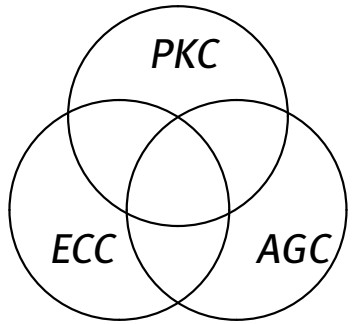
Evaluation of public-key cryptosystems based on algebraic geometry codes

Ruud Pellikaan
joint work with
Irene Márquez-Corbella and Edgar Martínez-Moro

Combinatorial, Algebraic and Algorithmic Aspects of Coding Theory
25 July 2011, EPFL Lausanne

Introduction



- ECC = Error-correcting codes
- AGC = Algebraic geometry curves
- PKC = Public-key cryptosystems

Projective systems and codes

Katsman-Tsfasman-Vladut:

Let $\mathbb{F}$ be a field
A projective system $\mathcal{P} = (P_1, \ldots, P_n)$ in $\mathbb{P}^r(\mathbb{F})$
is an n-tuple of points P_j in this projective space
such that not all these points lie in a hyperplane

Let $P_j = (p_{0j} : p_{1j} : \ldots : p_{rj})$
Let $G_{\mathcal{P}}$ be the $(r+1) \times n$ matrix with $(p_{0j}, p_{1j}, \ldots, p_{rj})^T$ as j-th column
Then $G_{\mathcal{P}}$ has rank $r+1$, since not all points lie in a hyperplane.

Code of a curve in projective space

If $\mathbb{F}$ is a finite field, then $G_{\mathcal{P}}$ is the generator matrix of a nondegenerate $[n, r+1, d]$ code over $\mathbb{F}$ where $n-d$ is the maximal number of points of $\mathcal{P}$ that lie in a hyperplane of $\mathbb{P}^{k-1}(\mathbb{F})$

Example

Let $\mathcal{X}$ be an irreducible projective curve over $\mathbb{F}_q$ of degree m in $\mathbb{P}^{k-1}$
Let $\mathcal{P}$ be an enumeration of n points of $\mathcal{X}(\mathbb{F}_q)$
Then $G_{\mathcal{P}}$ is the generator matrix of a code with parameters $n, k, d]$

$$d \geq n - m$$

Codes and projective systems

Conversely:
Let G be a generator matrix of a nondegenerate $[n, k, d]$ code over $\mathbb{F}_q$
Then G has no zero columns
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

One-to-one correspondence between:
generalized equivalence classes of nondegenerate $[n, k]$ codes over $\mathbb{F}_q$
and
equivalence classes of projective systems of n points in $\mathbb{P}^{k-1}(\mathbb{F}_q)$

Generalized Reed-Solomon codes

$\mathbf{a} = (a_1, \ldots, a_n)$ an n-tuple of mutually distinct elements of $\mathbb{F}_q$
$\mathbf{b} = (b_1, \ldots, b_n)$ an n-tuple of nonzero elements of $\mathbb{F}_q$

$$GRS_k(\mathbf{a}, \mathbf{b}) = \{ (f(a_1)b_1, \ldots, f(a_n)b_n) \mid f(X) \in \mathbb{F}_q[X], \deg(f(X) < k \}$$

parameters: $[n, k, n-k+1]$ if $k \leq n$

generator matrix:

$$G_k(\mathbf{a}, \mathbf{b}) = \begin{pmatrix} b_1 & \cdots & b_j & \cdots & b_n \\ a_1 b_1 & \cdots & a_j b_j & \cdots & a_n b_n \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_1^{k-1} b_1 & \cdots & a_j^{k-1} b_j & \cdots & a_n^{k-1} b_n \end{pmatrix}$$

Normal rational curves and GRS codes

The projective system of the the code $GRS_k(\mathbf{a}, \mathbf{b})$ with generator matrix $G_k(\mathbf{a}, \mathbf{b})$ is

$$\mathcal{P}_k(\mathbf{a}) = ((1 : a_j : \cdots : a_j^i : \cdots : a_j^{k-1}) \mid j = 1, \ldots, n)$$

Consider the embedding $\mathbb{P}^1 \to \mathbb{P}^r$ by the degree r map given by

$$(y_0 : y_1) \mapsto (y_0^r : y_0^{r-1} y_1 : \cdots : y_0^{r-i} y_1^i : \cdots : y_0 y_1^{r-1} : y_1^r)$$

The image of this map in $\mathbb{P}^r$ is the NRC (normal rational curve) $\mathcal{X}_r$

Every hyperplane intersects $\mathcal{X}_r$ in at most r points and

$$\mathcal{P}_k(\mathbf{a}) \subseteq \mathcal{X}_{k-1}(\mathbb{F}_q)$$

Vanishing ideal of rational normal curve

The vanishing ideal $I(\mathcal{X}_r)$ of $\mathcal{X}_r$ is generated by the quadratic polynomials:

$$X_i X_{r-i} - X_j X_{r-j}, \quad \text{for } 0 \leq i < j \leq r$$

that is the determinantal ideal of the 2×2 minors of the $2 \times r$ matrix

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

since the rows of the matrix

$$\begin{pmatrix} 1 & y & \cdots & y^i & \cdots & y^{r-1} \\ y & y^2 & \cdots & y^{i+1} & \cdots & y^r \end{pmatrix}$$

are dependent for all y and ....

Algebraic geometry codes

Let $\mathcal{X}$ be an algebraic variety over $\mathbb{F}_q$
with a subset $\mathcal{P}$ of $\mathcal{X}(\mathbb{F}_q)$ enumerated by $P_1, \ldots, P_n$

Suppose that we have a vector space L over $\mathbb{F}_q$
of functions on $\mathcal{X}$ with values in $\mathbb{F}_q$
So $f(P_i) \in \mathbb{F}_q$ for all i and $f \in L$
In this way we have an evaluation map

$$ev_{\mathcal{P}} : L \longrightarrow \mathbb{F}_q^n$$

defined by $ev_{\mathcal{P}}(f) = (f(P_1), \ldots, f(P_n))$

This evaluation map is linear, so its image is a linear code

Codes on the affine line

The classical example:
Generalized Reed-Solomon codes

The geometric object $\mathcal{X}$ is the affine line over $\mathbb{F}_q$
The points are n distinct elements of $\mathbb{F}_q$
L is the vector space of polynomials of degree at most $k-1$
and with coefficients in $\mathbb{F}_q$

This vector space has dimension k
Such polynomials have at most $k-1$ zeros
so nonzero codewords have at least $n-k+1$ nonzeros

This code has parameters $[n, k, n-k+1]$ if $k \leq n$

Codes on curves-function fields

Let $\mathcal{X}$ be an algebraic curve over $\mathbb{F}_q$ of genus g
$\mathbb{F}_q(\mathcal{X})$ is the function field of the curve $\mathcal{X}$ with field of constants $\mathbb{F}_q$

Let f be a nonzero rational function on the curve
The divisor of zeros and poles of f is denoted by (f)

Let E be a divisor of $\mathcal{X}$ of degree m
Then

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

The dimension of the space $L(E)$ is denoted by $l(E)$
Then $l(E) \geq m + 1 - g$ and equality holds if $m > 2g - 2$
by the Theorem of Riemann-Roch

Codes on curves

Let $\mathcal{P} = (P_1, \ldots, P_n)$ an n-tuple of mutual distinct points of $\mathcal{X}(\mathbb{F}_q)$ with divisor $D = P_1 + \cdots + P_n$

If the support of E is disjoint from D, then the evaluation map

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

where $\mathrm{ev}_{\mathcal{P}}(f) = (f(P_1), \ldots, f(P_n))$, is well defined.

The algebraic geometry code $C_L(\mathcal{X}, \mathcal{P}, E)$ is the image of $L(E)$ under the evaluation map $\mathrm{ev}_{\mathcal{P}}$
If $m < n$, then $C_L(\mathcal{X}, \mathcal{P}, E)$ is an $[n, k, d]$ code with

$$k \geq m + 1 - g \text{ and } d \geq n - m$$

Dual codes on curves

Let ω be a differential form with
a simple pole at P_j with residue 1 for all $j = 1, \ldots, n$

Let K be the canonical divisor of ω
Let m be the degree of the divisor E on $\mathcal{X}$
with disjoint support from $\mathcal{P}$

Let $E^\perp = D - E + K$ and $m^\perp = \deg(E^\perp)$
Then $m^\perp = 2g - 2 - m + n$ and

$$C_L(\mathcal{X}, \mathcal{P}, E)^\perp = C_L(\mathcal{X}, \mathcal{P}, E^\perp)$$

Codes on curves

Embedding of $\mathcal{X}$ in linear system of E of degree m
Let $f_1, f_2, \ldots, f_k$ be a basis of $L(E)$

$$\varphi : \mathcal{X} \longrightarrow \mathbb{P}^{k-1}$$

$$P \mapsto (f_1(P), f_2(P), \ldots, f_k(P))$$

$\mathcal{Y} = \varphi(\mathcal{X})$ is a curve of degree m in $\mathbb{P}^{k-1}$
$\mathcal{Q} = (\varphi(P_1), \ldots, \varphi(P_n))$ projective system

$$G_{\mathcal{Q}} = \begin{pmatrix} f_1(P_1) & \cdots & f_1(P_j) & \cdots & f_1(P_n) \\ f_2(P_1) & \cdots & f_2(P_j) & \cdots & f_2(P_n) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ f_k(P_1) & \cdots & f_k(P_j) & \cdots & f_k(P_n) \end{pmatrix}$$ generator matrix

minimum distance $\geq n - m$

Decoding linear codes

Decoding problem

Input: $(G, \mathbf{y})$
where G is a $k \times n$ a matrix G over $\mathbb{F}_q$ of rank k, and $\mathbf{y}$ in $\mathbb{F}_q^n$

Output: A closest codeword $\mathbf{c}$
so $d(\mathbf{c}, \mathbf{y})$ is minimal for all $\mathbf{c}$ in the code C with generator matrix G

This problem is NP-hard
Berlekamp-McEliece-Van Tilborg

Code based PKC systems

McEliece:
Let $\mathcal{C}$ be a class of codes that have
efficient decoding algorithms correcting t errors with $t \leq (d-1)/2$

Secret key: (S, G, P)
S an invertible $k \times k$ matrix
G a $k \times n$ generator matrix of a code C in $\mathcal{C}$.
P an $n \times n$ permutation matrix

Public key: $G' = SGP$

Message: $\mathbf{m}$ in $\mathbb{F}_q^k$
Encryption: $\mathbf{y} = \mathbf{m}G' + \mathbf{e}$ with random chosen $\mathbf{e}$ in $\mathbb{F}_q^n$ of weight t
Decryption: $\mathbf{y}P^{-1} = \mathbf{m}SG + \mathbf{e}P^{-1}$ and $\mathbf{e}P^{-1}$ has weight t
Decoder gives $\mathbf{c} = \mathbf{m}SG$ as closest codeword

McEliece PKC system with Goppa codes

Binary Goppa codes with parameters $[n, k, d]$ where

$$n = 2^m, \quad k \geq 2^m - mr, \quad d \geq 2r + 1$$

Are subfield subcodes of GRS codes over $\mathbb{F}_{2^m}$ with parameters

$$[2^m, 2^m - mr, r + 1]$$

Testcase: binary Goppa code with $m = 10$, $r = 50$

$$[1024, \geq 524, \geq 101]$$

Corrects 50 errors

Attack by Sidelnikov-Shestakov

Suppose $\mathcal{C}$ is the class of Generalized Reed-Solomon codes

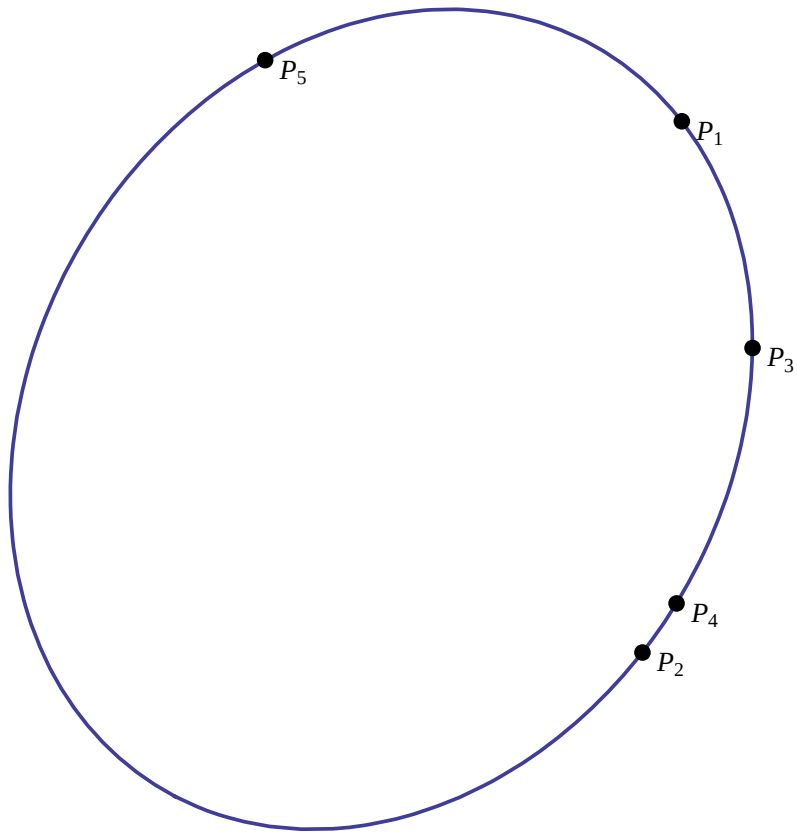
A GRS code of length n and dimension $k = r + 1$
gives a projective system of n points in general position
on a NRC of degree r in projective space of dimension r

Special case: $k = 3$ and $r = 2$:
a NRC of degree 2 in the projective plane is a conic
5 points in general position determine this conic

Steiner: parametrization of this conic in the plane given these 5 points

Algorithm of Sidelnikov-Shestakov for arbitrary k
Complexity: linear algebra $\mathcal{O}(n^3)$

Conic determined by 5 points



NRC of degree r determined by $r + 3$ points

Veronese 1882, Bordiga 1885, Castelnuovo 1885:

Let $\mathcal{P}$ be a collection of $r + 3$ points in general position in $\mathbb{P}^r$
Then there is a unique NRC of degree r passing through the points of $\mathcal{P}$

Twisted cubic, r=3:

(spiraal)

PKC system using subcodes of GRS codes

Berger-Loidreau: use subcodes of GRS codes

Attack: Wieschebrink, Marquez-Martinez-P

Projection of NRC to linear subspace
Singular rational curve in projective space
Interpolation problem

Goppa codes are specific subfield subcodes of GRS codes

PKC system using algebraic geometry codes

Janwa-Moreno: use algebraic geometry codes

Problem for the attacker:

Input: a generator matrix of an AG code $C_L(\mathcal{X}, \mathcal{P}, E)$
Output: $(\mathcal{X}, \mathcal{P}, E)$ if this triple is unique or $(\mathcal{X}', \mathcal{P}', E')$ otherwise

This system was broken for codes on curves of genus $g \leq 2$
by Faure-Minder

AG, WAG and SAG codes

A code C over $\mathbb{F}$ is called weakly algebraic-geometric (WAG) if $C = C_L(\mathcal{X}, \mathcal{P}, E)$ for some triple $(\mathcal{X}, \mathcal{P}, E)$ where:

- $\mathcal{X}$ is an algebraic curve over $\mathbb{F}_q$
- $\mathcal{P}$ is an n-tuple of mutually distinct points of $\mathcal{X}(\mathbb{F}_q)$
- E is divisor of degree m on $\mathcal{X}$

Then $(\mathcal{X}, \mathcal{P}, E)$ is called a WAG representation of C
If $m < n$, then it is called AG
If $2g - 2 < m < n$, then it is called strongly algebraic-geometric (SAG)

Theorem[P-Shen-van Wee]: Every code has a WAG representation

Equivalent representations

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 $\varphi(E) \equiv F$

They are called strict equivalent or strict isomorphic if
moreover $\varphi(E) \equiv_{\mathcal{Q}} F$

Proposition

Let $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{Y}, \mathcal{Q}, F)$ be WAG representations of C and D, resp.
Then:

(1) If $(\mathcal{X}, \mathcal{P}, E)$ and $(\mathcal{Y}, \mathcal{Q}, F)$ are equivalent, then $C \equiv D$

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

Retrieving E from $C_L(\mathcal{X}, \mathcal{P}, E)$ given $(\mathcal{X}, \mathcal{P})$

Theorem[Munuera-P]:

Let $\mathcal{X}$ be a curve of genus g and $D = P_1 + \cdots + P_n$

Let E and F be divisors of degree m with $2g - 1 < m < n - 1$

Then

$$C_L(\mathcal{X}, \mathcal{P}, E) = C_L(\mathcal{X}, \mathcal{P}, F) \text{ if and only if } E \equiv_{\mathcal{P}} F$$

Strict equivalent representations

Let $(\mathcal{X}, \mathcal{P}, E)$ be a *WAG* representation of C such that $m > 2g$
Let $r = l(E) - 1$ and $\{f_0, \ldots, f_r\}$ be a basis of $L(E)$
Consider the following map:

$$\varphi_E : \mathcal{X} \longrightarrow \mathbb{P}^r$$

defined by $\varphi_E(P) = (f_0(P), \ldots, f_r(P))$
If $m > 2g$, then $r = m - g$ and φ_E defines an embedding of the curve $\mathcal{X}$ in $\mathbb{P}^r$ of degree m with image $\mathcal{Y} = \varphi_E(\mathcal{X})$

Let $Q_j = \varphi_E(P_j)$ and $\mathcal{Q} = (Q_1, \ldots, Q_n)$ then $\varphi_E(E) = \mathcal{X} \cdot H = F$ for some hyperplane H of $\mathbb{P}^r$ that is disjoint from $\mathcal{Q}$

Furthermore $(\mathcal{Y}, \mathcal{Q}, F)$ is also a WAG representation of the code C that is strict isomorphic with $(\mathcal{X}, \mathcal{P}, E)$

Curves defined by quadrics

Normal rational normal curve is defined by quadratic equations.

The canonical model of a 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
Enriques 1919, Petri 1923 and Babbage 1939

This result for the canonical divisor was generalized for
arbitrary divisors E under certain constraints on the degree
Mumford 1970, Saint-Donat 1972 and Arbarello 1978

Curves defined by quadrics

Let $\mathcal{X}$ be an absolutely irreducible and nonsingular curve of genus g over the perfect field $\mathbb{F}$

Let E be a divisor on $\mathcal{X}$ of degree m

If $m \geq 2g + 2$
then $\mathcal{Y} = \varphi_E(\mathcal{X})$ is a normal curve in $\mathbb{P}^{m-g}$
which is the intersection of quadrics

More precisely:
$I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$ the ideal generated by the homogeneous elements of degree two in $I(\mathcal{Y})$

Retrieving $(\mathcal{X}, D, E)$ from $C_L(\mathcal{X}, D, E)$

Let $\mathcal{Y}$ be a curve embedded in projective r-space of degree m
Let $I(\mathcal{Y})$ be the vanishing ideal of $\mathcal{Y}$
Let $\mathcal{Q}$ be a subset of $\mathcal{Y}$ of n points
Then

$$I(\mathcal{Y}) \subseteq I(\mathcal{Q})$$

Let $I_2(\mathcal{Y})$ be the ideal generated by
the homogeneous elements of degree two in $I(\mathcal{Y})$

Suppose $I_2(\mathcal{Y}) = I(\mathcal{Y})$

$$\text{If } n > 2m, \text{ then } I(\mathcal{Y}) = I_2(\mathcal{Q})$$

By Bézout

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

Let $\mathcal{Q}$ be an n-tuple of points in $\mathbb{P}^{k-1}$ over $\mathbb{F}$ not in a hyperplane
Let $G_{\mathcal{Q}}$ be the $k \times n$ matrix associated to $\mathcal{Q}$ with basis $\mathbf{g}_1, \ldots, \mathbf{g}_k$
Denote the second symmetric power of C by by $S^2(C)$

If $x_i = \mathbf{g}_i$ then $S^2(C)$ has basis $\left\{x_i x_j \mid 1 \leq i \leq j \leq n\right\}$
and dimension $\binom{k+1}{2}$

Denote by $\langle C * C \rangle$ or $C^{(2)}$ the square of C
that is, the linear subspace in $\mathbb{F}^n$ generated by $\left\{\mathbf{a} * \mathbf{b} | \mathbf{a}, \mathbf{b} \in C\right\}$

Consider the linear map $\sigma : S^2(C) \longrightarrow C^{(2)}$
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 $K^2(C)$

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

Then

$$0 \longrightarrow K^2(C) \longrightarrow S^2(C) \longrightarrow C^{(2)} \longrightarrow 0$$

is an exact sequence and

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

Proposition

Let $\mathcal{Q}$ be an n-tuple of points in $\mathbb{P}^r$ over $\mathbb{F}$ not in a hyperplane
Then the complexity of the computation of $I_2(\mathcal{Q})$ is at most $\mathcal{O}(n^2\binom{r}{2})$

VSAG codes

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

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

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

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

Note that the dual of a VSAG code is again VSAG

Main theorem

Let C be a VSAG code

Then a VSAG representation can be obtained efficiently
from its generator matrix

Moreover all VSAG representations of C are strict isomorphic

shortening SAG codes

The dimension of a VSAG code satisfies the following bound

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

Let $R = k/n$ and $\gamma = g/n$
Then for $n \to \infty$ and $\gamma \leq \frac{1}{4}$ (so $q \geq 25$):

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

Proposition
If $\gamma \leq \frac{1}{4}$ and C is a SAG code in the range

$$\gamma \leq R \leq 1 - 3\gamma \quad \text{or} \quad 3\gamma \leq R \leq 1 - \gamma$$

then by shortening C sufficiently many times one gets a VSAG code

Conclusion

SAG codes are not secure for a code based PKC

1. If $\frac{1}{6} \leq \gamma \leq \frac{1}{4}$ in the range:

$$\gamma \leq R \leq 1 - 3\gamma \quad \text{or} \quad 3\gamma \leq R \leq 1 - \gamma$$

2. If $\gamma \leq \frac{1}{6}$ in the range:

$$\gamma \leq R \leq 1 - \gamma$$