# A Distinguisher for High Rate McEliece Cryptosystems

Jean-Charles Faugère[1], Valérie Gauthier[4], Ayoub Otmani[2,3], Ludovic Perret[1], and Jean-Pierre Tillich[2]

[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
[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
[3] GREYC - Université de Caen - Ensicaen
Boulevard Maréchal Juin, 14050 Caen Cedex, France.
[4] Department of Mathematics
Technical University of Denmark
Matematiktorvet, Building 303 S, 2800 Kgs. Lyngby, Denmark
v.g.umana@mat.dtu.dk

**Abstract.** The Goppa Code Distinguishing (GD) problem consists in distinguishing the matrix of Goppa code from a random matrix. Up to now, it was widely believed that this problem is computationally hard. The hardness of this problem was a mandatory assumption to prove the security of code-based cryptographic primitives like McEliece's cryptosystem. We present a polynomial time distinguisher for alternant and Goppa codes of high rate over any field. The key ingredient is an algebraic technique already used to asses the security McEliece's cryptosystem. Our method permits to indeed distinguish public keys of the CFS signature scheme for all parameters considered as secure and some realistic secure parameters of McEliece. The idea is to consider the dimension of the solution space of a linearized system deduced from a polynomial one. It turns out that this dimension depends on the type of code considered. We provide explicit formulas for the value of the dimension for "generic" random, alternant, and Goppa code over any alphabet.

**Keywords:** public-key cryptography, McEliece cryptosystem, CFS signature, algebraic cryptanalysis, distinguisher, Goppa code distinguishing.

# 1 Introduction

The purpose of this paper is to investigate the difficulty of the Goppa Code Distinguishing (GD) problem that is encountered in code-based public key cryptography. Note that GD is a variant of the Code Equivalence [24]. This problem appeared in [10] several years after McEliece's pioneering work [20] 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 a polynomial-time decoding algorithm [3, 23]. The first code-based signature scheme came out in [10] almost twenty years after McEliece's proposal. The only difference between encryption and signature lies in the choice of the parameters of the binary Goppa codes. For signature, Goppa codes have to correct very few errors. This leads to a very high rate $R = k/n$ with $n$ the length and $k$ the dimension of the code. It holds that $k = n - rm$, where $r$ is the number of errors and $n$ is usually chosen equal to $2^m$.

These two cryptographic primitives base their security under two computational assumptions: the intractability of decoding random linear codes [2], 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 [16, 17, 26, 7, 4] 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 [18], 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) [25] to check whether the Goppa code candidate is *permutation-equivalent* to the code defined by the public generator matrix. The time complexity of this method is $\mathcal{O}\left(2^{mr}\right)$ assuming that the cost of the SSA algorithm is negligible which is a reasonable assumption for Goppa codes, but not for all linear codes.

The authors of [10] alleviated the McEliece assumption 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 [10], 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 [25], as explained below.

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 currently a mandatory assumption to prove the semantic and CCA2 security of McEliece in the random oracle model and in the standard model [22, 13, 5], the security in the random oracle model against existential forgery [10, 11] of the CFS signature scheme [10], the provable security of several primitives such as a threshold ring signatures scheme [12], an identity-based identification scheme [8], 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 theoretical weakness observed on this scheme since thirty years.

In this paper, we present a deterministic polynomial-time distinguisher for solving it for codes of high rate. Along the way, we also solve the GD problem for alternant codes. The key ingredient is a new algebraic technique introduced in [14] to attack two variants [1, 21] of McEliece. It has been observed [14] that a key recovery attack against these cryptosystems, as well as the genuine McEliece's system, can be reduced to solving a set of polynomial equations. In the cases of [1, 21], additional structures permit to drastically reduce the number of variables and solve it efficiently using dedicated Gröbner bases techniques [14]. For McEliece's cryptosystem, solving this polynomial system seems to be out of the scope of such dedicated techniques.

This algebraic approach can be used to construct an efficient *distinguisher*. To do so, we consider the dimension of the solution space of a linear system deduced from the polynomial system by a linearization technique which introduces many new unknowns. It turns out that the linearized system is not of full rank and depends on the kind of code considered. This particular feature permits to construct an efficient distinguisher for alternant codes and Goppa codes over any field by basically computing the rank of the linearized system. Our technique permits to indeed distinguish a public key of the CFS signature scheme for all parameters proposed in [15], and some realistic parameters of McEliece like a 90-bit security scheme based on a binary Goppa code of length $n = 2^{13}$ that corrects $r = 19$ errors. We provide explicit formulas for "generic" random, alternant, and Goppa code over any alphabet. We performed extensive experiments to compare our theoretical results to confirm that the generic formulas are accurate.

## 2 Algebraic Cryptanalysis of McEliece-like Cryptosystems

The reader who is not aware of basic notions on coding and code-based cryptography can find a brief introduction to the subject in Appendix A. 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 $\mathbf{H}$ with coefficients in $\mathbb{F}_{q^m}$ and the associated alternant code $\mathscr{A}$ is the set of vectors of $\mathbb{F}_q^n$ which belong to the right kernel of $\mathbf{H}$ *i.e.*

$$\mathscr{A} = \{\mathbf{c} \in \mathbb{F}_q^n \mid \mathbf{H}\mathbf{c}^T = \mathbf{0}\}. \tag{1}$$

$r$ satisfies in this case the condition $r \geqslant \frac{n-k}{m}$ where $k$ is the dimension of $\mathscr{A}$. For alternant codes, there exists a parity-check matrix with a very special form related to Vandermonde matrices of the form:

$$\boldsymbol{V}_r(\mathbf{x}, \mathbf{y}) \overset{\text{def}}{=} \begin{pmatrix} y_1 & \cdots & y_n \\ y_1 x_1 & \cdots & y_n x_n \\ \vdots & & \vdots \\ y_1 x_1^{r-1} & \cdots & y_n x_n^{r-1} \end{pmatrix} \tag{2}$$

where $\mathbf{x} = (x_1, \dots, x_n) \in (\mathbb{F}_{q^m})^n$ and $\mathbf{y} = (y_1, \dots, y_n)$ in $(\mathbb{F}_{q^m})^n$.

**Definition 1 (Alternant code).** *The alternant code of order $r$ over $\mathbb{F}_q$ associated to $\mathbf{x} = (x_1, \dots, x_n) \in (\mathbb{F}_{q^m})^n$ where all $x_i$'s are distinct and $\mathbf{y} = (y_1, \dots, y_n) \in \left(\mathbb{F}_{q^m}^*\right)^n$, denoted by $\mathscr{A}_r(\mathbf{x}, \mathbf{y})$, is*

$$\mathscr{A}_r(\mathbf{x}, \mathbf{y}) = \{\mathbf{c} \in \mathbb{F}_q^n | \boldsymbol{V}_r(\mathbf{x}, \mathbf{y})\mathbf{c}^T = \mathbf{0}\}. \tag{3}$$

We recall a key feature about 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 of the form $\mathbf{H} = \boldsymbol{V}_r(\mathbf{x}, \mathbf{y})$ is given for it.*

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

**Definition 2 (Goppa codes).** *The Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ over $\mathbb{F}_q$ associated to a polynomial $\Gamma(x)$ of degree $r$ over $\mathbb{F}_{q^m}$ and a certain $n$-tuple $\mathbf{x} = (x_1, \dots, x_n)$ of distinct elements of $\mathbb{F}_{q^m}$ satisfying $\Gamma(x_i) \neq 0$ for all $i, 1 \leqslant i \leqslant n$, is the alternant code $\mathscr{A}_r(\mathbf{x}, \mathbf{y})$ of order $r$ with $y_i$ being defined by $y_i = \Gamma(x_i)^{-1}$.*

Goppa codes, viewed as alternant codes, naturally inherit a decoding algorithm that corrects up to $\frac{r}{2}$ errors. But in the case of *binary* Goppa codes, we can correct twice as many errors (Fact 2). The starting point is the following result, which is essentially derived from a discussion in a paragraph about Goppa codes in [19, p. 341].

**Theorem 1.** *A binary Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ associated to a Goppa polynomial $\Gamma(X)$ of degree $r$ without multiple roots is equal to the alternant code $\mathscr{A}_{2r}(\mathbf{x}, \mathbf{y})$, with $y_i = \Gamma(x_i)^{-2}$.*

**Fact 2 ([23]).** *There exists a polynomial time algorithm decoding all errors of Hamming weight at most $r$ in a Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ when $\Gamma$ has degree $r$ and has no multiple roots, if $\mathbf{x}$ and $\Gamma$ are known.*

We explain now how we can construct an algebraic system for the McEliece cryptosystem [14]. This algebraic system is the main ingredient of the distinguisher. According to Fact 1, the knowledge of $\boldsymbol{V}_r(\boldsymbol{x}^*, \boldsymbol{y}^*)$ permits to efficiently decode the public code, *i.e.* to recover $\boldsymbol{u}$ from $\boldsymbol{u}\boldsymbol{G} + \boldsymbol{e}$. By the very definition of the public code $\boldsymbol{G}$, we have: $\boldsymbol{V}_r(\boldsymbol{x}^*, \boldsymbol{y}^*)\boldsymbol{G}^T = \boldsymbol{0}$.

Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ 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 [14] of the following system:

$$\left\{ g_{i,1} Y_1 X_1^j + \ldots + g_{i,n} Y_n X_n^j = 0 \;\middle|\; i \in \{1, \dots, k\}, j \in \{0, \dots, r-1\} \right\} \tag{4}$$

where the $g_{i,j}$'s are the entries of the known matrix $\boldsymbol{G}$.

Solving this system is equivalent to find a key equivalent to the secret key. For McEliece's scheme, the system is too large. For compact variants of McEliece [1,21] as described in [1,21], additional structures permit to drastically reduce the number of variables; allowing to solve (4) for a large set of parameters in polynomial-time using dedicated Gröbner bases techniques [14]. But the general case is still exponential. Note that 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. Otherwise, as pointed in Fact 2, the decoder will be able to decode only $\frac{r}{2}$ errors.

## 3 A Distinguisher of Alternant and Goppa Codes

We present in this part the algebraic distinguisher. which is based on the non-linear system (4). Let $\mathbf{G} = (g_{ij})_{\substack{1 \leqslant i \leqslant k \\ 1 \leqslant j \leqslant n}}$ be a generator matrix of the public code. We can assume without loss of generality that $\mathbf{G}$ is systematic in its $k$ first positions. Such a form can be easily obtained by a Gaussian elimination and by a suitable permutation of the columns. We describe now a simple way of using this particular form for solving (4). The strategy is as follows. Let $\mathbf{P} = (p_{ij})_{\substack{1 \leqslant i \leqslant k \\ k+1 \leqslant j \leqslant n}}$ be the submatrix of

$\mathbf{G}$ formed by its last $n-k=mr$ columns. For any $i \in \{1,\dots,k\}$ and $e \in \{0,\dots,r-1\}$, we can rewrite (4) as

$$Y_i X_i^e = \sum_{j=k+1}^{n} p_{i,j} Y_j X_j^e. \tag{5}$$

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

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

It is possible to reorder this to obtain $\sum_{j=k+1}^{n-1} \sum_{j'>j}^{n} p_{i,j} p_{i,j'} \left( Y_j Y_{j'} X_{j'}^2 + Y_{j'} Y_j X_j^2 \right) = 0$. 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 a system $\mathcal{L}_{\mathbf{P}}$ of $k$ linear equations involving the $Z_{jj'}$'s:

$$\mathcal{L}_{\mathbf{P}} \overset{\text{def}}{=} \left\{ \sum_{j=k+1}^{n-1} \sum_{j'>j}^{n} p_{i,j} p_{i,j'} Z_{jj'} = 0 \;\middle|\; i \in \{1,\dots,k\} \right\}. \tag{6}$$

To solve this system it is necessary that the number of equations is greater than the number of unknowns *i.e.* $k \geqslant \binom{mr}{2}$ with the hope that the rank of $\mathcal{L}_{\mathbf{P}}$ denoted by $\operatorname{rank}(\mathcal{L}_{\mathbf{P}})$ is almost equal to the number of variables. Observe that the linear systems (6) have coefficients in $\mathbb{F}_q$ whereas solutions are sought in the extension field $\mathbb{F}_{q^m}$. But the dimension $D$ of the vector space solution of $\mathcal{L}_{\mathbf{P}}$ does not depend on the underlying field because $\mathcal{L}_{\mathbf{P}}$ can always be seen as a system over $\mathbb{F}_{q^m}$. Remark that we obviously have $D = \binom{mr}{2} - \operatorname{rank}(\mathcal{L}_{\mathbf{P}})$. It appears that $D$ is amazingly large. It even depends on whether or not the code with generator matrix $\mathbf{G}$ is chosen as a (generic) alternant code or as a Goppa code. Interestingly enough, when $\mathbf{G}$ is chosen at random, $\operatorname{rank}(\mathcal{L}_{\mathbf{P}})$ is equal to $\min\left\{k, \binom{mr}{2}\right\}$ with very high probability. In particular, the dimension of the solution space is typically $0$ when $k$ is larger than the number of variables $\binom{mr}{2}$.

Although this *defect* in the rank is an obstacle to break the McEliece cryptosystem, it can be used to distinguish the public generator from a random code. Moreover, since the linear system $\mathcal{L}_{\mathbf{P}}$ is defined over $\mathbb{F}_q$, there exist two vector spaces solution depending on whether the underlying field is $\mathbb{F}_{q^m}$ or $\mathbb{F}_q$. This duality leads to the following definition.

**Definition 3.** *For any integer $r \geqslant 1$ and $m \geqslant 1$, let us denote by $N \overset{def}{=} \binom{mr}{2}$ the number of variables in the linear system $\mathcal{L}_{\mathbf{P}}$ as defined in (6) and $D$ the dimension of the vector space solution of $\mathcal{L}_{\mathbf{P}}$. The* normalized dimension *of $\mathcal{L}_{\mathbf{P}}$ denoted by $\varDelta$ is defined as:*

$$\varDelta \overset{def}{=} \frac{D}{m}.$$

Throughout the paper we consider three cases corresponding to the possible choices for the entries $p_{i,j}$'s. We denote by $\varDelta_{\text{random}}$ the normalized dimension when the $p_{ij}$'s are chosen uniformly and independently at random in $\mathbb{F}_q$. When $\mathbf{G}$ is chosen as a generator matrix of a random alternant (*resp.* Goppa) code of degree $r$, we denote the normalized dimension by $\varDelta_{\text{alternant}}$ (*resp.* $\varDelta_{\text{Goppa}}$). Note that in our probabilistic model, a random alternant code is obtained by picking uniformly and independently at random two vectors $(x_1,\dots,x_n)$ and $(y_1,\dots,y_n)$ from $(\mathbb{F}_{q^m})^n$ such that the $x_i$'s are all different and the $y_i$'s are all nonzero. A random Goppa code is obtained by also taking in the same way a random vector $(x_1,\dots,x_n)$ in $(\mathbb{F}_{q^m})^n$ with all the $x_i$'s different and a random *irreducible* polynomial $\Gamma(z)=\sum_i \gamma_i z^i$ of degree $r$.

A thorough experimental study (see Appendix E) through intensive computations with Magma [6] by randomly generating 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 $\{3,\dots,50\}$ and several $m$ revealed that the (normalized) dimension of the vector space over $\mathbb{F}_q$ of the solutions of (6) follows the following formulas. Recall that by definition $N=\binom{mr}{2}$ and $k=n-rm$ where $n\leqslant q^m$.

**Experimental Fact 1 (Alternant Case).** *As long as $N-m\Delta_{\mathit{alternant}}<k$, with very high probability the normalized dimension $\Delta_{\mathit{alternant}}$ has the following value $T_{\mathit{alternant}}$:*

$$T_{\mathit{alternant}}=\frac{1}{2}(r-1)\left((2e+1)r-2\frac{q^{e+1}-1}{q-1}\right)\ \textit{for}\ e\overset{\mathit{def}}{=}\left\lfloor\log_q(r-1)\right\rfloor. \qquad (7)$$

As for the case of random Goppa codes we also obtain formulas different from those of alternant codes. Note however that the Goppa codes are generated by means of a random irreducible $\Gamma(z)$ of degree $r$ and hence $\Gamma(z)$ has no multiple roots. In particular, we can apply Theorem 1 in the binary case.

**Experimental Fact 2 (Goppa Case).** *As long as $N-m\Delta_{\mathit{Goppa}}<k$, with very high probability the normalized dimension $\Delta_{\mathit{Goppa}}$ has the following value $T_{\mathit{Goppa}}$:*

$$T_{\mathit{Goppa}}=\begin{cases}\frac{1}{2}(r-1)(r-2)=T_{\mathit{alternant}} & \textit{for}\ \ r<q-1\\[2ex] \frac{1}{2}r\Big((2e+1)r-2q^e+2q^{e-1}-1\Big) & \textit{for}\ \ r\geqslant q-1\end{cases} \qquad (8)$$

*where $e$ is the unique integer such that: $q^e-2q^{e-1}+q^{e-2}<r\leqslant q^{e+1}-2q^e+q^{e-1}$.*

Based upon these experimental observations, we are now able to define a *distinguisher* between random codes, alternant codes and Goppa codes. This distinguisher will be in particular useful to distinguish between McEliece public keys and random matrices.

**Definition 4 (Random Code Distinguisher).** *Let $m$ and $r$ be integers such that $m\geqslant 1$ and $r\geqslant 1$. Let $\mathbf{G}$ be a $k\times n$ matrix whose entries are in $\mathbb{F}_q$ with $n\leqslant q^m$ and $k\overset{\mathit{def}}{=}n-rm$. Without loss of generality, we assume that $\mathbf{G}$ is systematic i.e. $\mathbf{G}=(\mathbf{I}_k\mid\mathbf{P})$. Let $\mathcal{L}_{\mathbf{P}}$ be the linear system associated to $\mathbf{G}$ as defined in* (6)*, and $\Delta$ the normalized dimension of $\mathcal{L}_{\mathbf{P}}$. We define the* Random Code Distinguisher *$\mathcal{D}$ as the mapping which takes in input $\mathbf{G}$ and outputs $b$ in $\{-1,0,1\}$ such that:*

$$\mathcal{D}(\mathbf{G})=\begin{cases}-1\ \ \textit{if}\ \Delta=T_{\mathit{alternant}}\\ \ \ \ 0\ \ \textit{if}\ \Delta=T_{\mathit{Goppa}}\\ \ \ \ 1\ \ \textit{otherwise.}\end{cases} \qquad (9)$$

## 4 The Random Case

The purpose of this section is to study the behavior of $D_{\text{random}}$, namely the dimension of the solution space of $\mathcal{L}_{\mathbf{P}}$ when the entries of the matrix $\mathbf{P}$ are drawn independently from the uniform distribution over $\mathbb{F}_q$. In this case, when can show that:

**Theorem 2.** *Assume that $N \leq k$ and that the entries of $\mathbf{P}$ are drawn independently from the uniform distribution over $\mathbb{F}_q$. Then for any function $\omega(x)$ tending to infinity as $x$ goes to infinity, we have*

$$\mathbf{prob}\left(D_{\textit{random}} \geq mr\omega(mr)\right) = o(1),$$

*as $mr$ goes to infinity.*

This theorem will be proved in Appendix B. Notice that if choose $\omega(x) = \log(x)$ for instance, then asymptotically the dimension $D_{\text{random}}$ of the solution space is with very large probability smaller than $mr\log(mr)$. When $m$ and $r$ are of the same order – which is generally chosen in practice – this quantity is smaller than $D_{\text{alternant}}$ or $D_{\text{Goppa}}$ which are of the form $\Omega(mr^2)$.

The main ingredient for proving Theorem 2 consists in analyzing a certain (partial) Gaussian elimination process on the matrix $\boldsymbol{M} \overset{\text{def}}{=} (p_{ij}p_{ij'})_{\substack{1\leq i\leq k\\ j,j':k+1\leq j<j'\leq n}}$. Basically it amounts to view the matrix $\boldsymbol{M}$ in block form, each block consisting in the matrix $\boldsymbol{B}_j = (p_{ij}p_{ij'})_{\substack{1\leq i\leq k\\ j<j'\leq n}}$. Each block $\boldsymbol{B}_j$ is of size $k \times (rm - j)$. Notice that in $\boldsymbol{B}_j$, the rows for which $p_{i,k+j} = 0$ consist only of zeros.

To start the Gaussian elimination process with $\boldsymbol{B}_1$, we will therefore pick up $rm - 1$ rows for which $p_{i,k+1} \neq 0$. This gives a square matrix $\boldsymbol{M}_1$. We perform Gaussian elimination on $\boldsymbol{M}$ by adding rows involved in $\boldsymbol{M}_1$ to put the first block $\boldsymbol{B}_1$ in standard form. We carry on this process with $\boldsymbol{B}_2$ by picking now $rm - 2$ rows which have not been chosen before and which correspond to $p_{i,k+2} \neq 0$. This yields a square submatrix $\boldsymbol{M}_2$ of size $rm - 2$ and we continue this process until reaching the last block. The key observation is that:

$$\operatorname{rank}(\boldsymbol{M}) \geq \operatorname{rank}(\boldsymbol{M}_1) + \operatorname{rank}(\boldsymbol{M}_2) + \cdots + \operatorname{rank}(\boldsymbol{M}_{rm-1}).$$

A rough analysis of this process yields the theorem above. The important point is what happens for different blocks are independent processes, it corresponds to looking at different rows of the matrix $\mathbf{P}$. A more detailed analysis would probably yield a stronger result that $\mathbf{prob}(D_{\text{random}} \geq \omega(mr)) = o(1)$, for any function $\omega$ going to infinity with $mr$ or allowing to treat the case $N \geq k$ where we would like to show that $\mathbf{prob}(D_{\text{random}} \geq N - k + \omega(mr)) = o(1)$. But, this is beyond the scope of this paper.

## 5 Interpretation of the Normalized Dimension – The Alternant Case

We first consider the case of alternant codes over $\mathbb{F}_q$ of degree $r$. The goal of this section is to identify a set of vectors which, after decomposition according to a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, provides a basis of the solution space of $\mathcal{L}_{\mathbf{P}}$. First observe that to set up the linear system $\mathcal{L}_{\mathbf{P}}$ as defined in (6), we have used the trivial identity $Y_iY_iX_i^2 = (Y_iX_i)^2$. Actually, we can use any identity $Y_iX_i^aY_iX_i^b = Y_iX_i^cY_iX_i^d$ 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 the same algebraic system $\mathcal{L}_{\mathbf{P}}$ with:

$$\sum_{j=k+1}^{n}\sum_{j'>j} p_{i,j}p_{i,j'}\left(Y_jX_j^aY_{j'}X_{j'}^b + Y_{j'}X_{j'}^aY_jX_j^b + Y_jX_j^cY_{j'}X_{j'}^d + Y_{j'}X_{j'}^cY_jX_j^d\right) = 0. \tag{10}$$

So, the fact that *there are many different ways of combining the equations of the algebraic system together yielding the same linearized system $\mathcal{L}_{\mathbf{P}}$* explains why the dimension of the vector space solution $V_{q^m}$ is large.

For larger values of $r$, the automorphisms of $\mathbb{F}_{q^m}$ of the kind $x \mapsto x^{q^\ell}$ for some $\ell \in \{0, \ldots, m-1\}$ can be used to obtain the identity $Y_i^{q^{\ell'}}X_i^{aq^{\ell'}}Y_i^{q^\ell}X_i^{bq^\ell} = Y_i^{q^{\ell'}}X_i^{cq^{\ell'}}Y_i^{q^\ell}X_i^{dq^\ell}$ for any integers $a$, $b$, $c$,

$d$, $\ell$, $\ell'$ such that $aq^{\ell'} + bq^{\ell} = cq^{\ell'} + dq^{\ell}$. We get again the linear system $\mathcal{L}_{\mathbf{P}}$ but the decomposition over $\mathbb{F}_q$ of the entries of vectors obtained from such equations give vectors that are dependent of those coming from the identity $Y_i X_i^a Y_i^{q^{\ell-\ell'}} X_i^{bq^{\ell-\ell'}} = Y_i X_i^c Y_i^{q^{\ell-\ell'}} X_i^{dq^{\ell-\ell'}}$ if we assume $\ell' \leqslant \ell$. Therefore, we are only interested to vectors that satisfy equations obtained with $0 \leqslant a, b, c, d < r$, $0 \leqslant \ell < m$ and $a + q^{\ell} b = c + q^{\ell} d$.

**Definition 5.** *Let $a$, $b$, $c$ and $d$ be integers in $\{0, \dots, r-1\}$ and an integer $\ell$ in $\{0, \dots, \lfloor \log_q(r-1) \rfloor\}$ such that $a + q^{\ell} b = c + q^{\ell} d$. We define $\boldsymbol{Z}_{a,b,c,d,\ell} \overset{def}{=} \left(\boldsymbol{Z}_{a,b,c,d,\ell}[j,j']\right)_{k+1 \leqslant j < j' \leqslant n}$ where*

$$\boldsymbol{Z}_{a,b,c,d,\ell}[j,j'] \overset{def}{=} Y_j X_j^a Y_{j'}^{q^\ell} X_{j'}^{q^\ell b} + Y_{j'} X_{j'}^a Y_j^{q^\ell} X_j^{q^\ell b} + Y_j X_j^c Y_{j'}^{q^\ell} X_{j'}^{q^\ell d} + Y_{j'} X_{j'}^c Y_j^{q^\ell} X_j^{q^\ell d},$$

*for any $j$ and $j'$ satisfying $k + 1 \leqslant j < j' \leqslant n$.*

Without loss of generality, we can assume that $d > b$ and set $\delta = d - b$. Moreover, as we have $a + q^{\ell} b = c + q^{\ell} d$, it implies that $a = c + q^{\ell}\delta$. Note that any vector $\boldsymbol{Z}_{a,b,c,d,\ell}$ is uniquely described by the tuple $(b, c, \delta, \ell)$ by setting $d = b + \delta$ and $a = c + q^{\ell}\delta$ provided that $1 \leqslant \delta \leqslant r - 1 - b$ and $0 \leqslant c + q^{\ell}\delta \leqslant r - 1$.

The next proposition shows that some vectors $\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}$ can be expressed as a linear combination of vectors defined with $\delta = 1$. All the proofs of this section are gathered in Appendix C.

**Proposition 1.** *Let $\ell$, $\delta$, $b$ and $c$ be integers such that $\ell \geqslant 0$, $\delta \geqslant 1$, $1 \leqslant b + \delta \leqslant r - 1$ and $1 \leqslant c + q^{\ell}\delta \leqslant r - 1$. Let us assume that $\delta \geqslant 2$ and let $b_i \overset{def}{=} b + i - 1$ and $c_i \overset{def}{=} c + q^{\ell}(\delta - i)$. We have*

$$\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell} = \sum_{i=1}^{\delta} \boldsymbol{Z}_{c_i+q^\ell,b_i,c_i,b_i+1,\ell}. \tag{11}$$

From Proposition 1, we deduce that the set of vectors $\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+1,\ell}$ *i.e.* $\delta = 1$ form a spanning set for the vector space generated by all the vectors $\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}$. We can characterize more precisely this set.

**Definition 6.** *Let $\mathcal{B}_r$ be the set of* nonzero *vectors $\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}$ obtained with tuples $(\delta, b, c, \ell)$ such that $\delta = 1$ and satisfying the following conditions:*

$$\begin{cases} 0 \leqslant b \leqslant r-2 \text{ and } 0 \leqslant c \leqslant r - 1 - q^{\ell} & \text{if } 1 \leqslant \ell \leqslant \lfloor \log_q(r-1) \rfloor \\ 0 \leqslant b < c \leqslant r-2 & \text{if } \ell = 0. \end{cases}$$

**Proposition 2.** *Let $r$ be an integer such that $r \geqslant 3$. The cardinality of $\mathcal{B}_r$ is equal to $T_{alternant}$.*

Proposition 2 gives an explanation of the value of $D_{\text{alternant}}$. To see this, let us define

**Definition 7.** *Consider a certain decomposition of the elements of $\mathbb{F}_{q^m}$ in a $\mathbb{F}_q$ basis. Let $\pi_i : \mathbb{F}_{q^m} \mapsto \mathbb{F}_q$ be the function giving the $i$-th coordinate in this decomposition. By extension we denote for a vector $\boldsymbol{z} = (z_j)_{1 \leq j \leq n} \in \mathbb{F}_{q^m}^n$ by $\pi_i(\boldsymbol{z})$ the vector $(\pi_i(z_j))_{1 \leq j \leq n} \in \mathbb{F}_q^n$.*

We have the following heuristic.

**Heuristic 1.** *For random choices of $x_i$'s and $y_i$'s with $1 \leqslant i \leqslant n$ the set $\{\pi_i(\boldsymbol{Z}) | 1 \leq i \leq m, \boldsymbol{Z} \in \mathcal{B}_r\}$ forms a basis of $\mathcal{L}_{\mathbf{P}}$.*

## 6 Interpretation of the Normalized Dimension – The Binary Goppa Case

In this section we will explain Experimental Fact 2 in the case of a binary Goppa code. We denote by $r$ the degree of the Goppa polynomial. In this case, it is readily seen that the theoretical expression $T_{\text{Goppa}}$ has a simpler expression given by

**Proposition 3.** *Let us define* $e \overset{def}{=} \lceil \log_2 r \rceil + 1$ *and* $N \overset{def}{=} \binom{mr}{2}$. *When* $q = 2$, *the formula in Equation* (8) *can be simplified to* $T_{\mathit{Goppa}} = \frac{1}{2} r \Big( (2e+1)r - 2^e - 1 \Big)$.

Theorem 1 shows that a binary Goppa code of degree $r$ can be regarded as a binary alternant code of degree $2r$. This seems to indicate that we should have

$$D_{\text{Goppa}}(r) = m T_{\text{alternant}}(2r).$$

This is not the case however. It turns out that $D_{\text{Goppa}}(r)$ is significantly smaller than this. In our experiments, we have found out that the vectors of $\mathcal{B}_{2r}$ still form a generating set for $\mathcal{L}_{\mathbf{P}}$, but that they are not independent anymore. Our goal is here to identify the dependencies between the elements of $\mathcal{B}_{2r}$.

We are really interested in the dependencies over the binary field $\mathbb{F}_2$, but we are first going to find linear relations over the extension field $\mathbb{F}_{2^m}$. There are many of them, as shown by the following proposition which exploits that the $Y_i$'s are derived from the Goppa polynomial $\Gamma(z)$ by $Y_i = \Gamma(X_i)^{-1}$. Again, the proofs of this section are postponed to Appendix D.

**Proposition 4.** *Let* $t$, $\ell$ *and* $c$ *be integers such that* $0 \leqslant t \leqslant r-2$, $1 \leqslant \ell \leqslant \lfloor \log_2(2r-1) \rfloor$ *and* $0 \leqslant c \leqslant 2r - 2^\ell - 1$. *We set* $c^* \overset{def}{=} c + 2^{\ell-1}$. *It holds that:*

$$\sum_{b=0}^{r} \gamma_b^{2^\ell} \boldsymbol{Z}_{c+2^\ell,t+b,c,t+b+1,\ell} = \boldsymbol{Z}_{c^*+2^{\ell-1},2t,c^*,2t+1,\ell-1} + \boldsymbol{Z}_{c+2^{\ell-1},2t+1,c,2t+2,\ell-1}. \quad (12)$$

As a consequence of Proposition 4, $\mathcal{B}_{2r}$ can not be a basis of the linearized system in the Goppa case. We count the number of such equations in the following proposition.

**Proposition 5.** *The number* $N_L$ *of equations of the form* (12) *is equal to* $2(r-1)\,(ru + 1 - 2^u)$ *where* $u \overset{def}{=} \lfloor \log_2(2r-1) \rfloor$.

Notice that each equation of the form (12) involves one vector of $\mathcal{B}_{2r}$ that does not satisfy the other equations. These equations are therefore independent and by denoting by $< \mathcal{B}_{2r} >_{\mathbb{F}_{2^m}}$ the vector space over $\mathbb{F}_{2^m}$ generated by the vectors of $\mathcal{B}_{2r}$ we should have

$$\dim < \mathcal{B}_{2r} >_{\mathbb{F}_{2^m}} \leq |\mathcal{B}_{2r}| - N_L.$$

The experiments we have made indicate that actually equality holds here. However, this does not mean that the dimension of the vector space over $\mathbb{F}_2$ generated by the set $\{\pi_i(\boldsymbol{Z}), \boldsymbol{Z} \in \mathcal{B}_{2r}, 1 \leq i \leq m, \boldsymbol{Z} \in \mathcal{B}_{2r}\}$ is equal to $m \dim < \mathcal{B}_{2r} >_{\mathbb{F}_{2^m}}$. It turns out that there are still other dependencies among the $\pi_i(\boldsymbol{Z})$'s. The following proposition gives an explanation of how such dependencies occur.

**Proposition 6.** *Let* $\boldsymbol{Q}_{a,b,c,d,\ell} \overset{def}{=} \left(\boldsymbol{Q}_{a,b,c,d,\ell}[j,j']\right)_{k+1 \leqslant j < j' \leqslant n}$, *with* $\boldsymbol{Q}_{a,b,c,d,\ell}[j,j'] = (\boldsymbol{Z}_{a,b,c,d,\ell}[j,j'])^2$. *For any integers* $b \geqslant 0$, $t \geqslant 0$, $\delta \geqslant 1$ *and* $\ell$ *such that* $0 \leqslant \ell \leqslant \lfloor \log_2(2r-1) \rfloor - 1$, $b + \delta \leq 2r - 1$ *and* $t + 2^\ell \delta \leqslant r - 1$, *we have*

$$\boldsymbol{Z}_{2t+2^{\ell+1}\delta,b,2t,b+\delta,\ell+1} = \sum_{c=0}^{r} \gamma_c^2 \boldsymbol{Q}_{c+2^\ell\delta,b,t+c,b+\delta,\ell}. \quad (13)$$

**Proposition 7.** *The number $N_Q$ of vectors of $\mathcal{B}_{2r}$ satisfying Equation* (13) *is equal to* $(2r-1)(ru - 2^u + 1)$ *where* $u \overset{def}{=} \lfloor \log_2(2r-1) \rfloor$.

Each of such equation gives rise to $m$ linear equations over $\mathbb{F}_2$ involving the $\pi_i(\boldsymbol{Z})$ for $\boldsymbol{Z}$ in $\mathcal{B}_{2r}$. Therefore, it could be expected that $\Delta_{\text{Goppa}} = |\mathcal{B}_{2r}| - N_L - N_Q$. But, some vectors in $\mathcal{B}_{2r}$ appear both in linear relations of the form (12) and "quadratic" equations of the form (13). More precisely, let $\mathcal{B}_{2r}^{\text{quad}}$ be the subset of vectors of $\mathcal{B}_{2r}$ which are involved in an Equation of type (13). There are equations of type (12) which involve only vectors of $\mathcal{B}_{2r}^{\text{quad}}$. Let $N_1$ be their numbers. Moreover, it is possible by adding two equations of type (12) involving at least one vector which is not in $\mathcal{B}_{2r}^{\text{quad}}$ to obtain an equation which involves only vectors of $\mathcal{B}_{2r}^{\text{quad}}$. Let $N_0$ be the number of such sums. Finally, let $N_{L\cap Q} \overset{\text{def}}{=} N_1 + N_0$. It is possible to count such equations to obtain

**Proposition 8.** $N_{L\cap Q} = (r-1)\Big((u - \frac{1}{2})r - 2^u + 2\Big)$ *where* $u \overset{def}{=} \lfloor \log_2(2r-1) \rfloor$.

**Proposition 9.** *For any integer $r \geqslant 2$, we have $T_{Goppa}(r) = |\mathcal{B}_{2r}| - N_L - N_Q + N_{L\cap Q}$.*

## 7 Conclusion and Cryptographic Implications

The existence of a distinguisher for the specific case of binary Goppa codes has consequences for code-based cryptographic primitives because it represents, and by far, the favorite choice in such primitives. We focus in this part on secure parameters that are within the range of validity of our distinguisher. The simple expression given in Proposition 3 is not valid for any value of $r$ and $m$ but tends to be true for codes that have a code rate $\frac{n-mr}{n}$ that is close to one. This kind of codes are mainly encountered with the public keys of the CFS signature scheme [10]. If we assume that the length $n$ is equal to $2^m$ and we denote by $r_{\min}$ the smallest integer $r$ such that $N - mT_{\text{Goppa}} \geqslant 2^m - mr$ then any binary Goppa code defined of degree $r \geqslant r_{\min}$ cannot be distinguished from a random linear code by our technique. This value is gathered in Table 1. One can notice for instance that the binary Goppa code obtained with $m = 13$ and $r = 19$ corresponding to a McEliece public key of 90 bits of security, fits in the range of validity of our distinguisher. The values of $r_{\min}$ in Table 1 are checked by experimentations for $m \leqslant 16$ whereas those for $m \geqslant 17$ are obtained by solving the equation $\frac{mr}{2}\Bigg((2e+1)r - 2^e - 1\Bigg) = \frac{1}{2}mr(mr-1) - 2^m + mr$. Eventually, all the keys proposed in [15] (See therein Table 4) for the CFS scheme can be distinguished.

**Table 1.** Smallest order $r$ of a binary Goppa code of length $n = 2^m$ for which our distinguisher does not work.

| $m$ | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r_{\min}$ | 5 | 8 | 8 | 11 | 16 | 20 | 26 | 34 | 47 | 62 | 85 | 114 | 157 | 213 | 290 | 400 |

## References

1. T. P. Berger, P.L. Cayrel, P. Gaborit, and A. Otmani. Reducing key length of the McEliece cryptosystem. In Bart Preneel, editor, *Progress in Cryptology - Second International Conference on Cryptology in Africa (AFRICACRYPT 2009)*, volume 5580 of *Lecture Notes in Computer Science*, pages 77–97, Gammarth, Tunisia, June 21-25 2009.

2. E. Berlekamp, R. McEliece, and H. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24(3):384–386, May 1978.
3. E. R. Berlekamp. Factoring polynomials over finite fields. In E. R. Berlekamp, editor, *Algebraic Coding Theory*, chapter 6. McGraw-Hill, 1968.
4. D. J. Bernstein, T. Lange, and C. Peters. Attacking and defending the McEliece cryptosystem. In *PQCrypto*, volume 5299 of *LNCS*, pages 31–46, 2008.
5. Bhaskar Biswas and Nicolas Sendrier. McEliece cryptosystem implementation: Theory and practice. In *PQCrypto*, pages 47–62, 2008.
6. W. Bosma, J. J. Cannon, and Catherine Playoust. The Magma algebra system I: The user language. *J. Symb. Comput.*, 24(3/4):235–265, 1997.
7. A. Canteaut and F. Chabaud. A new algorithm for finding minimum-weight words in a linear code: Application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511. *IEEE Transactions on Information Theory*, 44(1):367–378, 1998.
8. Pierre-Louis Cayrel, Philippe Gaborit, David Galindo, and Marc Girault. Improved identity-based identification using correcting codes. *CoRR*, abs/0903.0069, 2009.
9. Colin Cooper. On the distribution of rank of a random matrix over a finite field. *Random Struct. Algorithms*, 17(3-4):197–212, 2000.
10. N. T. Courtois, M. Finiasz, and N. Sendrier. How to achieve a McEliece-based digital signature scheme. *Lecture Notes in Computer Science*, 2248:157–174, 2001.
11. Léonard Dallot. Towards a concrete security proof of Courtois, Finiasz and Sendrier signature scheme. In *WEWoRC*, pages 65–77, 2007.
12. Léonard Dallot and Damien Vergnaud. Provably secure code-based threshold ring signatures. In *IMA Int. Conf.*, pages 222–235, 2009.
13. Rafael Dowsley, Jörn Müller-Quade, and Anderson C. A. Nascimento. A CCA2 secure public key encryption scheme based on the McEliece assumptions in the standard model. In *CT-RSA*, pages 240–251, 2009.
14. Jean-Charles Faugère, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. Algebraic cryptanalysis of mceliece variants with compact keys. In Henri Gilbert, editor, *EUROCRYPT*, volume 6110 of *Lecture Notes in Computer Science*, pages 279–298. Springer, 2010.
15. M. Finiasz and N. Sendrier. Security bounds for the design of code-based cryptosystems. In M. Matsui, editor, *Asiacrypt 2009*, volume 5912 of *LNCS*, pages 88–105. Springer, 2009.
16. P. J. Lee and E. F. Brickell. An observation on the security of McEliece's public-key cryptosystem. In *Advances in Cryptology - EUROCRYPT'88*, volume 330/1988 of *Lecture Notes in Computer Science*, pages 275–280. Springer, 1988.
17. J. S. Leon. A probabilistic algorithm for computing minimum weights of large error-correcting codes. *IEEE Transactions on Information Theory*, 34(5):1354–1359, 1988.
18. P. Loidreau and N. Sendrier. Weak keys in the mceliece public-key cryptosystem. *IEEE Transactions on Information Theory*, 47(3):1207–1211, 2001.
19. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North–Holland, Amsterdam, fifth edition, 1986.
20. R. J. McEliece. *A Public-Key System Based on Algebraic Coding Theory*, pages 114–116. Jet Propulsion Lab, 1978. DSN Progress Report 44.
21. R. Misoczki and P. S. L. M. Barreto. Compact McEliece keys from Goppa codes. In *Selected Areas in Cryptography (SAC 2009)*, Calgary, Canada, August 13-14 2009.
22. Ryo Nojima, Hideki Imai, Kazukuni Kobara, and Kirill Morozov. Semantic security for the McEliece cryptosystem without random oracles. *Des. Codes Cryptography*, 49(1-3):289–305, 2008.
23. N. Patterson. The algebraic decoding of Goppa codes. *IEEE Transactions on Information Theory*, 21(2):203–207, 1975.
24. Erez Petrank and Ron M. Roth. Is code equivalence easy to decide? *IEEE Transactions on Information Theory*, 43(5):1602–1604, 1997.
25. N. Sendrier. Finding the permutation between equivalent linear codes: The support splitting algorithm. *IEEE Transactions on Information Theory*, 46(4):1193–1203, 2000.
26. J. Stern. A method for finding codewords of small weight. In G. D. Cohen and J. Wolfmann, editors, *Coding Theory and Applications*, volume 388 of *Lecture Notes in Computer Science*, pages 106–113. Springer, 1988.

## A Code-Based Public-Key Cryptography

The main cryptographic primitives in code-based public-key cryptography are the McEliece encryption and the CFS signature [10]. We recall that a linear *code* over a finite field $\mathbb{F}_q$ of $q$ elements defined by a $k \times n$ matrix $\boldsymbol{G}$ (with $k \leq n$) over $\mathbb{F}_q$ is the vector space $\mathscr{C}$ spanned by its rows *i.e.* $\mathscr{C} \overset{\text{def}}{=} \left\{ \boldsymbol{uG} \mid \boldsymbol{u} \in \mathbb{F}_q^k \right\}$. $\boldsymbol{G}$ 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 $r$ errors if it recovers $\boldsymbol{u}$ from the knowledge of $\boldsymbol{uG} + \boldsymbol{e}$ for all possible $\boldsymbol{e} \in \mathbb{F}_q^n$ of weight at most $r$.

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

*Public key:* the matrix $\boldsymbol{G} = \boldsymbol{SG_s}\mathbf{P}$.

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

*Decryption:* $\mathbf{c}' = \mathbf{c}\mathbf{P}^{-1}$ is computed from the ciphertext $\mathbf{c}$. Notice that $\mathbf{c}' = (\boldsymbol{uSG_s}\mathbf{P} + \boldsymbol{e})\mathbf{P}^{-1} = \boldsymbol{uSG_s} + \boldsymbol{e}\mathbf{P}^{-1}$ and that $\boldsymbol{e}\mathbf{P}^{-1}$ is of Hamming weight at most $r$. Therefore the aforementioned decoding algorithm can recover in polynomial time $\boldsymbol{uS}$ and therefore the plaintext $\boldsymbol{u}$ by multiplication by $\boldsymbol{S}^{-1}$.

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, namely the alternant code family ([19, 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 [10], 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 $\boldsymbol{G}$ and who wishes to sign a message $\mathbf{x} \in \mathbb{F}_2^k$ has to compute a string $\boldsymbol{u}$ such that the Hamming weight of $\mathbf{x} - \boldsymbol{uG}$ is at most $r$. Anyone (a *verifier*) can publicly check the validity of a signature. Unfortunately, this approach can only provide signatures for messages $\mathbf{x}$ that are within distance $t$ from a codeword $\boldsymbol{uG}$. 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 [10] to choose as in the McEliece cryptosystem, binary Goppa codes for this purpose with the following parameters $n = 2^m$ and $k = n - mr$. The number of trials is of order $r!$ 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 $\frac{2^m - rm}{2^m} = 1 - \frac{mr}{2^m}$ which is for large $n$ (that is for large values of $2^m$) and moderate values of $r$ 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 $r = 10$ whereas the McEliece cryptosystem for the same security needs $n = 2^{11}$ and $r = 32$ ([15]). 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.

## B Proof of Theorem 2

It will be convenient to assume that the columns of $\boldsymbol{M}$ are ordered lexicographically. The index of the first column is $(j, j') = (k+1, k+2)$, the second one is $(j, j') = (k+1, k+3)$, while the last one is $(j, j') = (n-1, n)$. The matrices $\boldsymbol{M}_i$'s which are involved in the Gaussian elimination process mentioned in Section 4 are defined inductively as follows.

Let $E_1$ be the subset of $\{1, \ldots, k\}$ of indices $s$ such that $p_{s,k+1} \neq 0$. Let $F_1$ be the subset of $E_1$ formed by its first $rm-1$ elements (if these elements exist). Now , we set

$$\boldsymbol{M}_1 \overset{\text{def}}{=} (p_{s,k+1} p_{s,j})_{\substack{s \in F_1 \\ k+1 < j \leq n}}. \tag{14}$$

Let $r_1$ be the rank of $\boldsymbol{M}_1$. To simplify the discussion, we assume that:

1. $F_1 = \{1, 2, \ldots, rm-1\}$,
2. the submatrix $\boldsymbol{N}_1$ of $\boldsymbol{M}_1$ formed by its first $r_1$ rows and columns is of full rank.

Note that we can always assume this by performing suitable row and column permutations. In other words $\boldsymbol{M}$ has the following block structure:

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{N}_1 & \boldsymbol{B}_1 \\ \boldsymbol{A}_1 & \boldsymbol{C}_1 \end{pmatrix}.$$

We denote:

$$\boldsymbol{M}^{(1)} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{N}_1^{-1} & \mathbf{0} \\ -\boldsymbol{A}_1 \boldsymbol{N}_1^{-1} & \mathbf{I} \end{pmatrix} \boldsymbol{M},$$

where $\mathbf{0}$ is a matrix of size $r_1 \times (k - r_1)$ with only zero entries and $\mathbf{I}$ is the identity matrix of size $k - r_1$. Notice that $\boldsymbol{M}^{(1)}$ takes the block form:

$$\boldsymbol{M}^{(1)} = \begin{pmatrix} \mathbf{I} & \boldsymbol{B}_1' \\ \mathbf{0} & \boldsymbol{C}_1' \end{pmatrix}.$$

This basically amounts to perform Gaussian elimination on $\boldsymbol{M}$ to put the first $r_1$ columns in standard form. We then define inductively the $E_i, F_i, \boldsymbol{M}_i, \boldsymbol{M}^{(i)}$ and $\boldsymbol{N}_i$ as follows:

$$E_i \overset{\text{def}}{=} \{s | 1 \leq s \leq k, p_{s,k+i} \neq 0\} \setminus F_{i-1},$$
$$F_i \overset{\text{def}}{=} \text{the first } rm - i \text{ elements of } E_i.$$

$\boldsymbol{M}_i$ is the submatrix of $\boldsymbol{M}^{(i-1)}$ obtained from the rows in $F_i$ and the columns associated to the indices of the form $(k+i, j')$ where $j'$ ranges from $k+i+1$ to $n$. $\boldsymbol{M}^{(i)}$ is obtained from $\boldsymbol{M}^{(i-1)}$ by first choosing a square submatrix $\boldsymbol{N}_i$ of $\boldsymbol{M}_i$ of full rank and with the same rank as $\boldsymbol{M}_i$ and then by performing Gaussian elimination on the rows in order to put the columns of $\boldsymbol{M}^{(i-1)}$ involved in $\boldsymbol{N}_i$ in standard form (i.e. the submatrix of $\boldsymbol{M}^{(i-1)}$ corresponding to $\boldsymbol{N}_i$ becomes the identity matrix while the other entries in the columns involved in $\boldsymbol{N}_i$ become zero). It is clear that the whole process leading to $\boldsymbol{M}^{(rm-1)}$ amounts to perform (partial) Gaussian elimination to $\boldsymbol{M}$. Hence:

**Lemma 1.** *When $|E_i| \geq rm - i$, for all $i \in \{1, \ldots, rm-1\}$, we have:*

$$\operatorname{rank}(\boldsymbol{M}) \geq \sum_{i=1}^{rm-1} \operatorname{rank}(\boldsymbol{M}_i).$$

Another observation is that $\boldsymbol{M}_i$ is equal to the sum of the submatrix $(p_{s,k+i}p_{s,j})_{\substack{s\in F_i\\ k+i<j\leq n}}$ of $\boldsymbol{M}$ and a certain matrix which is some function on the entries $p_{t,k+i}p_{t,j}$ where $t$ belongs to $F_1 \cup \dots F_{i-1}$ and $j$ ranges over $\{k+i+1, n\}$. Since by definition of $F_i$, $p_{s,k+i}$ is different from 0 for $s$ in $F_i$. In addition, the rank of $\boldsymbol{M}_i$ does not change by multiplying each row of index $s$ by $p_{s,k+i}^{-1}$. Then, it turns out that the rank of $\boldsymbol{M}_i$ is equal to the rank of a matrix which is the sum of the matrix $(p_{s,j})_{\substack{s\in F_i\\ k+i<j\leq n}}$, another matrix depending on the $p_{t,k+i}p_{t,j}$'s (where $t$ ranges over $F_1 \cup \dots F_{i-1}$) and the $p_{s,k+1}$'s with $s \in F_i$. This proves that:

**Lemma 2.** *Assume that $|E_i| \geq rm - i$ for all $i \in \{1, \dots, rm-1\}$. Then, the random variables $\operatorname{rank}(\boldsymbol{M}_i)$ are independent and $\operatorname{rank}(\boldsymbol{M}_i)$ is distributed as the rank of a square matrix of size $rm - i$ with entries drawn independently from the uniform distribution on $\mathbb{F}_q$.*

Another essential ingredient for proving Theorem 2 is the following well known lemma. (see for instance [9][Theorem 1])

**Lemma 3.** *There exist two positive constants $K_1$ and $K_2$ depending on $q$ such that the probability $p(s,\ell)$ that a random $\ell \times \ell$ matrix over $\mathbb{F}_q$ is of rank $\ell - s$ (where the coefficients are drawn independently from each other from the uniform distribution on $\mathbb{F}_q$) satisfies*

$$\frac{A}{q^{s^2}} \leq p(s,\ell) \leq \frac{B}{q^{s^2}}.$$

This enables to control the exponential moments of the defect of a random matrix. For a square matrix $\boldsymbol{M}$ of size $\ell \times \ell$, we define the defect $d(\boldsymbol{M})$ by $d(\boldsymbol{M}) \stackrel{\text{def}}{=} \ell - \operatorname{rank}(\boldsymbol{M})$.

**Lemma 4.** *If $\boldsymbol{M}$ is random square matrix whose entries are drawn independently from the uniform distribution over $\mathbb{F}_q$, then there exists some constant $K$ such that for every $\lambda > 0$,*

$$\mathbb{E}\left(q^{\lambda d(\boldsymbol{M})}\right) \leq K q^{\frac{\lambda^2}{4}},$$

*$\mathbb{E}(.)$ denoting the expectation.*

*Proof.* By using Lemma 3, we obtain:

$$\mathbb{E}\left(q^{\lambda d(\boldsymbol{M})}\right) \leq \sum_{d=0}^{\infty} q^{\lambda d} \frac{B}{q^{d^2}} \leq B \sum_{d=0}^{\infty} q^{\lambda d - d^2}.$$

Observe that the maximum of the function $d \mapsto q^{\lambda d - d^2}$ is reached for $d_0 = \frac{\lambda}{2}$ and is equal to $q^{\frac{\lambda^2}{4}}$. Then, we can write the sum above as:

$$\sum_{d=0}^{\infty} q^{\lambda d - d^2} = \sum_{d \leq d_0} q^{\lambda d - d^2} + \sum_{d > d_0} q^{\lambda d - d^2}$$

Finally, we notice that:

$$\begin{aligned} \frac{q^{\lambda(d+1)-(d+1)^2}}{q^{\lambda d - d^2}} &\geq \frac{q^{\lambda(d_0+1)-(d_0+1)^2}}{q^{\lambda d_0 - d_0^2}} = \frac{1}{q} \text{ for } d > d_0, \\ \frac{q^{\lambda(d-1)-(d-1)^2}}{q^{\lambda d - d^2}} &\geq \frac{q^{\lambda(d_0-1)-(d_0-1)^2}}{q^{\lambda d_0 - d_0^2}} = \frac{1}{q} \text{ for } d \leq d_0. \end{aligned}$$

This leads to:

$$\begin{aligned}\sum_{d=0}^{\infty} q^{\lambda d - d^2} &\leq \sum_{d \leq d_0}^{\infty} q^{d - \lfloor d_0 \rfloor} q^{\frac{\lambda^2}{4}} + \sum_{d \geq d_0}^{\infty} q^{\lceil d_0 \rceil - d} q^{\frac{\lambda^2}{4}} \\ &= O\left(q^{\frac{\lambda^2}{4}}\right).\end{aligned}$$

We can use now the previous lemma together with Lemma 1 and Lemma 2 to derive

**Lemma 5.** *Assuming that $|E_i| \geq rm - i$ for all $i \in \{1, \dots, t\}$, we get:*

$$\mathbf{prob}\left(\sum_{i=1}^{t} d(\boldsymbol{M}_i) \geq u\right) \leq K^t q^{-\frac{u^2}{t}}$$

*where $K$ is the constant appearing in the previous lemma.*

*Proof.* Let $D \stackrel{\text{def}}{=} \sum_{i=1}^{t} d(\boldsymbol{M}_i)$. Using Markov's inequality:

$$\mathbf{prob}(D \geq u) \leq \frac{\mathbb{E}(q^{\lambda D})}{q^{\lambda u}} \tag{15}$$

for some well chosen $\lambda > 0$. The exponential moment appearing at the numerator is upper-bounded with the help of the previous lemma and by using the independence of the random variables $q^{\lambda d(\boldsymbol{M}_i)}$, i.e.:

$$\begin{aligned}\mathbb{E}(q^{\lambda D}) &= \mathbb{E}\left(q^{\lambda \sum_{i=1}^{t} d(\boldsymbol{M}_i)}\right) \\ &= \prod_{i=1}^{t} \mathbb{E}\left(q^{\lambda d(\boldsymbol{M}_i)}\right) \\ &\leq K^t q^{\frac{t\lambda^2}{4}}. \qquad (16)\end{aligned}$$

Using now (16) in (15), we obtain $\mathbf{prob}(D \geq \alpha t) \leq K^t \frac{q^{\frac{t\lambda^2}{4}}}{q^{\lambda u}} = K^t q^{\frac{t\lambda^2}{4} - \lambda u}$. We choose $\lambda = \frac{2u}{t}$ to minimize this upper-bound, leading to:

$$\mathbf{prob}(D \geq u) \leq K^t q^{-\frac{u^2}{t}}.$$

□

The last ingredient for proving heorem 2 is a bound on the probability that $E_i$ is too small to construct $F_i$.

**Lemma 6.** *Let $u_i \stackrel{def}{=} \binom{mr}{2} - \frac{(2rm-i)(i-1)}{2}$, then*

$$\mathbf{prob}\left(|E_i| < rm - i \mid \; |F_1| = rm - 1, \dots, |F_{i-1}| = rm - i + 1\right) \leq e^{-2\frac{\left(\frac{q-1}{q} u_i - rm - i + 1\right)^2}{u_i}}$$

*Proof.* When all the sets $F_j$ are of size $rm - j$ for $j$ in $\{1, \dots, i-1\}$, it remains $N - \sum_{j=1}^{i-1}(rm - j) = N - \frac{(2rm-i)(i-1)}{2} = u_i$ rows which can be picked up for $E_i$. Let $S_t$ be the sum of $t$ Bernoulli variables of parameter $\frac{q-1}{q}$. We obviously have

$$\mathbf{prob}\left(|E_i| < rm - i \| F_1| = rm - 1, \dots, |F_{i-1}| = rm - i + 1\right) = \mathbf{prob}(S_{u_i} < rm - i)\,.$$

It remains to use the Hoeffding inequality on the binomial tails to finish the proof. □

We are ready now to prove Theorem 2

*Proof (of Theorem 2).* Let $u = \lceil \sqrt{mr\omega(mr))} \rceil$. We observe now that if all $E_j$'s are of size at least $rm - j$ for $j \in \{1, \dots, u\}$, we can write

$$\begin{aligned} D &= N - \operatorname{rank}(\boldsymbol{M}) \\ &\leq N - \sum_{i=1}^{rm-u} \operatorname{rank}(\boldsymbol{M}_i) \text{ (by Lemma 1)} \\ &= \sum_{i=1}^{rm-1} (rm - i) - \sum_{i=1}^{rm-u} \operatorname{rank}(\boldsymbol{M}_i) \\ &= \sum_{i=1}^{rm-u} d(\boldsymbol{M}_i) + \sum_{i=rm-u+1}^{rm-1} (rm - i) \\ &= \sum_{i=1}^{rm-u} d(\boldsymbol{M}_i) + \frac{u(u-1)}{2} \\ &< \sum_{i=1}^{rm-u} d(\boldsymbol{M}_i) + \frac{mr\omega(mr)}{2}. \end{aligned}$$

From this we deduce that

$$\mathbf{prob}(D_{\text{random}} \geq mr\omega(mr)) \leq \mathbf{prob}(A \cup B) \leq \mathbf{prob}(A) + \mathbf{prob}(B)$$

where $A$ is the event "$\sum_{i=1}^{rm-u} d(\boldsymbol{M}_i) \geq \frac{mr\omega(mr)}{2}$" and $B$ is the event "for at least one $E_j$ with $j \in \{1, \dots, rm-u\}$ we have $|E_j| < rm - j$". We use now Lemma 5 to prove that $\mathbf{prob}(A) = o(1)$ as $rm$ goes to infinity. We finish the proof by noticing that the probability of the complementary set of $B$ satisfies

$$\begin{aligned} \mathbf{prob}(\bar{B}) &= \mathbf{prob}\left( \bigcap_{i=1}^{rm-u} |E_i] \geq rm - i \right) \\ &= \prod_{i=1}^{rm-u} \mathbf{prob}\left( |E_i| \geq rm - i \,||F_1| = rm - 1, \dots, |F_{i-1}| = rm - i + 1 \right) \\ &= 1 - o(1) \text{ (by Lemma 6).} \end{aligned}$$

□

## C The Alternant Case

In this appendix, we are going to prove Proposition 1 and Proposition 2.

### C.1 Proof of Proposition 1

We require the following lemma.

**Lemma 7.** *For any integers a, b, c, d, e, f in* $\{0,\dots,r-1\}$*, and an integer* $\ell$ *in* $\left\{0,\dots,\lfloor\log_q(r-1)\rfloor\right\}$ *such that* $a+q^\ell b = c+q^\ell d$ *we have:*

$$\boldsymbol{Z}_{a,b,c,d,\ell} + \boldsymbol{Z}_{c,d,e,f,\ell} = \boldsymbol{Z}_{a,b,e,f,\ell} \tag{17}$$

*Proof (Proposition 1).* Let $b^* \overset{\text{def}}{=} b+1$, $\delta^* \overset{\text{def}}{=} \delta - 1$ and $c^* \overset{\text{def}}{=} c+q^\ell\delta^*$. Then $c^*$ is the integer such that $c^*+q^\ell = c+q^\ell\delta$, one can see that $c+q^\ell\delta^* = c+q^\ell(\delta-1) = c^*$ and by Lemma 7 we have:

$$\boldsymbol{Z}_{c^*+q^\ell,b,c^*,b+1,\ell} + \boldsymbol{Z}_{c+q^\ell\delta^*,b^*,c,b^*+\delta^*,\ell} = \boldsymbol{Z}_{c^*+q^\ell,b,c,b^*+\delta^*,\ell} = \boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}$$

which means that

$$\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell} = \boldsymbol{Z}_{c^*+q^\ell,b,c^*,b+1,\ell} + \boldsymbol{Z}_{c+q^\ell\delta^*,b^*,c,b^*+\delta^*,\ell} \tag{18}$$

The proof follows by induction. □

### C.2 Proof of Proposition 2

*Proof (Proposition 2).* Let us set $e \overset{\text{def}}{=} \lfloor\log_q(r-1)\rfloor$. Then the number of elements in $\mathcal{B}_r$ is given by the number of tuples $(b,c,\ell)$. Therefore we get:

$$\begin{aligned}|\mathcal{B}_r| &= \frac{1}{2}(r-1)(r-2) + \sum_{\ell=1}^{e}\sum_{b=0}^{r-2}(r-q^\ell) = \frac{1}{2}(r-1)\left(r-2+2er-2\sum_{\ell=1}^{e}q^\ell\right)\\ &= \frac{1}{2}(r-1)\left((2e+1)r-2\sum_{\ell=0}^{e}q^\ell\right) = T_{\text{alternant}}\end{aligned}$$

□

## D The Binary Goppa Case

### D.1 Proof of Proposition 4

Propositon 4 which needs Lemma 8 is actually a particular case of Proposition 10.

**Lemma 8.** *Let $\ell$, $\delta$, $b$ and $c$ be integers such that $\ell \geqslant 0$, $\delta \geqslant 1$, $1 \leqslant b+\delta \leqslant r-1$, $1 \leqslant c+q^\ell\delta \leqslant r-1$. We have for any $j$ and $j'$ such that $k+1 \leqslant j < j' \leqslant n$:*

$$\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}[j,j'] = \left(X_j^\delta + X_{j'}^\delta\right)^{q^\ell} \left(Y_j X_j^c \left(Y_{j'} X_{j'}^b\right)^{q^\ell} + Y_{j'} X_{j'}^c \left(Y_j X_j^b\right)^{q^\ell}\right) \tag{19}$$

*Proof.* Let $d = b+\delta$ and $a = c + q^\ell\delta$. We can write that:

$$\begin{aligned}\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}[j,j'] = \boldsymbol{Z}_{a,b,c,d,\ell}[j,j'] &= Y_j Y_{j'}^{q^\ell}\left(X_j^a X_{j'}^{q^\ell b} + X_j^c X_{j'}^{q^\ell d}\right) + Y_{j'} Y_j^{q^\ell}\left(X_{j'}^a X_j^{q^\ell b} + X_{j'}^c X_j^{q^\ell d}\right)\\ &= Y_j Y_{j'}^{q^\ell} X_{j'}^{q^\ell b}\left(X_j^a + X_j^c X_{j'}^{q^\ell \delta}\right) + Y_{j'} Y_j^{q^\ell} X_j^{q^\ell b}\left(X_{j'}^a + X_{j'}^c X_j^{q^\ell\delta}\right)\end{aligned}$$

Using the identity $a = c + q^\ell\delta$, we also have:

$$\begin{aligned}\boldsymbol{Z}_{c+q^\ell\delta,b,c,b+\delta,\ell}[j,j'] &= Y_j Y_{j'}^{q^\ell} X_{j'}^{q^\ell b} X_j^c \left(X_j^{q^\ell\delta} + X_{j'}^{q^\ell\delta}\right) + Y_{j'} Y_j^{q^\ell} X_j^{q^\ell b} X_{j'}^c\left(X_{j'}^{q^\ell\delta} + X_j^{q^\ell\delta}\right)\\ &= \left(X_j^{q^\ell\delta} + X_{j'}^{q^\ell\delta}\right)\left(Y_j Y_{j'}^{q^\ell} X_{j'}^{q^\ell b} X_j^c + Y_{j'} Y_j^{q^\ell} X_j^{q^\ell b} X_{j'}^c\right)\end{aligned}$$

□

**Proposition 10.** *Let $t$, $\ell$, $\delta$ and $c$ be integers such that $t \geqslant 0$, $\ell \geqslant 1$, $\delta \geqslant 1$, $t+\delta \leqslant r-1$, $c \geqslant 0$ and $c + 2^\ell\delta \leqslant 2r-1$. We have:*

$$\sum_{b=0}^{r} \gamma_b^{2^\ell} \boldsymbol{Z}_{c+2^\ell\delta,t+b,c,t+b+\delta,\ell} = \boldsymbol{Z}_{c'+2^{\ell'}\delta',b',c',b'+\delta',\ell'} \tag{20}$$

*where $\ell' = \ell - 1$, $\delta' = 2\delta$, $b' = 2t$, $c' = c$.*

*Proof.* By Lemma 8, we have that:

$$\boldsymbol{Z}_{c+2^\ell\delta,t+b,c,t+b+\delta,\ell}[j,j'] = \left(X_j^\delta + X_{j'}^\delta\right)^{2^\ell}\left(Y_j X_j^c Y_{j'}^{2^{\ell-1}} X_{j'}^{2^\ell t}\left(Y_{j'} X_{j'}^{2b}\right)^{2^{\ell-1}} + Y_{j'} X_{j'}^c Y_j^{2^{\ell-1}} X_j^{2^\ell t}\left(Y_j X_j^{2b}\right)^{2^{\ell-1}}\right)$$

Using the fact that $Y_j \sum_{b=0}^r \gamma_b^2 X_j^{2b} = 1$ and $Y_{j'} \sum_{b=0}^r \gamma_b^2 X_{j'}^{2b} = 1$ we also have:

$$\sum_{b=0}^{r} \gamma_b^{2^\ell} \boldsymbol{Z}_{c+2^\ell\delta,t+b,c,t+b+\delta,\ell}[j,j'] =$$

$$\left(X_j^\delta + X_{j'}^\delta\right)^{2^\ell}\left(Y_j X_j^c Y_{j'}^{2^{\ell-1}} X_{j'}^{2^\ell t}\left(Y_{j'}\sum_{b=0}^r \gamma_b^2 X_{j'}^{2b}\right)^{2^{\ell-1}} + Y_{j'} X_{j'}^c Y_j^{2^{\ell-1}} X_j^{2^\ell t}\left(Y_j \sum_{b=0}^r \gamma_b^2 X_j^{2b}\right)^{2^{\ell-1}}\right)$$

$$= \left(X_j^{2\delta} + X_{j'}^{2\delta}\right)^{2^{\ell-1}}\left(Y_j X_j^c \left(Y_{j'} X_{j'}^{2t}\right)^{2^{\ell-1}} + Y_{j'} X_{j'}^c\left(Y_j X_j^{2t}\right)^{2^{\ell-1}}\right) = \boldsymbol{Z}_{c'+2^{\ell'}\delta',b',c',b'+\delta',\ell'}[j,j']$$

with $\ell' = \ell - 1$, $\delta' = 2\delta$, $b' = 2t$, $c' = c$. Since $c' + 2^{\ell'}\delta' = c + 2^\ell\delta$ and $c + 2^\ell\delta \leqslant 2r-1$ we have $c' + 2^{\ell'}\delta' \leqslant 2r - 1$. Moreover, we require $b' + \delta' \leqslant 2r-1$ which means $2(t+\delta) \leqslant 2r-1$. This last inequality implies $t + \delta \leqslant r - 1$. □

*Proof (of Proposition 4).* By Proposition 10 when $\delta = 1$ we have the following equality:

$$\sum_{b=0}^{r} \gamma_b^{2^\ell} \boldsymbol{Z}_{c+2^\ell,t+b,c,t+b+1,\ell} = \boldsymbol{Z}_{c+2^\ell,2t,c,2(t+1),\ell-1}$$

Moreover by Proposition 1, we also have:

$$\boldsymbol{Z}_{c+2^\ell,2t,c,2(t+1),\ell-1} = \boldsymbol{Z}_{c^*+2^{\ell-1},2t,c^*,2t+1,\ell-1} + \boldsymbol{Z}_{c+2^{\ell-1},2t+1,c,2t+2,\ell-1}$$

where by definition $c^*$ is equal to $c + 2^{\ell-1}$. □

### D.2 Proof of Proposition 5

*Proof (of Proposition 5).* Each equation is defined by a triple $(t, c, \ell)$. As $0 \leqslant t \leqslant r-2$, $1 \leqslant \ell \leqslant u$ and $0 \leqslant c \leqslant 2r - 2^\ell - 1$, we therefore have:

$$N_L = \sum_{t=0}^{r-2} \sum_{\ell=1}^{u} (2r - 2^\ell).$$

One can easily check that this expression is exactly the same as given in the proposition. □

### D.3 Proof of Proposition 6 and Proposition 7

*Proof (of Proposition 6).* For any $j$ and $j'$ such that $k+1 \leqslant j < j' \leqslant n$, we have:

$$\begin{aligned}
&\sum_{c=0}^{r} \gamma_c^2 \left(\boldsymbol{Z}_{c+2^\ell\delta,b,t+c,b+\delta,\ell}\right)^2 [j,j'] = \\
&\left(X_j^\delta + X_{j'}^\delta\right)^{2^{\ell+1}} \left(\left(Y_{j'} X_{j'}^b\right)^{2^{\ell+1}} X_j^{2t} Y_j^2 \sum_{c=0}^{r} \gamma_c^2 X_j^{2c} + \left(Y_j X_j^b\right)^{2^{\ell+1}} X_{j'}^{2t} Y_{j'}^2 \sum_{c=0}^{r} \gamma_c^2 X_{j'}^{2c}\right) \\
= &\left(X_j^\delta + X_{j'}^\delta\right)^{2^{\ell+1}} \left(\left(Y_{j'} X_{j'}^b\right)^{2^{\ell+1}} Y_j X_j^{2t} + \left(Y_j X_j^b\right)^{2^{\ell+1}} Y_{j'} X_{j'}^{2t}\right) = \boldsymbol{Z}_{c'+2^{\ell'}\delta,b',c',b'+\delta',\ell'}[j,j']
\end{aligned}$$

with $\ell' = \ell + 1$, $\delta' = \delta$, $b' = b$, $c' = 2t$ and $c' + 2^{\ell'}\delta' = 2t + 2^{\ell+1}\delta$. In particular, one can easily check that the necessary conditions are $b + \delta \le 2r - 1$ and $t + 2^\ell\delta \leqslant r - 1$ in order for this equation to hold. □

*Proof (of Proposition 7).* By Proposition 6 we know that $N_Q$ is the number of vectors $\boldsymbol{Z}_{2t+2^{\ell+1}\delta,b,2t,b+\delta,\ell+1}$ obtained with $\delta = 1$, $b \geqslant 0$, $t \geqslant 0$ and satisfying $0 \leqslant \ell \leqslant u-1$, $b + \delta \le 2r - 1$ and $t + 2^\ell\delta \leqslant r - 1$. Therefore we have:

$$N_Q = \sum_{l=0}^{u-1} \sum_{t=0}^{r-1-2^\ell} (2r-1) \tag{21}$$

which is equal to the desired expression. □

### D.4 Proof of Proposition 8

*Proof (of Proposition 8).* We will consider vectors $\boldsymbol{Z}_{c+2^\ell,b,c,b+1,\ell}$ of $\mathcal{B}_{2r}$ that satisfy Equation (13) and such that there exists a linear relation that link them. In other words, we consider all the linear relations of the form $\sum_i \alpha_i \boldsymbol{Z}_{c_i+2^{\ell_i},b_i,c_i,b_i+1,\ell_i} = 0$ with $\alpha_i$ in $\mathbb{F}_{2^m}$ and where each $\boldsymbol{Z}_{c_i+2^{\ell_i},b_i,c_i,b_i+1,\ell_i}$ is equal to a linear relation of the form (13). We will see that the number of *independent* equations is equal to $N_{L\cap Q}$. Firstly, one can observe that for any such vectors we necessary have $c_i$ even and $1 \leqslant \ell_i \leqslant u$. We also know by Proposition 4 that for any integers $t$, $\ell$ and $c$ such that $0 \leqslant t \leqslant r-2$, $1 \leqslant \ell \leqslant u$ and $0 \leqslant c \leqslant 2r - 2^\ell - 1$, we have the following linear relation:

$$\sum_{b=0}^{r} \gamma_b^{2^\ell} \boldsymbol{Z}_{c+2^\ell,t+b,c,t+b+1,\ell} = \boldsymbol{Z}_{c^*+2^{\ell-1},2t,c^*,2t+1,\ell-1} + \boldsymbol{Z}_{c+2^{\ell-1},2t+1,c,2t+2,\ell-1}$$

where by definition $c^* = c + 2^{\ell-1}$. Note in particular that whenever $c$ is even then $c^*$ is also even and if $\ell \geqslant 2$ then we obtain a linear relation between some vectors that also satisfy quadratic equations of the form (13). Each equation enables to remove one quadratic equation. So if we denote by $N_1$ the number of equations of the form (12) with $c$ even and $\ell \geqslant 2$, we have then:

$$N_1 = \sum_{t=0}^{r-2}\sum_{\ell=2}^{u}\left(\frac{1}{2}(2r-2^\ell)\right) = (r-1)\sum_{\ell=1}^{u-1}(r-2^\ell) = (r-1)\Big((u-1)r - 2^u + 2\Big). \tag{22}$$

Moreover in the case $\ell = 1$ Equation (20) becomes

$$\sum_{b=0}^{r} \gamma_b^{2} \boldsymbol{Z}_{c+2,t+b,c,t+b+1,1} = \boldsymbol{Z}_{c+2,2t,c,2t+2,0}.$$

In particular when $c$ is even, say for instance $c = 2t'$ for some integer, then this last equation can be rewritten as:

$$\sum_{b=0}^{r} \gamma_b^{2} \boldsymbol{Z}_{2t'+2,t+b,2t',t+b+1,1} = \boldsymbol{Z}_{2t'+2,2t,2t',2t+2,0}. \tag{23}$$

We know that when $t' = t$ then $\boldsymbol{Z}_{2t'+2,2t,2t',2t+2,0}$ is zero. In that case we obtain new relations between vectors satisfying quadratic equations that are independent even from those obtained with $\ell \geqslant 2$. As for the case when $t \neq t'$ we also have $\boldsymbol{Z}_{2t'+2,2t,2t',2t+2,0} = \boldsymbol{Z}_{2t+2,2t',2t,2t'+2,0}$. From this identity and from Equation (23) we then obtain new relations of the following form:

$$\sum_{b=0}^{r} \gamma_b^{2} \boldsymbol{Z}_{2t'+2,t+b,2t',t+b+1,1} = \sum_{b=0}^{r} \gamma_b^{2} \boldsymbol{Z}_{2t+2,t'+b,2t,t'+b+1,1} \tag{24}$$

This last equation involve only vectors that satisfy also quadratic equations. So the number $N_0$ of equations of the form (24) is given by the number of sets $\{t, t'\}$. But by assumption $t$ and $t'$ should satisfy $0 \leqslant t \leqslant r-2$ and $c = 2t'$ with $0 \leqslant c \leqslant 2r-3$, which implies that $0 \leqslant t' \leqslant r-2$. Therefore, $N_0$ is equal to the number $(t, t')$ such that $t \leqslant t'$ and thus we get:

$$N_0 = \sum_{t=0}^{r-2}\sum_{t'=t}^{r-2} = \frac{1}{2}(r-1)r. \tag{25}$$

Finally, by gathering all the cases we therefore obtain that:

$$N_{L\cap Q} = N_1 + N_0 = (r-1)\Big((u-1)r - 2^u + 2\Big) + \frac{1}{2}(r-1)r.$$

□

### D.5 Proof of Proposition 9

*Proof (of Proposition 9).* Set $u \overset{\text{def}}{=} \lfloor \log_2(2r-1) \rfloor$. From Equation (7), we have $|\mathcal{B}_{2r}| = (2r-1)\big((2u+1)r - 2^{u+1} + 1\big)$ which implies from Proposition 7

$$|\mathcal{B}_{2r}| - N_Q = (2r-1)\left((2u+1)r - 2^{u+1} + 1 - (ru - 2^u + 1)\right) \tag{26}$$

$$= (2r-1)((u+1)r - 2^u). \tag{27}$$

Moreover, from Proposition 5 and Proposition 8, we can write:

$$N_L - N_{L\cap Q} = (r-1)\Big(2ur + 2 - 2^{u+1} - (ur - \frac{r}{2} - 2^u + 2)\Big) \tag{28}$$

$$= (r-1)\Big((u + \frac{1}{2})r - 2^u\Big) \tag{29}$$

Therefore by gathering all these equalities we obtain:

$$|\mathcal{B}_{2r}| - (N_L + N_Q - N_{L\cap Q}) = r\left((u + \frac{3}{2})r - 2^u - \frac{1}{2}\right) \tag{30}$$

On the other hand from Proposition 3, we have $T_{\text{Goppa}}(r) = \frac{1}{2}r\left((2e+1)r - 2^e - 1\right)$ where $e = \lceil \log_2 r \rceil + 1$. Using the basic inequality $2r - 1 < 2r < 2(2r-1)$, we have therefore $\log_2(2r-1) < \log_2(r) + 1 < \log_2(2r-1) + 1$ which finally implies $\lceil \log_2 r \rceil = u$. Thus, $T_{\text{Goppa}}(r) = \frac{1}{2}r\left((2u+3)r - 2^{u+1} - 1\right)$ and the proposition is proved. □

## E Experimental Results

We gathered samples of results we obtained through intensive computations with the Magma system [6] 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 $\Gamma(z)$ of degree $r$ and hence $\Gamma(z)$ has no multiple roots. In particular, we can apply Theorem 1 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 2 and Table 3 give figures for the binary case with $m = 14$. We define $T_{\mathsf{alternant}}$ and $T_{\mathsf{Goppa}}$ respectively as the expected normalized dimensions for an alternant and a Goppa code deduced from the formulas (7) and (8). 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 $T_{\mathsf{alternant}}$ fit as long as $k \geq N - mT_{\mathsf{alternant}}$ which correspond to $r \leq 15$. This is also the case for binary Goppa codes since we have $mT_{\mathsf{Goppa}} = D_{\text{Goppa}}$ as long as $k \geq N - mT_{\mathsf{Goppa}}$ *i.e.* $r \leq 25$. We also give in Table 10 and Table 11 the examples 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. 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 $r \leq 50$.

**Table 2.** $q = 2$ and $m = 14$

| $r$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 861 | 1540 | 2415 | 3486 | 4753 | 6216 | 7875 | 9730 | 11781 | 14028 | 16471 | 19110 | 21945 | 24976 |
| $k$ | 16342 | 16328 | 16314 | 16300 | 16286 | 16272 | 16258 | 16244 | 16230 | 16216 | 16202 | 16188 | 16174 | 16160 |
| $D_{\text{random}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 269 | 2922 | 5771 | 8816 |
| $D_{\text{alternant}}$ | 42 | 126 | 308 | 560 | 882 | 1274 | 1848 | 2520 | 3290 | 4158 | 5124 | 6188 | 7350 | 8816 |
| $mT_{\mathsf{alternant}}$ | 42 | 126 | 308 | 560 | 882 | 1274 | 1848 | 2520 | 3290 | 4158 | 5124 | 6188 | 7350 | 8610 |
| $D_{\text{Goppa}}$ | 252 | 532 | 980 | 1554 | 2254 | 3080 | 4158 | 5390 | 6776 | 8316 | 10010 | 11858 | 13860 | 16016 |
| $mT_{\mathsf{Goppa}}$ | 252 | 532 | 980 | 1554 | 2254 | 3080 | 4158 | 5390 | 6776 | 8316 | 10010 | 11858 | 13860 | 16016 |

**Table 3.** $q = 2$ and $m = 14$

| $r$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 28203 | 31626 | 35245 | 39060 | 43071 | 47278 | 51681 | 56280 | 61075 | 66066 | 71253 | 76636 | 82215 | 87990 |
| $k$ | 16146 | 16132 | 16118 | 16104 | 16090 | 16076 | 16062 | 16048 | 16034 | 16020 | 16006 | 15992 | 15978 | 15964 |
| $D_{\text{random}}$ | 12057 | 15494 | 19127 | 22956 | 26981 | 31202 | 35619 | 40232 | 45041 | 50046 | 55247 | 60644 | 66237 | 72026 |
| $D_{\text{alternant}}$ | 12057 | 15494 | 19127 | 22956 | 26981 | 31202 | 35619 | 40232 | 45041 | 50046 | 55247 | 60644 | 66237 | 72026 |
| $mT_{\mathsf{alternant}}$ | 10192 | 11900 | 13734 | 15694 | 17780 | 19992 | 22330 | 24794 | 27384 | 30100 | 32942 | 35910 | 39004 | 42224 |
| $D_{\text{Goppa}}$ | 18564 | 21294 | 24206 | 27300 | 30576 | 34034 | 37674 | 41496 | 45500 | 50046 | 55247 | 60644 | 66237 | 72026 |
| $mT_{\mathsf{Goppa}}$ | 18564 | 21294 | 24206 | 27300 | 30576 | 34034 | 37674 | 41496 | 45500 | 49686 | 54054 | 58604 | 63336 | 68250 |

**Table 4.** $q = 2$ and $m = 15$

| $r$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 990 | 1770 | 2775 | 4005 | 5460 | 7140 | 9045 | 11175 | 13530 | 16110 | 18915 | 21945 | 25200 | 28680 |
| $k$ | 32723 | 32708 | 32693 | 32678 | 32663 | 32648 | 32633 | 32618 | 32603 | 32588 | 32573 | 32558 | 32543 | 32528 |
| $D_{\text{random}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $D_{\text{Goppa}}$ | 270 | 570 | 1050 | 1665 | 2415 | 3300 | 4455 | 5775 | 7260 | 8910 | 10725 | 12705 | 14850 | 17160 |
| $mT_{\text{Goppa}}$ | 270 | 570 | 1050 | 1665 | 2415 | 3300 | 4455 | 5775 | 7260 | 8910 | 10725 | 12705 | 14850 | 17160 |

**Table 5.** $q = 2$ and $m = 15$

| $r$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 32385 | 36315 | 40470 | 44850 | 49455 | 54285 | 59340 | 64620 | 70125 | 75855 | 81810 | 87990 | 94395 | 101025 |
| $k$ | 32513 | 32498 | 32483 | 32468 | 32453 | 32438 | 32423 | 32408 | 32393 | 32378 | 32363 | 32348 | 32333 | 32318 |
| $D_{\text{random}}$ | 0 | 3817 | 7987 | 12382 | 17002 | 21847 | 26917 | 32212 | 37732 | 43477 | 49447 | 55642 | 62062 | 68707 |
| $D_{\text{Goppa}}$ | 19890 | 22815 | 25935 | 29250 | 32760 | 36465 | 40365 | 44460 | 48750 | 53235 | 57915 | 62790 | 67860 | 73125 |
| $mT_{\text{Goppa}}$ | 19890 | 22815 | 25935 | 29250 | 32760 | 36465 | 40365 | 44460 | 48750 | 53235 | 57915 | 62790 | 67860 | 73125 |

**Table 6.** $q = 2$ and $m = 15$

| $r$ | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 107880 | 114960 | 122265 | 129795 | 137550 | 145530 | 153735 | 162165 | 170820 | 179700 | 188805 | 198135 | 207690 | 217470 |
| $k$ | 32303 | 32288 | 32273 | 32258 | 32243 | 32228 | 32213 | 32198 | 32183 | 32168 | 32153 | 32138 | 32123 | 32108 |
| $D_{\text{random}}$ | 75577 | 82672 | 89992 | 97537 | 105307 | 113302 | 121522 | 129967 | 138637 | 147532 | 156652 | 165997 | 175567 | 185362 |
| $D_{\text{Goppa}}$ | 78585 | 84240 | 90585 | 97537 | 105307 | 113302 | 121522 | 129967 | 138637 | 147532 | 156652 | 165997 | 175567 | 185362 |
| $mT_{\text{Goppa}}$ | 78585 | 84240 | 90585 | 97155 | 103950 | 110970 | 118215 | 125685 | 133380 | 141300 | 149445 | 157815 | 166410 | 175230 |

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

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

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

| $r$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 36856 | 41328 | 46056 | 51040 | 56280 | 61776 | 67528 | 73536 | 79800 | 86320 | 93096 | 100128 | 107416 | 114960 |
| $k$ | 65264 | 65248 | 65232 | 65216 | 65200 | 65184 | 65168 | 65152 | 65136 | 65120 | 65104 | 65088 | 65072 | 65056 |
| $D_{\text{random}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 2360 | 8384 | 14664 | 21200 | 27992 | 35040 | 42344 | 49904 |
| $D_{\text{Goppa}}$ | 21216 | 24336 | 27664 | 31200 | 34944 | 38896 | 43056 | 47424 | 52000 | 56784 | 61776 | 66976 | 72384 | 78000 |
| $mT_{\text{Goppa}}$ | 21216 | 24336 | 27664 | 31200 | 34944 | 38896 | 43056 | 47424 | 52000 | 56784 | 61776 | 66976 | 72384 | 78000 |

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

| $r$ | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 122760 | 130816 | 139128 | 147696 | 156520 | 165600 | 174936 | 184528 | 194376 | 204480 | 214840 | 225456 | 236328 |
| $k$ | 65040 | 65024 | 65008 | 64992 | 64976 | 64960 | 64944 | 64928 | 64912 | 64896 | 64880 | 64864 | 64848 |
| $D_{\text{random}}$ | 57720 | 65792 | 74120 | 82704 | 91544 | 100640 | 109992 | 119600 | 129464 | 139584 | 149960 | 160592 | 171480 |
| $D_{\text{Goppa}}$ | 83824 | 89856 | 96624 | 103632 | 110880 | 118368 | 126096 | 134064 | 142272 | 150720 | 159408 | 168336 | 177504 |
| $mT_{\mathsf{Goppa}}$ | 83824 | 89856 | 96624 | 103632 | 110880 | 118368 | 126096 | 134064 | 142272 | 150720 | 159408 | 168336 | 177504 |

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

| $r$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 153 | 276 | 435 | 630 | 861 | 1128 | 1431 | 1770 | 2145 | 2556 | 3003 | 3486 | 4005 | 4560 |
| $k$ | 4078 | 4072 | 4066 | 4060 | 4054 | 4048 | 4042 | 4036 | 4030 | 4024 | 4018 | 4012 | 4006 | 4000 |
| $D_{\text{random}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 560 |
| $D_{\text{alternant}}$ | 6 | 18 | 60 | 120 | 198 | 294 | 408 | 540 | 690 | 858 | 1044 | 1248 | 1470 | 1710 |
| $mT_{\mathsf{alternant}}$ | 6 | 18 | 60 | 120 | 198 | 294 | 408 | 540 | 690 | 858 | 1044 | 1248 | 1470 | 1710 |
| $D_{\text{Goppa}}$ | 18 | 60 | 120 | 198 | 294 | 408 | 540 | 750 | 990 | 1260 | 1560 | 1890 | 2250 | 2640 |
| $mT_{\mathsf{Goppa}}$ | 18 | 60 | 120 | 198 | 294 | 408 | 540 | 750 | 990 | 1260 | 1560 | 1890 | 2250 | 2640 |

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

| $r$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | 5151 | 5778 | 6441 | 7140 | 7875 | 8646 | 9453 | 10296 | 11175 | 12090 | 13041 | 14028 | 15051 | 16110 |
| $k$ | 3994 | 3988 | 3982 | 3976 | 3970 | 3964 | 3958 | 3952 | 3946 | 3940 | 3934 | 3928 | 3922 | 3916 |
| $D_{\text{random}}$ | 1157 | 1790 | 2459 | 3164 | 3905 | 4682 | 5495 | 6344 | 7229 | 8150 | 9107 | 10100 | 11129 | 12194 |
| $D_{\text{alternant}}$ | 2064 | 2448 | 2862 | 3306 | 3905 | 4682 | 5495 | 6344 | 7229 | 8150 | 9107 | 10100 | 11129 | 12194 |
| $mT_{\mathsf{alternant}}$ | 2064 | 2448 | 2862 | 3306 | 3780 | 4284 | 4818 | 5382 | 5976 | 6600 | 7254 | 7938 | 8652 | 9396 |
| $D_{\text{Goppa}}$ | 3060 | 3510 | 3990 | 4500 | 5040 | 5610 | 6210 | 6840 | 7500 | 8190 | 9107 | 10100 | 11129 | 12194 |
| $mT_{\mathsf{Goppa}}$ | 3060 | 3510 | 3990 | 4500 | 5040 | 5610 | 6210 | 6840 | 7500 | 8190 | 8910 | 9660 | 10440 | 11250 |