{tail}(b) \) which contradicts the fact that \( R(C) \) is reduced.

Therefore the set \( \text{Min}_{\text{red}}(C) \) is a minimal test set w.r.t. the order \( \prec \) given by minimal codewords that allow the ts-GDD algorithm stated in Algorithm 3. It can be also seen as a test set for the modular integer program in Equation (2).

**Conclusions**

We have shown an unified approach via the Gröbner presentation of a code to two gradient descent decoding algorithms, one that the search is done changing the coset representative (l-GDDA) and the one given by descending within the same coset (ts-GDDA) that were claimed to be of different nature. This two algorithms come from two ways of computing the reduction of a monomial modulo a binomial ideal associated to the code. Unfortunately there are some obstructions for generalizing this approach in a straightforward way to non binary codes mainly motivated by the non-admissibility of the ordering needed for decoding (see [4]). Further research lines of the authors point to generalizing the border basis for the non binary case in order to describe the set of minimal codewords of a code.

**References**

[1] A. Ashikhmin and A. Barg, Minimal vectors in linear codes IEEE Trans. Inform. Theory 44 (1998), 2010–2017.
[2] A. Barg, Complexity issues in coding theory, In Handbook of Coding Theory, Elsevier Science, Vol. 1, (1998), 649–754
[3] E.R. Berlekamp, R.J. McEliece, H.C.A. van Tilborg, On the Inherent Intractability of Certain Coding Problems IEEE Trans. Inform. Theory, IT-24, no. 3, (1978), 384–386.
[4] M. Borges-Quintana, M.A. Borges-Trenard and E. Martínez-Moro, On a Gröbner bases structure associated to linear codes, J. Discrete Math. Sci. Cryptogr. 10 (2007), no. 2, 151–191.
[5] M. Borges-Quintana, M.A. Borges-Trenard, E. Martínez-Moro, A general framework for applying FGLM techniques to linear codes, AAECC 16, Lecture Notes in Comput. Sci., 3857, (2006), 76–86.
[6] M. Borges-Quintana, M.A. Borges-Trenard, E. Martínez-Moro, A Gröbner representation of linear codes, In: T. Shaska, W.C. Huffman, D. Joyner, V. Ustimenko (eds.) Advances in Coding Theory and Cryptography, World Scientific (2007), 17–32.
[7] M. Borges-Quintana, M.A. Borges-Trenard, E. Martínez-Moro, GBLA-LC: Gröbner Bases by Linear Algebra and Linear Codes, In: ICM 2006. Mathematical Software, EMS, (2006), 604–605.
[8] M. Borges-Quintana, M.A. Borges-Trenard, P. Fitzpatrick, E. Martínez-Moro, On a Gröbner bases and combinatorics for binary codes, Appl. Algebra Engrrg. Comm. Comput. 19 (2008) 393–411.
[9] M. Borges-Quintana, M.A. Borges-Trenard, I. Márquez-Corbellá, E. Martínez-Moro, On the Border of a Binary Code. Submitted to Jour. Comp. Applied Maths.(2009).
[10] J. Bruck, M. Naor, The Hardness of Decoding Linear Codes with Preprocessing, IEEE Trans. Inform. Theory 36, no. 2, (1990)
[11] The GAP Group, GAP – Groups, Algorithms, and Programming, Version 4.12 (2009), [http://www.gap-system.org].
[12] D. Ikegami and Y. Kaji, Maximum likelihood decoding for linear block codes using Gröbner bases, IEICE Trans. Fund. Electron. Commun. Comput. Sci. E86-A , 3 (2003) 643–651.
[13] A. Kehrein and M. Kreuzer, Characterizations of border bases.Journal of Pure and Applied Algebra, 196 (2005), 251–270.
[14] A. Kehrein and M. Kreuzer, Computing border bases. Journal of Pure and Applied Algebra, 205, (2006) 279–295.
[15] R. Liebler, Implementing gradient descent decoding, Michigan Math. J. 58 , Issue 1 (2009), 285–291.
[16] I. Márquez-Corbella, E. Martínez Moro, Combinatorics of minimal codewords of some linear codes, Submitted to Advances in Mathematics of Communications (2010).
[17] T. Mora, Solving polynomial equation systems. II. Macaulay’s paradigm and Gröbner technology, Encyclopedia of Mathematics and its Applications, 99. Cambridge University Press, Cambridge, (2005).