

# On the structure of the Schur squares of Twisted Generalized Reed-Solomon codes and application to cryptanalysis

Alain Couvreur<sup>1,2</sup>, Rakhi Pratihar<sup>1,2</sup>, Nihan Tanisali<sup>1,2</sup>, and Iliaria Zappatore<sup>3</sup>

<sup>1</sup> Inria

<sup>2</sup> Laboratoire LIX, CNRS UMR 7161, École Polytechnique, Institut Polytechnique de Paris, 1 rue Honoré d'Estienne d'Orves, 91120 Palaiseau Cedex

<sup>3</sup> XLIM, CNRS UMR 7252, Université de Limoges, 123, avenue Albert Thomas - 87060 Limoges Cedex

{alain.couvreur,rakhi.pratihar,nihan.tanisali}@inria.fr  
ilaria.zappatore@unilim.fr

**Abstract.** Twisted generalized Reed-Solomon (TGRS) codes constitute an interesting family of evaluation codes, containing a large class of maximum distance separable codes non-equivalent to generalized Reed-Solomon (GRS) ones. Moreover, the Schur squares of TGRS codes may be much larger than those of GRS codes with same dimension. Exploiting these structural differences, in 2018, Beelen, Bossert, Puchinger and Rosenkilde proposed a subfamily of Maximum Distance Separable (MDS) Twisted Reed-Solomon (TRS) codes over  $\mathbb{F}_q$  with  $\ell$  twists  $q \approx n^{2\ell}$  for McEliece encryption, claiming their resistance to both Sidelnikov Shestakov attack and Schur products-based attacks. In short, they claimed these codes to resist to classical key recovery attacks on McEliece encryption scheme instantiated with Reed-Solomon (RS) or GRS codes. In 2020, Lavauzelle and Renner presented an original attack on this system based on the computation of the subfield subcode of the public TRS code.

In this paper, we show that the original claim on the resistance of TRS and TGRS codes to Schur products based-attacks is wrong. We identify a broad class of codes including TRS and TGRS ones that is distinguishable from random by computing the Schur square of some shortening of the code. Then, we focus on the case of single twist (*i.e.*,  $\ell = 1$ ), which is the most efficient one in terms of decryption complexity, to derive

---

This work is partially funded by the French Agence Nationale de la Recherche project ANR-21-CE39-0009-BARRACUDA, by Plan France 2030 ANR-22-PETQ-0008 and by Horizon-Europe MSCA-DN project ENCODE.

an attack. The technique is similar to the distinguisher-based attacks of RS code-based systems given by Couvreur, Gaborit, Gauthier-Umaña, Otmani, Tillich in 2014.

**Keywords:** Twisted generalised Reed-Solomon codes, Schur products, Code-based Cryptography, McEliece encryption scheme, Cryptanalysis,

## 1 Introduction

*McEliece's encryption scheme* dates back to the early ages of public key cryptography. For a long time it has been considered unusable because of the significant size of the public key. However, with the recent and growing interest of post-quantum cryptographic primitives, Classic McEliece [1] could be standardized in the near future. Besides the seminal proposal by McEliece himself based on classical Goppa codes, there have been many attempts to replace Goppa codes by other families of codes with more efficient decoding algorithms in order to reduce the key size. In 1986, Niederreiter [23] suggested generalized Reed-Solomon codes (GRS) to replace Goppa codes, but it was shown to be insecure by Sidelnikov and Shestakov in [27]. Thereafter, several instantiations using codes “close to GRS” codes appeared. Berger and Loidreau replaced the GRS code by a subcode of low codimension [7]; Wieschebrink [33] included some random columns in a generator matrix of a GRS code; this approach was further enhanced in [30,31] by additionally “mixing” the random columns with the original ones via specific linear transformations; finally, in [2] Baldi, Bianchi, Chiaraluce, Rosenthal, and Schipani proposed to mask the structure of a GRS code by right multiplying it by a “partially weight-preserving” matrix. All these proposals have been partially or fully broken using attacks derived from a square code distinguisher [9,11,15,34].

*Twisted Reed-Solomon codes* (TRS) are evaluation codes in the Hamming metric. These codes were introduced in [5], adapting to the Hamming setting Sheekey's construction of *twisted Gabidulin codes* [25] in rank metric. Unlike Reed-Solomon codes, TRS codes are not always maximum distance separable (MDS) codes, but this family of codes contains MDS codes that are not equivalent to generalized Reed-Solomon (GRS) codes. Recently, in [3], TRS codes have been proposed as an alternative to Goppa codes for McEliece cryptosystem, and example parameters are given that provide shorter keys compared to the original McEliece cryptosystem for the same security level. The authors also singled out a subfamily of twisted Reed-Solomon codes, which they “provably” claimed to be resistant against several known structural attacks on the McEliece cryptosystem

based on RS-like codes: Sidenlikov–Shestakov [27], Wieschebrink [32,34], Schur square-distinguishing [9].

More recently, Lavauzelle and Renner [20] gave an efficient key-recovery attack on the TRS variant proposed in [3] based on identifying some specific structure of the *subfield subcode*. Lavauzelle and Renner assumed the security claims of [3] on the resistance of TRS codes to Schur squares to be true and identified another weakness that was very specific to the chosen public keys (*i.e.* the underlying TRS code is MDS) coming from some tower of extensions of finite fields. It is worth mentioning that their attack restricts to a limited subfamily of TRS codes.

In this paper, *our contributions* are threefold. First, we show the claims of [3] on the resistance of TRS codes to Schur square distinguishing are wrong. This holds even for their generalized version : TGRS codes. Hence we are able to prove that such codes are distinguishable from random as soon as their number  $\ell$  of twists is in  $O(1)$ . The latter assumption is reasonable since the decoding of such codes is exponential in  $\ell$ . In short, we prove that any TGRS codes that could be proposed for the McEliece scheme are actually distinguishable from random.

Second, for the case of TGRS codes with a single twist ( $\ell = 1$ ) we show how to derive a polynomial time attack from the Schur square distinguisher. This attack is in the very same flavour as the attack of [9,15] on BBCRS scheme [2] and runs in  $O(q^3n^4)$  operations in  $\mathbb{F}_q$ . Note that the family of codes for which such an attack applies is much larger than the family broken by [20], which was restricted to a restricted family of TRS codes.

Third, the attacks in [9,15] involved a heuristic argument that was claimed to hold “with a high probability”. In the present article, we provide a detailed analysis of the success probability of the algorithm, providing a proven attack of the scheme. Note that we are able to estimate this success probability for a range of parameters that is **strictly included** in the range of parameters we can actually attack. In short : this attack works on a broad range of parameters and we can prove its success without involving any heuristic arguments in some subrange of parameters.

## Outline of the article

Section 2 provides the basic notation and the necessary prerequisites on GRS codes, TGRS codes, Schur products, and McEliece encryption scheme. In Section 3, we introduce a new class of codes called *quasi-GRS codes* which strictly includes TGRS codes, which turns out to be the class of codes we succeed distinguishing from random by computing the Schur square of some of their shortening.

In Section 4, we show how to distinguish quasi-GRS codes from random. Section 5 is dedicated to the presentation of an attack on the McEliece instantiated with TGRS codes with a single twist and a *SageMath* implementation of the attack together with timing results are given in Section 6. Finally, Section 7 is dedicated to a probability analysis yielding the proof of a crucial theorem for the attack.

## 2 Preliminaries

Let  $q$  be a prime power,  $\mathbb{F}_q$  be a finite field of order  $q$ , and  $\mathbb{F}_q^n$  the  $\mathbb{F}_q$  vector space of dimension  $n$ . In this article, vectors are represented by lowercase bold letters:  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  and matrices by uppercase bold letters  $\mathbf{G}, \mathbf{H}$ . Given a positive integer  $n$ , we denote  $[n]$  to be the set  $[n] \stackrel{\text{def}}{=} \{1, \dots, n\}$ . We denote by  $\mathbb{F}_q[x]_{<k}$  the space of polynomials over  $\mathbb{F}_q$  of degree strictly less than  $k$ . Finally, given elements  $a_1, \dots, a_s$  of a given vector space  $V$ , we denote by  $\langle a_1, \dots, a_s \rangle$  the vector space spanned by these elements. Note that all the considered vector spaces in this paper are over  $\mathbb{F}_q$ , hence the field is not specified when mentioning dimension or vector span.

Sometimes, since we work with *Twisted Generalized Reed Solomon codes*, it would be useful to remove one monomial  $x^h$  (for a certain  $0 \leq h \leq k-1$ ) from the previously defined subspace  $\mathbb{F}_q[x]_{<k}$ . In that case, we denote:

$$\text{Mon}_{<k, \hat{h}} \stackrel{\text{def}}{=} \langle x^0, \dots, \widehat{x^h}, \dots, x^{k-1} \rangle, \quad (1)$$

where the hat notation means that the monomial  $x^h$  is removed. In a more general situation, when we remove more than one monomial, for instance  $x^{h_1}, \dots, x^{h_\ell}$  for  $\ell \geq 1$ , we denote,

$$\text{Mon}_{<k, \hat{\mathbf{h}}} \stackrel{\text{def}}{=} \langle x^i : i \in [n] \setminus \{h_1, \dots, h_\ell\} \rangle \quad \text{where} \quad \mathbf{h} \stackrel{\text{def}}{=} (h_1, \dots, h_\ell). \quad (2)$$

Finally, the present article crucially uses the notion of *shortening* of codes:

**Definition 1.** *Given a code  $\mathcal{C} \subseteq \mathbb{F}_q^n$  and a subset  $I = \{i_1, \dots, i_{|I|}\} \subseteq [n]$ , the shortening of  $\mathcal{C}$  at  $I$ , denoted as  $\mathcal{C}_I$ , is defined as*

$$\mathcal{C}_I \stackrel{\text{def}}{=} \{(x_i)_{i \in [n] \setminus I} : \mathbf{x} = (x_1, \dots, x_n) \in \mathcal{C} \text{ such that } \forall i \in I, x_i = 0\}.$$

*Remark 1.* In short, the shortened code is the subcode of vectors whose entries indexed by  $I$  are zero and whose prescribed zero entries are removed. Then at some places in the paper – we will mention it when needed – we will not remove the prescribed zero entries from the shortening and hence take as a definition:

$$\{\mathbf{x} \in \mathcal{C} : \forall i \in I, x_i = 0\}.$$

## 2.1 Schur products of linear spaces and evaluation codes

One of the most important tools in the cryptanalysis of the McEliece encryption scheme and its variants is the Schur product of codes. For this reason, we introduce the following definitions.

**Definition 2 (Componentwise product).** Given  $\mathbf{a} = (a_1, \dots, a_n)$  and  $\mathbf{b} = (b_1, \dots, b_n)$  in  $\mathbb{F}_q^n$ , we denote by  $\mathbf{a} \star \mathbf{b}$  the componentwise or Schur product as

$$\mathbf{a} \star \mathbf{b} \stackrel{\text{def}}{=} (a_1 b_1, \dots, a_n b_n).$$

**Definition 3 (Schur product of linear codes and square code).** The Schur product of two linear codes  $\mathcal{A}, \mathcal{B} \subset \mathbb{F}_q^n$  is defined as,

$$\mathcal{A} \star \mathcal{B} \stackrel{\text{def}}{=} \{\mathbf{a} \star \mathbf{b} : \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}$$

When  $\mathcal{A} = \mathcal{B}$  then  $\mathcal{A} \star \mathcal{A}$  is called square of  $\mathcal{A}$  and is denoted by  $\mathcal{A}^2$ .

We can easily derive an upper bound on the dimension of the Schur product of two codes and of the square product code : if we consider a basis of  $\mathcal{A}$  and  $\mathcal{B}$ , the product space is generated by the componentwise products of their elements.

**Proposition 1.** If  $\mathcal{A}$  and  $\mathcal{B}$  two linear codes of length  $n$ . Then,

1.  $\dim(\mathcal{A} \star \mathcal{B}) \leq \dim(\mathcal{A}) \dim(\mathcal{B})$ ,
2.  $\dim(\mathcal{A}^2) \leq \binom{\dim(\mathcal{A})+1}{2}$

In particular, for a random code of dimension  $k$  and length  $n$ , it can be shown that the dimension of the square is  $\min \left\{ \frac{k(k+1)}{2}, n \right\}$  with high probability. Such a statement is proved in [8, Thm. 2.3] for binary codes and in [16, Thm. 2] for  $q$ -ary ones.

**Proposition 2.** Let  $k, n \geq 0$  such that  $\binom{k+1}{2} < n$  and  $\mathcal{A}$  be a random  $[n, k]$  code. We have

$$\text{Prob} \left[ \dim \mathcal{A}^2 < \binom{k+1}{2} \right] = o(1).$$

## 2.2 Generalized Reed-Solomon codes and the square code construction

Given a vector  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{F}_q^n$  for distinct  $\alpha_i$ 's and  $\mathbf{v} = (v_1, \dots, v_n)$  a nonzero vector in  $\mathbb{F}_q^n$ , we consider the following *evaluation map* associated to  $\boldsymbol{\alpha}$  and  $\mathbf{v}$ ,

$$\text{ev}_{\boldsymbol{\alpha}, \mathbf{v}} : \begin{cases} \mathbb{F}_q[x] \longrightarrow & \mathbb{F}_q^n \\ f(x) \longmapsto & (v_1 f(\alpha_1), \dots, v_n f(\alpha_n)). \end{cases} \quad (3)$$

**Definition 4 (Generalized Reed-Solomon Code).** Let  $k, n$  be positive integers,  $k < n \leq q$ . The  $[n, k]_q$  generalized Reed-Solomon (GRS) code associated with  $\alpha, \mathbf{v}$  is defined as

$$\mathbf{GRS}_k(\alpha, \mathbf{v}) \stackrel{\text{def}}{=} \{(\text{ev}_{\alpha, \mathbf{v}}(f) : f \in \mathbb{F}_q[x]_{<k})\}.$$

If  $\mathbf{v} = (1, \dots, 1)$ , then the code is a Reed-Solomon code denoted as  $\mathbf{RS}_k(\alpha)$ .

GRS codes lie at the core of algebraic coding theory. Many other algebraic constructions of codes, such as Alternant codes, BCH codes, Goppa codes, Srivastava codes, and so on, are derived from GRS ones. GRS codes are famous because of their optimal parameters: they are known to be *Maximum distance Separable* (MDS), *i.e.* they have the best possible minimum distance with respect to their length and dimension. In addition, such codes benefit from efficient decoding algorithms up to half their minimum distance [24, Chapter 6] and even beyond using list decoding [17, 29].

The Schur product (see Definition 3) plays a central role in the cryptanalysis of the McEliece cryptosystem and its variants. It indeed permits to distinguish algebraically *structured* codes such as GRS ones from random codes.

**Proposition 3.**  $\mathbf{GRS}_k(\alpha, \mathbf{v}) \star \mathbf{GRS}_\ell(\alpha, \mathbf{v}) = \mathbf{GRS}_{k+\ell-1}(\alpha, \mathbf{v} \star \mathbf{v})$ .

*Proof.* See [9]. □

Given a GRS code of dimension  $k$ , Proposition 3 entails that the dimension of its square is  $2k-1$  while Proposition 2 asserts that it would be  $\min\{n, \frac{k(k+1)}{2}\}$  for a random code. Thus, GRS codes of dimension  $k < \frac{n}{2}$  can be easily distinguished from random ones by computing their Schur square. For GRS codes of dimension  $\geq \frac{n}{2}$ , they can be distinguished from random by computing the square of their duals which is itself a GRS code.

Sometimes, the sole use of the Schur product is insufficient to provide a distinguisher, it is then useful to consider the square of a shortening of the code. Shortening then squaring turns out to be very efficient to distinguish codes “close” to GRS codes, as shown in [9, 10, 11, 12, 13, 14, 15]. For this reason, we introduce the following Lemma that will be useful later.

**Lemma 1.** *Shortening of an  $[n, k]$  GRS code at  $a \leq k$  positions gives an  $[n-a, k-a]$  GRS code.*

### 2.3 Twisted generalized Reed–Solomon codes

*Twisted Reed–Solomon codes* (TRS) and *Twisted Generalised Reed–Solomon* (TGRS) codes are slight variants of GRS codes. First inspired by a rank–metric counterpart [25], they have been first introduced in [4,5] as a possible alternative to GRS codes containing some MDS codes. Next, they have been proposed for cryptographic applications in [3] with the argument that their structure was better hidden with respect to classical attacks such as Sidelnikov Shestakov attack or Schur square attack. They are defined as follows.

**Definition 5 (Twisted generalized Reed–Solomon Codes).** *For positive integers  $n, k, \ell$  with  $\ell \leq k \leq n \leq q$ , suppose that  $\mathbf{h} = (h_1, \dots, h_\ell) \in \{0, \dots, k-1\}^\ell$ ,  $\mathbf{t} = (t_1, \dots, t_\ell) \in \{1, \dots, n-k\}^\ell$  and  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_\ell) \in \mathbb{F}_q^\ell$ . Then*

$$\mathcal{P}_{\mathbf{t}, \mathbf{h}, \boldsymbol{\eta}}^{n, k} \stackrel{\text{def}}{=} \left\{ f = \sum_{i=0}^{k-1} f_i x^i + \sum_{j=1}^{\ell} \eta_j f_{h_j} x^{k-1+t_j} : f_i \in \mathbb{F}_q \right\} \quad (4)$$

is a  $k$ -dimensional  $\mathbb{F}_q$ -subspace of  $\mathbb{F}_q[x]$ . Furthermore, let  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{F}_q^n$  where  $\alpha_i$ , for  $i = 1, \dots, n$  are distinct and  $\mathbf{v} = (v_1, \dots, v_n) \in (\mathbb{F}_q^\times)^n$ . Then the corresponding linear code is defined as

$$\mathcal{C}_{\boldsymbol{\alpha}, \mathbf{v}, \mathbf{t}, \mathbf{h}, \boldsymbol{\eta}}^{n, k} \stackrel{\text{def}}{=} \text{ev}_{\boldsymbol{\alpha}, \mathbf{v}}(\mathcal{P}_{\mathbf{t}, \mathbf{h}, \boldsymbol{\eta}}^{n, k}) \subset \mathbb{F}_q^n.$$

is called  $(\boldsymbol{\alpha}, \mathbf{v}, \mathbf{t}, \mathbf{h}, \boldsymbol{\eta})$ -twisted generalized Reed–Solomon (TGRS) code. For brevity, we will simply use  $\mathcal{P}$  and  $\mathcal{C}$  to denote the space of twisted polynomials  $\mathcal{P}_{\mathbf{t}, \mathbf{h}, \boldsymbol{\eta}}^{n, k}$  and the TGRS code  $\mathcal{C}_{\boldsymbol{\alpha}, \mathbf{v}, \mathbf{t}, \mathbf{h}, \boldsymbol{\eta}}^{n, k}$ .

The integer  $\ell$  is referred to as the *number of twists*, and we call the vectors  $\mathbf{t}, \mathbf{h}$  and  $\boldsymbol{\eta}$  the *twist vector*, *hook vector*, and *coefficient vector*, respectively.

In [3], the authors also proposed a *brute force* technique to decode Twisted Reed–Solomon codes, which can be easily generalized to TGRS. Given a received word  $\mathbf{y} = \mathbf{c} + \mathbf{e} \in \mathbb{F}_q^n$ , where  $\mathbf{c} \in \mathcal{C}$ , the main idea consists in guessing  $\ell$  elements  $g_1, \dots, g_\ell \in \mathbb{F}_q$  and then decoding  $\mathbf{y} - \text{ev}_{\boldsymbol{\alpha}, \mathbf{v}} \sum_{j=1}^{\ell} g_j \eta_j x^{k-1+t_j}$  as a received word of a  $\mathbf{GRS}_k(\boldsymbol{\alpha})$ . This strategy succeeds if  $g_j = f_{h_j}$  for all  $1 \leq j \leq \ell$ . The complexity of such a decoder is  $q^\ell$  times the complexity of a GRS decoder, and the decoding radius is the same as the GRS one.

In [5], the authors proposed a decoding strategy for TRS codes based on a *key equation* which can eventually achieve a better complexity than the brute force. In particular, they introduced a *partial unique decoder*, which means that there are some error patterns that cannot be corrected. They also provide heuristic estimations of the *failure probability* of the decoder and of the decoding radius,

which can be smaller than half of the minimum distance of the code. However, even if this algorithm achieves a better complexity than brute force, it is still exponential in the number of twists.

## 2.4 Reasoning on the level of polynomial spaces

All the codes involved are derived from polynomial evaluations. Note that, given the evaluation map  $\text{ev}_{\alpha, \mathbf{v}}$ , then by interpolation, any element  $\mathbf{c} \in \mathbb{F}_q^n$  can be seen as  $\mathbf{c} = \text{ev}_{\alpha, \mathbf{v}}(f)$  for a unique  $f \in \mathbb{F}_q[x]_{<n}$ . Consequently, we will often reason either at the level of codewords, or at the level of polynomials. The correspondence is as follows: for any code  $\mathcal{C} \subseteq \mathbb{F}_q^n$ , there exists a unique subspace  $\mathcal{P}_{\mathcal{C}} \subseteq \mathbb{F}_q[x]_{<n}$  such that

$$\mathcal{C} = \text{ev}_{\alpha, \mathbf{v}}(\mathcal{P}_{\mathcal{C}}). \quad (5)$$

This correspondence will allow us to move between codewords and polynomials as needed. Similarly to codes, we define products of polynomial spaces as follows. Given,  $\mathcal{P}, \mathcal{R} \subseteq \mathbb{F}_q[x]$  we define

$$\mathcal{P}\mathcal{R} \stackrel{\text{def}}{=} \langle fg : f \in \mathcal{P}, g \in \mathcal{R} \rangle.$$

When  $\mathcal{P} = \mathcal{R}$  we call this the *square* of  $\mathcal{P}$  and denote it  $\mathcal{P}^2$ . Next, given a polynomial  $f \in \mathbb{F}_q[X]$  we use  $f\mathcal{P}$  for  $\langle f \rangle \mathcal{P}$ .

As in the code setting (see Proposition 1), we have natural upper bounds.

**Proposition 4.** *Let  $\mathcal{P}$  and  $\mathcal{R}$  be two subspaces of  $\mathbb{F}_q[x]$ , then*

$$\dim \mathcal{P}\mathcal{R} \leq \dim \mathcal{P} \cdot \dim \mathcal{R} \quad \text{and} \quad \dim \mathcal{P}^2 \leq \binom{\dim \mathcal{P} + 1}{2}.$$

When considering space of polynomials of bounded degree, as in the case of GRS codes, we also have an explicit description of the product.

**Proposition 5.**  $(\mathbb{F}_q[x]_{<k})(\mathbb{F}_q[x]_{<\ell}) = \mathbb{F}_q[x]_{<k+\ell-1}$ .

**2.4.1 Relation with Schur squares.** The previously introduced map

$$\text{ev}_{\alpha, \mathbf{v}} : \begin{cases} \mathbb{F}_q[x] & \longrightarrow & \mathbb{F}_q^n \\ f & \longmapsto & (f(\alpha_1), \dots, f(\alpha_n)) \end{cases}$$

is a surjective ring morphism when  $\mathbb{F}_q^n$  is equipped with the Schur product  $\star$ . Consequently, from the setup in (5), we obtain a surjective map  $\mathcal{P}_{\mathcal{C}}^2 \rightarrow \mathcal{C}^2$  which yields:

**Proposition 6.** *Let  $\mathcal{C} = \text{ev}_{\alpha, \mathbf{v}}(\mathcal{P}_{\mathcal{C}})$  for  $\mathcal{P}_{\mathcal{C}} \subseteq \mathbb{F}_q[x]_{<n}$ . Then*

$$\dim \mathcal{C}^2 \leq \dim \mathcal{P}_{\mathcal{C}}^2.$$



**2.4.2 Relations with shortenings.** Given a code  $\mathcal{C} = \text{ev}_{\alpha,v}(\mathcal{P})$  for some  $\mathcal{P} \subseteq \mathbb{F}_q[x]$  and  $I \subseteq [n]$ , the shortening  $\mathcal{C}_I$  of  $\mathcal{C}$  at  $I$  corresponds to the evaluation of  $\mathcal{P} \cap (p_I)$  where  $(p_I)$  is the ideal spanned by

$$p_I(x) \stackrel{\text{def}}{=} \prod_{i \in I} (x - \alpha_i).$$

In the sequel, we denote

$$\mathcal{P}_{p_I} \stackrel{\text{def}}{=} \mathcal{P} \cap (p_I).$$

## 2.5 The McEliece cryptosystem and its variants

McEliece code-based original cryptosystem was introduced in 1978 by McEliece [21] and it used binary Goppa codes. However, it corresponds to very general framework that can be instantiated with many possible codes.

Consider a family of codes  $\mathcal{F}$  parameterized by a set  $\mathcal{S}$ . Hence, we are given a map

$$\mathcal{C} : \begin{cases} \mathcal{S} \rightarrow \mathcal{F} \\ s \mapsto \mathcal{C}(s). \end{cases}$$

The above map is the trapdoor:  $\mathcal{C}(s)$  should be easy to compute from the knowledge of  $s$  but given  $\mathcal{C} \in \mathcal{F}$ , recovering  $s \in \mathcal{S}$  such that  $\mathcal{C} = \mathcal{C}(s)$  should be hard. Moreover, suppose that for any  $s \in \mathcal{S}$  we are given a decoder  $D(s)$  that corrects up to  $t$  errors for the code  $\mathcal{C}(s)$ . Here again, decoding should not be possible without the knowledge of  $s$ . Then, McEliece encryption scheme can be described as follows:

*Key generation.* Draw  $s \in \mathcal{S}$  at random. The secret key is  $s$ .

The public key is a pair  $(\mathbf{G}_{pub}, t)$ , where  $\mathbf{G}_{pub}$  is a generator matrix of  $\mathcal{C}(s)$  and  $t$  is the number of errors that the algorithm  $D(s)$  can decode.

*Encryption.* To encrypt a plaintext  $\mathbf{m} \in \mathbb{F}_q^k$ , choose a random  $\mathbf{e} \in \mathbb{F}_q^n$  with Hamming weight  $wt_H(\mathbf{e}) = t$ . The ciphertext is  $\mathbf{c} \stackrel{\text{def}}{=} \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}$ .

*Decryption.* Apply the decoder  $D(s)$  to the ciphertext  $\mathbf{c}$  to recover  $\mathbf{m}$ .

**Examples of instantiations.** McEliece's original proposal [21] is instantiated with classical Goppa codes, with the secret being the pair  $(L, g)$  where  $L$  is the so-called *evaluation sequence* and  $g$  is the *Goppa Polynomial*. Later on, Niederreiter [23] proposed to instantiate it with GRS codes, where the secret is  $(\mathbf{x}, \mathbf{y})$ . This proposal was subsequently broken in [27]. Since then, several

alternative instantiations have been suggested (the following list is far from being exhaustive): Reed-Muller codes [26], Algebraic Geometry codes and their subfield subcodes [18], subcodes of GRS codes [7], MDPC codes [22], among others. Several other proposals rely on slightly modified GRS codes; see, for instance, [2,19,30,33]. However, many of the instantiations based on codes “close” to GRS codes have been vulnerable to attacks, with many such attacks involving the Schur product [9,10,11,12,13,14,34]

Recently, Beelen, Bossert, Puchinger, and Rosenkilde proposed TRS codes in [3] claiming that the corresponding cryptosystem is resistant to the well-known attacks [9,27,32].

## 2.6 Cryptanalysis of the McEliece system based on GRS codes and their variants

The TRS variant of the McEliece cryptosystem was already broken in a prior work by Lavauzelle and Renner [20], which examined a setup different from ours. More precisely,

1. they considered a prime power  $q_0$ , the integers  $k < n \leq q_0 - 1$  with  $2\sqrt{n} + 6 < k \leq n/2 - 2$  and a twist  $\ell$  such that,

$$\frac{n+1}{k-\sqrt{n}} < \ell + 2 < \min\{k+3, 2n/k, \sqrt{n}-2\}.$$

Further, they set  $q_i \stackrel{\text{def}}{=} q_{i-1}^2 = q_0^{2^i}$  for  $i = 1, \dots, \ell$ , such that  $\mathbb{F}_{q_0} \subset \mathbb{F}_{q_1} \subset \dots \subset \mathbb{F}_{q_\ell} = \mathbb{F}_q$  is a chain of subfields. And finally they take  $t_i = (i+1)(r-2) - k + 2$  and  $h_i = r - 1 + i$  for  $i = 1, \dots, \ell$ , where  $r \stackrel{\text{def}}{=} \lceil \frac{n+1}{\ell+2} \rceil + 2$ .

Note that this choice of parameters guarantees the underlying TRS to be MDS [3].

2. They assumed that the integers  $q_0, n, k, \ell$  and the hook and twist vector  $\mathbf{h}, \mathbf{t}$  are *public parameters*.

Lavauzelle and Renner’s approach **assumes the validity of claims in [3] regarding the indistinguishability of such codes** with respect to the Schur square even after shortening. So they use a strategy based on the **recovery of the subfield subcode** to attack the system. The novelty of their approach lies in the fact that, in this case, they use the subfield subcode structure of the TRS to attack the system, whereas usually the subfield subcode operation is used to hide the structure of the code used in encryption to improve the security. In particular, thanks to the previous choice of parameters, they describe the

structure of the subfield subcode (in  $\mathbb{F}_{q_0}$ ) as a subspace of low codimension of a classical RS code, and they exploit this code to recover the hidden TRS code.

In contrast, our work considers a **general TGRS family** without specific parameter assumptions, and we demonstrate that these codes, like GRS codes, can indeed be distinguished from random codes. Then, we extend the distinguisher-based attack of [9] to such a code family.

### 3 $\ell$ -quasi-GRS codes

In this section, we introduce a broader class of codes called *quasi-GRS* codes, which contains TGRS codes. Notably, the distinguisher we describe further, along with most of the cryptanalysis techniques presented, applies to quasi-GRS codes. The interest of this class is that it is closed under duality and “most of the times”, closed under shortening.

**Definition 6.** Let  $\alpha \in \mathbb{F}_q^n$  be a sequence of distinct elements and  $\mathbf{v} \in (\mathbb{F}_q^\times)^n$ . An  $\ell$ -quasi-GRS ( $\ell$ -qGRS) code is defined as a code  $\mathcal{C}$  such that

$$\mathcal{C} = \mathcal{C}_0 \oplus \mathcal{C}_1,$$

where  $\mathcal{C}_0$  is a subcode of codimension  $\ell$  of  $\mathbf{GRS}_k(\alpha, \mathbf{v})$  and  $\mathcal{C}_1$  has dimension  $\ell$  and satisfies  $\mathcal{C}_1 \cap \mathbf{GRS}_k(\alpha, \mathbf{v}) = 0$ .

**Proposition 7.** A TGRS code with  $\ell$  twists is an  $\ell$ -qGRS code.

*Proof.* Using notation from Definition 5, define

$$\begin{aligned} \mathcal{C}_0 &\stackrel{\text{def}}{=} \text{ev}(\langle x^i : i \in [n] \setminus \{h_1, \dots, h_\ell\} \rangle) = \text{ev}(\text{Mon}_{<k, \hat{\mathbf{h}}}) \\ \text{and } \mathcal{C}_1 &\stackrel{\text{def}}{=} \text{ev}(\langle x^{h_1} + \eta_1 x^{k-1+t_1}, \dots, x^{h_\ell} + \eta_\ell x^{k-1+t_\ell} \rangle). \end{aligned}$$

This yields the result.  $\square$

**Proposition 8.** The dual of an  $\ell$ -qGRS code is an  $\ell$ -qGRS code.

*Proof.* Let  $\mathcal{C} = \mathcal{C}_0 \oplus \mathcal{C}_1$  be an  $\ell$ -qGRS code with  $\mathcal{C}_0 \subseteq \mathbf{GRS}_k(\alpha, \mathbf{v})$  and  $\mathcal{C}_1 \cap \mathbf{GRS}_k(\alpha, \mathbf{v}) = 0$ . Then,  $\mathcal{C}$  has codimension  $\ell$  in  $\mathbf{GRS}_k(\alpha, \mathbf{v}) \oplus \mathcal{C}_1$ . Therefore,  $\mathbf{GRS}_k(\alpha, \mathbf{v})^\perp \cap \mathcal{C}_1^\perp$  has codimension  $\ell$  in  $\mathcal{C}^\perp$ . Denote by  $\mathcal{C}'_0 \stackrel{\text{def}}{=} \mathbf{GRS}_k(\alpha, \mathbf{v})^\perp \cap \mathcal{C}_1^\perp$  and let  $\mathcal{C}'_1$  be a complement subspace of  $\mathcal{C}'_0$  in  $\mathcal{C}^\perp$ . Therefore, it is clear that  $\mathcal{C}'_0 \oplus \mathcal{C}'_1 = \mathcal{C}^\perp$  and  $\dim(\mathcal{C}'_1) = \ell$ . Note that  $\mathcal{C}'_1 \cap \mathbf{GRS}_k(\alpha, \mathbf{v})^\perp = \{0\}$ , as otherwise,  $\mathbf{GRS}_k(\alpha, \mathbf{v})^\perp \cap \mathcal{C}'_1 \subseteq \mathbf{GRS}_k(\alpha, \mathbf{v})^\perp \cap \mathcal{C}^\perp = \mathcal{C}'_0$  will lead to a contradiction. Now, since  $\mathbf{GRS}_k(\alpha, \mathbf{v})^\perp$  is again a GRS code and  $\mathcal{C}'_0$  has codimension  $\ell$  in  $\mathbf{GRS}_k(\alpha, \mathbf{v})^\perp$ , we can conclude that  $\mathcal{C}^\perp$  is an  $\ell$ -qGRS code.  $\square$

**Proposition 9.** Let  $\mathcal{C}$  be an  $\ell$ -qGRS code of dimension  $k$  and  $I \subseteq [n]$  with  $|I| \leq k - \ell$ . If  $\dim \mathcal{C}_I = k - |I|$ , then  $\mathcal{C}_I$  is an  $\ell'$ -qGRS code for some  $\ell' \leq \ell$ .

The proof of Proposition 9 rests on the following lemma.

**Lemma 2.** Let  $\mathcal{D} = \mathbf{GRS}_k(\boldsymbol{\alpha}, \mathbf{v}) \oplus \mathcal{D}_1$  for some code  $\mathcal{D}_1$  of dimension  $\ell$ . Let  $I \subseteq [n]$  such that  $|I| < k$ . Denote  $s \stackrel{\text{def}}{=} |I|$ . Then,  $\dim \mathcal{D}_I = \dim \mathcal{D} - s$  and is of the form  $\mathcal{D}_I = \mathbf{GRS}_{k-s}(\boldsymbol{\alpha}', \mathbf{v}') \oplus \mathcal{D}'_1$  for some  $\boldsymbol{\alpha}', \mathbf{v}' \in \mathbb{F}_q^{n-s}$  and some code  $\mathcal{D}'_1 \subseteq \mathbb{F}_q^{n-s}$  of dimension  $\ell$ .

*Proof.* Without loss of generality, by possibly permuting the entries of codewords, one can assume that  $I = \{1, \dots, s\}$ . Since  $\mathbf{GRS}_k(\boldsymbol{\alpha}, \mathbf{v})$  is MDS, it has a systematic generator matrix:  $\left( \mathbf{I}_k \mid \mathbf{A} \right)$  for some  $\mathbf{A} \in \mathbb{F}_q^{k \times (n-k)}$ . Then, by elimination,  $\mathcal{D}$  has a generator matrix of the form

$$\begin{pmatrix} \mathbf{I}_k & \mathbf{A} \\ (0) & \mathbf{B} \end{pmatrix}$$

for some matrix  $\mathbf{B} \in \mathbb{F}_q^{\ell \times (n-k)}$  in row echelon form. Let  $\mathbf{G}$  be the matrix obtained by removing the first  $s$  rows and columns of the above matrix. The matrix  $\mathbf{G}$  is nothing but a generator matrix of  $\mathcal{D}_I$ . Since  $\mathbf{B}$  is in row echelon form, we have the result on the dimension of  $\mathcal{D}_I$ . Finally, the  $k - s$  first rows of  $\mathbf{G}$  give a generator matrix of the shortened GRS code, which is itself a GRS code. This yields the result on the structure of  $\mathcal{D}_I$ .  $\square$

*Proof of Proposition 9.* By definition,  $\mathcal{C}$  has codimension  $\ell$  in a code

$$\mathcal{D} \stackrel{\text{def}}{=} \mathbf{GRS}_k(\boldsymbol{\alpha}, \mathbf{v}) \oplus \mathcal{C}_1$$

for some code  $\mathcal{C}_1$  of dimension  $\ell$ . By Lemma 2,  $\mathcal{D}_I$  has dimension  $k + \ell - |I|$ . Since by assumption,  $\dim \mathcal{C}_I = \dim \mathcal{C} - |I|$ , then  $\mathcal{C}_I$  has codimension  $\ell$  in  $\mathcal{D}_I$ . By Lemma 2 again,  $\mathcal{D}_I = \mathbf{GRS}_{k-s}(\boldsymbol{\alpha}', \mathbf{v}') \oplus \mathcal{C}'_1$  for some  $\boldsymbol{\alpha}', \mathbf{v}' \in \mathbb{F}_q^{n-s}$  and some code  $\mathcal{C}'_1 \subseteq \mathbb{F}_q^{n-s}$ . Let  $\mathcal{C}'_0 \stackrel{\text{def}}{=} \mathcal{C}_I \cap \mathbf{GRS}_{k-s}(\boldsymbol{\alpha}', \mathbf{v}')$ . The code  $\mathcal{C}'_0$  has codimension  $\ell' \leq \ell$  in  $\mathbf{GRS}_{k-s}(\boldsymbol{\alpha}', \mathbf{v}')$  and since  $\dim \mathcal{C}_I = k - s$ , we deduce that  $\mathcal{C}_I$  is the direct sum of a codimension  $\ell'$  subspace of  $\mathbf{GRS}_{k-s}(\boldsymbol{\alpha}', \mathbf{v}')$  and a code of dimension  $\ell'$ . Hence it is an  $\ell'$ -qGRS code.  $\square$

*Remark 2.* Note that the typical scenario is that the shortening of an  $\ell$ -qGRS code remains an  $\ell$ -qGRS. Cases where  $\ell$  decreases are sporadic. For  $\ell = 1$ , we can observe the following.

**Lemma 3.** Let  $\mathcal{E}$  be a subspace of codimension 1 of  $\mathbb{F}_q[x]_{<k}$  and assume that

$$p(x) = \prod_{\alpha_i, i \in I} (x - \alpha_i) \quad \text{for } I \subsetneq [n] \quad \text{with } |I| = a.$$

Then  $\mathcal{E}_{p(x)}$  has codimension 0 or 1 in  $p(x)\mathbb{F}_q[x]_{<k-a}$ .

*Proof.* By definition,  $\mathcal{E}_{p(x)} \subseteq p(x)\mathbb{F}_q[x]_{<k-a}$  and  $\dim p(x)\mathbb{F}_q[x]_{<k-a} = k - a$ . Thus  $\dim \mathcal{E}_{p(x)} \leq k - a$ . On the other hand, note that

$$\mathcal{E}_{p(x)} = \bigcap_{i=1}^a \ker \begin{pmatrix} \mathcal{E} \rightarrow \mathbb{F}_q \\ f \mapsto f(\alpha_i) \end{pmatrix}. \quad (6)$$

This implies  $\dim(\mathcal{E}_{p(x)}) \geq \dim(\mathcal{E}) - a = k - 1 - a$ , and the result follows.  $\square$

*Remark 3.* It turns out that, in general, the shortening of a TGRS code is not a TGRS one. Therefore, Proposition 9 highlights an important interest of the notion of qGRS codes compared to TGRS codes. This will be particularly useful in the sequel, as the shortening operation plays a crucial role in the distinguishers and attacks to be discussed.

## 4 TGRS codes and a distinguisher based on Schur product

In this section, we show how shortening and squaring can be exploited to distinguish qGRS codes from random ones for suitable parameters. From Proposition 9, this straightforwardly leads a distinguisher on TGRS codes for suitable parameters. We start by discussing the general setting, *i.e.* an arbitrary number of  $\ell$  twists. We then focus specifically on the case where  $\ell = 1$ , for which we will derive an attack in the subsequent section.

### 4.1 A distinguisher for general $\ell$

**4.1.1 Polynomial setting.** Let  $(\alpha, \mathbf{v})$  such that  $\alpha$  is a sequence of distinct elements and  $\mathbf{v}$  is a sequence of nonzero elements both in  $\mathbb{F}_q$ . We have the corresponding evaluation map  $\text{ev}_{\alpha, \mathbf{v}} : \mathbb{F}_q[x] \rightarrow \mathbb{F}_q^n$ .

In what follows, we consider a qGRS code  $\mathcal{C}$  (see Definition 6) such that

$$\mathcal{C} = \mathcal{C}_0 \oplus \mathcal{C}_1$$

where  $\mathcal{C}_0$  has codimension  $\ell$  in  $\mathbf{GRS}_k(\alpha, \mathbf{v})$  (thus  $\mathcal{C}_0$  is a subcode of  $\mathbf{GRS}_k(\alpha, \mathbf{v})$ ),  $\mathcal{C}_1$  has dimension  $\ell$  and  $\mathcal{C}_1 \cap \mathbf{GRS}_k(\alpha, \mathbf{v}) = 0$ . In particular,  $\mathcal{C}$  has dimension  $k$ .

Via the correspondence discussed in § 2.4,  $\mathcal{C} = \text{ev}_{\alpha, v}(\mathcal{P})$  such that  $\mathcal{P} \subseteq \mathbb{F}_q[x]_{<n}$  and

$$\mathcal{P} = \mathcal{P}_0 \oplus \mathcal{P}_1,$$

where  $\mathcal{P}_0$  has codimension  $\ell$  in  $\mathbb{F}_q[x]_{<k}$  and  $\mathcal{P}_1 \cap \mathbb{F}_q[x]_{<k} = 0$ . Denote  $v_1, \dots, v_\ell$  a basis of  $\mathcal{P}_1$  whose elements have strictly increasing degrees and denote by

$$d_\ell \stackrel{\text{def}}{=} \deg(v_\ell) \tag{7}$$

which is the maximal possible degree for an element of  $\mathcal{P}$ .

**4.1.2 Inequalities for distinguishing.** For distinguishing  $\mathcal{C}$  from a random code, we examine the dimension of  $\mathcal{C}^2$  or  $\mathcal{C}_I^2$  for some  $I \subseteq [n]$ . For  $\mathcal{C}$  to be distinguishable from random codes, we need

$$\dim \mathcal{C}_I^2 \leq \min \left\{ n - |I|, \binom{\dim \mathcal{C}_I + 1}{2} \right\}.$$

*Remark 4.* For  $|I| < k$ , the typical situation is that  $\dim \mathcal{C}_I = k - |I|$  and we will assume this situation throughout this section. Note that if this condition does not hold for almost any such  $I$ , such an observation would yield a distinguisher on the code.

**Proposition 10.** *Let  $\mathcal{C}$  be an  $\ell$ -qGRS code of dimension  $k$  obtained by evaluation of  $\mathcal{P} \subseteq \mathbb{F}_q[x]$  with maximal degree  $d_\ell$ . Then,*

$$\dim \mathcal{C}^2 \leq \min \left\{ (\ell + 2)k - 1 - \frac{\ell(\ell - 1)}{2}, k + d_\ell + \frac{\ell(\ell + 1)}{2}, 2d_\ell + 1 \right\}.$$

*Proof.* Consider the notation introduced in § 4.1.1. Proposition 6 asserts that

$$\dim \mathcal{C}^2 \leq \dim \mathcal{P}^2.$$

Thus we will estimate  $\dim \mathcal{P}^2$ . Note that

$$\mathcal{P}^2 = \mathcal{P}_0^2 + \mathcal{P}_0\mathcal{P}_1 + \mathcal{P}_1^2. \tag{8}$$

Since  $\mathcal{P}_0 \subseteq \mathbb{F}_q[x]_{<k}$ , Proposition 5 asserts that  $\dim \mathcal{P}_0^2 \leq 2k - 1$ . Then, by Proposition 4,

$$\dim \mathcal{P}_0\mathcal{P}_1 \leq \ell(k - \ell) \quad \text{and} \quad \dim \mathcal{P}_1^2 \leq \frac{\ell(\ell + 1)}{2}.$$

Including the last estimates in (8) yields:

$$\dim \mathcal{P}^2 \leq 2k - 1 + \ell(k - \ell) + \frac{\ell(\ell + 1)}{2} = (\ell + 2)k - 1 - \frac{\ell(\ell - 1)}{2},$$

from which we can deduce the first upper bound.

Next observe that, from the definition of  $d_\ell$ ,  $\mathcal{P} \subseteq \mathbb{F}_q[x]_{<d_\ell+1}$ . Therefore, rewriting (8) as

$$\mathcal{P}^2 \subseteq \mathcal{P}_0 \mathbb{F}_q[x]_{<d_\ell+1} + \mathcal{P}_1^2,$$

then, since  $\mathcal{P}_0 \subseteq \mathbb{F}_q[x]_{<k}$ , from Proposition 5 we get,

$$\mathcal{P}^2 \subseteq \mathbb{F}_q[x]_{<k+d_\ell} + \mathcal{P}_1^2$$

which yields the second upper bound:

$$\dim \mathcal{C}^2 \leq k + d_\ell + \frac{\ell(\ell + 1)}{2}.$$

Finally, from the inclusion  $\mathcal{P} \subseteq \mathbb{F}_q[x]_{<d_\ell+1}$  together with Proposition 5, we can deduce that  $\mathcal{P}^2 \subseteq \mathbb{F}_q[x]_{<2d_\ell+1}$  yielding the third inequality:

$$\dim \mathcal{C}^2 \leq 2d_\ell + 1.$$

□

*Remark 5.* The third upper bound  $\dim \mathcal{C}^2 \leq 2d_\ell+1$  corresponds to the dimension of the square of the code referred to as the *outer code* in [3].

**Theorem 1.** *Let  $\mathcal{C}$  be an  $\ell$ -qGRS of dimension  $k$  and  $I \subseteq [n]$  be a set such that  $|I| < k$  such that  $\dim \mathcal{C}_I = k - |I|$ , then,*

$$\dim \mathcal{C}_I^2 \leq \min \left\{ (\ell + 2)(k - |I|) - 1 - \frac{\ell(\ell - 1)}{2}, \right. \\ \left. k + d_\ell - 2|I| + \frac{\ell(\ell + 1)}{2}, 2(d_\ell - |I|) + 1 \right\}.$$

*Proof.* According to Proposition 9, the shortening of  $\mathcal{C}$  is still an  $\ell'$ -qGRS code for  $\ell' \leq \ell$ . Thus, the result is a direct consequence of Proposition 10 where the following changes of variables, are applied:

$$\ell \leftarrow \ell' \quad k \leftarrow k - |I| \quad d_\ell \leftarrow d_\ell - |I|.$$

Two things may require clarifications. First, according to this change of variables the result should involve  $\ell'$  and not  $\ell$ . However, since the upper bounds of Proposition 10 are increasing functions in  $\ell$ , if they hold for  $\ell' \leq \ell$  they still hold for  $\ell$ . Thus  $\ell'$  can be replaced by  $\ell$  in the final bounds.

Second, the change of variable  $d_\ell \leftarrow d_\ell - |I|$  requires further details. From § 2.4.2, the shortened code  $\mathcal{C}_I$  corresponds to the polynomial space  $\mathcal{P}_{p_I}$  where  $p_I(x) \stackrel{\text{def}}{=} \prod_{i \in I} (x - \alpha_i)$ . Since  $\mathcal{P} \subseteq \mathbb{F}_q[x]_{<d_\ell+1}$ , then  $\mathcal{P}_{p_I} \subseteq p_I \mathbb{F}_q[x]_{<d_\ell-|I|+1}$ . Therefore, when shortening by forgetting the common factor  $p_I$ , we effectively deal with polynomials of degree  $\leq d_\ell - |I|$ . This concludes the proof.  $\square$

**Corollary 1.** *Let  $\mathcal{C}$  be a  $(\alpha, \mathbf{v}, \mathbf{t}, \mathbf{h}, \eta)$  TGRS code of dimension  $k$ . Denote by  $t_{max}$  the largest entry of  $\mathbf{t}$ . Let  $I \subseteq [n]$  with  $|I| < k$ . Then*

$$\dim \mathcal{C}_I^2 \leq \min \left\{ (\ell + 2)(k - |I|) - 1 - \frac{\ell(\ell - 1)}{2}, \right. \\ \left. 2k + t_{max} - 1 - 2|I| + \frac{\ell(\ell + 1)}{2}, 2(k + t_{max} - |I|) - 1 \right\}.$$

*Proof.* This is a straightforward consequence of Theorem 1 after observing that  $d_\ell = k + t_{max} - 1$ .  $\square$

**4.1.3 The range of the distinguisher.** There remains to analyze for which parameters qGRS codes can be distinguished from random ones. Since the class of qGRS codes is closed under duality, one can assume that  $k \leq \frac{n}{2}$  to estimate the range of the distinguisher.

Now our objective is to identify for which parameters  $(n, k, \ell)$  there exists  $a < k$  such that for  $I \subseteq [n]$  with  $|I| = a$  we have

$$\dim \mathcal{C}_I^2 < \min \left\{ n - a, \frac{(k - a)(k - a + 1)}{2} \right\}. \quad (9)$$

Theorem 1 yields 3 distinct bounds on  $\dim \mathcal{C}_I^2$ . The first one is that which decreases the fastest in  $a$ , thus, we will only consider this one, which permits a simpler analysis. Later on, when focusing on  $\ell = 1$ , the other bounds will be useful.

Now, we have to prove the existence of  $0 \leq a < k$  such that

$$(\ell + 2)(k - a) - 1 - \frac{\ell(\ell - 1)}{2} < \min \left\{ n - a, \frac{(k - a)(k - a + 1)}{2} \right\}.$$

A calculation proves that the upper bound  $< n - a$  holds for any  $a$  satisfying

$$a > \left( \frac{\ell + 2}{\ell + 1} \right) k - \frac{\ell(\ell - 1)}{2(\ell + 1)} - \frac{n}{\ell + 1}. \quad (10)$$

For the second inequality, we consider a stronger and simpler one:

$$(\ell + 2)(k - a) \leq \frac{(k - a)^2}{2},$$



which yields

$$a \leq k - 2(\ell + 2). \quad (11)$$

Now, putting (10) and (11) together, we can conclude that if there exists an  $a$  where  $a = |I|$  and such that

$$\left(\frac{\ell + 2}{\ell + 1}\right) k - \frac{\ell(\ell - 1)}{2(\ell + 1)} - \frac{n}{\ell + 1} < a \leq k - 2(\ell + 2)$$

then dimension of the shortened square code  $\mathcal{C}_I^2$  satisfies (9) and so we can distinguish these kind of codes from random ones.

**Theorem 2.** *An  $\ell$ -qGRS code of dimension  $k$  is distinguishable from random as soon as:*

$$n - k \geq \frac{3}{2}\ell^2 + \frac{5}{2}\ell + 4.$$

Recall here that we supposed  $k \leq \frac{n}{2}$ , the left-hand side is then larger than  $\frac{n}{2}$  while the right-hand side is in  $O(\ell^2)$ . For TGRS codes proposed for McEliece encryption,  $\ell$  is supposed to be  $O(1)$  since the decoding is exponential in  $\ell$ . So we can conclude that: **TGRS codes proposed for McEliece encryption are distinguishable from random.**

#### 4.2 The case $\ell = 1$

Let us apply the previous results to a single twisted TGRS code, *i.e.* in the case  $\ell = 1$ . Recall that  $\mathcal{C} = \text{ev}(\mathcal{P})$  where  $\mathcal{P} = \text{Mon}_{<k, \hat{h}} + \langle x^h + \eta x^{k-1+t} \rangle$ . The previous analysis provides.

**Lemma 4.** *Let  $\mathcal{C}$  be an  $(\alpha, \mathbf{v}, \mathbf{t}, \mathbf{h}, \eta)$  1-TGRS code of dimension  $k$  with  $\mathbf{t} = (t)$  for some positive integer  $t$ . Let  $I \subseteq [n]$  such that  $|I| < k$  and  $\dim \mathcal{C}_I = k - |I|$ . Then,*

$$\dim \mathcal{C}^2 \leq \begin{cases} 3k - 1 & \text{for } t \geq k, \\ 2k + t & \text{for } t < k. \end{cases}, \quad \dim \mathcal{C}_I^2 \leq \begin{cases} 3(k - |I|) - 1 & \text{for } t \geq k - |I|, \\ 2(k - |I|) + t & \text{for } t < k - |I|. \end{cases}$$

*Proof.* From Corollary 1, applied with  $\ell = 1$  and  $t_{max} = t$ , we get

$$\dim \mathcal{C}_I^2 \leq \min \{3(k - |I|) - 1, 2k + t - 2|I|\}.$$

This yields the proof. □

Next, we can redo the previous analysis in this case while making fewer approximations. We prove that the shortened code  $\mathcal{C}_I$  has a square whose dimension differs from the typical one as soon as

$$\frac{3k - n}{2} \leq |I| \leq k - 5. \quad (12)$$

Thus, such a code is distinguishable from random as soon as

$$n - k \geq 10.$$

Since we supposed  $k \leq \frac{n}{2}$ , any 1-qGRS code of length  $> 20$  is distinguishable from random.

## 5 A key-recovery attack on McEliece scheme with TGRS codes using Schur squares

In this section we present a key recovery attack for the McEliece cryptosystem instantiated with a TGRS code (see Section 2.5). Following the previous notations, let  $\mathcal{C}$  be the public code generated by  $\mathbf{G}_{pub}$ , a generator matrix of a  $q$ -ary single-twisted TGRS code  $\mathcal{C}_{\alpha, v, t, h, \eta}^{n, k} \subset \mathbb{F}_q^n$  of dimension  $k$ .

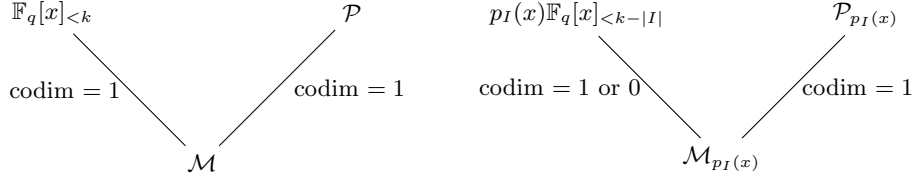
Note that, unlike [20], we do not impose any restriction on the parameters of the secret TGRS code.

### 5.1 Context and notation

Let  $\mathcal{C}$  or equivalently, the  $q$ -ary single-twisted TGRS code  $\mathcal{C}_{\alpha, v, t, h, \eta}^{n, k}$  have dimension  $k \leq \frac{n}{2}$  (otherwise, one can work with the dual code of  $\mathcal{C}$ ).

**Notation 3.** As  $\mathcal{C}$  has a single twist, we let  $t, h$ , and  $\eta$  to be the twist, the hook, and the coefficient of  $\mathcal{C}$ , respectively. Let  $\mathcal{P} = \mathcal{M} \oplus \langle x^h + \eta x^{k-1+t} \rangle$ , where  $\mathcal{M} = \text{Mon}_{< k, \hat{h}}$ . For a subset  $I \subseteq [n]$ , we define  $p_I(x) = \prod_{i \in I} (x - \alpha_i)$ .

Recall that the injectivity of  $\text{ev}_{\alpha, v}$  when restricted to  $\mathbb{F}_q[x]_{< n}$  implies  $\mathcal{C} = \text{ev}_{\alpha, v}(\mathcal{P})$  and the shortening of  $\mathcal{C}$  at  $I$  is  $\mathcal{C}_I = \text{ev}_{\alpha, v}(\mathcal{P}_{p_I(x)})$ . Note that,  $\mathcal{M}$  is a codimension 1 subspace of  $\mathbb{F}_q[x]_{< k}$ . By Lemma 3,  $\mathcal{M}_{p(x)}$  is a subspace of  $p(x)\mathbb{F}_q[x]_{< k-a}$  of codimension at most 1. Here is an inclusion diagram for the involved polynomial spaces, which are isomorphic to involved codes via the map  $\text{ev}_{\alpha, v}$ .



## 5.2 Key-recovery algorithm

We brief the main steps of our key recovery algorithm, which are given in detail in the following sections.

- Step 1 Compute a basis of  $\text{ev}_{\alpha,v}(\mathcal{M})$  or  $\text{ev}_{\alpha,v}(\mathcal{M}_{p_I(x)})$  using a distinguisher-based method (Section 5.2.1).
- Step 2 Use a Sidelnikov-Shestakov-like method given in [9, § 6.3] to recover the secret evaluation vectors and multipliers for  $\mathcal{C}$  (Section 5.2.2).
- Step 3 Compute the other secret parameters which are the twist  $t$ , the hook  $h$ , and the coefficient  $\eta$  (Section 5.2.3).

### 5.2.1 Step 1: recovery of a codimension one subcode of the public (shortened) code.

Following Lemma 4, we know that  $\mathcal{C}$  can be distinguished by considering its Schur square for the case  $3k < n$  and otherwise by taking the Schur square of a shortened code  $\mathcal{C}_I$  with  $(3k - n)/2 < |I| < k - 5$  (according to (12)). Now, depending on  $t$ , we consider the following two cases.

1. If  $2k - 2 + t < n$ , then we compute a basis of the subcode  $\text{ev}_{\alpha,v}(\mathcal{M})$  of  $\mathcal{C}$ .
2. Otherwise, we compute a basis of the subcode  $\text{ev}_{\alpha,v}(\mathcal{M}_{p_I(x)})$  of the shortened code  $\mathcal{C}_I$  for  $I$  such that  $2k - 2 + t - n < |I|$ .

Note that in Sec. 7, we work with polynomial spaces and to apply those results to the corresponding evaluation codes, we need  $\text{ev}_{\alpha,v}$  to be injective. Therefore, the choice of parameters in Case 1 has been made so that  $\text{ev}_{\alpha,v}$  is injective on the product space  $\mathcal{M}\mathcal{P}$ . In Case 2, to make sure the injectivity of  $\text{ev}_{\alpha,v}$  on  $\mathcal{M}_{p_I(x)}\mathcal{P}_{p_I(x)}$ ,  $I$  should be chosen such that  $2(k - |I|) - 2 + t < n - |I|$ , or equivalently,  $2k + t - 2 - n < |I|$ . Injectivity of  $\text{ev}_{\alpha,v}$  is crucial to apply the results on polynomial spaces, for instance, Proposition 11 and Theorem 4 to the corresponding codes.

First, we observe that, in Case 1, for a basis  $\text{ev}_{\alpha,v}(e_1), \dots, \text{ev}_{\alpha,v}(e_k)$  of  $\mathcal{C}$ , if we take three random elements  $\text{ev}_{\alpha,v}(f_1), \text{ev}_{\alpha,v}(f_2), \text{ev}_{\alpha,v}(f_3)$  from its monomial

subcode  $\text{ev}_{\alpha, \mathbf{v}}(\mathcal{M})$ , then the linear space spanned by

$$\{\text{ev}_{\alpha, \mathbf{v}}(f_i) \star \text{ev}_{\alpha, \mathbf{v}}(e_j) : i = 1, 2, 3, 1 \leq j \leq k\}$$

has low dimension, more precisely bounded by  $2k+2$ . This is due to the following statement.

**Proposition 11.** *Let  $f_1, f_2, f_3 \in \mathcal{M}$ , then*

$$\dim(\langle f_1, f_2, f_3 \rangle \mathcal{P}) \leq 2k + 2. \quad (13)$$

*Proof.* The result follows from the fact that  $\mathcal{M} \subseteq \mathbb{F}_q[x]_{<k}$  and  $\mathcal{P} \subseteq \mathcal{M} \oplus \langle c \rangle$  for  $c(x) = x^h + \eta x^{k-1+t}$ . Thus,

$$\langle f_1, f_2, f_3 \rangle \mathcal{P} \subseteq \langle f_1, f_2, f_3 \rangle \mathcal{M} + \langle f_1 c, f_2 c, f_3 c \rangle.$$

Since  $f_1, f_2, f_3 \in \mathcal{M}$  and  $\mathcal{M} \subseteq \mathbb{F}_q[x]_{<k}$ , the first term of the right-hand side is contained in  $\mathbb{F}_q[x]_{<2k-1}$ , which yields the result.  $\square$

It is important to note that Proposition 11 also holds in Case 2 if we replace  $\mathcal{M}$  by  $\mathcal{M}_{p_I(x)}$  and  $\mathcal{P}$  by  $\mathcal{P}_{p_I(x)}$ . More importantly, if we consider  $f_1, f_2, f_3$  randomly from  $\mathcal{P}$ , then due to the higher degree term  $x^{k-1+t}$ , the space  $\langle f_1, f_2, f_3 \rangle \mathcal{P}$  is very likely to have a much larger dimension than  $2k+2$ . This is indeed the case, as we prove the following result in Section 7. Note that in Section 7, we consider the case where  $\mathcal{M}_{p(x)}$  has codimension 1 in  $p(x)\mathbb{F}_q[x]_{<k-a}$ . For the case  $\mathcal{M}_{p(x)} = p(x)\mathbb{F}_q[x]_{<k-a}$ , the probabilistic estimates actually become even more favorable.

**Theorem 4.** *Let  $n, k, t$  be integers such that  $17 \leq k < n$  and  $17 \leq t \leq n - k$ . Suppose  $\mathcal{R} = \mathcal{E}' \oplus \langle c(x) \rangle \subseteq \mathbb{F}_q[x]$ , where  $\mathcal{E}' \underset{\text{codim } 1}{\subseteq} \mathbb{F}_q[x]_{<k}$  and  $c(x)$  has degree  $k - 1 + t$ . If  $f_1, f_2, f_3 \in \mathcal{R}$  and satisfy*

$$\dim(\langle f_1, f_2, f_3 \rangle \mathcal{R}) \leq 2k + 2,$$

*then  $f_1, f_2, f_3 \in \mathcal{E}'$  with probability  $\geq 1 - \frac{1}{q^2}$ . Furthermore, a basis of  $\mathcal{E}'$  is recovered with probability  $\geq (1 - \frac{(k-3)}{q^2})$ .*

*Remark 6.* Let us explain why exactly three functions are chosen in Theorem 4 and in Algorithm 1. If we draw  $s$  elements of the public at random code, the probability that each one lies in some fixed codimension one subcode is  $q^{-s}$ . Hence, in average, we will have  $O(q^s)$  trials to perform before succeeding. Thus,

the larger the  $s$  the larger the complexity of the attack. Next, observe that with  $s = 2$  we cannot distinguish elements of the subcodes since for any code  $\mathcal{D}$

$$\dim\langle f_1, f_2 \rangle * \mathcal{D} \leq 2 \dim \mathcal{D} - 1 < 2 \dim \mathcal{D} + 2.$$

Thus  $s = 3$  is the least number of elements that yields a successful attack. Because of the aforementioned complexity issues there is no interest in considering a higher  $s$ .

Now, we apply Theorem 4 by taking  $\mathcal{R} = \mathcal{P}$  and  $\mathcal{E}' = \mathcal{M}$  (Case 1) or  $\mathcal{R} = \mathcal{P}_{p_I(x)}$  and  $\mathcal{E}' = \mathcal{M}_{p_I(x)}$  (Case 2) to ensure the isomorphism between the polynomial spaces and the corresponding codes via the evaluation map  $\text{ev}_{\alpha, v}$ . Based on the above theorem, which is proved in Section 7, we recover the codimension 1 subcode  $\text{ev}_{\alpha, v}(\mathcal{M})$  of  $\mathcal{C}$  (in Case 1) or  $\text{ev}_{\alpha, v}(\mathcal{M}_{p(x)})$  of  $\mathcal{C}_I$  (in Case 2) with probability at least  $1 - \frac{k-3}{q^2}$  using a distinguisher based method explained in Algorithm 1. This algorithm was introduced in [9, Algorithm 1] in a quite different version: we adapted it to the framework of our attack. Note that the parameter  $k$  in Algorithm 1 refers to the dimension of  $\mathcal{R}$ . Therefore, when we apply the algorithm, it will be for  $k$  in Case 1 and for  $k - |I|$  in Case 2.

The recovery algorithm (Algorithm 1) proceeds as follows.

**In Case 1:** 1. Draw random triples  $(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3) \in \mathcal{C}^3$  until we find one such that

$$\dim(\langle \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3 \rangle \mathcal{C}) \leq 2k + 2.$$

According to Theorem 4 such a triple is likely to be in  $\text{ev}_{\alpha, v}(\mathcal{M})$ . Since  $\mathcal{M}$  has codimension 1 in  $\mathcal{C}$ , the probability of finding such a triple is  $\frac{1}{q^3}$  and hence in  $O(q^3)$  trials, we should find such a triple.

2. We collect other elements of  $\text{ev}_{\alpha, v}(\mathcal{M})$  until we get a basis of this monomial subcode. For this sake, draw  $\mathbf{c} \in \mathcal{C}$  and keep it if

$$\dim(\langle \mathbf{c}_1, \mathbf{c}_2, \mathbf{c} \rangle \mathcal{C}) \leq 2k + 2.$$

Such a  $\mathbf{c}$  is found in  $O(q)$  trials.

**In Case 2:** We do the same while replacing  $\mathcal{C}$  by  $\mathcal{C}_I$  and  $\mathcal{M}$  by  $\mathcal{M}_{p_I}$ . Next, we iterate this process by choosing subsets  $I_1, \dots, I_s \subset [n]$  satisfying  $\bigcap_{i=1}^s I_i = \emptyset$  and  $(3k - n)/2 < |I_i| < k - 5$  for all  $i = 1, \dots, s$ . The last condition being taken to recover the codimension 1 subcodes  $\text{ev}_{\alpha, v}(\mathcal{M}_{p_{I_i}})$  of  $\mathcal{C}_{I_i}$ . We collect shortenings  $\text{ev}_{\alpha, v}(\mathcal{M}_{p_{I_i}})$  and sum them up until the sum has dimension  $k - 1$ . This permits to recover the monomial subcode  $\text{ev}_{\alpha, v}(\mathcal{M})$ . This recovery of a subcode by summing up shortenings is already used and discussed in [14, § IV.F].

---

**Algorithm 1** Recovering  $\mathcal{C}_0$ .

---

**Input :** A basis  $\{\mathbf{c}_1, \dots, \mathbf{c}_k\}$  of  $\mathcal{C}$

**Output :** A basis  $\mathcal{B}$  of  $\mathcal{C}_0$ .

```
1: repeat
2:   for  $1 \leq i \leq 3$  do
3:     Randomly choose  $\mathbf{b}_i$  in  $\mathcal{C}$ 
4:   end for
5:    $\mathcal{D} \leftarrow \langle \{\mathbf{b}_i \star \mathbf{c}_j \mid 1 \leq i \leq 3 \text{ and } 1 \leq j \leq k\} \rangle$ 
6:   until  $\dim(\mathcal{D}) \leq 2k + 2$  and  $\dim(\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \rangle) = 3$ 
7:    $\mathcal{B} \leftarrow \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ 
8:    $s \leftarrow 4$ 
9:   while  $s \leq k - 1$  do
10:    repeat
11:      Randomly choose  $\mathbf{b}_s$  in  $\mathcal{C}$ 
12:       $\mathcal{E} \leftarrow \langle \{\mathbf{b}_i \star \mathbf{c}_j \mid i \in \{1, 2, s\} \text{ and } 1 \leq j \leq k\} \rangle$ 
13:    until  $\dim(\mathcal{E}) \leq 2k + 2$  and  $\dim(\langle \mathcal{B} \cup \{\mathbf{b}_s\} \rangle) = s$ 
14:     $\mathcal{B} \leftarrow \mathcal{B} \cup \{\mathbf{b}_s\}$ 
15:     $s \leftarrow s + 1$ 
16:  end while
17: return  $\mathcal{B}$ ;
```

---

In summary, whatever the parameters  $k, t$ , in the end of this step we have a basis of  $\text{ev}_{\alpha, \mathbf{v}}(\mathcal{M})$ .

**5.2.2 Step 2: recovery of the secret evaluation vector and the multipliers.** Now we have access to the space  $\text{ev}_{\alpha, \mathbf{v}}(\mathcal{M})$ . Since  $\mathcal{M} \subseteq \mathbb{F}_q[x]_{<k}$  and we assumed  $k \leq \frac{n}{2}$ , we see that  $\mathcal{M}^2 \subseteq \mathbb{F}_q[x]_{<2k-1}$  and hence  $\text{ev}_{\alpha, \mathbf{v}}(\mathcal{M}) \subsetneq \mathbb{F}_q^n$ . Moreover,  $\mathcal{M} = \langle 1, x, \dots, \widehat{x^h}, \dots, x^{k-1} \rangle$ . A classical argument from combinatorics permits to prove that if  $h \neq 1$  or  $k - 2$ , then  $\mathcal{M}^2 = \mathbb{F}_q[x]_{<2k-1}$ . In such a situation,  $\text{ev}_{\alpha, \mathbf{v}}(\mathcal{M})^2$  is a GRS code, and then the Sidelnikov–Shestakov attack can be applied to it as it is done in [34].

This permits to recover a candidate for the pair  $(\alpha, \mathbf{v})$ .

*Remark 7.* We did not consider here the pathological cases  $h = 1$  or  $k - 2$ . They should be subject to a separate development that we do not treat in this article.

**5.2.3 Step 3: recovery of the hook  $h$ , the twist  $t$ , and the coefficient  $\eta$  of the secret key.** The previous steps provide  $\alpha$  and  $\mathbf{v}$  which we use to recover the hook  $h$ , the twist  $t$  and the coefficient  $\eta$ .

**Recovering  $h$ .** For  $0 \leq i \leq k-1$ , we check if  $v \star \alpha^i \in \mathcal{C}$  or not. Our assumption on  $\mathbf{G}_{pub}$  implies that there is exactly one  $i$  for which  $v \star \alpha^i \notin \mathcal{C}$ , and that integer would be the hook  $h$ .

**Recovering  $t$ .** Take a codeword  $c \in \mathcal{C} \setminus \text{ev}_{\alpha,v}(\mathcal{M})$ . By interpolation we get  $f \in \mathbb{F}_q[x]_{<n}$  such that  $c = \text{ev}_{\alpha,v}(f)$  and  $\deg(f) = k-1+t$  for some positive  $t$ . Which yields  $t$ .

**Recovering  $\eta$ .** Still considering the same  $f$  and denoting by  $f_i$  its coefficient of degree  $i$  then  $\eta$  satisfies  $f_{k-1+t} = \eta f_h$ , which permits to deduce  $\eta$ .

**5.2.4 Attacked parameters (with a proof)** We provide the sets of parameters in Table 1, which includes the TRS codes considered in [20] for the single twist case (labeled as [LR]) as well as the parameters for the provable attack discussed in this article (labeled as [CPTZ]).

We emphasize that,

- In the table, we consider only parameters for which Theorem 4 holds. Note that, in Case 2 (as mentioned in the beginning of Subsection 5.2.1), we can apply Theorem 4 if  $k - |I| < 17$ . Thus considering the conditions  $2k + t - 2 - n < |I| \leq k - 17$ , we get  $t < n - k - 15$ .
- [20] restricts to a limited subclass of TRS codes while we consider the much broader class of TGRS codes (the attack can be extended to qGRS codes with no difficulty).
  - $q = q_0^2$  and  $k < n \leq q_0 - 1$  where  $q_0$  is a prime power,
  - $\frac{n}{3} + \sqrt{n} + \frac{1}{3} < k \leq \frac{n}{2} - 2$  (the left hand side inequality comes from the lower bound for  $\ell$  in the proposed set of parameters),
  - $r = \lceil \frac{n+1}{3} \rceil + 2$ ,
  - $t = 2\lceil \frac{n+1}{3} \rceil - k + 2$  and thus  $t \geq \frac{2n}{3} - \frac{n}{2} + 4 = \frac{n}{6} + 4$ .
  - $h = r$  which implies that  $h \neq 1, k - 2$ .

	[LR]	[CPTZ]
$q$	$n^2$	$n$
$k \in$	$[\frac{n}{3} + \sqrt{n} + \frac{1}{3}, \frac{n}{2} - 2]$	$[\sqrt{2n}, n - 14]$
$t$	$2\lceil \frac{n+1}{3} \rceil - k + 2$	$[17, n - k - 16]$
$h$	$\lceil \frac{n+1}{3} \rceil + 2$	$\neq 1, k - 2$

**Table 1.** The parameters for provable attacks in the case of single twist

We mention that in [3],  $t \geq \frac{2n}{3} - \frac{n}{2} + 4 = \frac{n}{6} + 4$ . Thus,  $t < 17$  will imply  $n < 78$ , which is really small compared to the practical parameters.

### 5.3 Complexity

According to [15, Prop. 2], the complexity of computing a Schur square is  $O(n^4)$ . Next, the bottleneck of the attack is the first step with the recovery of the first triple  $(c_1, c_2, c_3) \in \text{ev}_{\alpha, v}(\mathcal{M})^3$ . We claim that this step is the dominant one in the running of the attack. Since, this step requires  $O(q^3)$  trials this yields an overall complexity of  $O(q^3 n^4)$  operations in  $\mathbb{F}_q$ .

### 5.4 Some comments about the non covered cases

We observe the cases where the key-recovery attack might fail:

- if  $h = 1$  or  $k - 2$ , the square of the codimension 1 subcode we recover in Step 1, might not have Schur square equal to  $\mathbf{GRS}_{2k-1}$ .
- for  $k, t < 17$ , Theorem 4 is no longer valid.

Here we give some ideas on alternative ways to recover a valid secret key without going into detailed description.

**5.4.1 When  $t$  is small.** We consider in particular the case where  $t < k$  and that we still assume that  $k < \frac{n}{2}$ . Recall that  $\mathcal{C} = \text{ev}_{\alpha, v}(\mathcal{P})$ , where  $\mathcal{P} = \mathcal{M} \oplus \langle c(x) \rangle$  with

$$\mathcal{M} = \langle 1, x, \dots, \widehat{x^h}, \dots, x^{k-1} \rangle$$

for some  $h \leq k - 1$  and  $\deg(c) = k - 1 + t$ .

Suppose first that  $2k + t < n - 2$ . In this situation, observe that:

$$\mathcal{P}^2 = \mathcal{M}^2 + \mathcal{C}\mathcal{M} + \langle c^2 \rangle.$$

Since  $\deg(c) < k + t < 2k$  then  $\mathcal{M}^2 + \mathcal{C}\mathcal{M} \subseteq \mathbb{F}_q[x]_{2k+t}$  and it is very likely that this inclusion is an equality. In this case,  $\mathcal{C}^2$  is the direct sum of a GRS code and the one-dimensional space  $\text{ev}_{\alpha, v}(\langle c^2 \rangle)$ . In this situation, taking the dual of  $\mathcal{C}^2$  yields a codimension 1 subcode of a GRS code whose structure can be recovered using Wieschebrink's attack [34].

When  $2k + t > n$ , a similar approach can be applied to shortenings of the code.



## 6 Implementation

The first part of the attack, that is to say the computation of the monomial subcode (Step 1, Subsection 5.2.1) is implemented in the computer algebra system SageMath v9.5 [28]. It is available on <https://github.com/nihantanisali/TGRS>. Since the rest of the attack consists in applying Sidelnikov Shestakov attack together with other classical routines we considered that checking the validity of Step 1 was sufficient. Note that our implementation is a proof of concept that is far from being optimized. Hence, due to time and resource constraints we chose smaller  $q$  and  $k$ . The results are summarized in Table 6. Even though these small-scale instances do not fully capture the computational challenges posed by larger, standard parameters, they demonstrate that our method works in practice.

$q_0$	$n$	$k$	$t$	$l$	$h$	$a =  I $	Runtime (s)
53	51	17	17	1	8	–	240
53	51	19	8	1	8	–	261
61	51	25	17	1	19	8	3834

**Table 2.** Experimental results obtained by averaging several runtimes of Algorithm 1 on a 13th Gen Intel® Core™ i7-13800H  $\times$  20.

## 7 A result on the polynomial space

The aim of this section is to provide a proof of Theorem 4. We will consider a polynomial subspace

$$\mathcal{R} = \mathcal{E}' \oplus \langle c(x) \rangle \subseteq \mathbb{F}_q[x],$$

where  $\mathcal{E}' \subseteq_{\text{codim } 1} \mathbb{F}_q[x]_{<k}$ , and  $c(x) \in \mathbb{F}_q[x]$  such that  $\deg(c) = k - 1 + t$  with  $t > 0$ . We will determine the quantity

$$\dim \langle f_1, f_2, f_3 \rangle \mathcal{R}$$

for different choices of  $f_1, f_2, f_3 \in \mathcal{R}$  depending on admissible values of  $t$  and  $k$ , which will be specified later.

First, recall that Proposition 11 yields an upper bound on  $\dim \langle f_1, f_2, f_3 \rangle \mathcal{R}$  for  $f_1, f_2, f_3 \in \mathcal{E}'$ :

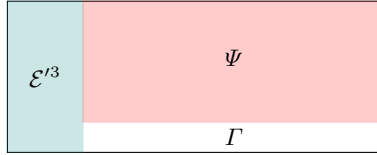
$$\dim \langle f_1, f_2, f_3 \rangle \mathcal{R} \leq 2k + 2.$$

Then, we study the general case of  $f_1, f_2, f_3 \in \mathcal{R}$  in order to prove Theorem 4. First we define the following sets.

$$\Psi \stackrel{\text{def}}{=} \{(f_1, f_2, f_3) \in \mathcal{R}^3 : \dim \langle f_1, f_2, f_3 \rangle \mathcal{R} > 2k + 2\}; \quad (14)$$

$$\Gamma \stackrel{\text{def}}{=} \{(f_1, f_2, f_3) \in \mathcal{R}^3 \setminus \mathcal{E}'^3 : \dim \langle f_1, f_2, f_3 \rangle \mathcal{R} \leq 2k + 2\}. \quad (15)$$

We get a partition of  $\mathcal{R}^3 = \mathcal{E}'^3 \sqcup \Gamma \sqcup \Psi$  as illustrated in Figure 1.



**Fig. 1.** Illustrates the partition  $\mathcal{E}'^3 \sqcup \Psi \sqcup \Gamma$  of  $\mathcal{R}^3$

Our main goal in this section is to prove that if  $(f_1, f_2, f_3) \in \mathcal{R}^3$  with  $\dim \langle f_1, f_2, f_3 \rangle \mathcal{R} \leq 2k + 2$ , then  $(f_1, f_2, f_3) \in \mathcal{E}'^3$  with probability  $\geq 1 - \frac{1}{q^2}$ . This probability corresponds to showing that

$$\frac{|\mathcal{E}'^3|}{|\mathcal{E}'^3 \sqcup \Gamma|} \geq 1 - \frac{1}{q^2}. \quad (16)$$

Let  $f_1, f_2, f_3 \in \mathcal{R}$ . If  $f_i \notin \mathcal{E}'$  for some  $1 \leq i \leq 3$ , then w.l.o.g. (after taking linear combinations), we can assume that the polynomials are of the form

$$\begin{aligned} f_1(x) &:= a_0 + a_1x + \cdots + a_{k-1}x^{k-1} + a_{k-1+t}c(x), \quad a_{k-1+t} \neq 0, \\ f_2(x) &:= b_0 + b_1x + \cdots + b_{k-1}x^{k-1}, \\ f_3(x) &:= c_0 + c_1x + \cdots + c_{k-1}x^{k-1}. \end{aligned}$$

First, note that  $\dim \langle f_1, f_2, f_3 \rangle \mathcal{R} \geq \dim \langle f_1, f_2, f_3 \rangle \mathcal{E}'$ . We now focus on  $\dim \langle f_1, f_2, f_3 \rangle \mathcal{E}'$ . Clearly, the dimension of the space  $f_1 \mathcal{E}'$  is  $k-1$ . We now fix a basis  $\{e_1(x), \dots, e_{k-1}(x)\}$  of  $\mathcal{E}'$  such that  $\deg(e_1) < \cdots < \deg(e_{k-1})$ .

For all polynomials  $e_i(x) \in \mathcal{E}'$ , we have that

$$k - 1 + t \leq \deg(e_i(x)f_1(x)) \leq 2k - 2 + t.$$

We define the subspace of  $f_1 \mathcal{E}'$  generated by  $e_i$ 's of *large degrees*, denoted by  $(f_1 \mathcal{E}')_t$ , as follows:

$$(f_1 \mathcal{E}')_t \stackrel{\text{def}}{=} \langle f_1 e_i : \deg(f_1) + \deg(e_i) \geq 2k - 1 \rangle.$$

We have

$$\dim((f_1\mathcal{E}')_l) = \begin{cases} k-1 & \text{if } t \geq k, \\ t-1 \text{ or } t & \text{if } t < k. \end{cases} \quad (17)$$

For  $t < k$  in the above (17),  $\dim((f_1\mathcal{E}')_l)$  has two possibilities depending on the missing integer  $t_1 \in \{0, \dots, k-1\} \setminus \{\deg(e_i) : i = 1, \dots, k-1\}$ . Indeed, if  $k-1+t+t_1 \geq 2k-1$ , then  $\dim((f_1\mathcal{E}')_l) = t-1$  or  $t$ , otherwise. Using the fact that the space  $f_2\mathcal{E}' + f_3\mathcal{E}'$  contains polynomials up to degree  $2k-2$  (as  $\deg(f_2), \deg(f_3) \leq k-1$ ), we get the equalities

$$\dim \langle f_1, f_2, f_3 \rangle \mathcal{E}' = \begin{cases} k-1 + \dim \langle f_2, f_3 \rangle \mathcal{E}' & \text{if } t \geq k, \\ \left. \begin{array}{l} t-1 + \dim \langle f_2, f_3 \rangle \mathcal{E}' \\ \text{or, } t + \dim \langle f_2, f_3 \rangle \mathcal{E}' \end{array} \right\} & \text{if } t < k \end{cases} \quad (18)$$

We now focus on estimating  $\dim \langle f_2, f_3 \rangle \mathcal{E}'$  where  $f_2$  and  $f_3$  are chosen uniformly at random from  $\mathcal{E}'$ . For this, we use a classical result on polynomials that we prove for the sake of self containedness.

**Lemma 5.** *Let  $f, g \in \mathbb{F}_q[x]_{<k}$  be two polynomials. Then*

$$\dim(f\mathbb{F}_q[x]_{<k} + g\mathbb{F}_q[x]_{<k}) = k + \max\{\deg f, \deg g\} - \deg(\gcd(f, g)).$$

*Proof.* Set  $s = \deg f$ ,  $u = \deg g$ ,  $h = \gcd(f, g)$  and  $d = \deg h$ . W.l.o.g. one can assume  $s = \max\{s, u\}$  and we set  $f_1(x) = f(x)/h(x)$  and  $g_1(x) = g(x)/h(x)$ . Consider the linear map

$$\phi : \begin{cases} \mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k} & \longrightarrow & f\mathbb{F}_q[x]_{<k} + g\mathbb{F}_q[x]_{<k} \\ (a, b) & \longmapsto & af + bg. \end{cases}$$

Any  $(a, b) \in \ker \phi$  satisfies  $af = -bg$ . By dividing both sides by  $h$  we get  $af_1 = -bg_1$  and since  $f_1, g_1$  are coprime, we deduce that there exists  $p \in \mathbb{F}_q[x]$  such that  $a = g_1p$  and  $b = -f_1p$ . Since  $a, b$  have degrees  $< k$ , we deduce that  $\deg f_1 + \deg p = s - d + \deg p < k$ . Therefore,

$$\ker \phi = \{(g_1p, -f_1p) : p \in \mathbb{F}_q[x]_{<k-s+d}\}.$$

Thus,  $\dim \ker \phi = k - s + d$  and the rank-nullity theorem permits to conclude.  $\square$

Observe that  $\dim \langle f_2, f_3 \rangle \mathcal{E}' \geq \dim(f_2\mathbb{F}_q[x]_{<k} + f_3\mathbb{F}_q[x]_{<k}) - 2$ . Now, (18) implies that in order to get  $\dim \langle f_1, f_2, f_3 \rangle \mathcal{E}' > 2k + 2$  it suffices to have

$$\dim \langle f_2, f_3 \rangle \mathcal{E}' > \begin{cases} k+3 & \text{if } t \geq k, \\ 2k-t+3 & \text{if } t < k. \end{cases} \quad (19)$$

As a straightforward consequence of Lemma 5, we reduce the dimension conditions to a condition on the degree of  $\gcd(f_2, f_3)$ . We give the condition more generally for any pair of polynomials  $(f, g) \in \mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}$  and we call it the *gcd condition* as defined below.

**Definition 7.** *We say that a polynomial pair  $(f, g) \in \mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}$  with degrees  $s$  and  $u$  satisfies the gcd condition if*

$$\deg(\gcd(f, g)) < \begin{cases} \max\{s, u\} - 5 & \text{if } t > k \\ \max\{s, u\} - (k - t) - 5 & \text{if } t \leq k. \end{cases} \quad (20)$$

*Remark 8.* Note that the bounds on the degree of the gcd in (20) are taken so that, if a pair  $(f_2, f_3) \in \mathcal{E}' \times \mathcal{E}'$  satisfies the gcd condition in Definition 7, then  $(f_1, f_2, f_3) \in \Psi$  for  $f_1 \in \mathcal{R} \setminus \mathcal{E}'$  (see (14) for the definition of  $\Psi$ ). Therefore, it helps to obtain a bound on the size of  $\Psi$ , that we determine in the subsequent part.

We aim to find the number of pairs  $(f_2, f_3) \in \mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}$  that satisfy the gcd condition given in Definition 7. For that, we introduce the required notations and notions.

**Definition 8.** *Let  $i, j, s, u$  be non-negative integers such that  $s, u < k$ . Define*

1.  $\mathcal{G} := \{(f, g) \in \mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k} : (f, g) \text{ satisfies the gcd condition}\};$
2.  $\mathcal{G}_{\mathcal{E}'} := \{(f, g) \in \mathcal{E}' \times \mathcal{E}' : (f, g) \text{ satisfies the gcd condition}\};$
3.  $\mathcal{A}(s, u) := \{(f, g) : \deg(f) = s, \deg(g) = u\};$
4.  $\mathcal{B}(s, u, i) := \{(f, g) : \deg(f) = s, \deg(g) = u, \deg(\gcd(f, g)) = i\};$
5.  $\mathcal{B}_j(s, u) := \{(f, g) : \deg(f) = s, \deg(g) = u, \deg(\gcd(f, g)) \leq j\}.$

In the rest of this section, we will make use of the following result in our counting arguments.

**Theorem 5 ([6, Theorem 3]).** *Let  $f$  and  $g$  be randomly chosen from the set of polynomials in  $\mathbb{F}_q[x]$  of degree  $s$  and  $u$  respectively, where  $s$  and  $u$  are not both zero. Then the probability of  $f$  and  $g$  being coprime is  $1 - \frac{1}{q}$ .*

We start counting the number of polynomials in  $\mathbb{F}_q[x]$  of fixed degrees.

**Lemma 6.** *Let  $s, u, i$  be nonnegative integers such that  $s, u < k$ . Then  $|\mathcal{A}(s, u)| = (q - 1)^2 q^{s+u}$  and moreover, among the polynomials in  $\mathcal{A}(s, u)$ ,*

1.  $(q - 1)^3 q^{s+u-1}$  many of them are coprime, i.e.  $|\mathcal{B}(s, u, 0)| = (q - 1)^3 q^{s+u-1};$
2.  $(q - 1)^2 q^{s+u-1}$  many of them are not coprime;

3. For  $i \leq \min\{s, u\}$ ,  $|\mathcal{B}(s, u, i)| = (q-1)^3 q^{s+u-i-1}$ .

*Proof.* The first two items follow directly from Theorem 5. The number of monic polynomials  $h$  of degree  $i$  is  $q^i$ . Observe that the number of polynomial pairs  $(f, g)$  with  $\gcd(f, g) = h$  is equal to the number of polynomial pairs  $(f', g')$  with  $\deg f' = s - i$ ,  $\deg g' = u - i$  and  $\deg(\gcd(f', g')) = 0$ . By the first item, there are  $(q-1)^3 q^{s+u-2i-1}$  such pairs. Since it is the same number for any  $\gcd(f, g) = h$ , and we have  $q^i$  many choices for  $h$ , there are  $(q-1)^3 q^{s+u-2i-1} q^i = (q-1)^3 q^{s+u-i-1}$  many polynomial pairs with  $\deg(\gcd(f, g)) = i$  where  $i \leq \min\{s, u\}$ .  $\square$

To count the sets of polynomials  $\mathcal{G}$  and  $\mathcal{G}'$ , we first count the number of polynomial pairs in  $\mathcal{A}(s, u)$  having gcd less than or equal to  $j$  where  $j \leq \min\{s, u\}$ .

**Lemma 7.** For any positive integers  $s, u$  and  $j \leq s, u < k$ , we have

$$\frac{|\mathcal{B}_j(s, u)|}{|\mathcal{A}(s, u)|} = 1 - \frac{1}{q^{j+1}}.$$

*Proof.* Note that  $\mathcal{B}(s, u, i)$ 's are non intersecting for distinct  $i$  and thus  $\mathcal{B}_j(s, u) = \bigsqcup_{i \leq j} \mathcal{B}(s, u, i)$ . Following Lemma 6, we have  $|\mathcal{B}(s, u, i)| = (q-1)^3 q^{s+u-i-1}$ . Hence,

$$\begin{aligned} \sum_{i \leq j} |\mathcal{B}(s, u, i)| &= \sum_{i \leq j} (q-1)^3 q^{s+u-i-1} \\ &= (q-1)^3 q^{s+u-j-1} (1 + \dots + q^j) \\ &= (q-1)^3 q^{s+u-j-1} \frac{(q^{j+1} - 1)}{(q-1)} \\ &= (q-1)^2 q^{s+u-j-1} (q^{j+1} - 1). \end{aligned}$$

Note that  $|\mathcal{A}(s, u)| = (q-1)^2 q^{s+u}$ , and therefore,

$$\frac{|\mathcal{B}_j(s, u)|}{|\mathcal{A}(s, u)|} = \frac{(q-1)^2 q^{s+u-j-1} (q^{j+1} - 1)}{(q-1)^2 q^{s+u}} = \frac{q^{s+u} - q^{s+u-j-1}}{q^{s+u}} = 1 - \frac{1}{q^{j+1}}.$$

$\square$

Now, we are ready to count the polynomials in  $\mathcal{G}$ . The lemma to follow yields a density estimate for  $\mathcal{G}$  for  $k, t$  large enough. The lower bound on  $k, t$  are chosen to further obtain a reasonable estimate for the density of  $\mathcal{E}'$ .

**Lemma 8.** For integers  $k, t \geq 17$ , we have

$$\frac{|\mathcal{G}|}{|\mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}|} \geq 1 - \frac{1}{q^7}.$$

*Proof.* Note that  $\mathcal{G} = \bigcup_{s,u < k} \mathcal{B}_j(s, u)$ . Let  $j_{s,u}$  be the greatest possible degree for  $\gcd(f, g)$  such that  $(f, g)$  satisfies the gcd condition where  $\deg f = s$  and  $\deg g = u$ . Using this, we get

$$\frac{|\mathcal{G}|}{|\mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}|} = \sum_{s,u < k} \frac{|\mathcal{A}(s, u)|}{|\mathbb{F}_q[x]_{<k}|^2} \frac{|\mathcal{B}_{j_{s,u}}(s, u)|}{|\mathcal{A}(s, u)|} \quad (21)$$

Observe that if we consider only the pairs  $(s, u)$  such that  $s + u = r \in \{2k - 10, \dots, 2k - 2\}$ , we get the following lower bound of the sum in (21),

$$\sum_{s,u} \frac{|\mathcal{A}(s, u)|}{|\mathbb{F}_q[x]_{<k}|^2} \frac{|\mathcal{B}_{j_{s,u}}(s, u)|}{|\mathcal{A}(s, u)|} \geq \sum_{r \geq 2k-10} \sum_{s+u=r} \frac{|\mathcal{A}(s, u)|}{|\mathbb{F}_q[x]_{<k}|^2} \frac{|\mathcal{B}_{j_{s,u}}(s, u)|}{|\mathcal{A}(s, u)|}. \quad (22)$$

We replace  $r$  by  $r' := (2k - 2) - r$  in the limits of the summation in the above equation (22). Note that for a fixed  $r'$ , there are exactly  $(r' + 1)$ -choices for the pairs  $(s, u)$  such that  $s + u = r$ , and following the value of  $|\mathcal{A}(s, u)|$  in Lemma 6, we have

$$\frac{|\mathcal{A}(s, u)|}{|\mathbb{F}_q[x]_{<k}|^2} = \left( \frac{q-1}{q} \right)^2 \frac{1}{q^{r'}}. \quad (23)$$

For  $r' = (2k - 2) - r \leq 8$ , and  $s + u = r$ ,  $\max\{s, u\} \geq k - 5$ . For any pair  $(s, u)$  such that  $(2k - 2) - r' = r = s + u \geq 2k - 10$ , this implies

$$\begin{cases} k - 10 \leq j_{s,u} < k - 6 & \text{if } t > k, \\ t - 10 \leq j_{s,u} < t - 6 & \text{if } t \leq k. \end{cases}$$

Since  $k, t \geq 17$ , then  $j_{s,u} + 1 \geq 8$  in both cases. Therefore, following Lemma 7, we get

$$\frac{|\mathcal{B}_{j_{s,u}}(s, u)|}{|\mathcal{A}(s, u)|} = 1 - \frac{1}{q^{j_{s,u}+1}} \geq 1 - \frac{1}{q^8}. \quad (24)$$

Combining (21), (22), (23) and (24) all together, we get

$$\frac{|\mathcal{G}|}{|\mathbb{F}_q[x]_{<k}|^2} \geq \left( \frac{q-1}{q} \right)^2 \left( 1 - \frac{1}{q^8} \right) \sum_{r'=0}^8 (r' + 1) \frac{1}{q^{r'}}.$$

The leftmost term is the truncated Taylor expansion of the function  $F = (x - 1)^{-2}$  evaluated at  $1/q$ . Let us write this the Taylor expansion as

$$(x - 1)^{-2} = 1 + 2x + 3x^2 + \dots + 8x^7 + 9x^8 + R_8(x).$$

Moreover, for any  $x \in (0, 1/5)$ , we have

$$R_8(x) = \frac{F^{(9)}(z)x^9}{9!}$$

for some  $z \in (0, 1/5)$  and one can prove that for any such  $x$ , we have  $|R_8(x)| < 121x^9$ . Taking  $x = 1/q$  (since we are dealing with GRS codes, one can reasonably assume that  $q \geq 11$ ), we finally get

$$\frac{|\mathcal{G}|}{|\mathbb{F}_q[x]_{<k}|^2} \geq \left(\frac{q-1}{q}\right)^2 \left(1 - \frac{1}{q^8}\right) \left(\left(\frac{q-1}{q}\right)^{-2} + O\left(\frac{1}{q^9}\right)\right) \geq 1 - \frac{1}{q^7}.$$

□

*Remark 9.* Note that in Lemma 8, the lower bound  $1 - \frac{1}{q^7}$  can be improved for larger  $t, k$  such that  $k, t \gg 17$ . As we can consider more  $r$ 's in (22) for those cases, the error term in the Taylor expansion gets smaller.

The following Lemma will be useful to count polynomials in  $\mathcal{G}_{\mathcal{E}'}$ .

**Lemma 9.** *We have that,*

$$\frac{|\mathcal{G}|}{|\mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}|} \geq 1 - \frac{1}{q^w} \implies \frac{|\mathcal{G}_{\mathcal{E}'}|}{|\mathcal{E}' \times \mathcal{E}'|} \geq 1 - \frac{1}{q^{w-2}}.$$

for some  $w \in \mathbb{Z}_{>0}$ .

*Proof.* Note that,

$$\begin{aligned} |(\mathcal{E}' \times \mathcal{E}') \cup \mathcal{G}| &= |\mathcal{E}' \times \mathcal{E}'| + |\mathcal{G}| - |\mathcal{G} \cap (\mathcal{E}' \times \mathcal{E}')| \\ &\leq |\mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}| = q^2 |\mathcal{E}' \times \mathcal{E}'|. \end{aligned}$$

Now, since  $\mathcal{G}_{\mathcal{E}'} = \mathcal{G} \cap (\mathcal{E}' \times \mathcal{E}')$ , from the previous inequality we can deduce that,

$$|\mathcal{G}_{\mathcal{E}'}| \geq (1 - q^2) |\mathcal{E}' \times \mathcal{E}'| + |\mathcal{G}|.$$

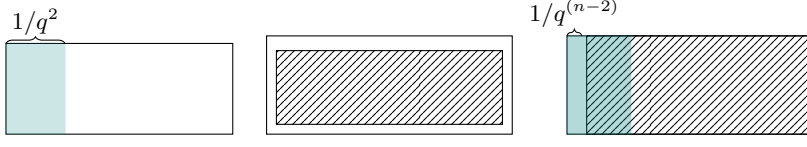
From the hypothesis of the lemma,

$$|\mathcal{G}| \geq q^2 |\mathcal{E}' \times \mathcal{E}'| \left(1 - \frac{1}{q^w}\right),$$

and so, putting all together, we can conclude the desired result,

$$|\mathcal{G}_{\mathcal{E}'}| \geq \left((1 - q^2) + \left(q^2 - \frac{1}{q^{w-2}}\right)\right) |\mathcal{E}' \times \mathcal{E}'| = \left(1 - \frac{1}{q^{w-2}}\right) |\mathcal{E}' \times \mathcal{E}'|.$$

□



**Fig. 2.** The first image illustrates the ratio of  $\mathcal{E}' \times \mathcal{E}'$  in  $\mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}$ , the second one is for the ratio of  $\mathcal{G}$  in  $\mathbb{F}_q[x]_{<k} \times \mathbb{F}_q[x]_{<k}$  and the third one represents the *worst case* intersection of the first two.

*Proof of Theorem 4.* We made the observation that the inequality

$$\dim \langle f_1, f_2, f_3 \rangle \mathcal{E}' > 2k + 2 \quad (25)$$

holds true if  $(f_2, f_3)$  satisfies the *gcd condition* as given in (20). Using Lemma 8, and then taking  $w = 7$  in Lemma 9, this probability is  $\frac{|\mathcal{G}_{\mathcal{E}'}|}{|\mathcal{E}' \times \mathcal{E}'|} \geq 1 - \frac{1}{q^5}$ . Following Remark 8, it means that  $\frac{|\Psi|}{|\Psi \cup \Gamma|} \geq 1 - \frac{1}{q^5}$ . Thus, if a triple  $(f_1, f_2, f_3) \in \Psi \cup \Gamma$ , then  $(f_1, f_2, f_3) \in \Gamma$  with probability  $\leq \frac{1}{q^5}$ . For a triple  $(f_1, f_2, f_3) \in \mathcal{R}^3$ , the probability of  $(f_1, f_2, f_3) \in \Psi \cup \Gamma$  is  $1 - \frac{1}{q^3}$ . Hence we get:

$$\frac{|\mathcal{E}'^3|}{|\mathcal{E}'^3 \cup \Gamma|} \geq \frac{\frac{1}{q^3}}{\left(1 - \frac{1}{q^3}\right) \frac{1}{q^5} + \frac{1}{q^3}} \geq 1 - \frac{1}{q^2}.$$

This, together with the subsequent statement complete the proof of the theorem.  $\square$

**Corollary 2.** *Let  $f_1, f_2, f_3 \in \mathcal{R}$  are three polynomials chosen uniformly at random. If  $f_1, f_2 \in \mathcal{E}'$  and  $(f_1, f_2, f_3) \in \mathcal{E}'^3 \cup \Gamma$ , then probability of  $(f_1, f_2, f_3) \in \mathcal{E}'^3$  is  $\geq 1 - \frac{1}{q^2}$ .*

*Proof.* Defining

$$\Gamma_0 := \{(f_1, f_2, f_3) \in \Gamma : \dim \langle f_1, f_2, f_3 \rangle \mathcal{R} \leq 2k + 2 \text{ and } f_1, f_2 \in \mathcal{E}'\},$$

and using  $\Gamma_0 \subseteq \Gamma$  the conditional probability described in the proposition can be lower bounded by

$$\frac{|\mathcal{E}'|}{|\mathcal{E}' \cup \Gamma_0|} \geq \frac{|\mathcal{E}'|}{|\mathcal{E}' \cup \Gamma|} \geq 1 - \frac{1}{q^2}.$$

$\square$

*Remark 10.* After iterating the process of appending a polynomial to primarily found  $(f_1, f_2, f_3) \in \mathcal{E}'^3 \sqcup \Gamma$   $k - 3$  times, we get a basis of the codimension 1 subspace  $\mathcal{E}'$  with probability  $(1 - \frac{(k-3)}{q^2})$ .



## 8 Conclusion

We showed that qGRS codes are easily distinguishable from random codes using classical cryptanalysis techniques on algebraic codes. Moreover, we proposed a complete key-recovery attack on TGRS and actually qGRS codes with  $\ell = 1$ . We left some questions open. First, our proof of the validity of the attack rests on a probability analysis which does not permit us to treat some cases where the attack still probably works. Second, for some very specific cases such as  $h = 1$  or  $k - 2$ , despite being able to distinguish the codes, we cannot conclude the attack. We hope such cases to be attacked but this would require an ad hoc manner to finish the attack. Finally the major remaining question is the case of  $\ell > 1$  twists. It is clear that a too high number of twist is not practical for cryptography since the decryption complexity is exponential in the number of twists. Still, the question of small  $\ell \geq 2$  remains open, since the codes are distinguishable from random ones, one can be optimistic on the possibility to extend the attack to larger  $\ell$ 's but this probably requires many technical adaptations both in the attack itself and in the probability analysis to prove the validity of the attack.

## References

1. Albrecht, M., Bernstein, D.J., Chou, T., Cid, C., Gilcher, J., Lange, T., Maram, V., von Maurich, I., Mizoczki, R., Niederhagen, R., Persichetti, E., Paterson, K., Peters, C., Schwabe, P., Sendrier, N., Szefer, J., Tjhai, C.J., Tomlinson, M., Wen, W.: Classic McEliece (merger of Classic McEliece and NTS-KEM). <https://classic.mceliece.org> (Nov 2022), fourth round finalist of the NIST post-quantum cryptography call
2. Baldi, M., Bianchi, M., Chiaraluce, F., Rosenthal, J., Schipani, D.: Enhanced public key security for the McEliece cryptosystem. *J. Cryptology* **29**, 1–27 (2016)
3. Beelen, P., Bossert, M., Puchinger, S., Rosenkilde, J.: Structural properties of twisted Reed-Solomon codes with applications to cryptography. In: Proc. IEEE Int. Symposium Inf. Theory - ISIT 2018. pp. 946–950. IEEE (2018)
4. Beelen, P., Puchinger, S., né Nielsen, J.R.: Twisted Reed-Solomon codes. In: Proc. IEEE Int. Symposium Inf. Theory - ISIT 2017. pp. 336–340. IEEE (2017)
5. Beelen, P., Puchinger, S., Rosenkilde, J.: Twisted Reed-Solomon codes. *IEEE Trans. Inform. Theory* **68**(5), 3047–3061 (2022)
6. Benjamin, A.T., Bennett, C.D.: The probability of relatively prime polynomials. *Mathematics Magazine* **80**(3), 196–202 (2007). <https://doi.org/10.1080/0025570X.2007.11953481>
7. Berger, T., Loidreau, P.: How to mask the structure of codes for a cryptographic use. *Des. Codes Cryptogr.* **35**, 63–79 (2005)

8. Cascudo, I., Cramer, R., Mirandola, D., Zémor, G.: Squares of random linear codes. *IEEE Trans. Inform. Theory* **61**(3), 1159–1173 (3 2015). <https://doi.org/10.1109/TIT.2015.2393251>
9. Couvreur, A., Gaborit, P., Gauthier-Umaña, V., Otmani, A., Tillich, J.P.: Distinguisher-based attacks on public-key cryptosystems using Reed–Solomon codes. *Des. Codes Cryptogr.* **73**, 641–666 (2014)
10. Couvreur, A., Lequesne, M.: On the security of subspace subcodes of Reed–Solomon codes for public key encryption. *IEEE Trans. Inform. Theory* **68**(1), 632–648 (2022). <https://doi.org/10.1109/TIT.2021.3120440>
11. Couvreur, A., Lequesne, M., Tillich, J.P.: Recovering short secret keys of RLCE in polynomial time. In: Ding, J., Steinwandt, R. (eds.) *Post-Quantum Cryptography 2019*. LNCS, vol. 11505, pp. 133–152. Springer, Chongqing, China (May 2019)
12. Couvreur, A., Márquez-Corbella, I., Pelikaan, R.: Cryptanalysis of McEliece cryptosystem based on algebraic geometry codes and their subcodes. *IEEE Trans. Inform. Theory* **63**(8), 5404–5418 (8 2017)
13. Couvreur, A., Otmani, A., Tillich, J.P.: Polynomial time attack on wild McEliece over quadratic extensions. In: Nguyen, P.Q., Oswald, E. (eds.) *Advances in Cryptology - EUROCRYPT 2014*. LNCS, vol. 8441, pp. 17–39. Springer Berlin Heidelberg (2014)
14. Couvreur, A., Otmani, A., Tillich, J.P.: Polynomial time attack on wild McEliece over quadratic extensions. *IEEE Trans. Inform. Theory* **63**(1), 404–427 (1 2017)
15. Couvreur, A., Otmani, A., Tillich, J.P., Gauthier-Umana, V.: A polynomial-time attack on the BBCRS scheme. In: *IACR International Workshop on Public Key Cryptography*. pp. 175–193. Springer (2015)
16. Faugère, J.C., Gauthier-Umaña, V., Otmani, A., Perret, L., Tillich, J.P.: A distinguisher for high-rate McEliece cryptosystems. *IEEE Trans. Inform. Theory* **59**(10), 6830–6844 (2013). <https://doi.org/10.1109/TIT.2013.2272036>
17. Guruswami, V., Sudan, M.: Improved decoding of reed-solomon and algebraic-geometric codes. In: *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*. pp. 28–37 (1998). <https://doi.org/10.1109/SFCS.1998.743426>
18. Janwa, H., Moreno, O.: McEliece public key cryptosystems using algebraic-geometric codes. *Des. Codes Cryptogr.* **8**(3), 293–307 (1996)
19. Khathuria, K., Rosenthal, J., Weger, V.: Encryption scheme based on expanded Reed–Solomon codes. *Adv. Math. Commun.* **15**(2), 207–218 (2021). <https://doi.org/10.3934/amc.2020053>, <http://aimsciences.org//article/id/0f055199-6fe4-404f-b206-517ce7d02a58>
20. Lavauzelle, J., Renner, J.: Cryptanalysis of a system based on twisted Reed–Solomon codes. *Des. Codes Cryptogr.* **88**(7), 1285–1300 (2020)
21. McEliece, R.: A public-key system based on algebraic coding theory. *Jet Propulsion Laboratory, California Institute of Technology* **3**, 85–86 (1978)

22. Misoczki, R., Tillich, J.P., Sendrier, N., Barreto, P.S.L.M.: MDPC-McEliece: New McEliece variants from moderate density parity-check codes. In: Proc. IEEE Int. Symposium Inf. Theory - ISIT. pp. 2069–2073 (2013). <https://doi.org/10.1109/ISIT.2013.6620590>
23. Niederreiter, H.: Knapsack-type cryptosystems and algebraic coding theory. Prob. Contr. Inform. Theory **15**(2), 157–166 (1986)
24. Roth, R.: Introduction to Coding Theory. Cambridge University Press, USA (2006)
25. Sheekey, J.: A new family of linear maximum rank distance codes. Adv. Math. Commun. **10**(3), 475–488 (2016). <https://doi.org/10.3934/amc.2016019>, <https://www.aims sciences.org/article/id/3ff311ae-18e0-47d8-8edc-ab169dbbd975>
26. Sidelnikov, V.M.: A public-key cryptosystem based on Reed-Muller codes. Discrete Math. Appl. **4**(3), 191–207 (1994)
27. Sidelnikov, V.M., Shestakov, S.: On the insecurity of cryptosystems based on generalized Reed-Solomon codes. Discrete Math. Appl. **1**(4), 439–444 (1992)
28. Stein, W., et al.: Sage Mathematics Software (Version 9.5). The Sage Development Team (2022), <http://www.sagemath.org>
29. Sudan, M.: Decoding of Reed-Solomon codes beyond the error-correction bound. Journal of Complexity **13**(1), 180–193 (1997). <https://doi.org/https://doi.org/10.1006/jcom.1997.0439>, <https://www.sciencedirect.com/science/article/pii/S0885064X97904398>
30. Wang, Y.: Quantum resistant random linear code based public key encryption scheme RLCE. In: Proc. IEEE Int. Symposium Inf. Theory - ISIT 2016. pp. 2519–2523. IEEE, Barcelona, Spain (Jul 2016). <https://doi.org/10.1109/ISIT.2016.7541753>, <http://dx.doi.org/10.1109/ISIT.2016.7541753>
31. Wang, Y.: RLCE-KEM. <http://quantumca.org> (2017), First round submission to the NIST post-quantum cryptography call
32. Wieschebrink, C.: An attack on a modified Niederreiter encryption scheme. In: Yung, M., Dodis, Y., Kiayias, A., Malk, T. (eds.) Public-Key Cryptography - PKC 2006. LNCS, vol. 3958, pp. 14–26. Springer (2006)
33. Wieschebrink, C.: Two NP-complete problems in coding theory with an application in code based cryptography. In: Proc. IEEE Int. Symposium Inf. Theory - ISIT 2006. pp. 1733–1737. IEEE (2006)
34. Wieschebrink, C.: Cryptanalysis of the Niederreiter public key scheme based on GRS subcodes. In: Post-Quantum Cryptography 2010. LNCS, vol. 6061, pp. 61–72. Springer (2010)