Oberwolfach
Preprints

OWP 2012 - 01

Irene Márquez-Corbella ; Edgar Martínez-Moro ;
Ruud Pellikaan

Cryptanalysis of Public-key Cryptosystems Based on Algebraic Geometry Codes

Mathematisches Forschungsinstitut Oberwolfach gGmbH
Oberwolfach Preprints (OWP) ISSN 1864-7596

Imprint:

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany

Tel +49 7834 979 50
Fax +49 7834 979 55
Email admin@mfo.de
URL www.mfo.de

The Oberwolfach Preprints (OWP, ISSN 1864-7596) are published by the MFO.

Noname manuscript No.
(will be inserted by the editor)

Cryptanalysis of public-key cryptosystems based on algebraic geometry codes

Irene Márquez-Corbella · Edgar Martínez-Moro · Ruud Pellikaan

Received: date / Accepted: date

Abstract This paper addresses the question of retrieving the triple $(\mathcal{X}, \mathcal{P}, E)$ from the algebraic geometry code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$, where $\mathcal{X}$ is an algebraic curve over the finite field $\mathbb{F}_q$, $\mathcal{P}$ is an n-tuple of $\mathbb{F}_q$-rational points on $\mathcal{X}$ and E is a divisor on $\mathcal{X}$. If $\deg(E) \geq 2g+1$ where g is the genus of $\mathcal{X}$, then there is an embedding of $\mathcal{X}$ onto $\mathcal{Y}$ in the projective space of the linear series of the divisor E. Moreover, if $\deg(E) \geq 2g+2$, then $I(\mathcal{Y})$, the vanishing ideal of $\mathcal{Y}$, is generated by $I_2(\mathcal{Y})$, the homogeneous elements of degree two in $I(\mathcal{Y})$. If $n > 2\deg(E)$, then $I_2(\mathcal{Y}) = I_2(\mathcal{Q})$, where $\mathcal{Q}$ is the image of $\mathcal{P}$ under the map from $\mathcal{X}$ to $\mathcal{Y}$. These two results imply that certain algebraic geometry codes are not secure if used in the McEliece public-key cryptosystem.

The first two authors are partially supported by Spanish MCINN under project MTM2007-64704. First author research is also supported by a FPU grant AP2008-01598 by Spanish MEC. Second author is also supported by Spanish MCINN under project MTM2010-21580-C02-02.

I. Márquez-Corbella
Department of Algebra, Geometry and Topology, University of Valladolid, Facultad de Ciencias, 47005 Valladolid, Spain
Tel.: +34 983 423046
Fax: +34 983 423788
E-mail: imarquez@agt.uva.es

E. Martínez-Moro
Department of Applied Mathematics, University of Valladolid, Campus Duques de Soria, E-42004 Soria, Spain.
Tel.: +34 975 129420
Fax: +34 975 129401
E-mail: edgar@maf.uva.es

R. Pellikaan
Department of Mathematics and Computing Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.
Tel.: +31 40 2474222
Fax: +31 40 2435810
E-mail: g.r.pellikaan@tue.nl

Keywords Algebraic geometry codes · public-key cryptosystems

Mathematics Subject Classification (2000) 14G50 · 94A60

1 Introduction

Algebraic geometry codes (AG codes) also known as Goppa codes were introduced in 1977 by V.D. Goppa. The interested reader is referred to [15, 19, 41, 42]. These codes have efficient decoding algorithms that correct up to half the designed minimum distance [20, 19]. And they are used not only for error-correction but also for public-key cryptography. Particularly, Janwa and Moreno [23] propose to use AG codes for the McEliece cryptosystem. This system was broken for codes on curves of genus $g \leq 2$ by Faure and Minder [14] but for higher genus the security status of this scheme was not known. Throughout this section we will briefly introduce some concepts of AG codes necessary to describe the motivations and objective of this article which will be discussed at the end of the section.

Let $\mathbb{F}_q$ be a finite field with q elements and $\mathbb{F}_q[X]$ be the polynomial ring in one variable over $\mathbb{F}_q$. By an algebraic curve we mean a curve that is absolutely irreducible, projective and nonsingular. The algebraic geometry code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ is constructed using an algebraic curve $\mathcal{X}$ defined over the finite field $\mathbb{F}_q$, an n-tuple $\mathcal{P} = (P_1, \ldots, P_n)$ of $\mathbb{F}_q$-rational points on $\mathcal{X}$ and a divisor E of $\mathcal{X}$ with disjoint support from $\mathcal{P}$ of degree m.

The *function field* of the curve $\mathcal{X}$ with field of constants $\mathbb{F}_q$ is denoted by $\mathbb{F}_q(\mathcal{X})$. Let f be a nonzero rational function on a curve $\mathcal{X}$ over $\mathbb{F}_q$, then the *principal divisor* of zeros and poles of f is denoted by (f). Two divisors D and E on a curve $\mathcal{X}$ are called *rational equivalent* if there exists a rational function f on $\mathcal{X}$ such that $E = D + (f)$, this is denoted by $D \equiv E$. Moreover the divisors D and E on a curve $\mathcal{X}$ with disjoint support with $\mathcal{P}$ are called *rational equivalent with respect to* $\mathcal{P}$ and denoted by $D \equiv_{\mathcal{P}} E$ if there exists a rational function f on $\mathcal{X}$ such that it has no poles at the points of $\mathcal{P}$, $E = D + (f)$ and $f(P_j) = 1$ for all $j = 1, \ldots, n$.

Let E be a divisor of $\mathcal{X}$ of degree m. Then we define the *vector space of rational functions associated to* E as the set

$$L(E) = \{f \in \mathbb{F}_q(\mathcal{X}) \mid f = 0 \text{ or } (f) \geq -E\},$$

and the *linear series* of E as the collection $|E| = \{ F \mid F \equiv E, F \geq 0 \}$.

If f is a nonzero element of $L(E)$ then $(f) \geq -E$. Let $F = (f) + E$ then $F \geq 0$ and $F \equiv E$ so $F \in |E|$. Conversely, every $F \in |E|$ comes from a nonzero rational function $f \in L(E)$ which is unique up to nonzero scalar multiplication since $(\lambda f) = (f)$ for all nonzero scalar multiple λ. Therefore we can define an isomorphism between the following spaces

$$\mathbb{P}(L(E)) = L(E)^*/\mathbb{F}_q^* \longrightarrow |E|.$$

The dimension of the space $L(E)$ is denoted by $l(E)$, thus the projective dimension of the linear series $|E|$ is equal to $l(E) - 1$. The *index of speciality* of

E is defined by $i(E) = l(K - E)$, where K is a canonical divisor. Furthermore, let g be the*genus* of the curve $\mathcal{X}$, then $l(E) = m + 1 - g + i(E) \geq m + 1 - g$. If $m > 2g - 2$ then, by Riemann-Roch Theorem, equality holds that is $i(E) = 0$.

Let $\mathcal{P} = (P_1, \ldots, P_n)$ be an n-tuple of mutual distinct $\mathbb{F}_q$-rational points on $\mathcal{X}$. Then the divisor $P_1 + \cdots + P_n$ will be denoted by P. If the support of E is disjoint from P, then the following evaluation map:

$$\mathrm{ev}_{\mathcal{P}} : L(E) \longrightarrow \mathbb{F}_q^n$$

is well defined by $\mathrm{ev}_{\mathcal{P}}(f) = (f(P_1), \ldots, f(P_n))$. The algebraic geometry code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ is the image of $L(E)$ under the evaluation map $\mathrm{ev}_{\mathcal{P}}$. The parameters of these codes satisfy the following bounds:

Proposition 1 *If $m < n$ then the dimension of the code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ is equal to $m + 1 - g + i(E) \geq m + 1 - g$ and its minimum distance is at least $n - m$. Moreover, if $m > 2g - 2$, then $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ has dimension $m + 1 - g$.*

Proof The statement about the minimum distance is a consequence of the fact that a principal divisor has degree zero. The dimension of the code follows from the Theorem of Riemann-Roch [41,42]. □

The following proposition is related with the dual code of an AG code.

Proposition 2 *Let $\mathcal{X}$ be an algebraic curve over $\mathbb{F}_q$ of genus g, $\mathcal{P}$ an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and E a divisor of $\mathcal{X}$ with disjoint support from $\mathcal{P}$ of degree m. Let ω be a differential form with a simple pole at P_j with residue 1 for all $j = 1, \ldots, n$ and let K be the canonical divisor of ω. Define $E^{\perp} = P - E + K$ and $m^{\perp} = \deg(E^{\perp})$. Then $m^{\perp} = 2g - 2 - m + n$ and $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)^{\perp} = \mathcal{C}_L(\mathcal{X}, \mathcal{P}, E^{\perp})$.*

Proof See [41, Proposition 2.2.10]. □

Definition 1 A code $\mathcal{C}$ over $\mathbb{F}_q$ is called *weakly algebraic-geometric* (WAG) if $\mathcal{C}$ is equal to $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ for some curve $\mathcal{X}$ over $\mathbb{F}_q$, an n-tuple $\mathcal{P} = (P_1, \ldots, P_n)$ of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and a divisor E with disjoint support from $\mathcal{P}$. In this case the triple $(\mathcal{X}, \mathcal{P}, E)$ is called a WAG *representation* of $\mathcal{C}$.

Proposition 3 *Every code has a WAG representation.*

Proof See [33, Theorem 2]. □

The codes C and D are called *(generalized) equivalent* if there exists a monomial matrix M such that $M(C) = D$. Two representations $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{Y}, \mathcal{Q}, F)$ are called *equivalent* or *isomorphic* if there is an isomorphism of curves $\varphi : \mathcal{X} \to \mathcal{Y}$ such that $\varphi(\mathcal{P}) = \mathcal{Q}$ and $\varphi(E) \equiv F$, and they are called *strict equivalent* or *strict isomorphic* if there is an isomorphism of curves $\varphi : \mathcal{X} \to \mathcal{Y}$ such that $\varphi(\mathcal{P}) = \mathcal{Q}$ and $\varphi(E) \equiv_{\mathcal{Q}} F$.

Proposition 4 *Let $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{Y}, \mathcal{Q}, F)$ be WAG representations of the codes $\mathcal{C}$ and $\mathcal{D}$, respectively. Then:*

(1) *If $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{Y}, \mathcal{Q}, F)$ are equivalent, then $\mathcal{C}$ and $\mathcal{D}$ are equivalent.*
(2) *If $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{Y}, \mathcal{Q}, F)$ are strict equivalent, then $\mathcal{C} = \mathcal{D}$.*

Proof The proof of the case $\mathcal{X} = \mathcal{Y}$ and $\mathcal{P} = \mathcal{Q}$ is given in [41, Prop. 2.2.14 (a)] for (1) and [40, Lemma 3.1] for (2) and both are generalized in a straight forward matter. □

Definition 2 Let $\mathcal{X}$ be an algebraic curve of genus g, $\mathcal{P}$ an n-tuple of n points and E a divisor of degree m. A WAG representation $(\mathcal{X}, \mathcal{P}, E)$ of the code $\mathcal{C}$ is called *algebraic-geometric* (AG) if $\deg(E) < n$.

Moreover if $2g - 2 + t < \deg(E) < n - t$ for some t then $(\mathcal{X}, \mathcal{P}, E)$ is called a *t-strong algebraic-geometric* (t-SAG) representation.

In particular, a 0-SAG representation is a SAG representation as defined in [33] where also several necessary conditions are given for a code to have an AG or a SAG representation. Next proposition shows that a converse of the second part of Proposition 4 holds for 1-SAG codes.

Proposition 5 *Let $\mathcal{X}$ be a curve over $\mathbb{F}_q$ of genus g and $\mathcal{P}$ be an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$. If E and F are divisors on $\mathcal{X}$ of degree m such that $2g - 1 < m < n - 1$ and $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E) = \mathcal{C}_L(\mathcal{X}, \mathcal{P}, F)$ then $E \equiv_{\mathcal{P}} F$.*

Proof See [31, Corollary 4.15]. □

Duality preserves the t-SAG representations as we will see in our next proposition.

Proposition 6 *If $\mathcal{C}$ has a t-SAG representation, then $\mathcal{C}^{\perp}$ has also a t-SAG representation.*

Proof The proof is similar to the one given in [33, Corollary 1] for $t = 0$. Let $(\mathcal{X}, \mathcal{P}, E)$ be a t-SAG representation of the code $\mathcal{C}$. Then $\mathcal{C} = \mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ where $\mathcal{X}$ is an absolutely irreducible non-singular curve over $\mathbb{F}_q$ of genus g, $\mathcal{P} = (P_1, \dots, P_n)$ is an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and E is a divisor with disjoint support from $\mathcal{P}$ such that $2g-2+t < \deg(E) < n-t$. Now, by Proposition 2, $\mathcal{C}^{\perp} = \mathcal{C}_L(\mathcal{X}, \mathcal{P}, E^{\perp})$ and $\deg(E^{\perp}) = 2g - 2 - m + n$. Hence $2g - 2 + t < \deg(E^{\perp}) < n - t$ and we conclude that $(\mathcal{X}, \mathcal{P}, E^{\perp})$ is a t-SAG representation of $C^{\perp}$. □

From Proposition 6 we deduce that an analogous statement of Proposition 5 holds for the dual codes.

Let $r = l(E) - 1$ and $\{f_0, \dots, f_r\}$ be a basis of $L(E)$. Consider the following map:

$$\varphi_E : \mathcal{X} \longrightarrow \mathbb{P}^r(\mathbb{F}_q)$$

defined by $\varphi_E(P) = (f_0(P), \dots, f_r(P))$.

If $m > 2g$ then $r = m - g$, so φ_E defines an embedding of the curve $\mathcal{X}$ of degree m in $\mathbb{P}^r$. More precisely, let $\mathcal{Y} = \varphi_E(\mathcal{X})$, $Q_j = \varphi_E(P_j)$ and $\mathcal{Q} = (Q_1, \dots, Q_n)$. Then $\mathcal{Y}$ is a curve in $\mathbb{P}^{m-g}$ of degree m, φ_E is an isomorphism from $\mathcal{X}$ to $\mathcal{Y}$ and $\varphi_E(E) = \mathcal{Y} \cdot H$ for some hyperplane H of $\mathbb{P}^{m-g}$ that is disjoint from Q. See [18, Theorems 7.33 and 7.40]. Let $F = \varphi_E(E) = \mathcal{Y} \cdot H$. Then $\mathcal{C} = \mathcal{C}_L(\mathcal{Y}, \mathcal{Q}, F)$, that is $(\mathcal{Y}, \mathcal{Q}, F)$ is also a representation of the code $\mathcal{C}$ which is strict isomorphic with $(\mathcal{X}, \mathcal{P}, E)$. Therefore we have shown the following proposition:

Proposition 7 *Let $(\mathcal{X}, \mathcal{P}, E)$ be a WAG representation of the code $\mathcal{C}$ such that $\deg(E) > 2g$. Let $\mathcal{Y} = \varphi_E(\mathcal{X})$, $\mathcal{Q} = \varphi_E(\mathcal{P})$ and $F = \varphi_E(E)$. Then $(\mathcal{Y}, \mathcal{Q}, F)$ is a representation of $\mathcal{C}$ that is strict isomorphic with $(\mathcal{X}, \mathcal{P}, E)$.*

Let us briefly mention the connection between linear codes and affine or projective varieties which is better explained by means of projective systems [42, §1.1.2]. An n-tuple of points $(P_1, \dots, P_n)$ in $\mathbb{P}^r(\mathbb{F}_q)$ is a *projective system* if not all these points lie in a hyperplane. The points of a projective system in $\mathbb{P}^r(\mathbb{F}_q)$ are in *general position* if no $r + 1$ points of them lie on a hyperplane.

Let $\mathcal{P} = (P_1, \dots, P_n)$ be a projective system in $\mathbb{P}^r(\mathbb{F}_q)$ where P_j is given by the homogeneous coordinates $(p_{0j} : p_{1j} : \dots : p_{rj})$. We define the $(r + 1) \times n$-matrix $G_{\mathcal{P}}$ as the matrix with P_j^T as j-th column. Then $G_{\mathcal{P}}$ has rank $r + 1$, since not all points lie in a hyperplane. That is $G_{\mathcal{P}}$ can be seen as a generator matrix of a nondegenerate $[n, r + 1]$ code over $\mathbb{F}_q$.

Conversely, let $\mathcal{C}$ be a nondegenerate $[n, k]$ code over $\mathbb{F}_q$ with generator matrix G. Take the columns of G as homogeneous coordinates of points in $\mathbb{P}^{k-1}(\mathbb{F}_q)$, this gives the projective system $\mathcal{P}_G$ over $\mathbb{F}_q$ of G. Furthermore the code has minimum distance d if an only if $n - d$ is the maximal number of points of $\mathcal{P}_G$ in a hyperplane of $\mathbb{P}^{k-1}(\mathbb{F}_q)$. Hence projective systems of points in general position correspond to MDS codes.

Now, using the notation introduced before Proposition 7, let G be the $k \times n$ matrix with entries $f_i(P_j)$. Then G is a generator matrix of the code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ and $\mathcal{Q}$ is the projective system $\mathcal{P}_G$ in $\mathbb{P}^r(\mathbb{F}_q)$.

The main task of this paper is recovering the triple $(\mathcal{X}, \mathcal{P}, E)$ from the code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$. To achieve this aim we use several facts. First of all a generator matrix of the code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ gives a projective system $\mathcal{Q} = \varphi_E(\mathcal{P})$ of points in the projective space $\mathbb{P}^r(\mathbb{F}_q)$. Furthermore under some assumptions not only the pair $(\mathcal{X}, E)$ gives an embedding $\varphi_E(\mathcal{X})$ of the curve in the projective r-space such that the embedded curve is defined by quadratic equations (see Section 3) but also the quadratic polynomials that vanish on $\mathcal{Q}$ generate the vanishing ideal of the embedded curve (see Section 4).

Finally we provide an understanding of the security of the McEliece PKC system based on algebraic geometry codes (see Section 6).

2 GRS codes and the rational normal curve

Our results are generalizations of the well-known case of generalized Reed-Solomon codes and its representation by rational normal curves.

Let $n,\ k$ be arbitrary integers such that $1 \leq k \leq n \leq q$. We define the set $L_k = \{f \in \mathbb{F}_q[X] : \deg(f(X)) \leq k-1\}$. Then for each $\mathbf{a},\ \mathbf{b} \in \mathbb{F}_q^n$ the evaluation map at these elements is given by:

$$\begin{array}{rcl} \mathrm{ev}_{\mathbf{a},\mathbf{b}} : L_k & \longrightarrow & \mathbb{F}_q^n \\ f & \mapsto & \mathrm{ev}_{\mathbf{a},\mathbf{b}}(f(X)) = (f(a_1)\cdot b_1, \ldots, f(a_n)\cdot b_n) \end{array}$$

We will denote the map $\mathrm{ev}_{\mathbf{a},\mathbf{b}}$ by $\mathrm{ev}_{\mathbf{a}}$ if $\mathbf{b}$ is the all-ones vector. If $\mathbf{a}$ is an n-tuple of mutually distinct elements of $\mathbb{F}_q \cup \{\infty\}$ and $\mathbf{b}$ an n-tuple of nonzero elements of $\mathbb{F}_q$, then this evaluation map is injective, since $f \in L_k$ has at most $k-1 < n$ zeros.

Definition 3 Let $\mathbf{a}$ be an n-tuple of mutually distinct elements of $\mathbb{F}_q$ and $\mathbf{b}$ an n-tuple of nonzero elements of $\mathbb{F}_q$. The *generalized Reed-Solomon* (GRS) code is defined by $\mathrm{GRS}_k(\mathbf{a},\mathbf{b}) := \{\mathrm{ev}_{\mathbf{a},\mathbf{b}}(f(X)) : f \in L_k\}$.

That is, for every codeword $\mathbf{c} \in \mathrm{GRS}_k(\mathbf{a},\mathbf{b})$ there exists a unique polynomial $f_{\mathbf{c}} \in L_k$, known as the *polynomial associated to* $\mathbf{c}$, such that $\mathbf{c} = \mathrm{ev}_{\mathbf{a},\mathbf{b}}(f_{\mathbf{c}}(X))$.

Theorem 1 *The code* $\mathrm{GRS}_k(\mathbf{a},\mathbf{b})$ *is an* $[n,k,n-k+1]$ *MDS code. Furthermore a generator matrix of* $\mathrm{GRS}_k(\mathbf{a},\mathbf{b})$ *is given by*

$$G_{\mathbf{a},\mathbf{b}} = \begin{pmatrix} b_1 & b_2 & \ldots & b_n \\ b_1a_1 & b_2a_2 & \ldots & b_na_n \\ \vdots & \vdots & \ddots & \vdots \\ b_1a_1^{k-1} & b_2a_2^{k-1} & \ldots & b_na_n^{k-1} \end{pmatrix}$$

Proof See [22, §5]. □

Consider the projective curve $\mathcal{X} = \mathbb{P}^1$ given by $z = 0$. This curve has genus zero and its points are $(x:y)$. Now, let $P_\infty = (1:0)$ and $P_j = (a_j : 1)$ for all $j = 1, \ldots, n$. We define $\mathcal{P} = (P_1, \ldots, P_n)$ and $E = (k-1)P_\infty$. Then $l(E) = k$, which is in accordance with $l(E) = \deg(E) + 1 - g$ for $\deg(E) = k-1 > 2g-2$, and a basis for $L(E)$ is given by

$$\mathcal{B}_1 = \left\{1, \frac{x}{y}, \frac{x^2}{y^2}, \ldots, \frac{x^{k-1}}{y^{k-1}}\right\}.$$

Then the code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ has $G_k(\mathbf{a},\mathbf{b})$ as generator matrix with $\mathbf{b}$ the all ones vector.

If D is an arbitrary divisor of degree $k-1$, then D is rational equivalent with $E=(k-1)P_\infty$, that is to say there exists a rational function g such that $D+(g)=E$. Hence $L(D)=gL(E)$ and a basis of $L(D)$ is given by

$$\mathcal{B}_2=\left\{g, g\frac{x}{y}, g\frac{x^2}{y^2}, \ldots, g\frac{x^{k-1}}{y^{k-1}}\right\}.$$

If D has disjoint support with $\mathcal{P}$, then $g(P_j)$ is well defined for all j and we deduce that $\mathcal{C}_L(\mathcal{X},\mathcal{P},D)$ has $G_k(\mathbf{a},\mathbf{b})$ as generator matrix with $b_j=g(P_j)$ for all $j\in\{1,\ldots,n\}$. See [41, §2.3].

For every divisor E on the projective line of degree $r>0$, the map φ_E gives an embedding of $\mathbb{P}^1$ in $\mathbb{P}^r$ of degree r defined by

$$(x:y)\longmapsto(x^r:x^{r-1}y:\ldots:x^{r-i}y^i:\ldots:xy^{r-1}:y^r),$$

the image of this map in $\mathbb{P}^r$ is called a *rational normal curve* in $\mathbb{P}^r$ of degree r, denoted by $\mathcal{X}_r$.

The columns of the matrix $G_{r+1}(\mathbf{a},\mathbf{b})$, considered as homogeneous coordinates, are points of the curve $\mathcal{X}_r$. Therefore a GRS code of dimension $r+1$ can be described as a projective system of points on a rational normal curve of degree r in $\mathbb{P}^r$.

Next Proposition shows that this rational normal curve is the intersection of $\binom{r}{2}$ quadrics.

Proposition 8 *The vanishing ideal $I(\mathcal{X}_r)$ of $\mathcal{X}_r$ is generated by the elements:*

$$X_iX_{r-i}-X_jX_{r-j}, \quad \textit{for } 0\leq i<j\leq r.$$

In other words, $I(\mathcal{X}_r)$ is the determinantal ideal of the 2×2 minors of the following $2\times r$ matrix

$$\begin{pmatrix} X_0 & X_1 & \ldots & X_i & \ldots & X_{r-1} \\ X_1 & X_2 & \ldots & X_{i+1} & \ldots & X_r \end{pmatrix}.$$

Proof See a proof of this classical result in [7,11,43] and [9,16]. For more details in determinantal ideals we refer the reader to [8,36]. □

Piggot and Steiner in [35] gave a construction of a parametrization of a conic in the plane given 5 points in general position. This construction is generalized in the following result.

Proposition 9 *Through any $r+3$ points in $\mathbb{P}^r(\mathbb{F}_q)$ in general position there passes a unique rational normal curve.*

Proof See [16, p. 530].

This classical result and the attack of Sidelnikov-Shestakov could be considered as an algorithmic implementation of the following well known property on rational normal curves:

Proposition 10 *Let r, n be integers such that $r\geq 2$ and $n\geq 2r+3$. If $\mathcal{Q}$ is an n-tuple of points in general position in $\mathbb{P}^r(\mathbb{F}_q)$ posing only $2r+1$ conditions on quadrics, then $\mathcal{Q}$ lies on a unique rational normal curve.*

Proof See [16, p. 528–531], [2, Chap. III, §2, p. 120], [17, p. 9–14] and [9].

3 Curves defined by quadratic equations

It was shown by Enriques [13], Babbage [4] and Petri [34] that the canonical model of a non-singular non-hyperelliptic projective curve of genus at least three is the intersection of quadrics and cubics, and of quadrics only except in case of a trigonal curve and a plane quintic. This result for the canonical divisor was generalized for arbitrary divisors E under certain constraints on the degree. See [3,25,29,30,37] and [16, p. 528–535] and [2, Chap. III, §2 and §3].

The polynomial ring $R = \mathbb{F}_q[X_0, X_1, \ldots, X_r]$ is graded by

$$R_d = \{ \ f(X) \in R \mid f(X) \text{ is zero or is homogeneous of degree d } \}.$$

So it can be expressed in the form $R = \oplus_{d=0}^{\infty} R_d$. We define $R_{\leq d} = \oplus_{e=0}^{d} R_e$. Let I be an ideal, in a similar way, we define the d graded part of homogenous elements of degree d in I as $I_d = I \cap R_d$ and, the set of homogeneous elements of degree at most d as $I_{\leq d} = I \cap R_{\leq d}$.

Proposition 11 *Let $\mathcal{X}$ be an algebraic curve of genus g over the perfect field $\mathbb{F}_q$. Let E be a divisor on $\mathcal{X}$ of degree m. If $m \geq 2g+1$, then $\mathcal{Y} = \varphi_E(\mathcal{X})$ is a normal curve in $\mathbb{P}^{m-g}$ which is the intersection of quadrics and cubics, in particular $I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$ and $I_3(\mathcal{Y})$. If $m \geq 2g+2$, then $I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$.*

Proof See [29,37,38]. □

Proposition 12 *Let $\mathcal{X}$ be an algebraic curve of genus $g \geq 5$ over the perfect field $\mathbb{F}_q$, that is not hyperelliptic, not triangular, not a double covering of an elliptic curve ($g \geq 6$) and not a non-singular plane quintic curve. Then "almost all" non-special divisors on $\mathcal{X}$ of degree $m = 2g$ define an embedding $\mathcal{Y}$ of $\mathcal{X}$ in $\mathbb{P}^{m-g}$ which is projectively normal and whose ideal is generated by $I_{\leq 3}(\mathcal{Y})$.*

Proof See [21]. □

Proposition 13 *Let s, g be integers such that $s \geq 4$ and $\binom{s-1}{2} \leq g \leq \binom{s}{2}$. Then for "almost all" pairs $(\mathcal{X}, E)$ where $\mathcal{X}$ is a curve of genus g over the complex numbers and E is a non-special divisor of degree $m \geq g + 2s - 1$ the embedding $\mathcal{Y}$ of $\mathcal{X}$ in $\mathbb{P}^{m-g}$ is projectively normal and its ideal is generated by $I_2(\mathcal{Y})$.*

Proof See [25]. □

4 Determination of $I_2(\mathcal{Q})$

Now suppose that an n-tuple $\mathcal{Q}$ of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{Y}$ in $\mathbb{P}^r$ is given such that $I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$. In this section we will deduce which hypothesis guarantees that $I_2(\mathcal{Q}) = I_2(\mathcal{Y})$ and how to compute $I_2(\mathcal{Q})$ efficiently.

Proposition 14 *Let m, r and n be integers such that $r \geq 2$ and $n > dm$. Let $\mathcal{Y}$ be an absolutely irreducible curve in $\mathbb{P}^r$ of degree m. If $\mathcal{Q}$ is an n-tuple of points that lies on the curve $\mathcal{Y}$, then $I_{\leq d}(\mathcal{Q}) = I_{\leq d}(\mathcal{Y})$.*

Proof $I(\mathcal{Y}) \subseteq I(\mathcal{Q})$, since $\mathcal{Q}$ lies on $\mathcal{Y}$. Hence $I_{\leq d}(\mathcal{Y}) \subseteq I_{\leq d}(\mathcal{Q})$.

Conversely, let f be a homogeneous polynomial in $I_{\leq d}(\mathcal{Q})$ of degree $e \leq d$ and let $\mathcal{Z}$ be the hypersurface of degree e defined by $f = 0$. If $\mathcal{Y} \cap \mathcal{Z}$ is a finite set over the algebraic closure of the field $\mathbb{F}_q$, then the intersection divisor $\mathcal{Y} \cdot \mathcal{Z}$ is well defined and has degree em by Bézout's theorem [18, Corollary 7.9], on account of the fact that $\mathcal{Y}$ has degree m and $\mathcal{Z}$ has degree e. Since $n > dm \geq em$, this contradicts the fact that the n points of $\mathcal{Q}$ are in the intersection of $\mathcal{Y}$ and $\mathcal{Z}$. Consequently $\mathcal{Y} \cap \mathcal{Z}$ is not finite and $\mathcal{Y}$ must be contained in $\mathcal{Z}$ as $\mathcal{Y}$ is an absolutely irreducible curve. Therefore f vanishes on $\mathcal{Y}$, that is $f \in I_{\leq d}(\mathcal{Y})$. □

The assumption $n > dm$ in the above proposition is tight. Take for instance an absolutely irreducible curve $\mathcal{Y}$ in $\mathbb{P}^r$ of degree m and a generic hypersurface $\mathcal{Z}$ of degree d defined by the homogeneous equation $f = 0$ such that the intersection consists of dm mutually distinct points $\mathcal{Q} = (Q_1, \ldots, Q_{dm})$. Then $\mathcal{Q}$ lies on $\mathcal{Y}$ and $f \in I_{\leq d}(\mathcal{Q})$ but f is not an element of $I_{\leq d}(\mathcal{Y})$.

Let $\mathcal{C}$ be a k dimensional subspace of $\mathbb{F}_q^n$ with basis $\{\mathbf{g}_1, \ldots, \mathbf{g}_k\}$. We denote by $S^2(\mathcal{C})$ the second symmetric power of $\mathcal{C}$, or equivalently the symmetrized tensor product of $\mathcal{C}$ with itself. If $x_i = \mathbf{g}_i$, then $S^2(\mathcal{C})$ has basis $\{x_i x_j \mid 1 \leq i \leq j \leq n\}$ and dimension $\binom{k+1}{2}$. Furthermore we denote by $\langle \mathcal{C} * \mathcal{C} \rangle$ or $\mathcal{C}^{(2)}$ the square of $\mathcal{C}$, that is the linear subspace in $\mathbb{F}_q^n$ generated by $\{\mathbf{a} * \mathbf{b} | \mathbf{a}, \mathbf{b} \in C\}$. See [10, §4 Definition 6] and [26,45]. Now we consider the linear map

$$\sigma : S^2(\mathcal{C}) \longrightarrow \mathcal{C}^{(2)},$$

where the element $x_i x_j$ is mapped to $\mathbf{g}_i * \mathbf{g}_j$. The kernel of this map will be denoted by $K^2(\mathcal{C})$.

Proposition 15 *Let $\mathcal{Q}$ be an n-tuple of points in $\mathbb{P}^r(\mathbb{F}_q)$ not in a hyperplane, $k = r + 1$, $G_\mathcal{Q}$ be the $k \times n$ matrix associated to $\mathcal{Q}$ and $\mathcal{C}$ be the subspace of $\mathbb{F}_q^n$ generated by the rows of $G_\mathcal{Q}$. Then*

$$I_2(\mathcal{Q}) = \{ \ \textstyle\sum_{1 \leq i \leq j \leq k} a_{ij} X_i X_j \ | \ \sum_{1 \leq i \leq j \leq k} a_{ij} x_i x_j \in K^2(\mathcal{C}) \ \}.$$

Proof Let g_{ij} be the entry of $G_\mathcal{Q}$ in the i-th row and the j-th column, then the points Q_t are given by the homogeneous coordinates $(g_{1t} : \cdots : g_{kt})$. Let $\sum_{1 \leq i \leq j \leq k} a_{ij} x_i x_j \in K^2(\mathcal{C})$ then:

$$0 = \sigma \left(\sum_{1 \leq i \leq j \leq k} a_{ij} x_i x_j \right) = \sum_{1 \leq i \leq j \leq k} a_{ij} g_i * g_j = \sum_{1 \leq i \leq j \leq k} a_{ij} \left(g_{i1} g_{j1}, \ldots, g_{in} g_{jn} \right)$$

So

$$\sum_{1 \leq i \leq j \leq k} a_{ij} g_{it} g_{jt} = 0 \quad \text{for all } t = 1, \ldots, n$$

Hence the evaluation $\sum_{1\leq i\leq j\leq k} a_{ij}X_iX_j$ at Q_t is zero for all $t = 1,\dots,n$. Therefore $\sum_{1\leq i\leq j\leq k} a_{ij}X_iX_j \in I_2(\mathcal{Q})$.

The converse inclusion is proved by reading the above backwards. □

Corollary 1 *Let $\mathcal{Q}$ be an n-tuple of points in $\mathbb{P}^r(\mathbb{F}_q)$ not in a hyperplane. Then $\mathcal{O}(n^2\binom{r}{2})$ is an upper bound on the complexity of the computation of $I_2(\mathcal{Q})$.*

Proof By Proposition 15, a basis of $K^2(\mathcal{C})$ gives directly a generating set of $I_2(\mathcal{Q})$. Recall that $\mathcal{C}^{(2)}$ is generated by the elements

$$\{g_i * g_j \mid 1 \leq i \leq j \leq k = r-1\}$$

which form an $\binom{k+1}{2} \times n$ matrix. Gaussian elimination of this matrix gives a matrix R in reduced row echelon form in $\mathcal{O}(n^2\binom{k+1}{2})$ elementary operations. A basis of $K^2(\mathcal{C})$ can be read of directly from R as the left kernel of R. □

In the general case we define the spaces $S^d(\mathcal{C})$, $\mathcal{C}^{(d)}$ and $K^d(\mathcal{C})$ for any positive integer d, then we have a similar result to that in Proposition 15 relating $I_d(\mathcal{Q})$ and $K^d(\mathcal{C})$. Furthermore we have that $\mathcal{O}(n^2\binom{k+d-1}{d})$ is an upper bound on the complexity of the computation of $I_d(\mathcal{Q})$.

The problem of the efficient computation of the vanishing ideal of a finite set of points was introduced by Buchberger and Möller in 1982 [28]. Then several generalization have been proposed, for instance, to the case of points with multiplicity, Lakshman in 1991 [24] and to the projective case, Cioffi in 1998 [12]. Lately, Abbott et al. [1] came with a variant of the classical BM Algorithm where they tame the problem of coefficient growth.

The following algorithm produces, in a finite number of steps, the reduced Gröbner basis of the ideal $I_2(\mathcal{Q})$ with respect to a desired ordering where $\mathcal{Q}$ is an n-tuple of points in $\mathbb{P}^r(\mathbb{F}_q)$. It is a straight-forward adaptation of the *Projective version of the classical Buchberger-Möller Algorithm* presented in [1] for the special case where we know that the elements of the reduced Gröbner basis have degree two.

Let $\mathbb{T}_2^{r+1}$ be the set of powers of degree two of the r variables $\{x_0,\dots,x_r\}$, let σ be a term ordering in $\mathbb{T}_2^{r+1}$ and let $\mathcal{Q} = \{Q_1,\dots,Q_n\}$ be an n-tuple of points in $\mathbb{P}^r(\mathbf{F}_q)$ where Q_j is given by the homogeneous coordinates $(q_{j0} : \dots : q_{jr})$.

```
Initialization: Let:
  I1 $L$ be the ordered list of the elements of $\mathbb{T}_2^{r+1}$ w.r.t. $\sigma$,
  I2 $G = []$ and $S = []$ be empty lists
  I3 and $M = (m_{ij})$ be an $0 \times n$ matrix over $\mathbb{F}_q$.
Main loop:
  L1 IF $L$ is empty then go to the End
     ELSE choose the power product $t = \min_{\prec}(L)$ and remove it from $L$.
```

```
L2 Compute the evaluation vector $(t(Q_1),\ldots,t(Q_n))$, and reduce it against
   the rows of the matrix $M$, to obtain
       $\mathbf{v} = (v_1,...,v_n) = (t(Q_1),\ldots,t(Q_n)) - \sum_i a_i(m_{i1},...,m_{in})$ with $a_i \in \mathbb{F}_q$.
L3 IF $\mathbf{v} = \mathbf{0}$ then add the polynomial $t - \sum_i a_i s_i$ to the list $G$, where $s_i$
   is the $i$-th element of the list $S$. Goto L1.
   ELSE add $\mathbf{v}$ as a new row of $M$ and $t - \sum_i a_i s_i$ as a new element to
   the list $S$. Goto L1.
End: Return $G$ the reduced Gr"obner basis of $I_2(\mathcal{Q})$ w.r.t. $\sigma$.
```

Let n be the number of points and let $r+1$ be the number of variables, that is r is the dimension of the ambient projective space. Then the complexity of the previous algorithm is $\mathcal{O}\left(\binom{r+1}{2}ns^2\right)$ where s is the maximum between the values $r+1$ and n.

In the above algorithm the matrix M is constructed in such a way that it is in reduced row-echelon form, after a permutation of the columns, if necessary. That is, we should consider that the matrix M has at most s linearly independent rows and therefore adding new rows to this matrix will not make any change in our algorithm. Then a vector $\mathbf{v} \in \mathbb{F}_q^n$ is *reduced against* M by continuously subtracting suitable multiple of the rows of M to cancel the first non-zero element of $\mathbf{v}$. The process stop when either the vector $\mathbf{v}$ becomes zero or there is no row in M that reduces its first non-zero entry.

To obtain the complexity of this algorithm there are several factors which have been taken into account. Firstly, the above remark about how the matrix M is constructed, then the fact that the cost that dominates the algorithm is the cost of step L2 which, for each evaluation vector $\mathbf{u}$, consist on applying Gauss elimination to the augmented matrix $\begin{pmatrix} M \\ \mathbf{u} \end{pmatrix}$. Furthermore there are $\binom{r+1}{2}$ elements in $\mathbb{T}_2^{r+1}$, therefore this step is executed at most $\binom{r+1}{2}$ times. To this we can also add the cost of the treatment of the list L but, since we are working with monomials of degree 2, this cost is often neglected.

5 Very strong algebraic-geometric codes

Definition 4 A code $\mathcal{C}$ over $\mathbb{F}_q$ is called *very strong algebraic-geometric* (VSAG) if $\mathcal{C}$ is equal to $\mathcal{C}_L(\mathcal{X},\mathcal{P},E)$ where the curve $\mathcal{X}$ over $\mathbb{F}_q$ has genus g, $\mathcal{P}$ consists of n points and E has degree m such that

$$2g+2 \leq m < \tfrac{1}{2}n \quad \text{or} \quad \tfrac{1}{2}n + 2g - 2 < m \leq n-4.$$

The dimension of a such a code is $k = m+1-g$. Thus the dimension satisfies the following bound

$$g+3 \leq k < \tfrac{1}{2}n - g + 1 \quad \text{or} \quad \tfrac{1}{2}n + g - 1 < k \leq n-g-3.$$

Note that if a code has a VSAG representation then its dual is also VSAG by an argument that is similar to the one given in the proof of Proposition 6.

Note that an $[n,k]$ VSAG code on a curve of genus g lies asymptotically in the range

$$\gamma \leq R \leq \tfrac{1}{2} - \gamma \quad \text{or} \quad \tfrac{1}{2} + \gamma \leq R \leq 1 - \gamma,$$

for $n \to \infty$, where $R = \frac{k}{n}$ is the *information rate* and $\gamma = \frac{g}{n}$ the *relative genus*. Note that $n \leq N = |\mathcal{X}(\mathbb{F}_q)|$, where $\mathcal{X}(\mathbb{F}_q)$ denotes the set of $\mathbb{F}_q$-rational points on $\mathcal{X}$. Hence $\gamma \geq \frac{g}{N}$.

Theorem 2 *Let $\mathcal{C}$ be a VSAG code then a VSAG representation can be obtained from its generator matrix. Moreover all VSAG representations of $\mathcal{C}$ are strict isomorphic.*

Proof Let $(\mathcal{X}, \mathcal{P}, E)$ be a VSAG representation of $\mathcal{C}$, that is $\mathcal{X}$ is an algebraic curve over $\mathbb{F}_q$ of genus g, $\mathcal{P}$ is an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and E is a divisor of $\mathcal{X}$ with disjoint support from $\mathcal{P}$ of degree m such that

$$2g + 2 \leq m < \tfrac{1}{2}n \quad \text{or} \quad \tfrac{1}{2}n + 2g - 2 < m \leq n - 4.$$

By duality we may assume that $2g+2 \leq m < \frac{1}{2}n$. Let $G \in \mathbb{F}_q^{k \times n}$ be a generator matrix of $\mathcal{C}$ and $\mathcal{Q} = \mathcal{P}_G$ be the associated projective system of G.

Since $m > 2g$, Proposition 7 confirms that there exists an embedding of the curve $\mathcal{X}$ in $\mathbb{P}^r$ of degree m:

$$\begin{array}{rcl} \varphi_E : \mathcal{X} & \longrightarrow & \mathbb{P}^r \\ \mathcal{P} & \longmapsto & \varphi_E(\mathcal{P}) = (f_0(\mathcal{P}), \ldots, f_r(\mathcal{P})) \end{array}$$

where $\{f_0, \ldots, f_r\}$ is a basis of $L(E)$ and $r = \dim(L(E)) - 1 = m - g$ satisfying that $\mathcal{Q} = \varphi_E(\mathcal{P})$ lies on the curve $\mathcal{Y} = \varphi_E(\mathcal{X})$ and $F = \varphi_E(E) = \mathcal{Y} \cdot H$ for some hyperplane H of $\mathbb{P}^{m-g}$ that is disjoint from $\mathcal{Q}$, such that $(\mathcal{Y}, \mathcal{Q}, F)$ is a representation of $\mathcal{C}$ that is strict isomorphic with $(\mathcal{X}, \mathcal{P}, E)$.

Furthermore, since $m \geq 2g+2$, Proposition 11 states that $I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$ and, since $n > 2m$, Proposition 15 affirms that $I_2(\mathcal{Y}) = I_2(\mathcal{Q})$. So the curve $\mathcal{Y}$ is determined by the n-tuple of points $\mathcal{Q}$. Hence starting with a generator matrix of $\mathcal{C}$ we get a representation $(\mathcal{Y}, \mathcal{Q}, F)$ of $\mathcal{C}$ that is strict isomorphic with $(\mathcal{X}, \mathcal{P}, E)$.

Let $(\mathcal{X}', \mathcal{P}', E')$ be another VSAG representation of $\mathcal{C}$ then $(\mathcal{Y}, \mathcal{Q}, F)$ is strict isomorphic with $(\mathcal{X}', \mathcal{P}', E')$ from the previous reasoning. Hence $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{X}', \mathcal{P}', E')$ are strict isomorphic. □

Theorem 2 does not apply in the range $\frac{1}{2}n \leq m \leq \frac{1}{2}n + 2g - 2$ but by the following shortening argument one can extend this result.

Let $\mathcal{C}$ be a code of length n and let t be an integer $0 \leq t \leq n$. The code $\mathcal{C}_t$ is obtained from $\mathcal{C}$ by *shortening* at the last t positions. It has length $n - t$ and is defined by

$$\mathcal{C}_t = \{(c_1, \ldots c_{n-t}) \mid (c_1, \ldots c_{n-t}, 0, \ldots, 0) \in \mathcal{C}\}.$$

Proposition 16 *Let $(\mathcal{X}, \mathcal{P}, E)$ be a WAG representation of $\mathcal{C}$ where $\mathcal{X}$ is an algebraic curve over $\mathbb{F}_q$ of genus g, $\mathcal{P}$ is an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and E is a divisor of $\mathcal{X}$ with disjoint support from $\mathcal{P}$ of degree m such that $\mathcal{C} = \mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$. Let $\mathcal{P}_t = (P_1, \ldots, P_{n-t})$ and let $E_t = E - (P_{n-t+1} + \cdots + P_n)$. If $\frac{1}{2}n \leq m \leq n - 2g - 3$ and $t = 2m - n + 1$, then $\mathcal{C}_t = \mathcal{C}_L(\mathcal{X}, \mathcal{P}_t, E_t)$ is a VSAG code.*

Proof If a point P of $\mathcal{X}$ is not in the support of E, then

$$L(E - P) = \{f \in L(E) | f(P) = 0\}.$$

So

$$L(E_t) = \{f \in L(E) \mid f(P_j) = 0 \text{ for all } n - t + 1 \leq j \leq n\},$$

since $P_{n-t+1}, \ldots, P_n$ are mutually distinct and they are not in the support of E. Hence $\mathcal{C}_t = \mathcal{C}_L(\mathcal{X}, \mathcal{P}_t, E_t)$.

Now assume that $\frac{1}{2}n \leq m \leq n - 2g - 3$ and let $t = 2m - n + 1$. Then $t \geq 1$ and E_t has degree $m - t = n - m - 1$ satisfying the following inequality

$$2g + 2 \leq \deg(E_t) \leq \frac{1}{2}n.$$

Therefore $\mathcal{C}_t$ is a VSAG code. □

Remark 1 The WAG code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ of Proposition 16 where the curve $\mathcal{X}$ has genus g, $\mathcal{P}$ consist of n points and E has degree m such that $\frac{1}{2}n \leq m \leq n - 2g - 3$ has dimension $k = m + 1 - g$ such that

$$\tfrac{1}{2}n + 1 - g \leq k \leq n - 3g - 2, \text{ so asymptotically } \tfrac{1}{2} - \gamma \leq R \leq 1 - 3\gamma \text{ for } n \to \infty.$$

Furthermore, if the dual code $\mathcal{C}(\mathcal{X}, \mathcal{P}, E)^{\perp}$ satisfies the condition of Proposition 16, i.e. $\frac{1}{2}n \leq m^{\perp} \leq n - 2g - 3$, then $4g + 1 \leq m \leq \frac{1}{2}n + 2g - 2$. Hence $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, E)$ has dimension k such that

$$3g + 2 \leq k \leq \tfrac{1}{2}n + g - 1, \text{ so asymptotically } 3\gamma \leq R \leq \tfrac{1}{2} + \gamma \text{ for } n \to \infty.$$

6 Cryptanalysis of PKC's using VSAG codes

In 1978, McEliece [27] introduced the first public key cryptosystem (PKC) based on the theory of error-correcting codes in particular he proposed to use a $[1024, 524, 101]$ classical binary Goppa code. The security of this scheme is based on the hardness of the decoding problem for general linear codes and the hardness of distinguishing a code with the prescribed structure from a random one. This property makes McEliece scheme an interesting candidate for post-quantum cryptography. An overview of the state of the art of cryptosystems that are secure against attacks by quantum computers is provided in [6]. Another advantage of this scheme is its fast encryption and decryption functions.

Many attempts to replace Goppa codes with different families of codes have been proven to be insecure as for example using GRS codes such as the original Niederreiter system [32] which was broken by Sidelnikov and Shestakov [39] in 1992. Recall that GRS codes can be seen as the special class of algebraic geometry codes on the projective line, that is the algebraic curve of genus zero. This result was generalized to curves of genus 1 and 2 by Faure and Minder [14] in 2008. These attacks can be viewed as retrieving the curve, n points on this curve and the divisor E.

Since the initial Niederreiter scheme is completely broken, Berger and Loidreau [5] proposed in 2005 another version which was designed to resist precisely the Sidelnikov-Shestakov attack. The main idea of this variant is to work with subcodes of the original GRS code rather than using the complete GRS code. However Wieschebrink [44] in 2006 presents the first feasible attack to the Berger-Loidreau cryptosystem that allows us to recover the secret key if the chosen subcode is large enough but which was impractical for small subcodes. Furthermore in 2010 Wieschebrink [45] noted that it seems that with high probability the square code of a subcode of a GRS code of parameters $[n, k]$ is itself a GRS code of dimension $2k - 1$. Therefore we can apply the Sidelnikov-Shestakov attack and thus reconstruct the secret key in polynomial time. Continuing this line of work, in [26], we characterized those subcodes which are weak keys for the Berger-Loidreau cryptosystem. That is, firstly those subcodes which are themselves GRS codes, we have seen that the probability of occurrence of this fact is very small, and secondly those subcodes whose square code is a GRS code of maximal dimension which has high probability of occurrence.

In 1996 Janwa and Moreno [23] proposed to use the collection of AG codes on curves for the McEliece cryptosystem. As we have already explained this system was broken for codes on curves of genus $g \leq 2$ by Faure and Minder [14]. But the security status of this proposal for higher genus was not known.

We could also attack codes with rates in the range $0 \leq R \leq \gamma$ by elaborating on Propositions 12 and 13 when $m \leq 2g$ as we did with Proposition 11 for $m > 2g$. A similar remark holds for the range $1 - \gamma \leq R \leq 1$ by duality. However, code-based PKC usually takes codes with rate very close to $\frac{1}{2}$, since all decoding algorithms for general codes have exponential complexity and the largest exponent is achieved for half-rate codes. Therefore from this point of view the intervals for the rate $[0, \gamma]$ and $[1 - \gamma, 1]$ are the least interesting.

Theorem 2 implies that one should not use VSAG codes for the McEliece PKC system in the range

$$\gamma \leq R \leq \tfrac{1}{2} - \gamma \quad \text{or} \quad \tfrac{1}{2} + \gamma \leq R \leq 1 - \gamma,$$

for $n \to \infty$, since there is an efficient attack by our result. By a shortening argument Proposition 16 and Remark 1 also codes in the range

$$\tfrac{1}{2} - \gamma \leq R \leq 1 - 3\gamma \quad \text{or} \quad 3\gamma \leq R \leq \tfrac{1}{2} + \gamma,$$

for $n \to \infty$, should be excluded.

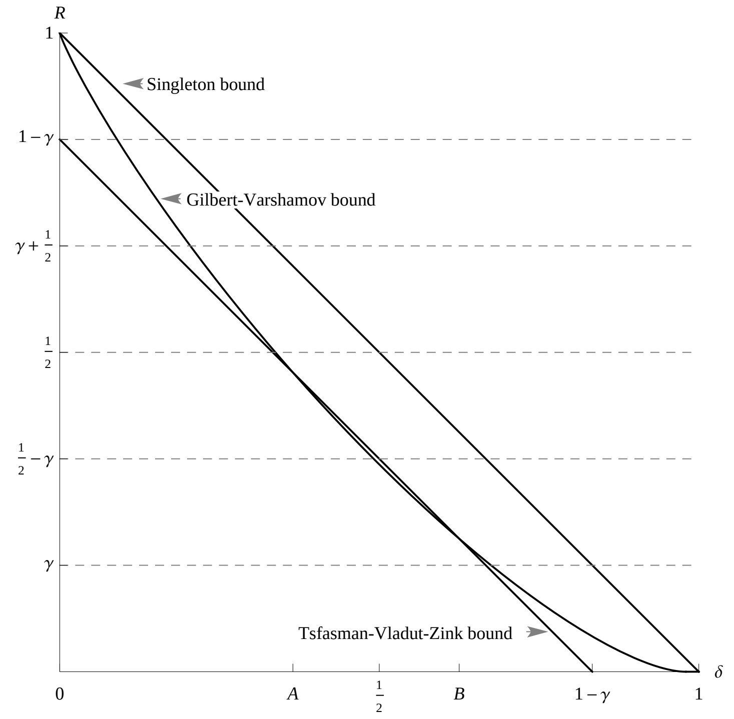

Fig. 1 Bounds on R as a function of the relative minimum distance δ for $q = 49$ and $\gamma = \frac{1}{6}$.

The above mentioned intervals $[\gamma, \frac{1}{2} - \gamma]$, $[\frac{1}{2} + \gamma, 1 - \gamma]$, $[\frac{1}{2} - \gamma, 1 - 3\gamma]$ and $[3\gamma, \frac{1}{2} + \gamma]$ are nonempty if and only if $\gamma \leq \frac{1}{4}$, and the union of these intervals cover the whole interval $[\gamma, 1 - \gamma]$ if and only if $\gamma \leq \frac{1}{6}$.

Notice that $\gamma = \frac{1}{\sqrt{q}-1}$ for AG codes on an asymptotically good sequence of algebraic curves over $\mathbb{F}_q$ attaining the Drinfeld-Vlădut bound and such that $n \approx |\mathcal{X}(\mathbb{F}_q)|$ where $n \to \infty$ for the length n of the codes. See [42]. So $\gamma \leq \frac{1}{4}$ if and only if $q \geq 25$, and $\gamma \leq \frac{1}{6}$ if and only if $q \geq 49$ for such sequences of codes.

In Figure 1, several bounds on the rate R of WAG codes are given as a function of their relative minimum distance δ for $q = 49$ and $\gamma = \frac{1}{6}$. For instance, the upper *Singleton* bound stated as $R + \delta \leq 1$ and the well known lower bounds: *Gilbert-Varshamov* bound given by $R \geq 1 - H_q(\delta)$ where H_q denotes the *q-ary entropy function* and *Tsfasman-Vlădut-Zink* bound which has the following form $R + \delta \geq 1 - \frac{1}{\sqrt{q}-1}$. Note that the last two curves intersect in exactly two points whenever $q \geq 49$, in our figure these points corresponds to the values of $\delta = A$ and B.

Acknowledgements This research was supported through the programme ”Research in Pairs” by the Mathematisches Forschungsinstitut Oberwolfach in 2011.

We like to thank Stanislav Bulygin and Xin-Wen Wu for their valuable discussions on the topics of this paper.

References

1. Abbott, J., Bigatti, A., Kreuzer, M., Robbiano, L.: Computing ideals of points. J. Symbolic Comput. **30**(4), 341–356 (2000). DOI 10.1006/jsco.2000.0411. URL http://dx.doi.org/10.1006/jsco.2000.0411
2. Arbarello, E., Cornalba, M., Griffiths, P.A., Harris, J.: Geometry of algebraic curves. Springer-Verlag, New York (1985)
3. Arbarello, E., Sernesi, E.: Petri's approach to the study of the ideal associated to a special divisor. Invent. Math. **49**, 99–119 (1978)
4. Babbage, D.: A note on the quadrics through a canonical curve. Journ. London Math. Soc. **14**, 310–315 (1939)
5. Berger, T., Loidreau, P.: How to mask the structure of codes for a cryptographic use. Designs, Codes and Cryptography **35**, 63–79 (2005)
6. Bernstein, D.: Introduction to post-quantum cryptography. In: J.B. D.J. Bernstein, E. Dahmen (eds.) Post-quantum cryptography, pp. 1–14. Springer-Verlag, Berlin (2009)
7. Bordiga, G.: Studio generale della quartica normale. Atti. R. Ist. Veneto di Sc. Lettere ed Arti **6**, 503–525 (1885–1886)
8. Bruns W. abd Vetter, U.: Determinantal rings. Lecture Notes Math. 1327, Springer, Berlin (1988)
9. Carlini, E., Catalisano, M.: Existence results for rational normal cuurves. J. London Math. Soc. **76**(2), 73–86 (2007)
10. Cascudo, I., Chen, H., Cramer, R., Xing, X.: Asymptotically good ideal linear secret sharing with strong multiplication overy any fixed finite field. In: S. Halevi (ed.) Advances in Cryptology - CRYPTO 2009, Lecture Notes in Computer Science, vol. 5677, pp. 466–486. Springer, Berlin (2009)
11. Castelnuovo, G.: Studio dellinvoluzione generale sulle curve razionali. Atti. R. Ist. Veneto di Sc. Lettere ed Arti **6**, 1167–1199 (1885–1886)
12. Cioffi, F.: Minimally generating ideals of points in polynomial time using linear algebra. Ricerche di Matematica XLVIII **1**, 55–63 (1999)
13. Enriques, F.: Sulle curve canoniche di genere p dello spazio a $p-1$ dimensioni. Rend. Accad. Sci. Ist. Bologna **23**, 80–82 (1919)
14. Faure, C., Minder, L.: Cryptanalysis of the McEliece cryptosystem over hyperelliptic codes. In: Proceedings 11th Int. Workshop on Algebraic and Combinatorial Coding Theory, ACCT 2008, pp. 99–107 (2008)
15. Goppa, V.: Codes associated with divisors. Probl. Inform. Transmission **13**, 22–26 (1977)
16. Griffiths, P., Harris, J.: Principles of algebraic geometry. Wiley-Interscience Publication, New York (1978)
17. Harris, J.: Algebraic geometry, a first course. Sprinber-Verlag, New York (1978)
18. Hirschfeld, J.W.P., Kochmáros, G., Torres, F.: Algebraic curves over a finite field. Princeton Univ. Press, Princeton (2008)
19. Hø holdt, T., Lint, J.v., Pellikaan, R.: Algebraic geometry codes. In: V. Pless, W. Huffman (eds.) Handbook of coding theory, vol. 1, pp. 871–961. North-Holland, Amsterdam (1998)
20. Hø holdt, T., Pellikaan, R.: On decoding algebraic-geometric codes. IEEE Transactions on Information **41**, 1589–1614 (1995)
21. Homma, M.: On the equations defining a projective curve embedded by a nonspecial divisor. Tsukuba Journ. Math. **3**(2), 31–39 (1979)
22. Huffman, W.C., Pless, V.: Fundamentals of error-correcting codes. Cambridge University Press, Cambridge (2003)
23. Janwa, H., Moreno, O.: McEliece public crypto system using algebraic-geometric codes. Designs, Codes and Cryptography **8**, 293–307 (1996)

24. Lakshman, Y.N.: A single exponential bound on the complexity of computing Gröbner bases of zero-dimensional ideals. In: Effective methods in algebraic geometry (Castiglioncello, 1990), *Progr. Math.*, vol. 94, pp. 227–234. Birkhäuser Boston, Boston, MA (1991)
25. Mancini, M.: Projectively normal curves defined by quadrics. Rend. Sem. Mat. Univ. Politec. Torino **59**(4), 269–275 (2001)
26. Márquez-Corbella, I., Martínez-Moro, E., Pellikaan, R.: The non-gap sequence of a subcode of a generalized Reed-Solomon code. In: WCC 2011 - Workshop on coding and cryptography, pp. 183–192. Paris, France (2001)
27. McEliece, R.J.: A public-key cryptosystem based on algebraic coding theory. DSN Progress Report **42–44**, 114–116 (1978)
28. Möller, H.M., Buchberger, B.: The construction of multivariate polynomials with preassigned zeros. In: Computer algebra (Marseille, 1982), *Lecture Notes in Comput. Sci.*, vol. 144, pp. 24–31. Springer, Berlin (1982)
29. Mumford, D.: Varieties defined by quadratic equations. In: Questions on algebraic varieties, C.I.M.E., III Ciclo, Varenna, 1969, pp. 29–100. Edizioni Cremonese, Rome (1970)
30. Mumford, D.: Curves and their Jacobians. Univ. Michigan Press, Ann Arbor (1975)
31. Munuera, C., Pellikaan, R.: Equality of geometric Goppa codes and equivalence of divisors. Journal of Pure and Applied Algebra **90**(3), 229–252 (1993)
32. Niederreiter, H.: Knapsack-type crypto systems and algebraic coding theory. Problems of Control and Information Theory **15**(2), 159–166 (1986)
33. Pellikaan, R., Shen, B.Z., van Wee, G.J.M.: Which linear codes are algebraic-geometric ? IEEE Trans. Inform. Theory **37**, 583–602 (1991)
34. Petri, K.: Über die invariante Darstellung algebraischer Funktionen einer Veränderlichen. Math. Ann. **88**(3-4), 242–289 (1923)
35. Piggott, H.E., Steiner, A.: Isogonal conjugates. A new approach to certain geometrical theorems and to a general theory of conics. Math. Gaz. **31**, 130–144 (1947)
36. Room, T.: The Geometry of Determinantal Loci. Cambridge University Press, Cambridge (1938)
37. Saint-Donat, B.: Sur les équations définissant une courbe algébrique. C. R. Acad. Sci. Paris **274**, 324–327, 487–489 (1972)
38. Saint-Donat, B.: On Petri's analysis of the linear system of quadrics through a canonical curve. Math. Ann. **206**, 157–175 (1973)
39. Sidelnikov, V.M., Shestakov, S.O.: On the insecurity of cryptosystems based on generalized Reed-Solomon codes. Discrete Math. Appl. **2**, 439–444 (1992)
40. Stichtenoth, H.: The automorphisms of geometric Goppa codes. Journ. Alg. **130**, 113–121 (1990)
41. Stichtenoth, H.: Algebraic function fields and codes. Springer, Berlin (1993)
42. Tsfasman, M.A., Vlădut, S.: Algebraic-geometric codes. Kluwer Academic Publishers, Dordrecht (1991)
43. Veronese, G.: Behandlung der projectivischen Verhältnisse der Räume von verschiedenen Dimensionen durch das Princip des Projectirens und Schneidens. Math. Ann. **19**, 161–234 (1882)
44. Wieschebrink, C.: An attack on the modified Niederreiter encryption scheme. In: PKC 2006, Lecture Notes in Computer Science, vol. 3958, pp. 14–26. Springer, Berlin (2006)
45. Wieschebrink, C.: Cryptanalysis of the Niederreiter public key scheme based on GRS subcodes. In: Post-Quantum Cryptography, Lecture Notes in Computer Science, vol. 6061, pp. 61–72. Springer, Berlin (2010)