A Distinguisher for High Rate McEliece Cryptosystems

Jean-Charles Faugère\textsuperscript{1}, Ayoub Otmani\textsuperscript{2,3}, Ludovic Perret\textsuperscript{1}, and Jean-Pierre Tillich\textsuperscript{2}

\textsuperscript{1} SALSA Project - INRIA (Centre Paris-Rocquencourt)
UPMC, Univ Paris 06 - CNRS, UMR 7606, LIP6
104, avenue du Président Kennedy 75016 Paris, France
jean-charles.faugere@inria.fr, ludovic.perret@lip6.fr

\textsuperscript{2} SECRET Project - INRIA Rocquencourt
Domaine de Voluceau, B.P. 105 78153 Le Chesnay Cedex - France
ayoub.otmani@inria.fr, jean-pierre.tillich@inria.fr

\textsuperscript{3} GREYC - Université de Caen - Ensicaen
Boulevard Maréchal Juin, 14050 Caen Cedex, France.

\textbf{Abstract.} The purpose of this paper is to study the difficulty of the so-called Goppa Code Distinguishing (GD) problem introduced by Courtois, Finiasz and Sendrier in Asiacrypt 2001. GD is the problem of distinguishing the public matrix in the McEliece cryptosystem from a random matrix. It is widely believed that this problem is computationally hard as proved by the increasing number of papers using this hardness assumption. To our point of view, disproving/mitigating this hardness assumption is a breakthrough in code-based cryptography and may open a new direction to attack McEliece cryptosystems. In this paper, we present an efficient distinguisher for alternant and Goppa codes of high rate over binary/non binary fields. Our distinguisher is based on a recent algebraic attack against compact variants of McEliece which reduces the key-recovery to the problem of solving an algebraic system of equations. We exploit a defect of rank in the (linear) system obtained by linearizing this algebraic system. It turns out that our distinguisher is highly discriminant. Indeed, we are able to precisely quantify the defect of rank for "generic" binary and non-binary random, alternant and Goppa codes. We have verified these formulas with practical experiments, and a theoretical explanation for such defect of rank is also provided. We believe that this work permits to shed some light on the choice of secure parameters for McEliece cryptosystems; a topic thoroughly investigated recently. Our technique permits to indeed distinguish a public key of the CFS signature scheme for all parameters proposed by Finiasz and Sendrier at Asiacrypt 2009. Moreover, some realistic parameters of McEliece scheme also fit in the range of validity of our distinguisher.

\textbf{Keywords:} public-key cryptography, McEliece cryptosystem, CFS signature, algebraic cryptanalysis, distinguisher.

1 Introduction

Code-based public key cryptography appeared with McEliece’s pioneering work [23] where the author proposed to use one-way trapdoor functions based on irreducible binary Goppa codes. The class of Goppa codes represents one of the most important example of linear codes having an efficient decoding algorithm [4, 28]. A binary Goppa code is defined by a polynomial $\Gamma$ of degree $r \geq 1$ with coefficients in some extension field $\mathbb{F}_{2^m}$ of degree $m > 1$ over $\mathbb{F}_2$, and a $n$-tuple $L = (x_1, \ldots, x_n)$ of distinct elements in $\mathbb{F}_{2^m}$ with $n \leq 2^m$. The trapdoor of the McEliece public-key scheme consists of a randomly picked $\Gamma$ which together with $L$ provide all the information to decode efficiently. The public key is a generator matrix of a randomly chosen Goppa code. A ciphertext is obtained by multiplying a plaintext with the public generator matrix and adding a random error vector of prescribed Hamming weight. The receiver decrypts the message thanks to the decoding algorithm that can be derived from the
secrets. Niederreiter [26] brings a significant modification of the McEliece cryptosystem by proposing to describe public linear codes through parity-check matrices. The resulting public key cryptosystem is as secure as McEliece’s one. The first code-based signature scheme came out in [12] almost twenty years McEliece’s proposal. The only difference between the encryption and the signature scheme lies in the choice of the parameters of the binary Goppa codes. For signature, Goppa codes have to be chosen such that they correct very few errors. This leads to a very high rate $R = k/n$ with $n$ is its length and $k$ being the dimension of the code. It holds that $k = n - rm$ where by definition $r$ is the number of errors and generally $n$ is chosen to be equal to $2^m$. For instance according to [18], an 80-bit security signature scheme imposes $r = 10$ and $m = 21$ which leads to $R = 0.9999$.

All these cryptographic primitives base their security under two assumptions: the intractability of decoding random linear codes [3], and the difficulty of recovering the private key or an equivalent one. The problem of decoding an unstructured code is a long-standing problem whose most effective algorithms [19, 20, 31, 10, 5] have an exponential time complexity. Thus, one may reasonably not expect much progress in this direction. On the other hand, no significant breakthrough has been observed during the last thirty years regarding the problem of recovering the private key. Indeed, although some weak keys have been identified in [21], the only known key-recovery attack is the exhaustive search of the secret polynomial $\Gamma$ of the Goppa code, and applying the Support Splitting Algorithm (SSA) [29] to check whether the Goppa code candidate is permutation-equivalent to the code defined by the public generator matrix. Despite the fact that there still does not exist a practical attack against McEliece’s proposal of using binary Goppa codes, one should not exclude the possibility of breakthrough in that field. The authors of [12] alleviated the McEliece assumptions by introducing the Goppa Code Distinguishing (GD) problem. They assume that no polynomial time algorithm exists that distinguishes a generator matrix of a Goppa code from a random generator matrix. This is a classical belief in code-based cryptography. For instance, according to [12], proving or disproving the hardness of the GD problem will have a significant impact: “Classification issues are in the core of coding theory since its emergence in the 50’s. So far nothing significant is known about Goppa codes, more precisely there is no known property invariant by permutation and computable in polynomial time which characterizes Goppa codes. Finding such a property or proving that none exists would be an important breakthrough in coding theory and would also probably seal the fate, for good or ill, of Goppa code-based cryptosystems”. Currently, the only known algorithm that solves GD problem is based on the enumeration of Goppa codes and the SSA algorithm [29], as explained below. The time complexity of this method is $O(2^{mr})$ assuming that the cost of the SSA algorithm is negligible (which is a reasonable assumption for Goppa codes, but not for all linear codes).

As a consequence, it is widely believed that distinguishing the public matrix in McEliece from a random matrix is computationally hard. Furthermore, the hardness of the Goppa Code Distinguishing (GD) problem is mandatory to prove the semantic and CCA2 security of McEliece in the random oracle model and in the standard model [27, 15, 8], the security in the random oracle model against existential forgery [12, 13] of the CFS signature [12] scheme, the provable security of several primitives such as a threshold ring signatures scheme [14], an identity-based identification scheme [11], which are build upon CFS. Therefore, showing that the Goppa Code Distinguishing problem is easier than expected will “unprove” most of the provable primitives based on McEliece, and more importantly will be the first serious cryptographic weakness observed on this scheme since thirty years. The purpose of this paper is to study the difficulty of the Goppa Code Distinguishing (GD) problem:

**Definition 1 (Goppa Code Distinguishing (GD) Problem).** Let $n$ and $k$ be two integers such that $k \leq n$. We denote by $\mathbf{Goppa}(n,k)$ the set of $k \times n$ generator matrices of Goppa codes. Similarly, $\mathbf{Random}(n,k)$ is the set of $k \times n$ random generator matrices. A distinguisher $\mathcal{D}$ is an algorithm that takes as input a matrix $G$ and returns a bit. We say that $\mathcal{D}$ solves the GD problem if it wins the following game:

- $b \leftarrow \{0,1\}$ if $b = 0$ then $G \overset{R}{\leftarrow} \mathbf{Goppa}(n,k)$ otherwise $G \overset{R}{\leftarrow} \mathbf{Random}(n,k)$
- If $\mathcal{D}(G) = b$ then $\mathcal{D}$ wins the games else $\mathcal{D}$ loses.
The probability that \( D \) outputs 1 when \( G \) is chosen as a random binary generator matrix of a Goppa code is denoted by \( \Pr[G \xrightarrow{R} \text{Random}(n,k) : D(G) = 1] \) and the probability that it outputs 1 when \( G \) is chosen randomly in \( \text{Random}(n,k) \) is denoted by \( \Pr[G \xrightarrow{R} \text{Random}(n,k) : D(G) = 1] \). We define the advantage of a distinguisher \( D \) as:

\[
\text{Adv}^{GD}(D) = \left| \Pr[G \xrightarrow{R} \text{Goppa}(n,k) : D(G) = 1] - \Pr[G \xrightarrow{R} \text{Random}(n,k) : D(G) = 1] \right|.
\]

In this paper, we present a deterministic polynomial-time distinguisher for solving the GD problem defined below with advantage close to 1 for codes of high rate. Along the way, we also solve the code distinguishing problem for alternant codes. The key ingredient is a new algebraic technique introduced in [17] to attack two variants [1, 24] of McEliece. It has been observed [17] that a key recovery attack can be used to construct an efficient distinguisher for alternant codes and Goppa codes over any alphabet. We performed extensive experiments to compare our theoretical results on valid McEliece public keys. They confirm that the generic formula are accurate. We emphasize that the Goppa Code Distinguishing problem has been widely considered as a hard problem in code-based cryptography as proved by the increasing number of papers using this assumption [27, 15, 8, 12–14, 11]. To our point of view, disproving/mitigating this hardness assumption is a breakthrough in code-based cryptography and may open a new direction to attack the McEliece cryptosystem. Although our attack remains theoretical, we believe that this work also permits to shed some light on the choice of secure parameters for McEliece cryptosystems; a topic thoroughly investigated recently [6, 7, 25, 18]. Our technique permits to indeed distinguish a public key of the CFS signature scheme for all parameters proposed by Finiasz and Sendrier [18]. Moreover, some realistic parameters of McEliece scheme also fit in the range of validity of our distinguisher like a binary Goppa code of length \( n = 2^{13} \) that corrects \( r = 19 \) errors. For these parameters, the scheme has a 90-bit security.

Organisation of Paper. In Section 2, we briefly recall the McEliece public-key cryptosystem as well as the Courtois-Finiasz-Sendrier CFS signature [12]. In Section 3, we recall several key features of Goppa and alternant codes. In Section 4, we precisely explain how we can mount an algebraic cryptanalysis against McEliece-like schemes i.e. namely how the algebraic system (1) is constructed. The distinguisher is presented in Section 5. Section 6 deals with the consequences of the existence of
a distinguisher in code-based cryptography. Finally, in Section 7 we explain how the formulas used in Section 5 have been obtained. To do so, we use together combinatorial properties of the linearized system and distinguishing features of Alternant/Goppa codes.

2 Code-Based Public-Key Cryptography

The main cryptographic primitives in code-based public-key cryptography are the McEliece encryption and the CFS signature [12]. We recall that a linear code over a finite field \( \mathbb{F}_q \) of \( q \) elements defined by a \( k \times n \) matrix \( G \) (with \( k \leq n \)) over \( \mathbb{F}_q \) is the vector space \( \mathcal{C} \) spanned by its rows \( i.e. \mathcal{C} \triangleq \{ uG \mid u \in \mathbb{F}_q^k \} \). \( G \) is chosen as a full-rank matrix, so that the code is of dimension \( k \). The rate of the code is given by the ratio \( \frac{k}{n} \). Code-based public-key cryptography focuses on linear codes that have a polynomial time decoding algorithm. The role of decoding algorithms is to correct errors of prescribed weight. We say that a decoding algorithm corrects \( t \) errors if it recovers \( u \) from the knowledge of \( uG + e \) for all possible \( e \in \mathbb{F}_q^n \) of weight at most \( t \).

**Secret key:** the triplet \( (S, G_s, P) \) of matrices defined over a finite field \( \mathbb{F}_q \) over \( q \) elements, with \( q \) being a power of two, that is \( q = 2^s \). \( G_s \) is a full rank matrix of size \( k \times n \), with \( k < n \), \( S \) is of size \( k \times k \) and is invertible. \( P \) is a permutation matrix of size \( n \times n \). \( G_s \) is chosen in such a way that its associated linear code (that is the set of all possible \( uG_s \) with \( u \) ranging over \( \mathbb{F}_q^k \)) has a decoding algorithm which corrects in polynomial time \( t \) errors.

**Public key:** the matrix \( G = SG_sP \).

**Encryption:** A plaintext \( u \in \mathbb{F}_q^k \) is encrypted by choosing a random vector \( e \) in \( \mathbb{F}_q^n \) of weight at most \( t \). The corresponding ciphertext is \( c = uG + e \).

**Decryption:** \( c' = cP^{-1} \) is computed from the ciphertext \( c \). Notice that \( c' = (uSG_sP + e)P^{-1} = uSG_s + eP^{-1} \) and that \( eP^{-1} \) is of Hamming weight at most \( t \). Therefore the aforementioned decoding algorithm can recover in polynomial time \( uS \) and therefore the plaintext \( u \) by multiplication by \( S^{-1} \).

What is generally referred to as the McEliece cryptosystem is this scheme together with a particular choice of the code, which consists in taking a binary Goppa code. This class of codes belongs to a more general class of codes (see Section 3, namely the alternant code family ([22, Chap. 12, p. 365])). The main feature of this last class of codes is that they can be decoded in polynomial time.

Another important code-based cryptographic primitive is the CFS scheme [12], which is the first signature scheme based on the security of the McEliece cryptosystem. In this kind of scheme, a user whose public key is \( P \) and who wishes to sign a message \( x \in \mathbb{F}_q^n \) has to compute a string \( u \) such that the Hamming weight of \( x - uG \) is at most \( t \). Anyone (a verifier) can publicly check the validity of a signature. Unfortunately, this approach can only provide signatures for messages \( x \) that are within distance \( t \) from a codeword \( uG \). The CFS scheme suggests to modify the message by appending a counter incremented until the decoding algorithm can find such a signature. The efficiency of this scheme heavily depends on the number of trials. It is suggested in [12] to choose as in the McEliece cryptosystem, binary Goppa codes for this purpose with the following parameters \( n = 2^m \) and \( k = n - mt \). The number of trials is of order \( t! \) in this case, which leads to choose a very small \( t \) and therefore to take a very large \( n \) in order to be secure. Notice that the code rate is then equal to \( \frac{2^m - tm}{2^m} = 1 - \frac{mt}{2^m} \) which is for large values of \( 2^m \) and moderate values of \( t \) quite close to 1. Thus, the major difference between the McEliece cryptosystem and the CFS scheme lies in the choice of the parameters. An 80-bit security CFS scheme requires \( n = 2^{21} \) and \( t = 10 \) whereas the McEliece cryptosystem for the same security needs \( n = 2^{11} \) and \( t = 32 \) ([18]). The code of the CFS scheme is of rate \( 1 - \frac{10 \times 21}{2^{21}} \approx 0.9999 \). We see here that the CFS scheme depends on very high rate binary Goppa codes.
3 Basic Facts about Alternant and Goppa Codes

As explained in the previous section, the McEliece cryptosystem relies on Goppa codes which belong to the class of alternant codes and inherit an efficient decoding algorithm from this. It is convenient to describe this class through a parity-check matrix over an extension field $\mathbb{F}_{q^m}$ of $\mathbb{F}_q$ over which the code is defined. In other words, the parity check matrix is an $r \times n$ matrix $H$ with coefficients in $\mathbb{F}_{q^m}$ and the associated alternant code $\mathcal{A}$ is the set of vectors of $\mathbb{F}_q^n$ which belong to the right kernel of $H$, i.e.

$$\mathcal{A} = \{ c \in \mathbb{F}_q^n \mid Hc^T = 0 \}. \quad (2)$$

$r$ satisfies in this case the condition $r \geq \frac{n-k}{m}$ where $k$ is the dimension of $\mathcal{A}$. For alternant codes, there exists a parity-check matrix with a very special form related to Vandermonde matrices. For reasons which will be made clear in Section 4, it will be convenient to work with the projective plane $\mathbb{F}_{q^m} \triangleq \mathbb{F}_{q^m} \cup \{ \infty \}$ and to consider the class of projective alternant codes (which are slightly more general than classical alternant codes). More precisely, any projective alternant code has a parity check matrix which is of the form

$$V_r(x, y) \overset{\text{def}}{=} \begin{pmatrix} y_1 & \cdots & y_n \\ y_1 x_1 & \cdots & y_n x_n \\ \vdots & \vdots & \vdots \\ y_1 x_1^{r-1} & \cdots & y_n x_n^{r-1} \end{pmatrix}. \quad (3)$$

where $x = (x_1, \ldots, x_n) \in (\mathbb{F}_{q^m})^n$ and $y = (y_1, \ldots, y_n)$ in $(\mathbb{F}_{q^m})^n$. When $x_i = \infty$ we use the convention that the $i$-th column of $V_r(x, y)$ is equal to $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ y_i \end{pmatrix}$.

**Definition 2 (Projective and classical alternant code).** The projective alternant code of order $r$ over $\mathbb{F}_q$ associated to $x = (x_1, \ldots, x_n) \in (\mathbb{F}_{q^m})^n$ (where all $x_i$’s are distinct) and $y = (y_1, \ldots, y_n) \in (\mathbb{F}_{q^m})^n$, denoted by $\mathcal{A}_r(x, y)$, is defined by

$$\mathcal{A}_r(x, y) = \{ c \in \mathbb{F}_q^n \mid V_r(x, y)c^T = 0 \}. \quad (4)$$

A classical alternant code corresponds to the case where all $x_i$’s are different from $\infty$.

The class of Goppa codes is a subfamily of alternant codes which are given by:

**Definition 3 (Projective and classical Goppa codes).** The projective Goppa code $G(x, \Gamma)$ over $\mathbb{F}_q$ associated to a polynomial $\Gamma(x)$ of degree $r$ over $\mathbb{F}_{q^m}$ and a certain $n$-tuple $x = (x_1, \ldots, x_n)$ of distinct elements of $\mathbb{F}_{q^m}$ satisfying $\Gamma(x_i) \neq 0$ for all $i$, $1 \leq i \leq n$, is the alternant code $\mathcal{A}_r(x, y)$ of order $r$ with $y_i$ being defined by $y_i = \Gamma(x_i)^{-1}$. A classical Goppa code corresponds to the case $x_i \in \mathbb{F}_q$ for all $i$ in $\{1, \ldots, n\}$.

It should be noted that the public code in the McEliece cryptosystem is also an alternant code. This is a simple consequence of the fact that $\{ uSG \gamma P \mid u \in \mathbb{F}_q^k \}$ is obtained from the secret code $\{ uG \gamma \mid u \in \mathbb{F}_q^k \}$ by permuting the coordinates in it with the help of $P$, since multiplying by an invertible matrix $S$ of size $k \times k$ leaves the code globally invariant.

---

4 We define $\Gamma(\infty) \overset{\text{def}}{=} \gamma_r$ for $\Gamma(X) = \sum_{i=0}^r \gamma_i X^i$. 

4 Algebraic Cryptanalysis of McEliece-like Cryptosystems

In this part, we explain more precisely how we construct the algebraic system described in (1). This algebraic system is the main ingredient of the distinguisher. We recall a key feature of alternant codes.

**Fact 1.** There exists a polynomial time algorithm decoding all errors of Hamming weight at most \( \frac{r}{2} \) for an alternant code of order \( r \) once a parity-check matrix \( H \) of the form \( H = V_r(x, y) \) is given for it.

The variants of McEliece’s cryptosystem based on general alternant codes or on non binary Goppa codes, such as [1, 24] for instance, add errors which are of weight smaller than or equal to \( r/2 \). In this case, it is possible to break these variants by finding \( x^* \) and \( y^* \) in \( \mathbb{F}_{q^m} \) such that:

\[
\{ xG \mid x \in \mathbb{F}_q^r \} = \{ y \in \mathbb{F}_q^n \mid V_r(x^*, y^*)y^T = 0 \}.
\]

According to Fact 1, the knowledge of \( V_r(x^*, y^*) \) permits to efficiently decode the public code, i.e. to recover \( u \) from \( uG + e \). By the very definition of the public code \( G \), we have:

\[
V_r(x^*, y^*)G^T = 0.
\]

This is the key observation of the algebraic approach used in [17] to cryptanalyze dyadic and quasi-cyclic variants of McEliece. Let \( X_1, \ldots, X_n \) and \( Y_1, \ldots, Y_n \) be \( 2n \) variables corresponding to the \( x_i^* \)'s and the \( y_i^* \)'s. Observe that such \( x_i^* \)'s and \( y_i^* \)'s are a particular solution of the following system:

\[
\left\{ g_{i,j}X_iY_j + \cdots + g_{i,n}Y_nX_n = 0 \mid i \in \{1, \ldots, k\}, j \in \{0, \ldots, r-1\} \right\}
\]

(6)

where the \( g_{i,j} \)'s are the entries of the known matrix \( G \). In the cases of [1, 24], additional structures permit to drastically reduce the number of variables allowing to solve (1) efficiently using dedicated Gröbner bases techniques [17].

For binary Goppa codes, it is essential to recover its description as a Goppa code and not only the \( x_i \)'s and the \( y_i \)'s giving its description as an alternant code. This is a consequence of the following result.

**Fact 2.** [28] There exists a polynomial time algorithm decoding all errors of Hamming weight at most \( r \) in a Goppa code \( \mathcal{G}(x, \Gamma) \) when \( \Gamma \) has degree \( r \) and has no multiple roots, if \( x \) and \( \Gamma \) are known.

If we recover only the \( x_i \)'s and the \( y_i \)'s we can decode only \( r/2 \) errors. The issue is now, once a possible description of a Goppa code has been found as an alternant code, that is once a solution \( x = (x_i)_{1 \leq i \leq n} \) and \( y = (y_i)_{1 \leq i \leq n} \) of the system (6) has been found, does there exist a polynomial \( \Gamma(X) \) of degree \( r \) such that \( y_i = \Gamma(x_i)^{-1} \) for all \( i \in \{1, \ldots, n\} \)? If such a polynomial exists, it can be easily found by interpolation. The problem is that a Goppa code has multiple descriptions as an alternant code, i.e., there are several \( x,y \)'s for which \( \mathcal{G} = \mathcal{A}(x, y) \). The solutions we are interested in are the ones for which \( y_i = \Gamma(x_i)^{-1} \) for all \( i \), and for some polynomial \( \Gamma \) of degree \( r \).

This raises the fundamental issue of finding all possible descriptions of the form (4) of an alternant code \( \mathcal{A} \), that is find all \( x,y \)'s such that \( \mathcal{A} = \mathcal{A}(x, y) \). When the extension field \( \mathbb{F}_{q^m} \) is the same as the definition\(^5\) field \( \mathbb{F}_q \), i.e. if \( m = 1 \), the problem was solved in [16]. This was the key of the cryptanalysis of McEliece’s variant based on generalized Reed-Solomon codes [30].

The general case is still unsolved. However, the results of [16] basically show that the we have at least one degree of freedom for \( Y_1 \) and three degrees of freedom for the \( X_i \)'s in the system (6). First of all it is straightforward to notice that if \( (X_i)_{1 \leq i \leq n}, (Y_i)_{1 \leq i \leq n} \) is a solution of the algebraic Equation (6) then \( (\alpha X_i)_{1 \leq i \leq n}, (\beta Y_i)_{1 \leq i \leq n} \) is also a solution for any \( \alpha, \beta \) in \( \mathbb{F}_{q^m} \). Therefore, we can specialize one \( (X_i, Y_i) \) arbitrarily. It turns out we can fix more variables thanks to the following proposition.

\(^5\) This means that the resulting code is a slight generalization of a generalized Reed-Solomon code known under the name of a Cauchy code.
Proposition 1. Let \( x = (x_i)_{1 \leq i \leq n} \in (\mathbb{F}_{q^m})^n \) be an \( n \)-tuple formed by distinct elements and let \( y = (y_i)_{1 \leq i \leq n} \in (\mathbb{F}_{q^m})^n \) be an \( n \)-tuple of nonzero elements. Let \( a, b, c, d \) be elements of \( \mathbb{F}_{q^m} \) such that that \( ad - bc \neq 0 \). Then

\[
\mathcal{A}_r \left( \frac{ax + b}{cx + d}, y \right) = \mathcal{A}_r (x, y),
\]

where

\[
\frac{ax + b}{cx + d} \overset{\text{def}}{=} (x_i')_{1 \leq i \leq n} \text{ with } x_i' = \frac{ax_i + b}{cx_i + d}, \ y'_i = (y_i')_{1 \leq i \leq n} \text{ with } y_i' = y_i(cx_i + d)^{-1}.
\]

Remark 1. The proof is in Appendix A. Notice that either \( x_i \) or \( x_i' \) might be infinite. We used here the usual rules to evaluate the homography \( z \mapsto \frac{az + b}{cz + d} \), namely \( \frac{a}{0} = \infty \), \( \frac{0}{0} = \infty \), \( \frac{a}{\infty} = 0 \), \( \beta + \infty = \infty \), \( 0 \times \infty = 0 \), \( \frac{a \times \infty + b}{c \times \infty + d} = \frac{a}{c} \), where \( \alpha \neq 0, \beta \) belong to \( \mathbb{F}_{q^m} \).

This result explains that there is (at least) one degree of freedom for the \( Y_i \)'s and three degrees of freedom for the \( X_i \)'s. It is quite helpful to allow here \( x_i \) which can be infinite since even all of them are in \( \mathbb{F}_{q^m} \), it might happen that \( cx_i + d \) is equal to zero. Therefore the corresponding image by the homography will be infinite. Finally, since the set of homographies acts 3-transitively over \( \mathbb{F}_{q^m} \cup \{ \infty \} \), we have:

Corollary 1. We can specialize (almost) randomly one \( Y_i \) and three \( X_i \)'s in (1). As long as the \( X_i \)'s are distinct, we still have a non-empty set of solutions for such modified system (1).

At first glance, the degree of freedom should be less for Goppa codes. Indeed, there is an additional crucial constraint for binary Goppa codes: a solution must verify \( Y_i = \Gamma(X_i)^{-1} \) for a certain polynomial of degree \( r \). Surprisingly, we can keep the same degree of freedom by considering a slight change of \( x_i \). Let \( \mathcal{G}(x, \Gamma) \) be the subcode of the Goppa code \( \mathcal{G}(x, \Gamma) \) formed by all codewords of even Hamming weight. Let \( \tilde{G} = (\tilde{g}_{i,j})_{1 \leq i \leq \tilde{k}} \) be a generator matrix of \( \mathcal{G}(x, \Gamma) \), that is a matrix of full rank whose rows generate \( \mathcal{G}(x, \Gamma) \). The dimension \( \tilde{k} \) of this subspace is either \( k \) or \( k - 1 \), where \( k \) is the dimension of the Goppa code \( \mathcal{G}(x, \Gamma) \). This subcode is itself an alternate code.

Proposition 2. \([2]\) It holds that:

\[
\tilde{G}(x, \Gamma) = \mathcal{A}_{r+1}(x, y)
\]

for \( \deg(\Gamma) = r \) and where \( y = (y_i)_i \) with \( y_i = \Gamma(x_i)^{-1} \).

This implies that the \( x_i \)'s and \( y_i \)'s are a particular solution of:

\[
\begin{align*}
\tilde{g}_{i,1} Y_1 X_i^1 + \cdots + \tilde{g}_{i,n} Y_n X_i^n &= 0 & i \in \{1, \ldots, \tilde{k}\}, j \in \{0, \ldots, r\}
\end{align*}
\]

(7)

where the \( \tilde{g}_{i,j} \)'s are the entries of the known matrix \( \tilde{G} \). Notice that this system is very similar to (6) with the exception that the powers of the \( X_i \)'s can now be equal to \( r \). The crucial result is now that

Proposition 3. \([2]\) Let \( x = (x_i)_{1 \leq i \leq n} \) be an \( n \)-tuple of distinct elements of \( \mathbb{F}_{q^m} \) and \( \Gamma \) be a polynomial of degree \( r \) such that \( \Gamma(x_i) \neq 0 \) for all \( i \in \{1, \ldots, n\} \). Let \( \psi(z) = \frac{az + b}{cz + d} \) be an homography with \( ad - bc \neq 0 \) and \( a, b, c, d \in \mathbb{F}_{q^m} \). Let \( x^\psi \overset{\text{def}}{=} (x_i^\psi)_{1 \leq i \leq n} \) with \( x_i^\psi \overset{\text{def}}{=} \psi^{-1}(x_i) \), \( \Gamma^\psi(X) \overset{\text{def}}{=} (cx + d)^{-1} \Gamma(x) = \sum_{i=0}^{r} \gamma_i X_i^i \). Then

\[
\mathcal{G}(x, \Gamma) = \tilde{G}(x^\psi, \Gamma^\psi).
\]

Once again, we can use that homographies have a 3-transitive action on \( \mathbb{F}_{q^m} \).

Corollary 2. We can specialize in (7) one of the \( Y_i \) and three of the \( X_i \)'s almost arbitrarily (with \( Y_i \neq 0 \) and such that the three \( X_i \)'s are distinct) and still obtain a solution for which there exists a polynomial \( \Gamma \) of degree \( r \) such that \( Y_i = \Gamma(X_i)^{-1} \) for all \( i \in \{1, \ldots, n\} \).
To finish this discussion, it will be helpful to notice that in the case of binary Goppa codes, we have even more algebraic equations than the ones given in System (6). The starting point is the following result, which is essentially derived from a discussion in a paragraph about Goppa codes in [22, p.341].

**Theorem 3.** A binary Goppa code \( G(x, \Gamma) \) associated to a Goppa polynomial \( \Gamma(X) \) of degree \( r \) without multiple roots is equal to the alternant code \( \mathcal{A}_{2r}(x, y) \), with \( y_i = \Gamma(x_i)^{-2} \).

In other words, \( x \) and \( y \) are solutions of the following algebraic system

\[
\begin{align*}
g_{i,1}Y_1X_1^j + \cdots + g_{i,n}Y_nX_n^j &= 0 \quad i \in \{1, \ldots, k\}, j \in \{0, \ldots, 2r-1\} \\
\end{align*}
\]

where \((g_{il})\) is a generator matrix of the Goppa code. Notice that the powers \( j \) are now in the range \( \{0, 1, \ldots, 2r-1\} \) and not in \( \{0, 1, \ldots, r-1\} \), as was the case before.

### 5 A Distinguisher of Alternant and Goppa Codes

We present in this part the algebraic distinguisher. Let \( G = (g_{ij})_{1 \leq i \leq k \atop 1 \leq j \leq n} \) be a generator matrix of the public code. We can assume without loss of generality that \( G \) is systematic in its \( k \) first positions. Such a form can be easily obtained by Gaussian elimination and by a suitable permutation of the columns. We describe now a simple way of using this particular form for solving (6). We assume that the rate of the public code is close to 1, i.e. \( \frac{k-mr}{n} \approx 1 \), which implies \( mr \ll n \). From a cryptographic point of view, this means that the expansion ratio between the size of the ciphertext and the size of the message is close to 1. This kind of rate has been proposed in [18]. The strategy is as follows.

#### 5.1 First step – expressing the \( Y_iX_i^d \)'s in terms of the \( Y_jX_j^d \)'s for \( j \in \{k+1, \ldots, n\} \).

Let \( P = (p_{ij})_{1 \leq i \leq k \atop k+1 \leq j \leq n} \) be the submatrix of \( G \) formed by its last \( mr \) columns (as in Figure 1). We can rewrite (6) as

\[
\begin{align*}
Y_i &= \sum_{j=k+1}^{n} p_{ij}Y_j \\
Y_iX_i &= \sum_{j=k+1}^{n} p_{ij}Y_jX_j \\
\cdots \\
Y_iX_i^{r-1} &= \sum_{j=k+1}^{n} p_{ij}Y_jX_j^{r-1} \\
\end{align*}
\]

for all \( i \in \{1, \ldots, k\} \).
5.2 Second step – using the trivial identity $Y_i Y_i X_i^2 = (Y_i X_i)^2$ and linearization.

Thanks to the trivial identity $Y_i Y_i X_i^2 = (Y_i X_i)^2$ for all $i$ in $\{1, \ldots, k\}$, we get:

$$\sum_{j=k+1}^{n} p_{i,j} Y_j \sum_{j=k+1}^{n} p_{i,j} Y_j X_j^2 = \left( \sum_{j=k+1}^{n} p_{i,j} Y_j X_j \right)^2, \text{ for all } i \in \{1, \ldots, k\}.$$  

It is possible to reorder this a little bit to obtain the following equations:

$$\sum_{j=k+1}^{n} \sum_{j'>j} p_{i,j} p_{i,j'} (Y_j Y_{j'} X_j^2 + Y_{j'} Y_j X_j^2) = 0$$  

(10)

We can now linearize this system by letting $Z_{jj'} \overset{\text{def}}{=} Y_j Y_{j'} X_j^2 + Y_{j'} Y_j X_j^2$. We obtain $k$ linear equations involving the $Z_{jj'}$’s:

$$\left\{ \sum_{j=k+1}^{n} \sum_{j'>j} p_{i,j} p_{i,j'} Z_{jj'} = 0, i = 1 \ldots k \right\}.$$  

(11)

To solve this system it is necessary that the number of equations is greater than the number of unknowns, i.e.:

$$k \geq \left( \frac{mr}{2} \right).$$

This approach works for alternant codes in general. However, for Goppa codes, it will be interesting to consider also a related system. It is obtained by applying the same approach described before but to the generator matrix $G$ of the subcode of the public code consisting in codewords of even Hamming weight. The reason which makes this new system interesting will be explained in Subsection 7.2, it is related to Proposition 2. We denote by $\tilde{k}$ the dimension of this code. We have either $\tilde{k} = k$ or $\tilde{k} = k - 1$.

As previously, we can suppose that $\tilde{G}$ is in systematic form: $\tilde{G} = (\tilde{I} | \tilde{P})$ where $\tilde{I}$ is the identity matrix of size $k$ or $k - 1$ (depending on the dimension of the subcode). Finally, let $\tilde{p}_{ij}$ be the coefficient in the $i$-th row and $j$-th column of $\tilde{P}$. We can proceed similarly and obtain a new linear system of equations:

$$\left\{ \sum_{j=k+1}^{n} \sum_{j'>j} \tilde{p}_{i,j} \tilde{p}_{i,j'} Z_{jj'} = 0, i = 1 \ldots \tilde{k} \right\}.$$  

(12)

When $\tilde{k} = k - 1$, the number of equations is smaller. It might be $k - 1$ instead of $k$ and the number of variables is also larger. It is equal to $\binom{n-k}{2} = \binom{mr+1}{2}$. However, we will see that due to Proposition 2, this system has also nice properties in the Goppa case.

5.3 Experimental behavior

Observe that the linear systems (11) and (12) have coefficients in $\mathbb{F}_q$, whereas solutions are sought in the extension field $\mathbb{F}_{q^n}$. In addition, the freedom of choosing three $X_i$’s and one $Y_i$ in order to reduce the number of unknowns in the linearized systems is not used. However, even if this additional knowledge is taken into account, the rank of the linear systems remains insufficient to solve the system. More precisely, the problem is that the dimension of the vector space solution of (11) is amazingly large. It even depends on whether or not the code with generator matrix $G$ is chosen as a (generic) alternant code or as a Goppa code. Interestingly enough, when $G$ is chosen at random, the dimension of the solution space is typically 0 when $k$ is larger than the number of variables. Although these facts are an obstacle to break the McEliece cryptosystem, it can be used to distinguish the public generator from a random code. Let us denote by:
\(- N \overset{\text{def}}{=} \binom{mr}{2}\) the number of variables in (11), \(\tilde{N}\) the number of variables of (12),
- \(D_{\text{random}}\), respectively \(\tilde{D}_{\text{random}}\), the dimension of the vector space solution of (11), respectively (12) when the \(p_{ij}\)'s are chosen uniformly at random in \(\mathbb{F}_q\),
- \(D_{\text{alternant}}\), respectively \(\tilde{D}_{\text{alternant}}\), the dimension of the vector space solution of (11), respectively (12) when \(G\) is chosen as a generator matrix of a random alternant code of degree \(r\),
- \(D_{\text{Goppa}}\), respectively \(\tilde{D}_{\text{Goppa}}\) the dimension of the vector space solution of (11), respectively (12) when \(G\) is chosen as a generator matrix of a random Goppa code of degree \(r\).

A thorough experimental study revealed that the dimension of the vector space over \(\mathbb{F}_q\) of the solutions of (11) follows typically the following formulas:

**Experimental fact 1** Let \(D\) be in \(\{D_{\text{alternant}}, \tilde{D}_{\text{alternant}}, D_{\text{Goppa}}, \tilde{D}_{\text{Goppa}}\}\). With very high probability and as long as \(N - D < k\), the dimension \(D\) has the following value:

\[
D_{\text{alternant}} = \frac{m(r - 1)(r - 2)}{2} \left( (2\ell + 1)r - 2q^{\ell+1} - \frac{q - 1}{q - 1} \right) \quad \text{for } \ell \overset{\text{def}}{=} \lfloor \log_q(r - 1) \rfloor \quad (13)
\]

\[
\tilde{D}_{\text{alternant}} = D_{\text{alternant}} \text{ for } q > 2
\]

\[
D_{\text{Goppa}} = \frac{mr(2\ell + 1)r - 2q^{\ell+1} + 2q^{\ell-1} - 1}{2} \quad (15)
\]

\[
\tilde{D}_{\text{Goppa}} = \frac{mr}{2} \left( (2\ell + 1)r - 2q^{\ell+1} + 2q^{\ell-1} + 1 \right) \quad (16)
\]

For \(r < q - 1\), it holds that 

\[
D_{\text{Goppa}} = \frac{m(r - 1)(r - 2)}{2} = D_{\text{alternant}} \quad (17)
\]

\[
\tilde{D}_{\text{Goppa}} = \frac{mr}{2} \left( (2\ell + 1)r - 2q^{\ell+1} + 2q^{\ell-1} + 1 \right)
\]

wheras for \(r \geq q - 1\), by denoting by \(\ell\) the unique integer such that \(q^\ell - 2q^{\ell+1} + 2q^{\ell-1} < r \leq q^{\ell+1} - 2q^{\ell+1} + q^{\ell-1}\), we obtain

\[
D_{\text{Goppa}} = \frac{mr}{2} \left( (2\ell + 1)r - 2q^{\ell+1} + 2q^{\ell-1} - 1 \right)
\]

\[
\tilde{D}_{\text{Goppa}} = \frac{mr}{2} \left( (2\ell + 1)r - 2q^{\ell+1} + 2q^{\ell-1} + 1 \right)
\]

We gathered samples of results we obtained through intensive computations with the Magma system [9] in order to confirm the formulas. We randomly generated alternant and Goppa codes over the field \(\mathbb{F}_q\) with \(q \in \{2, 4, 8, 16, 32\}\) for values of \(r\) in the range \(\{3, \ldots, 50\}\) and several \(m\). The Goppa codes are generated by means of an irreducible \(\Gamma\) of degree \(r\) and hence \(\Gamma\) has no multiple roots. In particular, we can apply Theorem 3 in the binary case. We compare the dimensions of the solution space against the dimension \(D_{\text{random}}\) of the system derived from a random linear code. Table 1 and Table 2 give figures for the binary case with \(m = 14\). We define \(T_{\text{alternant}}\) and \(T_{\text{Goppa}}\) respectively as the expected dimensions for an alternant and a Goppa code deduced from the formulas (13) and (15)-(17). We can check that \(D_{\text{random}}\) is equal to \(0\) for \(r \in \{3, \ldots, 12\}\) and \(D_{\text{random}} = N - k\) as expected. We remark that \(D_{\text{alternant}}\) is different from \(D_{\text{random}}\) whenever \(r \leq 15\), and \(D_{\text{Goppa}}\) is different from \(D_{\text{random}}\) as long as \(r \leq 25\). Finally we observe that our formulas for \(T_{\text{alternant}}\) fit as long as \(k \geq N - T_{\text{alternant}}\) which correspond to \(r \leq 15\). This is also the case for binary Goppa codes since we have \(T_{\text{Goppa}} = D_{\text{Goppa}}\) as long as \(k \geq N - T_{\text{Goppa}}\) i.e. \(r \leq 25\). We also give in Table 10 and Table 11 in Appendix B the examples that we obtained for \(q = 4\) and \(m = 6\) to check that the arguments also apply. We also compare binary Goppa codes and random linear codes for \(m = 15\) in Table 4-6 and \(m = 16\) in Table 7-9 (See Appendix B). We see that \(D_{\text{random}}\) and \(D_{\text{Goppa}}\) are different for \(r \leq 33\) when \(m = 15\) and for \(m = 16\) they are different even beyond our range of experiment \(r \leq 50\).

### 6 Cryptographic Implications

The existence of a distinguisher for the specific case of binary Goppa codes has consequences for code-based cryptographic primitives because it is represents, and by far, the favorite choice in such
Table 1. \( q = 2 \) and \( m = 14 \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>861</td>
<td>1540</td>
<td>2415</td>
<td>3486</td>
<td>4753</td>
<td>6216</td>
<td>7875</td>
<td>9730</td>
<td>11781</td>
<td>14028</td>
<td>16471</td>
<td>19110</td>
<td>21945</td>
<td>24976</td>
</tr>
<tr>
<td>( k )</td>
<td>16342</td>
<td>16232</td>
<td>16314</td>
<td>16300</td>
<td>16286</td>
<td>16272</td>
<td>16258</td>
<td>16244</td>
<td>16230</td>
<td>16216</td>
<td>16202</td>
<td>16188</td>
<td>16174</td>
<td>16160</td>
</tr>
<tr>
<td>( D_{\text{random}} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( D_{\text{alternant}} )</td>
<td>42</td>
<td>126</td>
<td>308</td>
<td>560</td>
<td>882</td>
<td>1274</td>
<td>1848</td>
<td>2520</td>
<td>3290</td>
<td>4158</td>
<td>5124</td>
<td>6188</td>
<td>7350</td>
<td>8816</td>
</tr>
<tr>
<td>( T_{\text{alternant}} )</td>
<td>42</td>
<td>126</td>
<td>308</td>
<td>560</td>
<td>882</td>
<td>1274</td>
<td>1848</td>
<td>2520</td>
<td>3290</td>
<td>4158</td>
<td>5124</td>
<td>6188</td>
<td>7350</td>
<td>8816</td>
</tr>
<tr>
<td>( D_{\text{Goppa}} )</td>
<td>252</td>
<td>532</td>
<td>980</td>
<td>1554</td>
<td>2254</td>
<td>3080</td>
<td>4158</td>
<td>5390</td>
<td>6776</td>
<td>8316</td>
<td>10010</td>
<td>11858</td>
<td>13860</td>
<td>16161</td>
</tr>
<tr>
<td>( T_{\text{Goppa}} )</td>
<td>252</td>
<td>532</td>
<td>980</td>
<td>1554</td>
<td>2254</td>
<td>3080</td>
<td>4158</td>
<td>5390</td>
<td>6776</td>
<td>8316</td>
<td>10010</td>
<td>11858</td>
<td>13860</td>
<td>16161</td>
</tr>
</tbody>
</table>

Table 2. \( q = 2 \) and \( m = 14 \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>28263</td>
<td>31626</td>
<td>35249</td>
<td>39006</td>
<td>43071</td>
<td>47278</td>
<td>51681</td>
<td>56280</td>
<td>61075</td>
<td>66066</td>
<td>71253</td>
<td>76636</td>
<td>82215</td>
<td>87990</td>
</tr>
<tr>
<td>( k )</td>
<td>16146</td>
<td>16132</td>
<td>16118</td>
<td>16104</td>
<td>16090</td>
<td>16062</td>
<td>160048</td>
<td>16068</td>
<td>16024</td>
<td>16006</td>
<td>15992</td>
<td>15978</td>
<td>15964</td>
<td></td>
</tr>
<tr>
<td>( D_{\text{random}} )</td>
<td>12057</td>
<td>15494</td>
<td>19127</td>
<td>22956</td>
<td>26981</td>
<td>31202</td>
<td>35619</td>
<td>40232</td>
<td>45041</td>
<td>50046</td>
<td>55247</td>
<td>60644</td>
<td>66237</td>
<td>72026</td>
</tr>
<tr>
<td>( D_{\text{alternant}} )</td>
<td>12057</td>
<td>15494</td>
<td>19127</td>
<td>22956</td>
<td>26981</td>
<td>31202</td>
<td>35619</td>
<td>40232</td>
<td>45041</td>
<td>50046</td>
<td>55247</td>
<td>60644</td>
<td>66237</td>
<td>72026</td>
</tr>
<tr>
<td>( T_{\text{alternant}} )</td>
<td>10192</td>
<td>11900</td>
<td>13734</td>
<td>15694</td>
<td>17780</td>
<td>19992</td>
<td>22330</td>
<td>24794</td>
<td>27384</td>
<td>30100</td>
<td>32942</td>
<td>35910</td>
<td>39004</td>
<td>42224</td>
</tr>
<tr>
<td>( D_{\text{Goppa}} )</td>
<td>18564</td>
<td>21294</td>
<td>24206</td>
<td>27300</td>
<td>30576</td>
<td>34034</td>
<td>37674</td>
<td>41496</td>
<td>45500</td>
<td>50046</td>
<td>55247</td>
<td>60644</td>
<td>66237</td>
<td>72026</td>
</tr>
<tr>
<td>( T_{\text{Goppa}} )</td>
<td>18564</td>
<td>21294</td>
<td>24206</td>
<td>27300</td>
<td>30576</td>
<td>34034</td>
<td>37674</td>
<td>41496</td>
<td>45500</td>
<td>50046</td>
<td>55247</td>
<td>60644</td>
<td>66237</td>
<td>72026</td>
</tr>
</tbody>
</table>

primitives. One of the reasons for this, is the fact that this class has withstood many cryptographic attacks for more than thirty years now. We focus in this part on secure parameters that are within the range of validity of our distinguisher. In Section 5, we gave a general expression of the distinguisher for a Goppa code over any finite field \( \mathbb{F}_q \). This expression can be easily simplified in the binary case \((q = 2)\).

**Proposition 4.** Let us define \( \ell \) \(\text{def} \) \( \lceil \log_2 r \rceil + 1 \) and \( N' \) \(\text{def} \) \( \left( \frac{m r}{2} \right) \). The formula for \( D_{\text{Goppa}} \) given in Equation (17) can be simplified to \( D_{\text{Goppa}} = \frac{m r}{2} \left( 2^\ell + 1 \right) r - 2^\ell - 1 \) as long as \( N - D_{\text{Goppa}} < n - m r \).

This simple expression is therefore not true for any value of \( r \) and \( m \) but tends to be true for codes that have a code rate \( \frac{n-m r}{n} \) that is close to one. This kind of codes are mainly encountered with the public keys of the CFS signature scheme. We will show that there also exist public keys of the McEliece cryptosystem that can be distinguished for parameters considered as secure. We assume that the length \( n \) is equal to \( 2^m \) and we denote by \( r_{\text{min}} \) the smallest integer \( r \) such that \( N - D_{\text{Goppa}} \geq 2^m - m r \). Recall that given a degree extension \( m \) over \( \mathbb{F}_2 \), any binary Goppa code defined with a polynomial \( \Gamma(z) \) of
degree \( r \geq r_{\text{min}} \) cannot be distinguished from a random linear code by our technique. This value is gathered in Table 3 for different values of \( m \). It provides therefore a lower bound for \( r \) in the choice of secure parameters if being unable to distinguish the public code from a random linear code is required. One can notice for instance that the McEliece key obtained with \( m = 13 \) and \( r = 19 \) and which corresponds to 90-bit of security, fits in the range of validity of our distinguisher. The values of \( r_{\text{min}} \) in Table 3 are checked by experimentations for \( m \leq 16 \) whereas those for \( m \geq 17 \) are obtained by solving the equation \( \frac{mr}{2} (2\ell + 1)r - 2\ell - 1 = \frac{1}{2}mr(mr - 1) - 2^m + mr \). Additionally, all the keys proposed in [18] (See therein Table 4) for the CFS scheme can be distinguished.

<table>
<thead>
<tr>
<th>Table 3. Smallest order ( r ) of a binary Goppa code of length ( n = 2^m ) for which our distinguisher does not work.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m )</td>
</tr>
<tr>
<td>( r_{\text{min}} )</td>
</tr>
</tbody>
</table>

7 An Explanation for the Distinguisher

The goal of this section is to provide a theoretical explanation to the practical behavior observed in the previous section. We first consider the case of alternant codes and will explain the defect of rank observed in the linearized systems described previously.

7.1 The generic alternant case

As a general comment, we emphasize that it seems difficult to obtain a precise lower bound or upper bound on the dimension \( D \), respectively \( D \) of the vector space solution of (11), respectively (12) holding for all alternant codes. Indeed, it is always possible to have degenerate cases for particular \( x \) and \( y \) defining the alternant code \( \mathcal{O}_q^r(x, y) \). When \( x \) and \( y \) are chosen in a subfield \( F_q^{m'} \) with \( m' \) being a divisor of \( m \), then the dimension \( D \) of the system is much smaller than predicted in experimental Fact 1. We have typically the same formula as in (13), but with \( m' \) replacing \( m \) there. On the other hand, when \( y \) is chosen accordingly to a Goppa code, then the dimension can be much larger.

However, there is a simple fact explaining what happens in the generic case for Formula (13), i.e. for “random” choices of \( x \) and \( y \). Indeed, to set up the linear system (11) or (12) we have used the trivial identity \( Y_iY_jX_i^aX_j^b = (Y_iX_j)^2 \). More generally, we can use any identity of the form \( Y_iX_j^aX_j^d = Y_iX_j^bY_jX_j^d \) with \( a, b, c, d \in \{0, 1, \ldots, r - 1\} \) such that \( a + b = c + d \). It is straightforward to check that we obtain in the same way the algebraic system:

\[
\sum_{j=k+1}^n \sum_{j' > j} p_{k,j} \tilde{p}_{i,j'} (Y_iX_j^aY_j'X_j'^b + Y_jX_j'^aY_jX_j'^b + Y_jX_j'^aY_j'X_j^d + Y_jX_j'^bY_j'X_j^d) = 0
\]

and

\[
\sum_{j=k+1}^n \sum_{j' > j} \tilde{p}_{i,j} \tilde{p}_{i,j'} (Y_iX_j^aY_j'X_j'^b + Y_jX_j'^aY_jX_j'^b + Y_jX_j'^aY_j'X_j^d + Y_jX_j'^bY_j'X_j^d) = 0.
\]

In other words:

\[
Z_{a,b,c,d} \overset{\text{def}}{=} (Y_iX_j^aY_j'X_j'^b + Y_jX_j'^aY_jX_j'^b + Y_jX_j'^aY_j'X_j^d + Y_jX_j'^bY_j'X_j^d)_{1 \leq j < j' \leq mr}
\]
is a solution of (11) whereas
\[
\mathbf{Z}_{a,b,c,d} \stackrel{\text{def}}{=} (Y_j X_j^a Y_j^b X_j^c + Y_j X_j^a Y_j X_j^b + Y_j X_j^a Y_j Y_j X_j^b + Y_j Y_j X_j^a Y_j X_j^b + Y_j Y_j Y_j X_j^a Y_j X_j^b)_{1 \leq j \leq n - k \atop j' > j}
\]
is a solution of (12). This yields many (presumably) independent vectors which are solution of (11) or (12). In other words, large dimension of the vector space solution of (11) or (12) is explained by the fact that there are many different ways of combining the equations of the algebraic system (10) together yielding the same linearized systems (11) or (12).

Observe that there are some relations among solutions, such as \(\mathbf{Z}_{a,b,c,d} + \mathbf{Z}_{c,d,e,f} = \mathbf{Z}_{a,b,c,f}\). However, if we define
\[
S_t \stackrel{\text{def}}{=} \{ (a, b) | a + b = t \},
\]
then we expect to obtain \(\sum_{t} (|S_t| - 1)\) linearly independent solutions to (11) or (12) from this process. The term \(|S_t| - 1\) in the sum is a simple consequence of the following fact.

**Fact 4.** Assume that we have \(\ell\) independent (over \(\mathbb{F}_2\)) vectors \(e_1, \ldots, e_\ell\). Then the set \(\{ e_i + e_j : i, j \in \{1, \ldots, \ell\} \}\) generates a vector space of dimension \(\ell - 1\) over \(\mathbb{F}_2\).

Finally, the solutions have coefficients over \(\mathbb{F}_q^m\). By decomposing each coefficient over \(\mathbb{F}_q\) we may finally have \(m \sum_t (|S_t| - 1)\) (potentially) independent vectors over \(\mathbb{F}_q\). This accounts for a generating set of size:
\[
m(r - 1)(r - 2) \over 2
\]
which agrees with Formula (13) when \(r \leq q\).

For larger values of \(r\), the automorphisms of \(\mathbb{F}_q^m\) leaving \(\mathbb{F}_q\) invariant have to be used. They are of the form \(x \mapsto x^\ell\) for some \(\ell \in \{0, \ldots, m - 1\}\). Notice that if we raise the equation \(Y_i X_i = \sum p_{ij} Y_j X_j\) to the \(q\)-th power we get:
\[
Y_i^q X_i^q = \sum p_{ij} Y_j^q X_j^q.
\]
We can use the same trick for \(Y_i = \sum p_{ij} Y_j\). From the trivial identity \(Y_i (Y_i X_i)^q = Y_i^q Y_i X_i^q\), we obtain a new algebraic equation which is
\[
\sum_{j=k+1}^n \sum_{j' > j} p_{ij} p_{ij'} (Y_j Y_j^q X_j^q + Y_j Y_j^q X_j^q + Y_j Y_j^q Y_j^q Y_j^q + Y_j Y_j^q Y_j^q X_j^q) = 0. \tag{21}
\]
To use \(Y_i X_i = \sum p_{ij} Y_j X_j^q\), we need to have \(r \geq q + 1\). However it should be noticed that if \(a + b = c + d\) then \(\mathbf{Z}_{a,b,c,d}\) and \(\mathbf{Z}_{qa,qb,qc,qd}\) only give \(m\) (potentially) independent vectors over \(\mathbb{F}_q\) (and not \(2m\)) after decomposing their coefficients over \(\mathbb{F}_q\). This comes from the fact that the Frobenius map \(x \mapsto x^q\) is a \(\mathbb{F}_q\)-linear transform. Therefore, the only new vectors obtained in this way are of the form \(\mathbf{Z}_{a,qb,c,qd}\) with \(0 \leq a, b, c, d < r\), \(0 \leq j < m\) and \(a + q^j b = c + q^j d\). This whole discussion leads to

**Heuristic 1** Let \(S_t^q \stackrel{\text{def}}{=} \{(a, b) | 0 \leq a < r, 0 \leq b < r, a + b = t\}\). For \(j\) in \(\{1, \ldots, m - 1\}\), we set \(S_t^j \stackrel{\text{def}}{=} \{(a, q^j b) | 0 \leq a < r, 0 \leq b < r, a + q^j b = t\}\). Then, for most choices of \(x\) and \(y\), we have:
\[
D_{\text{alternant}} = m \sum_{\{i, j : S_t^j \neq \emptyset\}} (|S_t^j| - 1).
\]

The sum appearing in the right-hand side has a very simple expression which is given by

\[\text{The notation } \{a, b\} \text{ refers to a multiset here. We may have } a = b.\]
Proposition 5.

\[
\sum_{\{t,j,i\} \neq 0} (|S_{t}^{i}| - 1) = \frac{r - 1}{2} \left\{ (2\ell + 1)r - 2q^{\ell+1} - 1 \right\}
\]

with \(\ell \overset{\text{def}}{=} \lceil \log_{q}(r - 1) \rceil\).

This finishes to explain the first part of Experimental Fact 1. In order to prove Proposition 5, we first prove the following lemma.

Lemma 1.

\[
|S_{t}^{0}| = \left\lceil \frac{t + 1}{2} \right\rceil, \text{ for } 0 \leq t \leq r - 1,
\]

\[
|S_{t}^{0}| = \left\lceil \frac{2r - t - 1}{2} \right\rceil, \text{ for } r \leq t \leq 2r - 2,
\]

\[
|S_{t}^{0}| = 0 \text{ otherwise.}
\]

\[
|S_{i}^{j}| \leq 1, \text{ if } q^{j} \geq r,
\]

\[
|S_{i}^{j}| = \min \left( r, \left\lceil \frac{t}{q^{j}} \right\rceil + 1 \right) - \max \left( \left\lceil \frac{t - r + 1}{q^{j}} \right\rceil, 0 \right) \text{ otherwise.}
\]

Proof. The first three equations follow directly from the definition of \(S_{t}^{0}\). Equation (26) is an easy consequence of the definition of \(S_{i}^{j}\). Let us assume now that \(r > q^{j}\). We now prove Equation (27). Let \((a, b)\) be a couple of integers such that:

\[
0 \leq a \leq r - 1
\]

\[
0 \leq b \leq r - 1
\]

\[
a + q^{j}b = t.
\]

From (28), (29) and (30), we obtain \(t - q^{j}b \leq r - 1\), which implies \(b \geq \left\lceil \frac{t - r + 1}{q^{j}} \right\rceil\). Together with (29)

\[
b \geq \max \left( \left\lceil \frac{t - r + 1}{q^{j}} \right\rceil, 0 \right).
\]

On the other hand, we also have \(b \leq r - 1\) and \(b \leq \left\lfloor \frac{t}{q^{j}} \right\rfloor\) since \(a \geq 0\). This implies

\[
b \leq \min \left( r - 1, \left\lfloor \frac{t}{q^{j}} \right\rfloor \right).
\]

All the \(b\)'s between these upper and lower bounds are possible. Then, there is only one corresponding \(a\) each time. This yields Equation (27). \(\square\)

From this, we deduce:

Lemma 2. It holds that:

\[
\sum_{t:S_{t}^{j} \neq 0} (|S_{t}^{j}| - 1) = \frac{(r - 1)(r - 2)}{2},
\]

\[
\sum_{t:S_{t}^{j} \neq 0} (|S_{t}^{j}| - 1) = (r - 1)(r - q^{j}) \text{ for } r \geq q^{j},
\]

\[
\sum_{t:S_{t}^{j} \neq 0} (|S_{t}^{j}| - 1) = 0 \text{ otherwise.}
\]
Proof. Let us first prove (33). By using Lemma 1, we obtain

$$\sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) = \sum_{t=0}^{r-1} \left( \left\lceil \frac{t+1}{2} \right\rceil - 1 \right) + \sum_{t=r}^{2r-2} \left( \left\lceil \frac{2r-t-1}{2} \right\rceil - 1 \right)$$

For $r$ odd (say $r = 2r' + 1$), we notice that $\sum_{t=0}^{r-1} \left( \left\lceil \frac{t+1}{2} \right\rceil - 1 \right) = r'(r' - 1) + r' = r'^2$ and that $\sum_{t=r}^{2r-2} \left( \left\lceil \frac{2r-t-1}{2} \right\rceil - 1 \right) = r'(r' - 1)$. This implies that $\sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) = r'(2r' - 1) = \frac{(r-1)(r-2)}{2}$. On the other hand, for $r$ even, say $r = 2r'$, we obtain $\sum_{t=0}^{r-1} \left( \left\lceil \frac{t+1}{2} \right\rceil - 1 \right) = r'(r' - 1)$ and $\sum_{t=r}^{2r-2} \left( \left\lceil \frac{2r-t-1}{2} \right\rceil - 1 \right) = r' - 1 + (r' - 1)(r' - 2) = (r' - 1)^2$. From this, we deduce that $\sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) = (r' - 1)(2r' - 1) = \frac{(r-1)(r-2)}{2}$. This proves (33).

To prove (34), we first notice that $|S^0_t|$ is positive if and only if $t$ belongs to $\{0, 1, \ldots, (q^j + 1)(r - 1)\}$. Then, we use Lemma 1 again and we obtain

$$\sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) = \sum_{t=0}^{(q^j+1)(r-1)} (|S^0_t^t| - 1)$$

$$= \sum_{t=0}^{(q^j+1)(r-1)} \min \left( r, \left\lfloor \frac{t}{q^j} \right\rfloor + 1 \right) - \max \left( \left\lfloor \frac{t-r+1}{q^j} \right\rfloor, 0 \right) - 1$$

$$= \sum_{t=0}^{(q^j+1)(r-1)} \min \left( r - 1, \left\lfloor \frac{t}{q^j} \right\rfloor \right) - \max \left( \left\lfloor \frac{t-r+1}{q^j} \right\rfloor, 0 \right)$$

$$= \sum_{t=0}^{(q^j+1)(r-1)} \min \left( r - 1, \left\lfloor \frac{t}{q^j} \right\rfloor \right) - \sum_{t=0}^{(q^j+1)(r-1)} \max \left( \left\lfloor \frac{t-r+1}{q^j} \right\rfloor, 0 \right). \quad (37)$$

Observe now that

$$\sum_{t=0}^{(q^j+1)(r-1)} \min \left( r - 1, \left\lfloor \frac{t}{q^j} \right\rfloor \right) = q^j \sum_{t=0}^{(q^j+1)(r-1)} \left\lfloor \frac{t}{q^j} \right\rfloor = \sum_{t=0}^{(q^j+1)(r-1)} \frac{t}{q^j} - \sum_{t=0}^{(q^j+1)(r-1)} (r - 1)$$

$$= q^j (0 + 1 + \cdots + r - 2) + (r - 1)r.$$

The other term appearing in the right-hand side of (37) is handled as follows

$$\sum_{t=0}^{(q^j+1)(r-1)} \max \left( \left\lfloor \frac{t-r+1}{q^j} \right\rfloor, 0 \right) = \sum_{t=r}^{(q^j+1)(r-1)} \max \left( \left\lfloor \frac{t-r+1}{q^j} \right\rfloor, 0 \right)$$

$$= q^j (1 + 2 + \cdots + r - 1).$$

By plugging these two expressions in (37) we obtain

$$\sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) = q^j (0 + 1 + \cdots + r - 2) + (r - 1)r - q^j (1 + 2 + \cdots + r - 1) = (r - 1)r - q^j (r - 1) = (r - 1)(r - q^j).$$

\[\square\]

Finally, we can now finish with the proof of Proposition 5.

**Proof.**

$$\sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) = \sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1) + \sum_{j=1}^{m-1} \sum_{t:S^0_t \neq \emptyset} (|S^0_t| - 1)$$

$$= \frac{(r - 1)(r - 2)}{2} + \sum_{j:q^j < r} (r - 1)(r - q^j)$$
Let \( \ell \) be the largest integer such that \( r > q^\ell \). We obtain

\[
\sum_{t,j:S_t^j \neq \emptyset} (|S_t^j| - 1) = \frac{r - 1}{2} \left\{ 2\ell r + (r - 2) - 2 \sum_{j=1}^\ell q^j \right\}
= \frac{r - 1}{2} \left\{ 2(\ell + 1)r - 2 \sum_{j=0}^\ell q^j \right\}
= \frac{r - 1}{2} \left\{ 2(\ell + 1)r - 2 \frac{q^{\ell+1} - 1}{q - 1} \right\}.
\]

This concludes the proof. \( \Box \)

### 7.2 The Goppa case

The simplest way to understand why there is a difference between the generic alternant case and the Goppa case is to compare \( \hat{D}_{\text{Goppa}} \) with \( \hat{D}_{\text{alternant}} \). First of all, the same reasoning as in the previous subsection can be done for the subcode \( \mathcal{A}_r(x,y) \) of even weights of an alternant code \( \mathcal{A}_r(x,y) \). This leads in the same way to the conclusion that in general:

\[
\hat{D}_{\text{alternant}} = \frac{m(r - 1)}{2} \left\{ (2\ell + 1)r - 2 \frac{q^{\ell+1} - 1}{q - 1} \right\},
\]

with \( \ell \equiv \lfloor \log_q(r - 1) \rfloor \). Notice, that from Proposition 2, we know that \( \mathcal{G}(x,G) \) is an alternant code of degree \( r + 1 \), when \( G \) is of degree \( r \). Therefore, we have

\[
\hat{D}_{\text{Goppa}} \geq \frac{mr}{2} \left\{ (2\ell + 1)(r + 1) - 2 \frac{q^{\ell+1} - 1}{q - 1} \right\}.
\]

It should be added that for \( r \leq q - 2 \), we actually have \( \hat{D}_{\text{Goppa}}(r) = \hat{D}_{\text{alternant}}(r + 1) \).

We do not have a general explanation for the formula observed for \( D_{\text{Goppa}} \) of non-binary Goppa codes. However, in the case of binary Goppa codes, we can use Theorem 3. In this case, when the Goppa polynomial \( G \) has only simple roots, we know that \( \mathcal{G}(x,G) \) is \( \mathcal{A}_{2r}(x,y') \), where \( r \equiv \deg(G) \) and \( y'_i = G(x_i)^{-1} \) where the \( x_i \)'s are the coordinates of \( x \) and the \( y'_i \)'s are the coordinates of \( y' \). This basically explains why the vector space solution of (11) is much greater for a binary Goppa code than for a binary alternant code of the same degree. This would suggest that \( D_{\text{Goppa}}(r) \geq \hat{D}_{\text{alternant}}(2r) \). However, this is not true. Now, there are linear relations among the vectors \( Z_{a,b,c,d} \) which are solutions of (11). Providing a cleaner explanation of the formula obtained for binary Goppa codes is much more involved and is beyond the scope of this article.

### Acknowledgement

The work described in this paper has been supported in part by the European Commission through the ICT programme under contract ICT-2007-216676 ECRYPT II. The first and third authors were also supported in part by the french ANR under the Computer Algebra and Cryptography (CAC) project ANR-09-JCJCJ-0064-01. The second and fourth authors were also supported by the french ANR Defis program under contract ANR-08-EMER-003 (COCQ project).
References


A Proof of Proposition 1

Proof. Let \( c = (c_i)_{1 \leq i \leq n} \) be a codeword in \( \mathcal{A}_r(x,y) \). Consider a polynomial \( P(X) = \sum_{j=0}^{r-1} a_j X^j \in \mathbb{F}_{q^m}[X] \) of degree at most \( r - 1 \) and notice that

\[
y'_i P \left( \frac{ax_i + b}{cx_i + d} \right) = y_i (cx_i + d)^{r-1} \sum_{0 \leq j \leq r-1} a_j \left( \frac{ax_i + b}{cx_i + d} \right)^j = y_i \sum_{0 \leq j \leq r-1} a_j (ax_i + b)^j (cx_i + d)^{r-1-j} = y_i Q(x_i)
\]

where \( Q \) is a polynomial of degree at most \( r - 1 \) which depends on \( a, b, c, d \) but not on \( i \). By the very definition of \( \mathcal{A}_r(x,y) \), we know that

\[
0 = \sum_{i=1}^{n} c_i y'_i Q(x_i) = \sum_{i=1}^{n} c_i y'_i P \left( \frac{ax_i + b}{cx_i + d} \right).
\]

In other words, we have just proved that \( c \in \mathcal{A}_r \left( \frac{ax + b}{cx + d}, y' \right) \). This proves that

\[
\mathcal{A}_r(x,y) \subset \mathcal{A}_r \left( \frac{ax + b}{cx + d}, y' \right).
\]

The inclusion in the other direction is proved similarly.
B Experimental Results

Table 4. \( q = 2 \) and \( m = 15 \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>990</td>
<td>1770</td>
<td>2775</td>
<td>4005</td>
<td>5460</td>
<td>7140</td>
<td>9045</td>
<td>11175</td>
<td>13530</td>
<td>16110</td>
<td>18915</td>
<td>21945</td>
<td>25200</td>
<td>28680</td>
</tr>
<tr>
<td>( k )</td>
<td>32723</td>
<td>32708</td>
<td>32693</td>
<td>32678</td>
<td>32663</td>
<td>32648</td>
<td>32633</td>
<td>32618</td>
<td>32603</td>
<td>32588</td>
<td>32573</td>
<td>32558</td>
<td>32543</td>
<td>32528</td>
</tr>
<tr>
<td>( D_{\text{random}} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \tilde{D}_{\text{Goppa}} )</td>
<td>270</td>
<td>570</td>
<td>1050</td>
<td>1665</td>
<td>2415</td>
<td>3300</td>
<td>4455</td>
<td>5775</td>
<td>7260</td>
<td>8910</td>
<td>10725</td>
<td>12705</td>
<td>14850</td>
<td>17160</td>
</tr>
<tr>
<td>( T_{\text{Goppa}} )</td>
<td>270</td>
<td>570</td>
<td>1050</td>
<td>1665</td>
<td>2415</td>
<td>3300</td>
<td>4455</td>
<td>5775</td>
<td>7260</td>
<td>8910</td>
<td>10725</td>
<td>12705</td>
<td>14850</td>
<td>17160</td>
</tr>
</tbody>
</table>

| \( N \) | 1035 | 1830 | 2850 | 4095 | 5565 | 7260 | 9180 | 11325 | 13695 | 16290 | 19110 | 22155 | 25420 | 28920 |
| \( k \) | 32722 | 32707 | 32692 | 32677 | 32662 | 32647 | 32632 | 32617 | 32602 | 32587 | 32572 | 32557 | 32542 | 32527 |
| \( D_{\text{random}} \) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| \( \tilde{D}_{\text{Goppa}} \) | 315 | 630 | 1125 | 1755 | 2520 | 3420 | 4590 | 5925 | 7425 | 9090 | 10920 | 12915 | 15075 | 17400 |

Table 5. \( q = 2 \) and \( m = 15 \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>32385</td>
<td>36315</td>
<td>40470</td>
<td>44850</td>
<td>49455</td>
<td>54285</td>
<td>59340</td>
<td>64620</td>
<td>70125</td>
<td>75855</td>
<td>81810</td>
<td>87990</td>
<td>94395</td>
<td>101025</td>
</tr>
<tr>
<td>( k )</td>
<td>32513</td>
<td>32498</td>
<td>32483</td>
<td>32468</td>
<td>32453</td>
<td>32438</td>
<td>32423</td>
<td>32408</td>
<td>32393</td>
<td>32378</td>
<td>32363</td>
<td>32348</td>
<td>32333</td>
<td>32318</td>
</tr>
<tr>
<td>( D_{\text{random}} )</td>
<td>0</td>
<td>3817</td>
<td>7987</td>
<td>12382</td>
<td>17002</td>
<td>21847</td>
<td>26917</td>
<td>32212</td>
<td>37732</td>
<td>43477</td>
<td>49447</td>
<td>55642</td>
<td>62062</td>
<td>68707</td>
</tr>
<tr>
<td>( \tilde{D}_{\text{Goppa}} )</td>
<td>19890</td>
<td>22815</td>
<td>25935</td>
<td>29250</td>
<td>32760</td>
<td>36465</td>
<td>40365</td>
<td>44460</td>
<td>48750</td>
<td>53235</td>
<td>57915</td>
<td>62790</td>
<td>67860</td>
<td>73125</td>
</tr>
<tr>
<td>( T_{\text{Goppa}} )</td>
<td>19890</td>
<td>22815</td>
<td>25935</td>
<td>29250</td>
<td>32760</td>
<td>36465</td>
<td>40365</td>
<td>44460</td>
<td>48750</td>
<td>53235</td>
<td>57915</td>
<td>62790</td>
<td>67860</td>
<td>73125</td>
</tr>
<tr>
<td>( \tilde{N} )</td>
<td>32640</td>
<td>36585</td>
<td>40755</td>
<td>45150</td>
<td>49770</td>
<td>54615</td>
<td>59685</td>
<td>64980</td>
<td>70500</td>
<td>76245</td>
<td>82215</td>
<td>88410</td>
<td>94830</td>
<td>101475</td>
</tr>
<tr>
<td>( \tilde{k} )</td>
<td>32512</td>
<td>32497</td>
<td>32482</td>
<td>32467</td>
<td>32452</td>
<td>32437</td>
<td>32422</td>
<td>32407</td>
<td>32392</td>
<td>32377</td>
<td>32362</td>
<td>32347</td>
<td>32332</td>
<td>32317</td>
</tr>
<tr>
<td>( \tilde{D}_{\text{random}} )</td>
<td>128</td>
<td>4088</td>
<td>8273</td>
<td>12683</td>
<td>17318</td>
<td>22178</td>
<td>27263</td>
<td>32573</td>
<td>38108</td>
<td>43868</td>
<td>49853</td>
<td>56063</td>
<td>62498</td>
<td>69158</td>
</tr>
<tr>
<td>( \tilde{D}_{\text{Goppa}} )</td>
<td>20145</td>
<td>23085</td>
<td>26220</td>
<td>29550</td>
<td>33075</td>
<td>36795</td>
<td>40710</td>
<td>44820</td>
<td>49125</td>
<td>53625</td>
<td>58320</td>
<td>63210</td>
<td>68295</td>
<td>73575</td>
</tr>
</tbody>
</table>
Table 6. $q = 2$ and $m = 15$

<table>
<thead>
<tr>
<th>$r$</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>107780</td>
<td>114960</td>
<td>122266</td>
<td>129793</td>
<td>137550</td>
<td>145530</td>
<td>153735</td>
<td>162163</td>
<td>170820</td>
<td>179700</td>
<td>188865</td>
<td>198135</td>
<td>207690</td>
<td>217470</td>
</tr>
<tr>
<td>$k$</td>
<td>32303</td>
<td>32288</td>
<td>32273</td>
<td>32258</td>
<td>32243</td>
<td>32228</td>
<td>32213</td>
<td>32198</td>
<td>32183</td>
<td>32168</td>
<td>32153</td>
<td>32138</td>
<td>32123</td>
<td>32108</td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>75577</td>
<td>82672</td>
<td>89992</td>
<td>97337</td>
<td>105307</td>
<td>113302</td>
<td>121522</td>
<td>129967</td>
<td>138637</td>
<td>147532</td>
<td>156652</td>
<td>165997</td>
<td>175567</td>
<td>185362</td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>78585</td>
<td>84240</td>
<td>90585</td>
<td>97537</td>
<td>105307</td>
<td>113302</td>
<td>121522</td>
<td>129967</td>
<td>138637</td>
<td>147532</td>
<td>156652</td>
<td>165997</td>
<td>175567</td>
<td>185362</td>
</tr>
<tr>
<td>$T_{\text{Goppa}}$</td>
<td>78585</td>
<td>84240</td>
<td>90585</td>
<td>97537</td>
<td>105307</td>
<td>113302</td>
<td>121522</td>
<td>129967</td>
<td>138637</td>
<td>147532</td>
<td>156652</td>
<td>165997</td>
<td>175567</td>
<td>185362</td>
</tr>
</tbody>
</table>

Table 7. $q = 2$ and $m = 16$

<table>
<thead>
<tr>
<th>$r$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>1128</td>
<td>2016</td>
<td>3160</td>
<td>4560</td>
<td>6216</td>
<td>8128</td>
<td>10296</td>
<td>12720</td>
<td>15400</td>
<td>18336</td>
<td>21528</td>
<td>24976</td>
<td>28680</td>
<td>32640</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k$</td>
<td>65488</td>
<td>65472</td>
<td>65456</td>
<td>65440</td>
<td>65424</td>
<td>65408</td>
<td>65392</td>
<td>65376</td>
<td>65360</td>
<td>65344</td>
<td>65328</td>
<td>65312</td>
<td>65296</td>
<td>65280</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>288</td>
<td>608</td>
<td>1120</td>
<td>1776</td>
<td>2576</td>
<td>3520</td>
<td>4752</td>
<td>6160</td>
<td>7744</td>
<td>9504</td>
<td>11440</td>
<td>13552</td>
<td>15840</td>
<td>18304</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_{\text{Goppa}}$</td>
<td>288</td>
<td>608</td>
<td>1120</td>
<td>1776</td>
<td>2576</td>
<td>3520</td>
<td>4752</td>
<td>6160</td>
<td>7744</td>
<td>9504</td>
<td>11440</td>
<td>13552</td>
<td>15840</td>
<td>18304</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$r$</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>39856</td>
<td>41328</td>
<td>46056</td>
<td>51040</td>
<td>56280</td>
<td>61776</td>
<td>67528</td>
<td>73536</td>
<td>79800</td>
<td>86320</td>
<td>93096</td>
<td>101280</td>
<td>107416</td>
<td>114960</td>
</tr>
<tr>
<td>$k$</td>
<td>65264</td>
<td>65248</td>
<td>65232</td>
<td>65216</td>
<td>65200</td>
<td>65184</td>
<td>65168</td>
<td>65152</td>
<td>65136</td>
<td>65120</td>
<td>65104</td>
<td>65088</td>
<td>65072</td>
<td>65056</td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>21216</td>
<td>24336</td>
<td>27664</td>
<td>31200</td>
<td>34494</td>
<td>38896</td>
<td>43056</td>
<td>47424</td>
<td>52000</td>
<td>56784</td>
<td>61776</td>
<td>66976</td>
<td>72384</td>
<td>78000</td>
</tr>
<tr>
<td>$T_{\text{Goppa}}$</td>
<td>21216</td>
<td>24336</td>
<td>27664</td>
<td>31200</td>
<td>34494</td>
<td>38896</td>
<td>43056</td>
<td>47424</td>
<td>52000</td>
<td>56784</td>
<td>61776</td>
<td>66976</td>
<td>72384</td>
<td>78000</td>
</tr>
</tbody>
</table>

Table 8. $q = 2$ and $m = 16$
Table 9. $q = 2$ and $m = 16$

<table>
<thead>
<tr>
<th>$r$</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
<th>43</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>122760</td>
<td>130816</td>
<td>139128</td>
<td>147696</td>
<td>156520</td>
<td>165600</td>
<td>174936</td>
<td>184528</td>
<td>194376</td>
<td>204480</td>
<td>214840</td>
<td>225456</td>
<td>236328</td>
<td></td>
</tr>
<tr>
<td>$k$</td>
<td>65040</td>
<td>65024</td>
<td>65008</td>
<td>64992</td>
<td>64976</td>
<td>64960</td>
<td>64944</td>
<td>64928</td>
<td>64912</td>
<td>64896</td>
<td>64880</td>
<td>64864</td>
<td>64848</td>
<td></td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>57720</td>
<td>65792</td>
<td>74120</td>
<td>82704</td>
<td>91544</td>
<td>100640</td>
<td>109992</td>
<td>119600</td>
<td>129464</td>
<td>139684</td>
<td>150408</td>
<td>161744</td>
<td>173680</td>
<td></td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>83284</td>
<td>89856</td>
<td>96624</td>
<td>103362</td>
<td>110880</td>
<td>118368</td>
<td>126096</td>
<td>134064</td>
<td>142272</td>
<td>150720</td>
<td>159408</td>
<td>168336</td>
<td>177504</td>
<td></td>
</tr>
<tr>
<td>$T_{\text{Goppa}}$</td>
<td>83284</td>
<td>89856</td>
<td>96624</td>
<td>103362</td>
<td>110880</td>
<td>118368</td>
<td>126096</td>
<td>134064</td>
<td>142272</td>
<td>150720</td>
<td>159408</td>
<td>168336</td>
<td>177504</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>123256</td>
<td>131328</td>
<td>139656</td>
<td>148240</td>
<td>157080</td>
<td>166176</td>
<td>175528</td>
<td>185136</td>
<td>195000</td>
<td>205120</td>
<td>215496</td>
<td>226128</td>
<td>237016</td>
<td></td>
</tr>
<tr>
<td>$k$</td>
<td>65039</td>
<td>65023</td>
<td>65007</td>
<td>64991</td>
<td>64975</td>
<td>64959</td>
<td>64943</td>
<td>64927</td>
<td>64911</td>
<td>64895</td>
<td>64879</td>
<td>64863</td>
<td>64847</td>
<td></td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>58217</td>
<td>66305</td>
<td>74649</td>
<td>83249</td>
<td>92105</td>
<td>101217</td>
<td>110585</td>
<td>120209</td>
<td>130089</td>
<td>140225</td>
<td>150617</td>
<td>161265</td>
<td>172169</td>
<td></td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>84320</td>
<td>90368</td>
<td>97152</td>
<td>104176</td>
<td>114400</td>
<td>124844</td>
<td>136688</td>
<td>149672</td>
<td>162808</td>
<td>176160</td>
<td>190752</td>
<td>205584</td>
<td>221712</td>
<td></td>
</tr>
</tbody>
</table>

Table 10. $q = 4$ and $m = 6$

<table>
<thead>
<tr>
<th>$r$</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>153</td>
<td>276</td>
<td>435</td>
<td>630</td>
<td>861</td>
<td>1128</td>
<td>1431</td>
<td>1770</td>
<td>2145</td>
<td>2556</td>
<td>3003</td>
<td>3486</td>
<td>4005</td>
<td>4560</td>
</tr>
<tr>
<td>$k$</td>
<td>4078</td>
<td>4072</td>
<td>4066</td>
<td>4060</td>
<td>4054</td>
<td>4048</td>
<td>4042</td>
<td>4036</td>
<td>4030</td>
<td>4024</td>
<td>4018</td>
<td>4012</td>
<td>4006</td>
<td>4000</td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>360</td>
</tr>
<tr>
<td>$D_{\text{alternant}}$</td>
<td>6</td>
<td>18</td>
<td>60</td>
<td>120</td>
<td>198</td>
<td>294</td>
<td>408</td>
<td>540</td>
<td>690</td>
<td>858</td>
<td>1044</td>
<td>1248</td>
<td>1470</td>
<td>1710</td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>18</td>
<td>60</td>
<td>120</td>
<td>198</td>
<td>294</td>
<td>408</td>
<td>540</td>
<td>690</td>
<td>858</td>
<td>1044</td>
<td>1248</td>
<td>1470</td>
<td>1710</td>
<td>2058</td>
</tr>
<tr>
<td>$T_{\text{Goppa}}$</td>
<td>18</td>
<td>60</td>
<td>120</td>
<td>198</td>
<td>294</td>
<td>408</td>
<td>540</td>
<td>690</td>
<td>858</td>
<td>1044</td>
<td>1248</td>
<td>1470</td>
<td>1710</td>
<td>2058</td>
</tr>
</tbody>
</table>

Table 11. $q = 4$ and $m = 6$

<table>
<thead>
<tr>
<th>$r$</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>5151</td>
<td>5778</td>
<td>6441</td>
<td>7140</td>
<td>7875</td>
<td>8646</td>
<td>9453</td>
<td>10296</td>
<td>11175</td>
<td>12090</td>
<td>13041</td>
<td>14028</td>
<td>15051</td>
<td>16110</td>
</tr>
<tr>
<td>$k$</td>
<td>3994</td>
<td>3988</td>
<td>3982</td>
<td>3976</td>
<td>3970</td>
<td>3964</td>
<td>3958</td>
<td>3952</td>
<td>3946</td>
<td>3940</td>
<td>3934</td>
<td>3928</td>
<td>3922</td>
<td>3916</td>
</tr>
<tr>
<td>$D_{\text{random}}$</td>
<td>1157</td>
<td>1790</td>
<td>2459</td>
<td>3164</td>
<td>3905</td>
<td>4682</td>
<td>5495</td>
<td>6344</td>
<td>7229</td>
<td>8150</td>
<td>9107</td>
<td>10100</td>
<td>11129</td>
<td>12194</td>
</tr>
<tr>
<td>$D_{\text{alternant}}$</td>
<td>2064</td>
<td>2448</td>
<td>2862</td>
<td>3306</td>
<td>3789</td>
<td>4284</td>
<td>4818</td>
<td>5382</td>
<td>5976</td>
<td>6600</td>
<td>7254</td>
<td>7938</td>
<td>8652</td>
<td>9396</td>
</tr>
<tr>
<td>$D_{\text{Goppa}}$</td>
<td>3060</td>
<td>3510</td>
<td>3990</td>
<td>4500</td>
<td>5040</td>
<td>5610</td>
<td>6210</td>
<td>6840</td>
<td>7500</td>
<td>8190</td>
<td>8910</td>
<td>9660</td>
<td>10440</td>
<td>11250</td>
</tr>
<tr>
<td>$T_{\text{Goppa}}$</td>
<td>3060</td>
<td>3510</td>
<td>3990</td>
<td>4500</td>
<td>5040</td>
<td>5610</td>
<td>6210</td>
<td>6840</td>
<td>7500</td>
<td>8190</td>
<td>8910</td>
<td>9660</td>
<td>10440</td>
<td>11250</td>
</tr>
</tbody>
</table>

Jean-Charles Faugère, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich