Groebner bases and error correcting codes: from Cooper Philosophy to Degrobnerization

Michela Ceria, Teo Mora and Massimiliano Sala

Abstract. In the Late Nineties, the classical approach to decode BCH codes based on Berlekamp's *key equation* was upsetted by the application of Gröbner bases to the problem; it appeared a series of papers which terminated with two different proposals: Orsini-Sala general error locator polynomial [14] and Augot *et al.* Newton-Based decoder [1]; both approaches payed not only the hard pre-computation of a Gröbner basis but (mainly) the density of their decoders.

A recent work-in-progress [4, 5, 6, 7] reconsidered the same problem within the frame of *Grobner-free Solving*, an approach aiming to avoid the computation of a Gröbner basis of a (0-dimensional) ideal $J \subset \mathcal{P}$ in favour of combinatorial algorithms, describing instead the structure of the algebra $\mathcal{P}/J$. The consequence is a preprocessing which is quadratic (and a decoding which is linear) on the length of the code.

Extended abstract

In 1990 Cooper [10, 11] suggested to use Gröbner bases' computation in order to decode cyclic codes. Let C be a binary BCH code correcting up to t errors, $\bar{s} = (s_1, \dots, s_{2t-1})$ be the syndrome vector associated to a received word. Cooper's idea consisted in interpreting the error locations $z_1, \dots z_t$ of C as the roots of the syndrome equation system: $f_i := \sum_{j=1}^{t} z_j^{2i-1} - s_{2i-1} = 0,\ 1 \le i \le t$, and, consequently, the plain error locator polynomial as the monic generator $g(z_1)$ of the principal ideal $\left\{\sum_{i=1}^{t} g_i f_i, g_i \in \mathbb{F}_2(s_1, \dots, s_{2t-1})[z_1, \dots, z_t]\right\} \bigcap \mathbb{F}_2(s_1, \dots, s_{2t-1})[z_1]$, which was computed via the elimination property of lexicographical Gröbner bases.

In a series of papers Chen et al. improved and generalized Cooper's approach to decoding. In particular, for a q-ary $[n, k, d]$ cyclic code, with correction capability t, they made two alternative proposals.

First of all, denoting, for an error with weight μ, $z_1, \dots, z_\mu$ the error locations, $y_1, \dots, y_\mu$ the error values and $s_1, \dots, s_{n-k} \in \mathbb{F}_{q^m}$ the associated syndromes, they interpreted [8] the coefficients of the plain error locator polynomial as the elementary symmetric functions σ_j and the syndromes as the *Waring functions*, $s_i = \sum_{j=1}^{\mu} y_j z_j^i$. They suggested to deduce the σ_j's from the (known) s_i's via a Gröbner basis computation for the ideal generated by the Newton identities; a similar idea was later developed in [1].

Alternatively, they considered [9] the *syndrome variety*

$$\left\{(s_1, \ldots, s_{n-k}, y_1, \ldots, y_t, z_1, \ldots, z_t) \in (\mathbb{F}_{q^m})^{n-k+2t} : s_i = \sum_{j=1}^{\mu} y_j z_j^i,\ 1 \leq i \leq n-k\right\}$$

and proposed to deduce, via a Gröbner basis pre-computation in

$$\mathbb{F}_q[x_1, \ldots, x_{n-k}, y_1, \ldots, y_t, z_1, \ldots, z_t],$$

a series of polynomials $g_\mu(x_1, \ldots, x_{n-k}, Z), \mu \leq t$ such that, for any error with weight μ and associated syndromes $s_1, \ldots, s_{n-k} \in \mathbb{F}_{q^m}$, $g_\mu(s_1, \ldots, s_{n-k}, Z)$ in $\mathbb{F}_{q^m}[Z]$ is the plain error locator polynomial.

Their approach was improved in a series of papers which introduced further applications of groebnerian technologies and which culminated with [14] which stated

Theorem 0.1. [14] *In the Gröbner basis of the ideal vanishing in each point of the syndrome variety, there is a unique polynomial, the* general error locator polynomial*, with shape*

$$g = z_t^t + \sum_{l=1}^{t} a_{t-l}(s_1, \ldots, s_{n-k}) z_t^{t-l}.$$

Such polynomial satisfies the following property: given a syndrome vector $s = (s_1, \ldots, s_{n-k}) \in (\mathbb{F}_{q^m})^{n-k}$ corresponding to an error with weight $\mu \leq t$, its t roots are the μ error locations plus zero counted with multiplicity $t - \mu$.

For a survey of *Cooper Philosophy* see [13], see [3] for Sala-Orsini locator.

Recently the same problem has been reconsidered in a group of papers [4, 6, 5] within the frame of *Grobner-free Solving*, an approach aiming to avoid the Gröbner bases computation for (0-dimensional) ideals.

In particular, given the syndrome variety

$$\mathsf{Z} = \left\{(c+d, c^3+d^3, c, d), c, d \in \mathbb{F}_{2^m}^*, c \neq d\right\}$$

of a BCH $[2^m - 1, 2]$-code C over $\mathbb{F}_{2^m}$, and denoted $\mathcal{I}(\mathsf{Z})$ the ideal of points of Z, [4] is able with good complexity to produce, via Cerlienco-Mureddu Algorithm [2] and Lazard Theorem, the set $\mathbf{N} := \mathbf{N}(\mathcal{I}(\mathsf{Z}))$ and proves that the related Gröbner basis has the shape

$$G = (x_1^n - 1, g_2, z_2 + z_1 + x_1, g_4)$$

where (see [14]) $g_2 = \frac{x_2^{\frac{n+1}{2}} - x_1^{\frac{n+1}{2}}}{x_2 - x_1} = x_2^{\frac{n-1}{2}} + \sum_{i=1}^{\frac{n-1}{2}} \binom{\frac{n-1}{2}}{i} x_1^i x_2^{\frac{n-1}{2}-i}$ and $g_4 = z_1^2 - \sum_{t \in \mathbf{N}} c_t t$ is Sala-Orsini general error locator polynomial. Such result allowed [4] to remark (applying Marinari-Mora Theorem) that, for decoding, it is sufficient to compute the polynomial, *half error locator polynomial* (HELP)

$$h(x_1,x_2,z_1) := z_1 - \sum_{t\in\mathbf{H}} c_t t \text{ where } \mathbf{H} := \{x_1^i x_2^j, 0 \le i < n, 0 \le j < \frac{n-1}{2}\}$$

which satisfies

$$h(c(1+a^{2j+1}), c^3(1+a^{3(2j+1)}), z_1) = z_1 - c, \text{ for each } c \in \mathbb{F}_{2^m}^*, 0 \le j < \frac{n-1}{2},$$

the other error ca^{2j+1} been computable via the polynomial $z_2 + z_1 + x_1 \in G$ as $z_2 := x_1 - z_1 = (c + ca^{2j+1}) - c = ca^{2j+1}$.

Such polynomial can be easily obtained with good complexity via Lundqvist interpolation formula [12] on the set of points

$$\left\{(c + ca^{2j+1}, c^3 + c^3a^{3(2j+1)}, c), c \in \mathbb{F}_{2^m}^*, 0 \le j < \frac{n-1}{2}\right\}.$$

Experiments showed that, in that setting, HELP has a very sparse formula, which has been proved (see [4]):

$$h(x_1,x_2,z_1) = z_1 + \sum_{i=1}^{\frac{n-1}{2}} a_i x_1^{(4-3i) \mod n} x_2^{(i-1) \mod \frac{n-1}{2}}$$

where the unknown coefficient can be deduced by Lundqvist interpolation on the set of points $\{(1+a^{2j+1}, 1+a^{3(2j+1)}, 1), 0 \le j < \frac{n-1}{2}\}$ and on the terms $\{x_1^{(4-3i) \mod n} x_2^{(i-1) \mod \frac{n+1}{2}}, 1 \le i < \frac{n+1}{2}\}$.

This suggested [6] to consider a binary cyclic code C over $GF(2^m)$, with length $n \mid 2^m - 1$ and *primary* defining set $S_C = \{1, l\}$. Thus it denoted by a a primitive $(2^m-1)^{\text{th}}$ root of unity so that $\mathbb{F}_{2^m} = \mathbb{Z}_2[a]$, $\alpha := \frac{2^m-1}{n}$ and $b := a^\alpha$ a primitive n^{th} root of unity, $\mathcal{R}_n := \{e \in \mathbb{F}_{2^m} : e^n = 1\}$ and $\mathcal{S}_n := \mathcal{R}_n \sqcup \{0\}$; considered the following sets of points

$\mathsf{Z}_2 := \{(c+d, c^l+d^l, c, d), c, d \in \mathcal{R}_n, c \ne d\}, \#\mathsf{Z}_2^\times = n^2 - n;$
$\mathsf{Z}_+ := \{(c+d, c^l+d^l, c, d), c, d \in \mathcal{S}_n, c \ne d\}, \#\mathsf{Z}_+^\times = n^2 + n,$
$\mathsf{Z}_{ns} := \{(c+d, c^l+d^l, c, d), c, d \in \mathcal{S}_n\} \setminus \{(0,0,c,c), c \in \mathcal{R}_n\}, \#\mathsf{Z}_{ns}^\times = n^2+n+1,$
$\mathsf{Z}_e := \{(c+d, c^l+d^l, c, d), c, d \in \mathcal{S}_n\}, \#\mathsf{Z}_e^\times = (n+1)^2,$

and denoted, for $* \in \{e, ns, +, 2\}$,

- $J_* := \mathcal{I}(\mathsf{Z}_*)$, the ideal of all polynomials vanishing in Z_*,
- $\mathsf{N}_* := \mathbf{N}(J_*)$ the Gröbner escalier of J_* w.r.t. the lex ordering with $x_1 < x_2 < z_1 < z_2$ and
- $\Phi_* : \mathsf{Z}_* \to \mathsf{N}_*$ a Cerlienco-Mureddu correspondence [2].

Then it assumed to know

(a). the structure of the order ideal N_2, $\#\mathsf{N}_2 = n^2 - n$, i.e. a minimal basis $\{t_1, \dots, t_r\}, t_i := x_1^{a_i} x_2^{b_i}$, of the monomial ideal $\mathcal{T} \setminus \mathsf{N}_2 = \mathbf{T}(\mathfrak{I}(\mathsf{Z}_2))$,
(b). a Cerlienco Mureddu Correspodence $\Phi_2 : \mathsf{N}_2 \to \mathsf{Z}_2$

and deduced with elementary arguments N_* and Φ_* for $* \in \{e, ns, +\}$.

References

[1] D. Augot, M. Bardet, J.C. Faugere, On formulas for decoding binary cyclic codes, *Proc. IEEE Int. Symp. Information Theory 2007*, (2007) .

[2] L. Cerlienco, M. Mureddu, *From algebraic sets to monomial linear bases by means of combinatorial algorithms*, Discrete Math. **139**, (1995) 73-87.

[3] F. Caruso, E. Orsini, C. Tinnirello and M. Sala *On the shape of the general error locator polynomial for cyclic codes* IEEE Transactions on Information Theory 63.6 (2017): 3641-3657.

[4] M. Ceria, T. Mora, M. Sala, *HELP: a sparse error locator polynomial for BCH codes*, in preparation.

[5] M. Ceria *Half error locator polynomials for efficient decoding of binary cyclic codes*, in preparation.

[6] M. Ceria, *Macaulay, Lazard and the Syndrome Variety*, in preparation.

[7] M.Ceria, T. Mora, M.Sala, *Zech Tableaux as tools for sparse decoding.* accepted for publications in Rendiconti del Seminario Matematico.

[8] X. Chen, I. S. Reed, T. Helleseth, K. Truong, Use of Gröbner Bases to Decode Binary Cyclic Codes up to the True Minimum Distance, *IEEE Trans. on Inf. Th.*, **40** (1994), 1654–1661.

[9] X. Chen, I. S. Reed, T. Helleseth, K. Truong, General Principles for the Algebraic Decoding of Cyclic Codes, *IEEE Trans. on Inf. Th.*, **40** (1994), 1661–1663.

[10] A.B. III Cooper, Direct solution of BCH decoding equations, In E. Arikan (Ed.) *Communication, Control and Singal Processing*, 281–286, Elsevier (1990)

[11] A.B. III Cooper, Finding BCH error locator polynomials in one step *Electronic Letters*, **27** (1991) 2090–2091

[12] Lundqvist S., *Vector space bases associated to vanishing ideals of points.* J. Pure Appl. Algebra **214** (2010), 309-321.

[13] E. Orsini, T. Mora,. *Decoding cyclic codes: the Cooper Philosophy.* in M.Sala et al., *Groebner Bases, Coding, and Cryptography.* Springer (2009), 62–92

[14] E. Orsini, M. Sala, *Correcting errors and erasures via the syndrome variety*, J. Pure Appl. Algebra, **200** (2005), 191–226.

Michela Ceria
Department of Computer Science University of Milan Milano, Italy
e-mail: `michela.ceria@gmail.com`

Teo Mora
Department of Mathematics University of Genoa Genoa, Italy
e-mail: `theomora@disi.unige.it`

Massimiliano Sala
Department of Matheatics University of Trento Trento, Italy
e-mail: `maxsalacodes@gmail.com`