# Decoding linear codes via systems solving: complexity issues and generalized Newton identities

Stanislav Bulygin (joint work with Ruud Pellikaan)

University of Valladolid
Valladolid, Spain

March 14, 2008

# Introduction

## Decoding problem

- *Complete decoding*: Given $\mathbf{y} \in \mathbb{F}_q^n$ and a code $C \subseteq \mathbb{F}_q^n$, so that $\mathbf{y}$ is at distance $d(\mathbf{y}, C)$ from the code, find $\mathbf{c} \in C : d(\mathbf{y}, \mathbf{c}) = d(\mathbf{y}, C)$.
- *Bounded up to half the minimum distance*: Additional assumption $d(\mathbf{y}, C) \leq (d(C) - 1)/2$. Then a codeword with the above property is unique.

# Introduction

## Decoding via systems solving

One distinguishes between two concepts:

- *Generic decoding*: Solve some system $S(C)$ and obtain some "closed" formulas $F$. Evaluating these formulas at data specific to a received word $\mathbf{y}$ should yield a solution to the decoding problem. For example for $f \in F : f(syndrome(\mathbf{y}), x) = poly(x)$. The roots of $poly(x) = 0$ yield error positions – general error-locator polynomial $f$.
- *Online decoding*: Solve some system $S(C, \mathbf{y})$. The solutions should solve the decoding problem.

## Computational effort

- Generic decoding. Preprocessing: very hard. Decoding: relatively simple.
- Online decoding. Preprocessing: – . Decoding: hard.

# Quadratic system method

## Unknown syndrome

Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of $\mathbb{F}_q^n$ and let $B$ be the $n \times n$ matrix with $\mathbf{b}_1, \ldots, \mathbf{b}_n$ as rows. The *unknown syndrome* $\mathbf{u}(B, \mathbf{e})$ of a word $\mathbf{e}$ w.r.t $B$ is the column vector $\mathbf{u}(B, \mathbf{e}) = B\mathbf{e}^T$ with entries $u_i(B, \mathbf{e}) = \mathbf{b}_i \cdot \mathbf{e}$ for $i = 1, \ldots, n$.

## Structure constants

For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ define $\mathbf{x} * \mathbf{y} = (x_1 y_1, \ldots, x_n y_n)$. Then $\mathbf{b}_i * \mathbf{b}_j$ is a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_n$, so there are constants $\mu_l^{ij} \in \mathbb{F}_q$ such that $\mathbf{b}_i * \mathbf{b}_j = \sum_{l=1}^{n} \mu_l^{ij} \mathbf{b}_l$. The elements $\mu_l^{ij} \in \mathbb{F}_q$ are the *structure constants* of the basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$.

## MDS matrix

Let $B_s$ be the $s \times n$ matrix with $\mathbf{b}_1, \ldots, \mathbf{b}_s$ as rows ($B = B_n$). Then $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is an *ordered MDS basis* and $B$ an *MDS matrix* if all the $s \times s$ submatrices of $B_s$ have rank $s$ for all $s = 1, \ldots, n$.

# Quadratic system method

## Check matrix

Let $C$ be an $\mathbb{F}_q$-linear code with parameters $[n, k, d]$. W.l.o.g $n \leq q$. $H$ is a check matrix of $C$. Let $\mathbf{h}_1, \dots, \mathbf{h}_{n-k}$ be the rows of $H$. One can express $\mathbf{h}_i = \sum_{j=1}^{n} a_{ij}\mathbf{b}_j$ for some $a_{ij} \in \mathbb{F}_q$. In other words $H = AB$ where $A$ is the $(n-k) \times n$ matrix with entries $a_{ij}$.

## Known syndrome

Let $\mathbf{y} = \mathbf{c} + \mathbf{e}$ be a received word with $\mathbf{c} \in C$ and $\mathbf{e}$ an error vector. The syndromes of $\mathbf{y}$ and $\mathbf{e}$ w.r.t $H$ are equal and known: $s_i(\mathbf{y}) := \mathbf{h}_i \cdot \mathbf{y} = \mathbf{h}_i \cdot \mathbf{e} = s_i(\mathbf{e})$. They can be expressed in the unknown syndromes of $\mathbf{e}$ w.r.t $B$: $s_i(\mathbf{y}) = \sum_{j=1}^{n} a_{ij} u_j(\mathbf{e})$ since $\mathbf{h}_i = \sum_{j=1}^{n} a_{ij}\mathbf{b}_j$ and $\mathbf{b}_j \cdot \mathbf{e} = u_j(\mathbf{e})$.

# Quadratic system method

## Linear forms

Let $B$ be an MDS matrix with structure constants $\mu_l^{ij}$. Define $U_{ij}$ in the variables $U_1, \ldots, U_n$ by $U_{ij} = \sum_{l=1}^{n} \mu_l^{ij} U_l$.

## Quadratic system

The ideal $J(\mathbf{y})$ in $\mathbb{F}_q[U_1, \ldots, U_n]$ is generated by

$$\sum_{l=1}^{n} a_{jl} U_l \quad - \quad s_j(\mathbf{y}) \quad \text{for} \quad j = 1, \ldots, r$$

The ideal $I(t, \mathcal{U}, V)$ in $\mathbb{F}_q[U_1, \ldots, U_n, V_1, \ldots, V_t]$ is generated by

$$\sum_{j=1}^{t} U_{ij} V_j \quad - \quad U_{it+1} \quad \text{for} \quad i = 1, \ldots, n$$

Let $J(t, \mathbf{y})$ be the ideal in $\mathbb{F}_q[U_1, \ldots, U_n, V_1, \ldots, V_t]$ generated by $J(\mathbf{y})$ and $I(t, \mathcal{U}, V)$.

# Quadratic system method

## Main result

Let $B$ be an MDS matrix with structure constants $\mu_l^{ij}$. Let $H$ be a check matrix of the code $C$ such that $H = AB$ as above. Let $\mathbf{y} = \mathbf{c} + \mathbf{e}$ be a received word with $\mathbf{c} \in C$ the codeword sent and $\mathbf{e}$ the error vector. Suppose that $\mathrm{wt}(\mathbf{e}) \neq 0$ and $\mathrm{wt}(\mathbf{e}) \leq \lfloor (d(C) - 1)/2 \rfloor$. Let $t$ be the smallest positive integer such that $J(t, \mathbf{y})$ has a solution $(\mathbf{u}, \mathbf{v})$ over $\overline{\mathbb{F}_q}$. Then

- $\mathrm{wt}(\mathbf{e}) = t$ and the solution is unique satisfying $\mathbf{u} = \mathbf{u}(\mathbf{e})$.
- the reduced Gröbner basis $G$ for the ideal $J(t, \mathbf{y})$ w.r.t any monomial ordering is

$$\begin{cases} U_i - u_i(\mathbf{e}), i = 1, \dots, n, \\ V_j - v_j, j = 1, \dots, t, \end{cases}$$

  where $(\mathbf{u}(\mathbf{e}), \mathbf{v})$ is the unique solution.

# Quadratic system method

## Features

- No field equations.
- The same result holds for the complete decoding.
- The solution lies in the field $\mathbb{F}_q$.
- The equations are at most quadratic.
- After solving $J(t, \mathbf{y})$ decoding is simple:
  $B^{-1}\mathbf{u}(B, \mathbf{e}) = B^{-1}B\mathbf{e}^T = \mathbf{e}^T$.

# Quadratic system method

## Analysis

From $J(\mathbf{y})$ one can express some $n-k$ $U$-variables via $k$ others. Substitution of those in $I(t,\mathcal{U},V)$ yields a systems of $n$ quadratic equations in $k+t$ variables, thus obtaining *overdetermined* system. Easier to solve when

- With constant $k$ and $t$, $n$ increases.
- With constant $n$ and $t$, $k$ decreases.

## Simulations

For example for random binary codes with $n=120$ and $k=10,\dots,40$ one can correct $5-20$ errors in $\leq 1000$ sec. via computing the reduced Gröbner basis in SINGULAR or MAGMA.

# Quadratic system method

## Generic solving

Generic relation between known and unknown syndromes for arbitrary linear code:

$$g_{ij}(S_1, \dots, S_{n-k})U_i = f_{ij}(S_1, \dots, S_{n-k}),$$

where $g_{ij}, f_{ij} \in \mathbb{F}_q[X_1, \dots, X_{n-k}]$ for all $i, j$ are defined over an MDS extension $\mathbb{F}_q$. We conjecture that for cyclic codes the relation is

$$U_i = f_i(S_1, \dots, S_{n-k}).$$

# Complexity issues

## Diagonal representation (joint with S.Ovsienko)

Our system is equivalent to

$$\begin{aligned} HX^T &= \mathbf{s} \\ X_i Y_i &= 0, i = 1, \ldots, n \\ \hat{H}_t Y^T &= \hat{\mathbf{s}}, \end{aligned}$$

where $H$ is a check matrix of the code $C$, $\mathbf{s}$ a known syndrome, $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ are new variables, $\hat{H}_t$ is a check matrix of a code with the generator matrix $B_t$, $\hat{\mathbf{s}}$ is a syndrome of the vector $\mathbf{b}_{t+1}$ w.r.t to $\hat{H}_t$.

# Complexity issues

## Macaulay matrix

Like above one can obtain a system Sys with $n$ quadratic equations and $k+t$ variables, w.l.o.g $X_1,\ldots,X_k, Y_1,\ldots,Y_t$. The monomials that appear in the system are $X_iY_j, 1 \le i \le k, 1 \le j \le t$, $X_1,\ldots,X_k, Y_1,\ldots,Y_t$. The total number of monomials appearing in the system is $kt+k+t=(k+1)(t+1)-1$. One can consider the *Macaulay matrix* of Sys: rows are indexed by the equations, columns by the monomials. Denote the matrix by $M(\text{Sys})$.

## Linearization

If $n \ge kt+k+t$ and $M(\text{Sys})$ is full-rank, one can find $U_i$'s by applying Gaussian elimination to $M(\text{Sys})$.

# Complexity issues

## Macaulay matrix is full-rank

Let $C$ be a random $[n, k]$ code over $\mathbb{F}_q$, defined e.g. by a random full-rank $(n-k) \times n$ check matrix $H$ and let $\mathbf{e}$ be a random error vector over $\mathbb{F}_q$ of weight $t$. Let $\mathrm{Sys} = \mathrm{Sys}(n, k, t)$ be the corresponding system as above. Then the probability of the fact that $M(\mathrm{Sys})$ has full-rank tends to 1 as $n$ tends to infinity.

## Idea of the proof

Degeneracy of $M(\mathrm{Sys})$ is reduced to the fact that

$$\mathbf{e}_I + C_I \subseteq (\widetilde{B_{t+1}})^{\perp}$$

Here $\mathbf{e}_I$ and $C_I$ are the vector $\mathbf{e}$ and the code $C$ resp. restricted to some $I$ positions from $\{1, \dots, n\}$ and $\widetilde{B_{t+1}}$ is a code equivalent to the code $B_{t+1}$ restricted to the same $I$ positions as before.

# Complexity issues

## Nice behavior

Macaulay matrix $M(\text{Sys})$ is almost always full-rank already for the moderate values of $n$ and $k$, e.g. already for $n = 20, \dots, 30$ and $k = 3, \dots, 6$ the probability of being full-rank is $\geq 70\%$.

## Polynomial-time decoding

Suppose that $n \geq kt + k + t$. Then complexity of finding $U_1, \dots, U_k$ via Gaussian elimination applied to $M(sys)$ is

$$\max\{(kt + k + t + 1)^3, n(kt + k + t + 1)\lceil \log_2 n \rceil\},$$

due to the fact that $M(\text{Sys})$ has many non-degenerate square submatricies. As a consequence, if $k = \mathcal{O}(n^\alpha)$ and $t = \mathcal{O}(n^\beta)$ for $0 < \alpha + \beta \leq 1, \alpha > 0, \beta > 0$ then the complexity of the algorithm above is $\mathcal{O}(n^{3(\alpha+\beta)})$.

# Complexity issues

## Extended linearization

One can try to go further and apply *extended linearization*. Consider binary case, so $X_i^2 = X_i$ for all $i$. Multiply the system Sys with all monomials in $X_1, \ldots, X_k$ of degree $s < k$. A system, call it $\mathrm{Sys}_s$, obtained in this way has $n(1 + \binom{k}{1} + \cdots + \binom{k}{s})$ equations and $C_s := C_{s-1} + \binom{k}{s+1}(t+1)$ monomials. Denote $\binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{s} =: f(k,s)$. If we assume that $M(\mathrm{Sys}_s)$ is full-rank, then if

$$n(1+\binom{k}{1}+\cdots+\binom{k}{s}) = nf(k,s) \geq C_s - 1 = (t+1)f(k,s+1)-1,$$

then successfull application of Gaussian elimination to $M(\mathrm{Sys})$ is possible. Study this further!

# Complexity issues

## Comparing with different random systems

Consider different types of random systems

- $R_1$ is a system of $n$ quadratic equations that has the same monomials as Sys, but the corresponding coefficients are randomly taken from $\mathbb{F}_q$. Require that $R_1$ has a unique solution in $\overline{\mathbb{F}_q}$.
- $R_2$ is a system that has the same properties as $R_1$, but the requirement on uniqueness of a solution is dropped out.
- $R_3$ is a fully random system of $n$ quadratic equations, i.e. it has all possible monomials of degree $\leq 2$ and the corresponding coefficients are random from $\mathbb{F}_q$

Note that $R_2$ and $R_3$ do not have solutions in general.

# Complexity issues

## Experiments

Using some experimental evidence we **conjecture** that there are following relations between the complexities for solving *Sys*, $R_1$, $R_2$, and $R_3$ with "general methods"

$$Compl(Sys) \approx Compl(R_1) \approx Compl(R_2) \ll Compl(R_3).$$

## Semi-regular sequences

Solving $R_3$-systems has to do with the *semi-regular sequences* introduced by M.Bardet *et.al.* Complexity estimates for the F5 algorithm are available. Note that these estimates would only give a poor upper bound in our situation.

# Generalized Newton identities

## Background on cyclic codes

Assume $(q, n) = 1$. Let $\mathbb{F} = \mathbb{F}_{q^m}$ be the splitting field of $X^n - 1$ over $\mathbb{F}_q$. Let $a$ be a *primitive n-th root of unity* Denote by $S_C$ a defining set of a cyclic code $C$ of length $n$, so that $S_C = \{i_1, \ldots, i_r\} \subseteq \{1, \ldots, n\}$. Then a check matrix $H$ of $C$ can be represented as a matrix with entries in $\mathbb{F}$:

$$H = \begin{pmatrix} 1 & a^{i_1} & a^{2i_1} & \ldots & a^{(n-1)i_1} \\ 1 & a^{i_2} & a^{2i_2} & \ldots & a^{(n-1)i_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a^{i_r} & a^{2i_r} & \ldots & a^{(n-1)i_r} \end{pmatrix}.$$

# Generalized Newton identities

## Background on cyclic codes

Let $\mathbf{y} = \mathbf{c} + \mathbf{e}$, vectors are also seen as polynomials. Define $s_i = y(a^i)$ for all $i = 1, \ldots, n$. Then $s_i = e(a^i) \forall i \in S_C$, and these $s_i$ are the known syndromes. If the error vector is of weight $t$, then it is of the form

$$\mathbf{e} = (0, \ldots, 0, e_{j_1}, 0, \ldots, 0, e_{j_l}, 0, \ldots, 0, e_{j_t}, 0, \ldots, 0),$$

more precisely there are $t$ indices $j_l$ with $1 \leq j_1 < \cdots < j_t \leq n$ such that $e_{j_l} \neq 0$ for all $l = 1, \ldots, t$ and $e_j = 0$ for all $j$ not in $\{j_1, \ldots, j_t\}$. We obtain
$s_{i_m} = y(a^{i_m}) = e(a^{i_m}) = \sum_{l=1}^{t} e_{j_l} (a^{i_m})^{j_l}, 1 \leq m \leq n - k.$

- $a^{j_1}, \ldots, a^{j_t}$ and also the $j_1, \ldots, j_t$ are called the *error locations*
- $e_{j_1}, \ldots, e_{j_t}$ are called the *error values*.

# Generalized Newton identities

## GNI for cyclic codes

Define $z_l = a^{j_l}$ and $y_l = e_{j_l}$. Then $s_{i_m} = \sum_{l=1}^{t} y_l z_l^{i_m}$, $1 \leq m \leq r$.
*Error-locator polynomial:*

$$\sigma(Z) = \prod_{l=1}^{t}(Z - z_l) = Z^t + \sigma_1 Z^{t-1} + \cdots + \sigma_{t-1} Z + \sigma_t,$$

where

$$\sigma_i = (-1)^i \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq t} z_{j_1} z_{j_2} \ldots z_{j_i}, \; 1 \leq i \leq t,$$

Generalized Newton identities (GNI):

$$s_i + \sum_{j=1}^{t} \sigma_j s_{i-j} = 0, \quad \text{for all} \quad i \in \mathbb{Z}_n.$$

# Generalized Newton identities

## GNI for cyclic codes

GNI give rise to several decoding algorithms

- Polynomial-time up to designed minimum distance: APGZ, Berlekamp-Massey
- Exponential up to true minimum distance: Chen *et.al.*, Augot *et.al.*

It is of interest to find some analogue for arbitrary linear codes.

# Generalized Newton identities

## RS matrix as a special case of MDS

Suppose $n \leq q$. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be an $n$-tuple of mutually distinct elements in $\mathbb{F}_q$. Define $\mathbf{b}_i = (x_1^{i-1}, \ldots, x_n^{i-1})$. Then $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is an MDS basis. In particular, if $a \in \mathbb{F}_q^*$ is an element of order $n$ and $x_j = a^{j-1}$ for all $j$, then $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is called a *Reed-Solomon (RS) basis* and the corresponding matrix is called a *RS matrix* and denoted by $B(a)$.

## Structure relations for RS

The above construction gives an RS basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{F}_q^n$ over $\mathbb{F}_q$ such that

$$\mathbf{b}_i * \mathbf{b}_j = \mathbf{b}_{i+j-1} \quad \text{and} \quad u_{ij}(\mathbf{e}) = u_{i+j-1}(\mathbf{e}) \ \forall i, j \bmod n.$$

# Generalized Newton identities

## GNI for linear codes

Suppose that $(n, q) = 1$ and let $a$ be a primitive $n$-th root of unity in $\mathbb{F}$, where $\mathbb{F}$ is splitting field of $X^n - 1$ over $\mathbb{F}_q$. Note that $\mathbb{F} = \mathbb{F}_{q^m}$, where $m$ is the smallest positive integer such that $n|(q^m - 1)$. As an MDS matrix we choose an RS-matrix $B(a)$. Now $I(t, \mathcal{U}, V)$ is generated by

$$\sum_{j=1}^{t} U_{i+j-1} V_j - U_{i+t}, 1 \leq i \leq n,$$

where indices are taken modulo $n$. So $I(t, \mathcal{U}, V)$ has the form of GNI up to renumbering of indices.

# Generalized Newton identities

## Consistency with GNI for cyclic codes

For the cyclic code $C$ and received vector $\mathbf{y} = \mathbf{c} + \mathbf{e}$ let $s_i, i \in \mathbb{Z}_n$ be the syndromes (both known and unknown) and let $\sigma_j, 1 \leq j \leq t$ be the coefficients of $\sigma(Z)$. Let $J(t, \mathbf{y})$ be the ideal that corresponds to $C$ and $\mathbf{y}$ constructed w.r.t the RS-matrix $B(a)$. Assume $t \leq (d(C) - 1)/2$, so that $J(t, \mathbf{y})$ has a unique solution $(\mathbf{u}(\mathbf{e}), \mathbf{v})$. Then the following hold:

$$u_i(\mathbf{e}) = s_{i-1}, v_j = -\sigma_{t-j+1}, \forall i, j,$$

where $s_0 = s_n$.

## Linear part

We also have that $J(\mathbf{y})$ is $U_{i+1} - s_i, i \in S_C$.

# Generalized Newton identities

## Eliminating $U$-variables

For the case of binary codes it is possible to use Waring function to eliminate $U$-variables in $J(t, \mathbf{y})$. If $U-$ and $V-$variables are connected via GNI, we have

$$U_{i+1} = W_i(V_t, \dots, V_1), 1 \le i \le n-1, U_1 = W_n(V_t, \dots, V_1),$$

where $W_i$ are Waring functions (polynomials). Thus substituting the above to $J(\mathbf{y})$ we have the system purely in $V-$variables

$$a_{j1} W_n(V_t, \dots, V_1) + \sum_{l=2}^{n} a_{jl} W_{l-1}(V_t, \dots, V_1) = s_j(\mathbf{y}), j = 1, \dots, r.$$

# Generalized Newton identities

## General error-locator polynomial

Existence of the following polynomial $L_C$ from $\mathbb{F}_q[X_1, \ldots, X_r, Z]$ for a code $C$ is of interest (here $r = n - k$). $L_C$ should satisfy the following two properties:

- $L_C = Z^e + a_{t-1}Z^{e-1} + \cdots + a_0$ with $a_j \in \mathbb{F}_q[X_1, \ldots, X_r]$, $0 \leq j \leq e - 1$;
- given a syndrome $\mathbf{s} = (s_1, \ldots, s_r) \in \mathbb{F}_{q^m}^r$ corresponding to an error of weight $t \leq e$ and error locations $\{j_1, \ldots, j_t\}$, if we evaluate the $X_i = s_i$ for all $1 \leq i \leq r$, then the roots of $L_C(\mathbf{s}, Z)$ are exactly $a^{j_1}, \ldots, a^{j_t}$ and 0 of multiplicity $e - t$, in other words

$$L_C(\mathbf{s}, Z) = Z^{e-t} \prod_{i=1}^{t} (Z - a^{k_i})$$

Via an RS-extension $\mathbb{F}_{q^m}$ it is possible to prove the existence of $L_C$ over $\mathbb{F}_{q^m}$. Study further the possibility of generic decoding using GNI.

# Further research

## Further research

The possible directions of research:

- Study methods of solving $J(t, \mathbf{y})$ or its equivalents other than Gröbner basis (e.g. extended linearization).
- Complexity analysis of solving, e.g. via the analysis of $R_2$ systems.
- Algorithmic questions connected with the existence of GNI for arbitrary linear codes.
- Generic decoding and the existence of general error-locator polynomial.