# Distinguishing Goppa codes using higher-order vanishing

Tobias Hemmert and Andreas Wiemers

Federal Office for Information Security, Bonn
tobias.hemmert@bsi.bund.de

**Abstract.** We present a new algorithm to distinguish alternant and Goppa codes from general linear codes. Our approach is based on higher-order vanishing of polynomials and can be applied to a wide set of code parameters. It also applies to Goppa code parameters used in the Classic McEliece key encapsulation mechanism. While these parameters are out of reach in practice, we analyse the behaviour of our distinguisher and estimate its complexity for them, indicating that it is more efficient than previous distinguishing approaches. This is supported by concrete experiments that distinguish codes with larger parameters in practice.

## 1 Introduction

The McEliece encryption scheme [19] is one of the oldest public-key cryptosystems. Since its publication in 1978, its security has remained remarkably stable [2]. Unlike the RSA cryptosystem, which was published a few years earlier, it is deemed secure not only against classical algorithms but also against large-scale quantum computers. The Classic McEliece key encapsulation mechanism [1] was a fourth-round candidate in the NIST standardisation process for post-quantum cryptography. It was not selected for standardisation by NIST mostly due to its large public key size, which limits its usability in many applications [3].

The McEliece cryptosystem employs binary Goppa codes, a family of error-correcting codes for which efficient decoding algorithms are known. Roughly speaking, the security of the (Classic) McEliece scheme is based on the assumption that decoding random errors for a randomly selected binary Goppa code is hard if one is given only a random basis of the Goppa code. The best known algorithms to solve this problem are information-set decoding algorithms [22][25][18][6], whose complexity has remained very stable over the past 60 years.

One way to approach the study of the hardness assumption that underpins the security of the McEliece cryptosystem is to consider a related but potentially easier problem. It is known that the decoding problem for general linear codes is NP-complete [8] and it is also believed to be hard on average. A natural question to ask is: Given a random basis (a) of a randomly selected (binary) Goppa code,

© IACR 2026. This article is the final version submitted by the authors to the IACR and to Springer on 05/06/2026. The version published by Springer is available at <DOI>.

or (b) of a randomly selected linear code of a fixed dimension, can one determine if one has been given (a) or (b)? We will call this the *distinguishing problem for Goppa codes*. If this problem is hard, then the problem underlying the security of (Classic) McEliece is also hard. The converse need not be true, of course.

*Previous work.* Due to its relationship with the McEliece cryptosystem, the distinguishing problem has received some attention over the past 15 years. [12] first describes a polynomial-time distinguisher for alternant/Goppa codes that works if the codes have a sufficiently high rate. The distinguisher exploits the fact that a certain linear system obtained by linearising a polynomial system describing key recovery behaves differently when it is constructed from a random code or an alternant/Goppa code. [21] observes that this distinguisher has a simpler interpretation in terms of componentwise products of codes. We refer to this as the *square distinguisher* and will briefly sketch its main idea below. [20] proves tight bounds for the behaviour of this distinguisher in the Goppa and alternant case. Hence, the authors are able to determine precisely the parameter ranges in which Goppa or alternant codes can be efficiently distinguished from random codes with high probability using the square distinguisher. A new approach using quadratic forms [10] allows in principle to distinguish alternant/Goppa codes outside the high-rate regime including in parameter ranges used in the McEliece cryptosystem. Its complexity is estimated to be lower than the best known key recovery attack against the McEliece cryptosystem, but it still grows rapidly outside the high-rate regime. Finally, the syzygy distinguisher [23] uses higher-order Betti numbers as invariants to distinguish alternant/Goppa and random codes. It is the first distinguisher that is asymptotically sub-exponential in the error-correcting capability of the code.

*Outline of our approach.* We introduce a new approach to distinguish alternant/Goppa codes and random codes that is based on higher-order vanishing of polynomials. Similarly to previous distinguishers, our starting point is the square distinguisher, which we briefly recall. Given a linear code $C \subset \mathbb{F}_q^n$, we define the *square code* of $C$ as

$$C^{*2} \coloneqq \operatorname{span}_{\mathbb{F}_q} \{\boldsymbol{c} * \boldsymbol{d} \mid \boldsymbol{c}, \boldsymbol{d} \in C\}$$

where $\boldsymbol{c} * \boldsymbol{d} \coloneqq (c_1 d_1, \ldots, c_n d_n)$ denotes the componentwise product of two vectors and $\operatorname{span}_{\mathbb{F}_q} U$ denotes the $\mathbb{F}_q$-linear span of a subset $U \subset \mathbb{F}_q^n$. If $C$ is a linear code of dimension $k$, then clearly

$$\dim_{\mathbb{F}_q} \left(C^{*2}\right) \leq \min \left\{ n, \binom{k+1}{2} \right\}$$

If $C$ is randomly chosen of sufficiently large dimension $k$, then [9] shows that the dimension of its square code attains this maximal value with high probability. In contrast, if $C = \mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y})$ is a generalised Reed-Solomon code (see Definition 1), it is easy to see that

$$C^{*2} = \mathrm{GRS}_{2k-1}(\boldsymbol{x}, \boldsymbol{y} * \boldsymbol{y})$$

and so its dimension is at most $2k-1$. If $2k-1<n$, we can therefore detect a difference with high probability and thus distinguish GRS codes from random codes. It is known that the dual of a GRS code is again a GRS code. So, if $2k-1\geq n$, we can instead compute the dimension of the square of the *dual* of the given code. Thus, the square distinguisher allows to distinguish GRS codes from random codes of the same dimension with high probability.

Suppose now that $C=\mathrm{Alt}_r(\boldsymbol{x},\boldsymbol{y})^\perp$ is the dual of an alternant (or, more specifically, a Goppa) code over $\mathbb{F}_q$ with field extension degree $m$. The authors of [20] prove upper bounds on the dimension of $C^{*2}$ in this case. Note that by Proposition 1 below, after field extension to $\mathbb{F}_{q^m}$(which does not affect the dimension of square codes), $C$ contains GRS codes. This is the reason why the dimension of the square code of $C$ is expected to be smaller than for random linear codes. However, for $m>1$ the square distinguisher works only for Goppa (or alternant) codes of high rate (so that its dual has low rate). This is because if we lower the rate of the alternant code, then $C^{*2}=\mathbb{F}_q^n$ soon fills the whole space in both the alternant/Goppa and the random case, so they cannot be distinguished.

The square distinguisher may be described equivalently as follows: Let $\boldsymbol{A}\in\mathbb{F}_q^{k\times n}$ be a matrix. We are interested in the case that $\boldsymbol{A}$ is the parity check matrix of an alternant/Goppa or of a random linear code. Let $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\in\mathbb{F}_q^k$ denote the columns of $\boldsymbol{A}$. Denoting by $\mathbb{F}_q[X_1,\ldots,X_k]^{(p)}$ the homogeneous polynomials of degree $p$, we define the map

$$\varphi_{\boldsymbol{A}}^{(2)}\colon \mathbb{F}_q[X_1,\ldots,X_k]^{(2)}\to\mathbb{F}_q^n,\quad f\mapsto(f(\boldsymbol{a}_1),\ldots,f(\boldsymbol{a}_n))$$

Note that the image of $\varphi_{\boldsymbol{A}}^{(2)}$ is precisely the square of the code with generator matrix $\boldsymbol{A}$. Thus, the square distinguisher essentially considers the rank (or, equivalently, the dimension of the kernel) of the map $\varphi_{\boldsymbol{A}}^{(2)}$.

A natural generalisation of this map is to not only consider quadratic polynomials in the domain, but also polynomials of higher degree. This is equivalent to considering $C^{*m}$ for $m\geq 3$. However, for our problem of distinguishing Goppa/alternant codes from random codes, this does not yield additional information compared to the square distinguisher: If the code $C^{*3}$ is a proper subset of $\mathbb{F}_q^n$ (which is required to be able to distinguish the alternant code from a random code), then so is $C^{*2}$. Thus the square distinguisher is already sufficient.

Instead of only considering simple vanishing of homogeneous polynomials at columns of $\boldsymbol{A}$, our main idea is to consider homogeneous polynomials of higher degree whose vanishing at the columns of $\boldsymbol{A}$ is *of higher order*. For example, we compute homogeneous polynomials of degree 3 that do not only themselves vanish at all columns of $\boldsymbol{A}$, but so do their first-order partial derivatives. More formally, our distinguisher (in degree 3) computes the dimension of the kernel of the map

$$\varphi_{\boldsymbol{A}}^{(3)}\colon \mathbb{F}_q[X_1,\ldots,X_k]^{(3)}\to\mathbb{F}_q^{n(k+1)},\quad f\mapsto\left[f(\boldsymbol{a}_j),\frac{\partial f}{\partial X_i}(\boldsymbol{a}_j)\right]_{i=1,\ldots,k}^{j=1,\ldots,n}$$

in the cases that $\boldsymbol{A}$ is the parity check matrix of an alternant/Goppa or random linear code. This can be extended to higher degrees: In degree $p$, our distinguisher computes the dimension of the space of homogeneous polynomials of degree $p$ whose partial derivatives of order up to $p-2$ all vanish when evaluated at the columns of $\boldsymbol{A}$.

Let us illustrate our approach by applying it to GRS codes in order to understand the core reason why our distinguisher works. The considerations below can be generalised to alternant and Goppa codes as we shall see in section 5.

We first consider the square distinguisher. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_{11}^8$ be such that the $x_i$ are pairwise distinct and $y_i \neq 0$ for all $i$. A generator matrix of the code $\mathrm{GRS}_3(\boldsymbol{x}, \boldsymbol{y})$ is given by

$$\boldsymbol{A}_1 \coloneqq \begin{pmatrix} y_1 & y_2 & \dots & y_8 \\ y_1x_1 & y_2x_2 & \dots & y_8x_8 \\ y_1x_1^2 & y_2x_2^2 & \dots & y_8x_8^2 \end{pmatrix}$$

The kernel of $\varphi^{(2)}_{\boldsymbol{A}_1}$ is non-trivial since, for instance, we have

$$f \coloneqq \det \begin{pmatrix} X_1 & X_2 \\ X_2 & X_3 \end{pmatrix} = X_1X_3 - X_2^2 \in \ker \varphi^{(2)}_{\boldsymbol{A}_1}$$

This is because $f(y_i, y_ix_i, y_ix_i^2) = y_i \cdot y_ix_i^2 - (y_ix_i)^2 = 0$ for all $1 \leq i \leq 8$. On the other hand, if $\boldsymbol{A}$ is a randomly selected $3 \times 8$ matrix of rank 3 over $\mathbb{F}_{11}$, it is likely that the kernel of $\varphi^{(2)}_{\boldsymbol{A}}$ is trivial. Hence GRS codes of dimension 3 and length 8 over $\mathbb{F}_{11}$ can be distinguished from randomly selected codes of rank 3 with high probability using $\varphi^{(2)}$.

Let us now consider the code $\mathrm{GRS}_5(\boldsymbol{x}, \boldsymbol{y}) \subset \mathbb{F}_{11}^8$. A generator matrix is given by

$$\boldsymbol{A}_2 \coloneqq \begin{pmatrix} y_1 & y_2 & \dots & y_8 \\ y_1x_1 & y_2x_2 & \dots & y_8x_8 \\ y_1x_1^2 & y_2x_2^2 & \dots & y_8x_8^2 \\ y_1x_1^3 & y_2x_2^3 & \dots & y_8x_8^3 \\ y_1x_1^4 & y_2x_2^4 & \dots & y_8x_8^4 \end{pmatrix}$$

Note that we cannot use $\varphi^{(2)}$ as a distinguisher since the square code of both a GRS code and of a random linear code of dimension 5 most likely fills the whole space $\mathbb{F}_{11}^8$. Thus, the kernel of $\varphi^{(2)}$ is likely to have the same dimension in both cases. However, $\varphi^{(3)}$ does yield a distinguisher: Consider the polynomial

$$g \coloneqq \det \begin{pmatrix} X_1 & X_2 & X_3 \\ X_2 & X_3 & X_4 \\ X_3 & X_4 & X_5 \end{pmatrix} \in \mathbb{F}_{11}[X_1, \dots, X_8]^{(3)}$$

Note that if we substitute $X_i \mapsto y_jx_j^i$ for some fixed $1 \leq j \leq 8$, the matrix

$$\begin{pmatrix} X_1 & X_2 & X_3 \\ X_2 & X_3 & X_4 \\ X_3 & X_4 & X_5 \end{pmatrix} \mapsto \begin{pmatrix} y_j & y_jx_j & y_jx_j^2 \\ y_jx_j & y_jx_j^2 & y_jx_j^3 \\ y_jx_j^2 & y_jx_j^3 & y_jx_j^4 \end{pmatrix}$$

has rank 1 since a row is obtained from the previous row via multiplication by $x_j$. So its determinant as well as all $2 \times 2$ minors vanish. This implies that $g$ vanishes when evaluated at any column of $\boldsymbol{A}_2$, *and so do all its first-order partial derivatives*! This is because any first-order partial derivative of the given determinant is a linear combination of its $2 \times 2$ minors, which in turn vanish when evaluated at any column of the generator matrix as just observed. Thus $g \in \ker \varphi^{(3)}_{\boldsymbol{A}_2}$ and so the kernel of $\varphi^{(3)}_{\boldsymbol{A}_2}$ is non-trivial. On the other hand, one can experimentally verify that if $\boldsymbol{A}$ is a randomly selected $5 \times 8$ matrix of rank 5, then the kernel of $\varphi^{(3)}_{\boldsymbol{A}}$ is likely to be trivial. We will prove in Proposition 3 that, in fact, the dimension of the kernel of $\varphi^{(3)}_{\boldsymbol{A}}$ is independent of the choice of generator matrix $\boldsymbol{A}$. Hence $\varphi^{(3)}$ yields a distinguisher between GRS codes and randomly selected linear codes over $\mathbb{F}_{11}$ of dimension 5 and length 8.

*Summary of results.* We propose a new approach for distinguishing Goppa (and, more generally, alternant) codes from random linear codes using higher-order vanishing. We analyse our distinguisher and explain its behaviour in the Goppa (or alternant) and random case. Our experiments demonstrate the effectiveness of the distinguisher for concrete parameter sets. To our knowledge, this approach is capable of distinguishing larger parameter sets in practice than previous approaches. Based on our analysis, we predict the complexity of distinguishing binary Goppa codes with parameters proposed in the Classic McEliece key encapsulation mechanism using our distinguisher. For example, for the smallest Classic McEliece parameter set, we provide a (very rough) complexity estimate of $2^{310}$, compared to $2^{401}$ for the syzygy distinguisher [23]. Thus, although still too large to distinguish these Goppa codes in practice, the complexity is lower than previous approaches.

*Organisation of this paper.* In section 2, we list some notation that will be used throughout the paper as well as basic definitions and results in coding theory. In section 3, we introduce a map that encodes simultaneous higher-order vanishing of polynomials at a given set of points and prove some basic properties. This is the central object of study in this paper. In section 4 and section 5, we analyse its behaviour for random and alternant/Goppa codes, respectively. These results allow us to specify our distinguisher and predict its behaviour in section 6. In subsection 6.1, we describe experimental results for concrete parameter sets. In subsection 6.2, we estimate the complexity for Classic McEliece parameter sets. We close with some remarks on similarities and differences compared to the syzygy distinguisher from [23] in subsection 6.3.

## 2 Preliminaries

Throughout this paper, let $q$ be a prime power. In subsection 2.1, we introduce some notation for linear algebra and for (polynomial) rings and their ideals which will be used throughout the paper. In subsection 2.2, we recall the definitions of GRS codes as well as alternant and Goppa codes and state their basic properties that will be used in the analysis of our distinguisher for these codes in section 5.

### 2.1 Notation for linear algebra and (polynomial) rings

We denote matrices by boldface upper-case letters and vectors by boldface lower-case letters. We denote the set of $k \times n$ matrices over $\mathbb{F}_q$ by $\mathbb{F}_q^{k\times n}$. Given some subset $U$ of a $\mathbb{F}_q$-vector space, we denote its $\mathbb{F}_q$-linear span by $\mathrm{span}_{\mathbb{F}_q}(U)$.

For any commutative ring $S$ and elements $s_1, \ldots, s_\ell \in S$, we denote by $\langle s_1, \ldots, s_\ell \rangle \subset S$ the ideal generated by these elements. For $p \in \mathbb{N}$, we denote the $p$th power of an ideal $\mathfrak{I} \subset S$ by $\mathfrak{I}^p \coloneqq \langle s_1 s_2 \cdots s_p \mid s_1, \ldots, s_p \in \mathfrak{I} \rangle$. If $S$ is a graded ring and $\mathfrak{I} \subset S$ is a homogeneous ideal, for any $p \in \mathbb{N}_0$ we denote by $\mathfrak{I}^{(p)}$ the subset of elements in $\mathfrak{I}$ of degree $p$.

Given an integer $k$, we denote by $R_k \coloneqq \mathbb{F}_q[X_1, \ldots, X_k]$ the polynomial ring in $k$ indeterminates over $\mathbb{F}_q$, equipped with its natural grading where each indeterminate has degree 1. For any non-negative integer $p$, we denote by $R_k^{(p)} \subset R_k$ the vector space (over $\mathbb{F}_q$) of homogeneous polynomials of degree $p$. Sometimes it will be convenient to work in the quotient ring $\tilde{R}_k \coloneqq R_k / \langle X_1^2, \ldots, X_k^2 \rangle$. Note that it inherits a grading from $R_k$.

We will need differential operators acting on polynomials. Given a monomial $g = X_{i_1} \cdots X_{i_\ell} \in R_k$ and any $f \in R_k$, we denote by $\mathcal{D}_g f \coloneqq \frac{\partial^\ell f}{\partial X_{i_1} \cdots \partial X_{i_\ell}}$ the partial derivative of $f$ by the indeterminates occurring in $g$ (with multiplicity). It will be convenient to (linearly) extend this definition beyond monomials in the following way: For any $g = \sum c_{i_1,\ldots,i_\ell} X_{i_1} \cdots X_{i_\ell} \in R_k$ where $c_{i_1,\ldots,i_\ell} \in \mathbb{F}_q$, we define

$$\mathcal{D}_g f \coloneqq \sum c_{i_1,\ldots,i_\ell} \cdot (\mathcal{D}_{X_{i_1}\cdots X_{i_\ell}} f) = \sum c_{i_1,\ldots,i_\ell} \cdot \frac{\partial^\ell f}{\partial X_{i_1} \cdots \partial X_{i_\ell}}$$

For example, we have $\mathcal{D}_{X_1X_2 - X_3^2} f = \frac{\partial^2 f}{\partial X_1 \partial X_2} - \frac{\partial^2 f}{\partial X_3^2}$.

### 2.2 Linear codes

Recall that a linear $[n, k]$-code over $\mathbb{F}_q$ is a $k$-dimensional subspace $C \subset \mathbb{F}_q^n$. In this work, we are mainly interested in alternant and Goppa codes, which are closely related to generalised Reed-Solomon codes. For convenience, we recall their definitions.

**Definition 1.** *Suppose $x_i, y_i \in \mathbb{F}_q$ for $1 \le i \le n$ such that the $x_i$ are pairwise distinct and $y_i \neq 0$ for all $i$. The* generalised Reed-Solomon (GRS) code *of dimension $r$ is defined by*

$$\mathrm{GRS}_r(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \{(y_1 f(x_1), \ldots, y_n f(x_n)) \mid f \in \mathbb{F}_q[X],\ \deg(f) < r\}$$

Alternant codes are defined by taking a GRS code over a larger field $\mathbb{F}_{q^m}$ and restricting to those vectors whose entries all lie in the subfield $\mathbb{F}_q$. Thus, an alternant code is a subfield subcode of a GRS code:

**Definition 2.** *Let $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ have distinct entries and $\boldsymbol{y} \in (\mathbb{F}_{q^m}^*)^n$. The* alternant code *with* support $\boldsymbol{x}$ *and* multiplier $\boldsymbol{y}$ *and* degree $r$ *over* $\mathbb{F}_q$ *is*

$$\mathrm{Alt}_r(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \mathrm{GRS}_r(\boldsymbol{x}, \boldsymbol{y})^\perp \cap \mathbb{F}_q^n$$

*Such an alternant code is* proper *if it has dimension $n - mr$.*

Note that in the definition, we define alternant codes to be the subfield subcode of the *dual* of a GRS code. However, this is mere convention since duals of GRS codes are again GRS codes with the same support but a different multiplier [16, Ch. 10, Theorem 4].

For the purposes of our distinguisher, we only need the following property of alternant codes: If we take the dual of an alternant code and extend scalars to the extension field $\mathbb{F}_{q^m}$, the resulting code decomposes as a sum of GRS codes:

**Proposition 1.** *Let $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ have distinct entries, $\boldsymbol{y} \in (\mathbb{F}_{q^m}^*)^n$, and $\mathrm{Alt}_r(\boldsymbol{x}, \boldsymbol{y})$ be an alternant code over $\mathbb{F}_q$. Then*

$$\mathrm{span}_{\mathbb{F}_{q^m}}\left(\mathrm{Alt}_r(\boldsymbol{x}, \boldsymbol{y})^\perp\right) = \sum_{i=0}^{m-1} \mathrm{GRS}_r\left(\boldsymbol{x}^{q^i}, \boldsymbol{y}^{q^i}\right)$$

*where $\boldsymbol{c}^{q^i} \coloneqq (c_1^{q^i}, \ldots, c_n^{q^i})$ denotes component-wise powers for any $\boldsymbol{c} \in \mathbb{F}_{q^m}^n$.*

*Proof.* See [5, Proposition 14]. □

Goppa codes are alternant codes with a special choice of multiplier:

**Definition 3.** *Let $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ have distinct entries and $g \in \mathbb{F}_{q^m}[X]$ be a polynomial of degree $t$ with $g(x_i) \neq 0$ for all $1 \leq i \leq n$. The* Goppa code *with* support $\boldsymbol{x}$ *and* Goppa polynomial $g$ *is defined as* $\mathrm{Gop}(\boldsymbol{x}, g) \coloneqq \mathrm{Alt}_t(\boldsymbol{x}, \boldsymbol{y})$ *where* $\boldsymbol{y} \coloneqq \frac{1}{g(\boldsymbol{x})} \coloneqq \left(\frac{1}{g(x_1)}, \ldots, \frac{1}{g(x_n)}\right)$.

We will mostly be interested in *binary* Goppa codes, i.e. those with $q = 2$. To prove a refined version of Proposition 1 in this case, we need a lemma.

**Lemma 1.** *Suppose $g \in \mathbb{F}_{2^m}[X]$ is square-free of degree $t \leq \frac{n}{2}$ and let $\boldsymbol{x} \in \mathbb{F}_{2^m}^n$ such that the entries $x_i$ are pairwise distinct and $g(x_i) \neq 0$ for all $1 \leq i \leq n$. Let $\boldsymbol{y} = g(\boldsymbol{x})^{-1} \in \mathbb{F}_{2^m}^n$. Then*

$$\mathrm{GRS}_t(\boldsymbol{x}, \boldsymbol{y}) + \mathrm{GRS}_t(\boldsymbol{x}^2, \boldsymbol{y}^2) = \mathrm{GRS}_{2t}(\boldsymbol{x}, \boldsymbol{y}^2)$$

*Proof.* It is clear that $\mathrm{GRS}_t(\boldsymbol{x}^2, \boldsymbol{y}^2) \subset \mathrm{GRS}_{2t}(\boldsymbol{x}, \boldsymbol{y}^2)$. For any $\boldsymbol{y}f(\boldsymbol{x}) \in \mathrm{GRS}_t(\boldsymbol{x}, \boldsymbol{y})$ where $f \in \mathbb{F}_{2^m}[X]$ is of degree $< t$, we have $\boldsymbol{y}f(\boldsymbol{x}) = \boldsymbol{y}^2 \cdot g(\boldsymbol{x})f(\boldsymbol{x})$ which lies in $\mathrm{GRS}_{2t}(\boldsymbol{x}, \boldsymbol{y}^2)$ since $fg$ is a polynomial of degree $< 2t$. This proves that $\mathrm{GRS}_t(\boldsymbol{x}, \boldsymbol{y}) + \mathrm{GRS}_t(\boldsymbol{x}^2, \boldsymbol{y}^2) \subset \mathrm{GRS}_{2t}(\boldsymbol{x}, \boldsymbol{y}^2)$.

Now it is sufficient to show that $\mathrm{GRS}_t(\boldsymbol{x}, \boldsymbol{y}) \cap \mathrm{GRS}_t(\boldsymbol{x}^2, \boldsymbol{y}^2) = \{0\}$. So let $\boldsymbol{c} \in \mathrm{GRS}_t(\boldsymbol{x}, \boldsymbol{y}) \cap \mathrm{GRS}_t(\boldsymbol{x}^2, \boldsymbol{y}^2)$. Then there exist polynomials $A, B \in \mathbb{F}_{2^m}[X]$ of degree $< t$ such that

$$\frac{A(\boldsymbol{x})}{g(\boldsymbol{x})} = \boldsymbol{c} = \frac{B(\boldsymbol{x}^2)}{g(\boldsymbol{x})^2} = \frac{B(\boldsymbol{x})^2}{g(\boldsymbol{x})^2}$$

Hence $g$ divides $B^2$. Since $g$ is square free, it must divide $B$. But $\deg(g) = t$ and $\deg(B) < t$, so $B = 0$. Hence $\boldsymbol{c} = 0$ as required. □

We deduce a refined version of Proposition 1 for binary Goppa codes, which will be important in the analysis of the behaviour of our distinguisher for binary Goppa codes:

**Proposition 2.** *Let $\boldsymbol{x} \in \mathbb{F}_{2^m}^n$ have pairwise distinct entries and $C = \mathrm{Gop}(\boldsymbol{x}, g)^\perp$ be the dual of a proper binary Goppa code with square-free Goppa polynomial $g$ of degree $t$. Set $\boldsymbol{y} = g(\boldsymbol{x})^{-1}$. Then*

$$\mathrm{span}_{\mathbb{F}_{q^m}}(C) = \bigoplus_{i=0}^{m/2-1} \mathrm{GRS}_{2t}(\boldsymbol{x}^{4^i}, \boldsymbol{y}^{2\cdot 4^i}), \text{ or}$$

$$\mathrm{span}_{\mathbb{F}_{q^m}}(C) = \left( \bigoplus_{i=0}^{(m-1)/2-1} \mathrm{GRS}_{2t}(\boldsymbol{x}^{4^i}, \boldsymbol{y}^{2\cdot 4^i}) \right) \oplus \mathrm{GRS}_t(\boldsymbol{x}^{2^{m-1}}, \boldsymbol{y}^{2^{m-1}})$$

*depending on whether $m$ is even or odd.*

*Proof.* This follows immediately from Proposition 1 and Lemma 1. □

## 3 Higher-order vanishing of polynomials

In subsection 3.1, we introduce a linear map that encodes simultaneous higher-order vanishing of polynomials at points given by the columns of some matrix. (In section 4 and section 5, we will take this matrix to be a parity check matrix of a random general linear or alternant/Goppa code.) In subsection 3.2, we make some observations leading to a slightly simplified version of this map. This will be the central object of study for the rest of the paper and at the heart of our distinguisher. In subsection 3.3, we describe an alternative point of view of higher-order vanishing using homogeneous vanishing ideals. Whilst the material in subsection 3.3 is not strictly necessary to understand the rest of the paper, it points out a connection to objects that are quite common in commutative algebra.

### 3.1 Defining higher-order vanishing

Recall that we denote by $R_k^{(p)}$ the $\mathbb{F}_q$-vector space of homogeneous polynomials of degree $p$ in $R_k = \mathbb{F}_q[X_1, \ldots, X_k]$.

**Definition 4.** *Let $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ and denote by $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{F}_q^k$ the columns of $\boldsymbol{A}$. We define the $\mathbb{F}_q$-linear map*

$$\varphi_{\boldsymbol{A}}^{(p)} \colon R_k^{(p)} \to \prod_{j=1}^{n} \mathbb{F}_q^{\binom{k+p-2}{p-2}}, \quad f \mapsto [(\mathcal{D}_\nu f)(\boldsymbol{a}_j)]_\nu^{j=1,\ldots,n}$$

*where $\nu$ runs through all monomials in $k$ indeterminates of degree $0 \le d \le p-2$. (Note that $\binom{k+p-2}{p-2}$ is the number of such monomials.)*

In other words, $\varphi_{\boldsymbol{A}}^{(p)}$ computes all partial derivatives of order at most $p-2$ of a given homogeneous polynomial $f \in R_k^{(p)}$ and evaluates them at all columns of $\boldsymbol{A}$. Note that for $p = 2$, the map $\varphi_{\boldsymbol{A}}^{(2)}$ simply evaluates homogeneous quadratic polynomials at all columns of $\boldsymbol{A}$. Hence, computing the dimension of the kernel of $\varphi_{\boldsymbol{A}}^{(2)}$ in the cases where $\boldsymbol{A}$ is a random matrix or a parity check matrix of an alternant or Goppa code is equivalent to the square distinguisher [21]. For general $p$, the kernel of $\varphi_{\boldsymbol{A}}^{(p)}$ contains precisely those polynomials of degree $p$ that simultaneously vanish of order $p-1$ at all points $\boldsymbol{a}_i$.

*Remark 1.* For any homogeneous polynomial $f \in R_k^{(u)}$ of degree $u$, it is easy to see that $\sum_{i=1}^{k} X_i \cdot \frac{\partial f}{\partial X_i} = u \cdot f$. This is sometimes referred to as *Euler's theorem*. Suppose $\mathrm{char}(\mathbb{F}_q) > p$. The above identity implies that if all partial derivatives of order $p-2$ of $f \in R_k^{(p)}$ vanish at a given point, then so do all partial derivatives of order $0 \leq d < p-2$. Thus, since we will mainly be interested in the kernel, we may simplify the definition of $\varphi_{\boldsymbol{A}}^{(p)}$ by only computing the partial derivatives of order $p-2$ and evaluating them at all columns of $\boldsymbol{A}$. Similarly, in lower characteristic, some partial derivatives are superfluous, but others are not. For example, if $\mathrm{char}(\mathbb{F}_q) = 2$ and $f \in R_k^{(3)}$, then vanishing of first-order partial derivatives of $f$ at a given point implies vanishing of $f$ by Euler's theorem, but this need not be true e.g. for $f \in R_k^{(4)}$. We will analyse this more closely in section 4. For now, we remark that for concrete computations involving $\varphi_{\boldsymbol{A}}^{(p)}$, it makes sense to omit the computation of partial derivatives that are superfluous in the given characteristic.

*Remark 2.* Instead of using ordinary partial derivatives in Definition 4, one could also use *Hasse derivatives* [17]. These are a slightly adapted version of ordinary differentiation with nicer properties in positive characteristic. For example, the fact that a polynomial is constant if and only if all its derivatives vanish is true for Hasse derivatives regardless of the characteristic of the ground field, but not for ordinary derivatives: For example, in characteristic 2, all derivatives of $X^2$ vanish, but $X^2$ is clearly not a constant polynomial. However, we stick to more familiar ordinary differentiation here, which yields equivalent results in the end.

In subsequent sections, we will analyse the dimension of the kernel of $\varphi_{\boldsymbol{A}}^{(p)}$ in the cases where $\boldsymbol{A}$ is the parity check matrix of a random alternant (or Goppa) code or of a random general linear code. Proposition 3 shows that this dimension does not depend on the choice of the parity check matrix. This will be crucial for our distinguisher because it means that the dimension of the kernel of $\varphi_{\boldsymbol{A}}^{(p)}$ is an invariant of the underlying code.

**Proposition 3.** *Suppose $\boldsymbol{A}, \boldsymbol{A}' \in \mathbb{F}_q^{k \times n}$ are row-equivalent. Then*

$$\dim_{\mathbb{F}_q}\left(\ker\left(\varphi_{\boldsymbol{A}}^{(p)}\right)\right) = \dim_{\mathbb{F}_q}\left(\ker\left(\varphi_{\boldsymbol{A}'}^{(p)}\right)\right)$$

*Proof.* Denoting by $\boldsymbol{X} = (X_1, \ldots, X_k)$ the row vector with indeterminates as entries, we claim that

$$\mathcal{D}_{\nu(\boldsymbol{X})}(f(\boldsymbol{X} \cdot \boldsymbol{S}^T)) = (\mathcal{D}_{\nu(\boldsymbol{X}\boldsymbol{S})}(f))(\boldsymbol{X} \cdot \boldsymbol{S}^T) \tag{1}$$

holds for any $\boldsymbol{S} \in \mathbb{F}_q^{k \times k}$ and $\nu, f \in R_k$. Note that since $\mathcal{D}_{\nu_1 \cdot \nu_2} = \mathcal{D}_{\nu_1} \circ \mathcal{D}_{\nu_2}$ and $\mathcal{D}_{\alpha\nu_1 + \beta\nu_2} = \alpha \cdot \mathcal{D}_{\nu_1} + \beta \cdot \mathcal{D}_{\nu_2}$ for all $\nu_1, \nu_2 \in R_k$ and $\alpha, \beta \in \mathbb{F}_q$, it suffices to prove the claim in the case that $\nu = X_j$ for some $j$. In this case, (1) reduces to the ordinary chain rule for differentiation.

Letting $\boldsymbol{S}$ be an invertible matrix such that $\boldsymbol{A}' = \boldsymbol{S}\boldsymbol{A}$, we deduce from (1) that the map $\ker\left(\varphi^{(p)}_{\boldsymbol{A}'}\right) \to \ker\left(\varphi^{(p)}_{\boldsymbol{A}}\right), f(\boldsymbol{X}) \mapsto f(\boldsymbol{X}\boldsymbol{S}^T)$ is well-defined, and it is an isomorphism since $\boldsymbol{S}$ is invertible. □

### 3.2 Reduction to multilinear polynomials

By Proposition 3, the dimension of the kernel of $\varphi^{(p)}_{\boldsymbol{A}}$ is invariant under row operations on the matrix $\boldsymbol{A}$. Let $\boldsymbol{A}'$ be the (unique) reduced row-echelon form of $\boldsymbol{A}$. Assuming $\boldsymbol{A}$ has full rank $k$, there are $k$ pivot columns $\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_k}$ of $\boldsymbol{A}'$ that are equal to the standard basis vectors in $\mathbb{F}_q^k$. Lemma 2 below determines the polynomials of higher-order vanishing at the standard basis vectors. This will allow us to modify the definition of $\varphi^{(p)}_{\boldsymbol{A}}$ to decrease the size of its domain and range while still capturing the interesting part of its kernel. First, we need a definition.

**Definition 5.** *We define*

$$U_k^{(p)} = \operatorname{span}_{\mathbb{F}_q}\left\{ X_1^{d_1} \cdots X_k^{d_k} \mid d_i \leq 1 \forall i \quad \textit{and} \sum_{i=1}^{k} d_i = p \right\}$$

*to be the subspace spanned by monomials in $R_k^{(p)}$ where each indeterminate occurs to degree at most 1.*

**Lemma 2.** *Let $\boldsymbol{I}_k$ denote the $k \times k$ identity matrix over $\mathbb{F}_q$ and $c \coloneqq \operatorname{char}(\mathbb{F}_q)$. For any $p \geq 2$, we have*

$$\ker\left(\varphi^{(p)}_{\boldsymbol{I}_k}\right) = U_k^{(p)} \oplus \langle X_i^c X_j^c \mid 1 \leq i < j \leq k \rangle^{(p)}$$

*where we recall that $\langle \cdot \rangle^{(p)}$ denotes the $\mathbb{F}_q$-subspace of homogeneous elements of degree $p$ in the ideal of $R_k$ generated by the indicated monomials.*

*Proof.* It is clear that $U_k^{(p)} \cap \langle X_i^c X_j^c \mid 1 \leq i < j \leq k \rangle^{(p)} = \{0\}$ and easy to verify that $U_k^{(p)} \oplus \langle X_i^c X_j^c \mid 1 \leq i < j \leq k \rangle^{(p)} \subset \ker\left(\varphi^{(p)}_{\boldsymbol{I}_k}\right)$.

Now suppose that $f \in \ker\left(\varphi^{(p)}_{\boldsymbol{I}_k}\right)$. Consider a monomial of the form

$$X_{i_0}^{d_0} X_{i_1}^{d_1} \cdots X_{i_\ell}^{d_\ell} \in R_k^{(p)}$$

with $i_0, \ldots, i_\ell$ pairwise distinct, $d_0 \geq 2$ and $d_1, \ldots, d_\ell < c$. Note that any monomial which contains an indeterminate of degree at least 2 and contains at most one indeterminate of degree $\geq c$ is of this form. Thus, showing that the coefficient of any such monomial in $f$ is zero implies that $f \in U_k^{(p)} \oplus \langle X_i^c X_j^c \mid 1 \leq i < j \leq k \rangle^{(p)}$ and thus concludes our proof. Suppose such a monomial occurs with coefficient $\lambda \in \mathbb{F}_q$ in $f$. Then we have

$$0 = \frac{\partial^{d_1+\ldots+d_\ell} f}{\partial X_{i_1}^{d_1} \cdots \partial X_{i_\ell}^{d_\ell}}(\boldsymbol{e}_{i_0}) = \lambda \cdot d_1! \cdots d_\ell!$$

where we used the fact that $f \in \ker\left(\varphi_{\boldsymbol{I}_k}^{(p)}\right)$ for the first equality. Since $d_1, \ldots, d_\ell < c$, this implies that $\lambda = 0$ as required. □

We clearly have $\ker\left(\varphi_{\boldsymbol{A}'}^{(p)}\right) \subset \ker\left(\varphi_{\boldsymbol{I}_k}^{(p)}\right)$. The space $\langle X_i^c X_j^c \mid 1 \leq i < j \leq k \rangle^{(p)}$ consists of elements that lie in the kernel purely due to the characteristic of our ground field, irrespective of the value of $\boldsymbol{A}'$ and its entries. Thus, for our distinguisher, we are only interested in the part of the kernel that lies in $U_k^{(p)}$. Note that if we differentiate any polynomial in $U_k^{(p)}$ by some indeterminate more than once, this yields the zero polynomial. Hence it suffices to consider only partial derivatives that differentiate by any given indeterminate at most once. This motivates the following slightly adapted definition. It reduces the size of the domain and range compared to $\varphi_{\boldsymbol{A}}^{(p)}$, but still captures the interesting part of its kernel.

**Definition 6.** *Let $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ be a matrix of rank $k$. Let $\tilde{\boldsymbol{A}} \in \mathbb{F}_q^{k \times (n-k)}$ be the matrix obtained from $\boldsymbol{A}$ by first computing the (unique) reduced row-echelon form of $\boldsymbol{A}$ and then removing the $k$ columns that contain the pivot entries, i.e. that contain the leading 1 in a row. Let $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_{n-k}$ denote the columns of $\tilde{\boldsymbol{A}}$. We define*

$$\tilde{\varphi}_{\boldsymbol{A}}^{(p)} \colon U_k^{(p)} \to \prod_{j=1}^{n-k} \mathbb{F}_q^{M_{k,p-2}}, \quad f \mapsto [(\mathcal{D}_\nu f)(\tilde{\boldsymbol{a}}_j)]_\nu^{j=1,\ldots,n-k}$$

*where $\nu$ runs through all monomials of degrees $0 \leq d \leq p-2$ where each indeterminate has degree at most 1, and $M_{k,p-2} \coloneqq \sum_{d=0}^{p-2} \binom{k}{d}$ is the number of such monomials.*

Recall that as pointed out in Remark 1, depending on the characteristic of the ground field, for computing the kernel we may omit partial derivatives of certain orders in the definition above.

### 3.3 Definition of higher-order vanishing using vanishing ideals

An alternative way to define higher-order vanishing is via vanishing ideals. We define the *homogeneous vanishing ideal at a point* $\boldsymbol{c} \in \mathbb{F}_q^k$ by

$$\mathfrak{a}_{\boldsymbol{c}} = \langle c_j X_i - c_i X_j \mid 1 \leq i < j \leq k \rangle \subset R_k$$

Given a matrix $\boldsymbol{A} \in \mathbb{F}_q^{k\times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{F}_q^k$, we consider the $(p-1)$-st powers $\mathfrak{a}_{\boldsymbol{a}_i}^{p-1}$ of the vanishing ideals and let

$$\phi_{\boldsymbol{A}} \colon R_k \to R_k/\mathfrak{a}_{\boldsymbol{a}_1}^{p-1} \times \cdots \times R_k/\mathfrak{a}_{\boldsymbol{a}_n}^{p-1}$$

be the map induced by the canonical projections. Note that the kernel of $\phi_{\boldsymbol{A}}$ is given by the ideal $\mathfrak{a}_{\boldsymbol{a}_1}^{p-1} \cap \ldots \cap \mathfrak{a}_{\boldsymbol{a}_n}^{p-1}$ which we call the *vanishing ideal of order* $p-1$ *at* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Properties of this ideal have been studied in the literature, see for example [11]. The maps $\varphi_{\boldsymbol{A}}^{(p)}$ and $\phi_{\boldsymbol{A}}^{(p)}$ (by which we denote the degree $p$ component of the graded linear map $\phi_{\boldsymbol{A}}$) are closely related:

**Proposition 4.** *Suppose* $\mathrm{char}(\mathbb{F}_q) > p$ *and let* $\boldsymbol{c} \in \mathbb{F}_q^k \backslash \{0\}$. *The kernel of the map*

$$\varphi_{\boldsymbol{c}}^{(p)} \colon R_k^{(p)} \to \mathbb{F}_q^{\binom{k+p-3}{p-2}}, \quad f \mapsto [\mathcal{D}_\nu f(\boldsymbol{c})]_\nu$$

*where* $\nu$ *runs through all* $\binom{k+p-3}{p-2}$ *monomials of degree* $p-2$, *is given by* $\left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)}$, *i.e. the subspace of homogeneous elements of degree* $p$ *in the ideal* $\mathfrak{a}_{\boldsymbol{c}}^{p-1}$.

Recall that by Remark 1, since $\mathrm{char}(\mathbb{F}_q) > p$, it does not make a difference in the above statement whether we consider all partial derivatives of order $\leq p-2$ or only those of order equal to $p-2$.

*Proof.* Without loss of generality, we may assume $c_1 = 1$. The homogeneous linear polynomials $g_i = X_i - c_i X_1$ with $2 \leq i \leq k$ then generate the ideal $\mathfrak{a}_{\boldsymbol{c}}$. As a $\mathbb{F}_q$-vector space, $\left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)}$ is spanned by elements of the form $f g_{i_1} \cdots g_{i_{p-1}}$ where $2 \leq i_1, \ldots, i_{p-1} \leq k$ and $f \in R_k^{(1)}$. It follows directly from the product rule for differentiation and the fact that $g_i(\boldsymbol{c}) = 0$ for all $i$ that any partial derivative of order $p-2$ of such an element vanishes at $\boldsymbol{c}$. This implies that $\left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)} \subset \ker\left(\varphi_{\boldsymbol{c}}^{(p)}\right)$.

We now prove $\ker\left(\varphi_{\boldsymbol{c}}^{(p)}\right) \subset \left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)}$ by induction on $p$. For $p = 2$, the claim follows immediately from the definition of $\mathfrak{a}_{\boldsymbol{c}}$. Now let $f \in \ker\left(\varphi_{\boldsymbol{c}}^{(p)}\right)$ for some $p > 2$. Then $\frac{\partial f}{\partial X_i} \in \ker\left(\varphi_{\boldsymbol{c}}^{(p-1)}\right)$ for all $i$, so by induction, $\frac{\partial f}{\partial X_i} \in \mathfrak{a}_{\boldsymbol{c}}^{p-2}$ for all $i$. By Euler's identity, $\sum_i X_i \cdot \frac{\partial f}{\partial X_i} = p \cdot f$, so since $\mathrm{char}(\mathbb{F}_q) > p$, we deduce that $f \in \left(\mathfrak{a}_{\boldsymbol{c}}^{p-2}\right)^{(p)}$. Thus we can write

$$f = h_1 + X_1 \cdot h_2 + X_1^2 \cdot \sum_{2 \leq i_1 \leq \ldots \leq i_{p-2} \leq k} \lambda_{i_1,\ldots,i_{p-2}} \cdot g_{i_1} \cdots g_{i_{p-2}}$$

for some $h_1 \in \left(\mathfrak{a}_{\boldsymbol{c}}^{p}\right)^{(p)}$ and $h_2 \in \left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p-1)}$ and $\lambda_{i_1,\ldots,i_{p-2}} \in \mathbb{F}_q$. Then for any $2 \leq i_1 \leq \ldots \leq i_{p-2} \leq k$, we have

$$\frac{\partial^{p-2} f}{\partial X_{i_1} \cdots \partial X_{i_{p-2}}}(\boldsymbol{c}) = \lambda_{i_1,\ldots,i_{p-2}} \cdot \mu_2! \mu_3! \cdots \mu_k!$$

where $\mu_i$ is the number of elements among $i_1, \ldots, i_{p-2}$ that are equal to $i$. This must vanish because $f \in \ker\left(\varphi_{\boldsymbol{c}}^{(p)}\right)$. Since $\mathrm{char}\,(\mathbb{F}_q) > p$, this implies that $\lambda_{i_1,\ldots,i_{p-2}} = 0$ for all $2 \leq i_1 \leq \ldots \leq i_{p-2} \leq k$ and hence $f \in \left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)}$ as claimed. □

Proposition 4 implies that in large characteristic and for any matrix $\boldsymbol{A}$, the maps $\varphi_{\boldsymbol{A}}^{(p)}$ and $\phi_{\boldsymbol{A}}^{(p)}$ have the same kernel. In lower characteristic, one has to be a bit more careful. For example if $\mathrm{char}(\mathbb{F}_q) = 2$, the polynomial $X_1^{p-2} \cdot (X_2 - c_2 X_1)^2$ lies in the kernel of $\varphi_{\boldsymbol{A}}^{(p)}$ for any $p \geq 2$. But it is not an element of $\left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)}$ for $p \geq 4$. However, following an argument similar to the proof of Proposition 4, one can show the following: If we restrict $\varphi_{\boldsymbol{A}}^{(p)}$ to the subspace $U_k^{(p)}$ of homogeneous polynomials of degree $p$ where each indeterminate occurs to degree at most 1, its kernel is equal to $\left(\mathfrak{a}_{\boldsymbol{c}}^{p-1}\right)^{(p)} \cap U_k^{(p)}$.

We will analyse the case of characteristic 2 more closely further below. In the following, we will only work with the definition of higher-order vanishing via differentials, i.e. with $\varphi_{\boldsymbol{A}}^{(p)}$ and $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$. However, we will make use of ideas from works studying the higher-order vanishing ideal.

## 4 Higher-order vanishing at randomly selected points

In order to analyse the rank of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ for a randomly selected matrix $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ of rank $k$, we show in subsection 4.1 that it can be interpreted as the vector space dimension of the degree $p$ component of a certain homogeneous ideal of polynomials. This will yield an equivalent definition of our distinguisher that does not make reference to differentiation, and it will turn out in subsection 4.2 that this viewpoint is more convenient in order to analyse the behaviour in the random case. Our considerations culminate in Heuristic 1, which we will use in our distinguisher to predict the rank of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ for random matrices $\boldsymbol{A}$.

Results making this sort of connection have already been described in the literature. [11, Theorem I] establishes a connection between higher-order vanishing ideals and ideals of certain polynomials that we also consider in the following, using what is called *inverse systems of ideals*. This is basically the content of Proposition 5 below. We shall give a self-contained exposition of our results in this section, but remark that it is heavily inspired by this work.

Since we only consider polynomials in $U_k^{(p)}$, it will be convenient to work in the quotient ring $\tilde{R}_k = R_k / \langle X_1^2, \ldots, X_k^2 \rangle$.

### 4.1 Interpreting higher-order vanishing in terms of polynomial ideals

The following polynomials are central to making the link between our map $\varphi$, defined in terms of differentials, and polynomial ideals.

**Definition 7.** *For $\boldsymbol{c} = (c_1, \ldots, c_k) \in \mathbb{F}_q^k$ and every integer $a \geq 1$ we define the polynomial*

$$L_{\boldsymbol{c}}^{[a]} \coloneqq \sum_{1 \leq i_1 < \ldots < i_a \leq k} c_{i_1} \cdots c_{i_a} X_{i_1} \cdots X_{i_a} \in \tilde{R}_k^{(a)}$$

*For any $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{F}_q^k$, we define the ideal*

$$\mathfrak{L}_{\boldsymbol{A}} \coloneqq \langle L_{\boldsymbol{a}_i}^{[j]} \mid 1 \leq i \leq n,\ j \geq 2 \rangle \subset \tilde{R}_k$$

These polynomials are multiplicatively related in the following way:

**Lemma 3.** *For any $\boldsymbol{c} \in \mathbb{F}_q^k$ and $a, b \geq 1$, the following identity holds in $\tilde{R}_k$:*

$$L_{\boldsymbol{c}}^{[a]} \cdot L_{\boldsymbol{c}}^{[b]} = \binom{a+b}{a} \cdot L_{\boldsymbol{c}}^{[a+b]}$$

*Proof.* Note that all monomials in the product $L_{\boldsymbol{c}}^{[a]} \cdot L_{\boldsymbol{c}}^{[b]}$ that contain an indeterminate of degree greater than 1 vanish in $\tilde{R}_k$. Any monomial of degree $a+b$ such that every indeterminate is of degree at most 1 occurs as many times as there are partitions of the set of occurring indeterminates in a set of size $a$ and a set of size $b$. There are $\binom{a+b}{a}$ such partitions. This implies the claim. □

The following simple observation is central because it allows us to rephrase evaluation of a homogeneous polynomial at a point (occurring in the definition of $\varphi_{\boldsymbol{A}}^{(p)}$) as differentiation with respect to a polynomial defined in Definition 7.

**Lemma 4.** *For any $\boldsymbol{c} \in \mathbb{F}_q^k$ and $\nu \in \tilde{R}_k^{(p-a)}$ and $f \in U_k^{(p)}$, the following holds:*

$$\mathcal{D}_{\nu \cdot L_{\boldsymbol{c}}^{[a]}} f = (\mathcal{D}_\nu f)(\boldsymbol{c}) \tag{2}$$

*Remark 3.* Note that the above expressions are in fact well-defined: Since $f \in U_k^{(p)}$, we have $\mathcal{D}_{X_i^2} f = 0$ for all $i$ and hence the expression $\mathcal{D}_\mu f$ is well-defined for $\mu \in \tilde{R}_k$ (and not only for $\mu \in R_k$). Furthermore, note that the expression on the left in (2) is strictly speaking an element in $R_k^{(0)}$, which we identify with $\mathbb{F}_q$, and the quantity to the right of the equality is an element of $\mathbb{F}_q$ by definition.

*Proof (of Lemma 4).* By linearity, it suffices to prove the statement for the case that $f = \mu$ is a monomial. $\mathcal{D}_{\nu \cdot L_{\boldsymbol{c}}^{[a]}} \mu$ is equal to the coefficient of the monomial $\mu$ in the polynomial $\nu \cdot L_{\boldsymbol{c}}^{[a]}$ since both are square-free of the same degree. By definition of $L_{\boldsymbol{c}}^{[a]}$, this is precisely $(\mathcal{D}_\nu \mu)(\boldsymbol{c})$. □

Now let $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ have rank $k$ and let $\tilde{\boldsymbol{A}} \in \mathbb{F}_q^{k \times \tilde{n}}$ with $\tilde{n} \coloneqq n - k$ be the matrix obtained by transforming $\boldsymbol{A}$ to reduced row-echelon form and omitting the pivot columns. Let $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_{\tilde{n}} \in \mathbb{F}_q^k$ denote the columns of $\tilde{\boldsymbol{A}}$. Note that Lemma 4 implies that the kernel of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ is trivial if and only if $\mathfrak{L}_{\tilde{\boldsymbol{A}}}^{(p)} = \tilde{R}_k^{(p)}$. Indeed, one can show more:

**Proposition 5.** *With notation as above,*

$$\dim_{\mathbb{F}_q}\left(\mathfrak{L}^{(p)}_{\tilde{\boldsymbol{A}}}\right) = \dim_{\mathbb{F}_q}\left(\mathrm{Im}\left(\tilde{\varphi}^{(p)}_{\boldsymbol{A}}\right)\right) \tag{3}$$

*Proof.* We pick some ordering of monomials in $\tilde{R}_k$ of degree $p$ and consider the degree $p$ Macaulay matrix $\boldsymbol{M}$ associated with the $L^{[j]}_{\tilde{\boldsymbol{a}}_i}$. That is, each row of $\boldsymbol{M}$ corresponds to a product $\nu \cdot L^{[j]}_{\tilde{\boldsymbol{a}}_i}$ for some $i, j$ and for some square-free monomial $\nu$ of degree $p-j$, and each column corresponds to a square-free monomial $\mu$ of degree $p$. The corresponding entry contains the $\mathbb{F}_q$-coefficient of the monomial $\mu$ in the polynomial $\nu \cdot L^{[j]}_{\tilde{\boldsymbol{a}}_i}$. Since $\mu$ is square-free of the same degree as $\nu \cdot L^{[j]}_{\tilde{\boldsymbol{a}}_i}$, this is equal to $\mathcal{D}_{\nu \cdot L^{[j]}_{\tilde{\boldsymbol{a}}_i}}\mu$, and by Lemma 4, this is the same as $(\mathcal{D}_\nu \mu)(\tilde{\boldsymbol{a}}_i)$. But this means that $\boldsymbol{M}$ is also a matrix representation of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}}$. This implies the claim. □

By Lemma 3, depending on the characteristic of the ground field, some of the generators of the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ are superfluous:

**Lemma 5.** *Let $c \coloneqq \mathrm{char}(\mathbb{F}_q)$ and $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ be any matrix. Then*

$$\mathfrak{L}_{\boldsymbol{A}} = \langle L^{[2]}_{\boldsymbol{a}_i}, L^{[c^\ell]}_{\boldsymbol{a}_i} \mid 1 \le i \le n,\ \ell \ge 1 \rangle$$

*Proof.* Consider an integer $j$ such that $2 < j < c$. (In the case $c = 2$ or $c = 3$, this does not occur of course.) By Lemma 3 we have

$$L^{[2]}_{\boldsymbol{a}_i} \cdot L^{[j-2]}_{\boldsymbol{a}_i} = \binom{j}{2} \cdot L^{[j]}_{\boldsymbol{a}_i}$$

Since $2 < j < c$, the binomial coefficient is non-zero in $\mathbb{F}_q$. Hence $L^{[j]}_{\boldsymbol{a}_i} \in \langle L^{[2]}_{\boldsymbol{a}_i}, L^{[c^\ell]}_{\boldsymbol{a}_i} \mid 1 \le i \le n,\ \ell \ge 1 \rangle$.

Now suppose $j > c$ is not a power of $c$. Let $c^\ell$ be the largest power of $c$ less than $j$. By Lemma 3 we have

$$L^{[c^\ell]}_{\boldsymbol{a}_i} \cdot L^{[j-c^\ell]}_{\boldsymbol{a}_i} = \binom{j}{c^\ell} \cdot L^{[j]}_{\boldsymbol{a}_i}$$

Lucas's theorem [7, Theorem 4.71] implies that $\binom{j}{c^\ell}$ is not divisible by $c$. Hence $L^{[j]}_{\boldsymbol{a}_i} \in \langle L^{[2]}_{\boldsymbol{a}_i}, L^{[c^\ell]}_{\boldsymbol{a}_i} \mid 1 \le i \le n,\ \ell \ge 1 \rangle$. This implies the claim. □

*Remark 4.* From Lemma 5 and Proposition 5, we deduce that in the definition of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}}$, it is sufficient to take into account the partial derivatives of order $p-2$ and $p-c^\ell$ for $1 \le \ell \le \lfloor \log_c(p) \rfloor$ only since the kernel of this map (with this reduced number of partial derivatives) is the same as the kernel of $\tilde{\varphi}$ as defined in Definition 6.

We have reformulated the rank (or, equivalently, the dimension of the kernel) of the map $\tilde{\varphi}^{(p)}_{\boldsymbol{A}}$ in terms of the dimension of the degree $p$ component of a polynomial ideal. This yields an equivalent definition of our distinguisher that does not make reference to differentiation. It turns out (see the next section) that the analysis of the case of random matrices $\boldsymbol{A}$ is easier using this formulation.

### 4.2 Semi-regularity and behaviour in the random case

Let $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ be a randomly selected matrix of rank $k$. We want to study the dimension of the kernel of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$. Note that $\tilde{\boldsymbol{A}}$ (obtained by bringing $\boldsymbol{A}$ into reduced row-echelon form and removing the pivot columns) is then a randomly selected $k \times \tilde{n}$ matrix where $\tilde{n} \coloneqq n - k$. Denoting by $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_{\tilde{n}}$ the columns of $\tilde{\boldsymbol{A}}$, Proposition 5 implies that

$$\dim_{\mathbb{F}_q}\left(\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)\right) = \dim_{\mathbb{F}_q}\left(\tilde{R}_k^{(p)}/\mathfrak{L}_{\tilde{\boldsymbol{A}}}^{(p)}\right) \tag{4}$$

which is equal to the Hilbert function in degree $p$ of the graded $\mathbb{F}_q$-algebra $\tilde{R}_k/\mathfrak{L}_{\tilde{\boldsymbol{A}}}$. It is well-known that this can be explicitly computed in the case that polynomials generating the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ form a so-called *semi-regular sequence*. Intuitively, this means the polynomials have as few relations between them as possible. We recall:

**Definition 8.** *Let* $\tilde{R}_k = \mathbb{F}_q[X_1, \ldots, X_k]/\langle X_1^2, \ldots, X_k^2\rangle$. *Let* $d$ *be a positive integer. A sequence* $f_1, \ldots, f_n \in \tilde{R}_k$ *of homogeneous elements is* $d$-regular *if the following holds, depending on* $\mathrm{char}(\mathbb{F}_q)$*:*

- *If* $\mathrm{char}(\mathbb{F}_q) \neq 2$*: For any* $1 \leq i \leq n$*, if* $g \in \tilde{R}_k$ *such that* $gf_i \in \langle f_1, \ldots, f_{i-1}\rangle$ *and* $\deg(gf_i) < d$*, then* $g \in \langle f_1, \ldots, f_{i-1}\rangle$.
- *If* $\mathrm{char}(\mathbb{F}_q) = 2$*: For any* $1 \leq i \leq n$*, if* $g \in \tilde{R}_k$ *such that* $gf_i \in \langle f_1, \ldots, f_{i-1}\rangle$ *and* $\deg(gf_i) < d$*, then* $g \in \langle f_1, \ldots, f_{i-1}, f_i\rangle$.

*We define the* degree of regularity *of the ideal* $I = \langle f_1, \ldots, f_n\rangle$ *as*

$$d_{reg}(I) \coloneqq \min\left\{d \geq 0 \mid \dim_{\mathbb{F}_q}(I^{(d)}) = \binom{k}{d}\right\}$$

*The sequence* $f_1, \ldots, f_n$ *is* semi-regular *if it is* $d_{reg}$*-regular.*

*Remark 5.* The above definition of semi-regularity in characteristic 2 agrees with [4, Definition 9] which, like our definition, defines semi-regularity for elements in $\tilde{R}_k$. For characteristic not equal to 2, [4, Definition 5] defines semi-regularity for elements in $R_k$ while we defined it for elements in $\tilde{R}_k$. Note that our definition is equivalent to the semi-regularity of the sequence $X_1^2, \ldots, X_k^2, \hat{f}_1, \ldots, \hat{f}_n \in R_k$ according to [4, Definition 5], where $\hat{f}_i$ is a lift of $f_i$ to $R_k$.

For $(d+1)$-regular sequences, one can easily determine the Hilbert series up to degree $d$. Let us introduce some useful notation: Given a formal power series $S = \sum_{i=0}^{\infty} a_i t^i$ with integer coefficients and letting $d \in \mathbb{N}_0 \cup \{\infty\}$ denote the smallest index such that $a_d < 0$, we define the power series

$$[S]^+ \coloneqq \sum_{i=0}^{\infty} a_i' \cdot t^i \quad \text{where } a_i' = \begin{cases} a_i & \text{if } i < d \\ 0 & \text{otherwise} \end{cases}$$

**Proposition 6.** *Let $f_1, \ldots, f_n \in \tilde{R}_k$ be homogeneous with $\deg(f_i) = d_i > 0$ for all $i$ such that they form a semi-regular sequence, and set $I \coloneqq \langle f_1, \ldots, f_n \rangle$. We define the power series*

$$S^{\boldsymbol{d}}_{n,k} = (1+t)^k \cdot \prod_{i=1}^{n}(1-t^{d_i}), \quad T^{\boldsymbol{d}}_{n,k} = \frac{(1+t)^k}{\prod_{i=1}^{n}(1+t^{d_i})}$$

*Then the Hilbert series $HS_I(t) = \sum_{i \geq 0} \dim_{\mathbb{F}_q}\left(\tilde{R}^{(i)}_k / I^{(i)}\right) \cdot t^i$ of $I$ is given by $[S^{\boldsymbol{d}}_{n,k}]^+$ if $\mathrm{char}(\mathbb{F}_q) \neq 2$ and by $[T^{\boldsymbol{d}}_{n,k}]^+$ if $\mathrm{char}(\mathbb{F}_q) = 2$.*

*Proof.* This follows immediately from [4, Proposition 6] if $\mathrm{char}(\mathbb{F}_q) \neq 2$ and from [4, Proposition 10] otherwise. □

*Remark 6.* It is conjectured that if the length of a sequence of polynomials is sufficiently large compared to the number of variables, then most such random sequences are semi-regular. However, little has been proven in this direction. [15] and [13] characterise the conditions under which a sequence of *one* or *two* quadratic polynomials in $\mathbb{F}_2[X_1, \ldots, X_k]/\langle X_1^2, \ldots, X_k^2 \rangle$ is semi-regular. [13, Theorem 4.4] shows that if the length of a sequence is too small (namely at most $k/4-1$), it cannot be semi-regular. On the other hand, [24, Theorem 1.1] gives a lower bound on the proportion of sequences that are semi-regular if the length of the sequence is at least $\lceil (k-1)(k-2)/6 \rceil$, and the bound tends to 1 as $k \to \infty$. In between, not much has been rigorously proven. See also the introduction of [13] for a discussion.

Let us now consider the polynomials $L^{[j]}_{\tilde{a}_i}$ and the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$. The polynomials $L^{[j]}_{\tilde{a}_i}$ are of course not chosen uniformly at random, even in our case that the matrix $\tilde{\boldsymbol{A}}$ has been chosen uniformly at random. Furthermore, we saw in Lemma 3 that there are multiplicative relations between them so that the full set of them cannot form a semi-regular sequence. Lemma 5 provides a smaller set of generators. One may hope that, having excluded the trivial relations, these now form a semi-regular sequence. However, we need to be a bit careful about the characteristic of the ground field:

*Example 1.* Suppose $\mathrm{char}(\mathbb{F}_q) = 3$. Note that in this case, by Lemma 3 we have $L^{[1]}_{\boldsymbol{c}} \cdot L^{[2]}_{\boldsymbol{c}} = 0$ for any $\boldsymbol{c} \in \mathbb{F}_q^k$. Hence a sequence consisting of a single polynomial $L^{[2]}_{\boldsymbol{c}}$ is not 4-regular. Thus, in this case, the sequence of polynomials described in Lemma 5 is generally not semi-regular.

For now, let us restrict to the cases of characteristic 2 and of characteristic $> p$. Our experiments suggest that if $\tilde{\boldsymbol{A}}$ has been chosen uniformly at random and parameter sizes are sufficiently large, then with high probability, the sequences of polynomials $L^{[j]}_{\tilde{\boldsymbol{a}}_i}$ (excluding the superfluous ones) indeed form a semi-regular sequence if $\mathrm{char}(\mathbb{F}_q) = 2$, and they are at least regular up to degree $p$ if $\mathrm{char}(\mathbb{F}_q) > p$. See Remark 7 for some discussion and experimental evidence. We summarise this in the following heuristic assumptions:

**Heuristic Assumption 1.** Suppose $\mathrm{char}(\mathbb{F}_q) = 2$ and $\tilde{n} > k$ and let $\tilde{\boldsymbol{A}} \in \mathbb{F}_q^{k \times \tilde{n}}$ be a randomly selected matrix with columns $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_{\tilde{n}}$. If $k$ is sufficiently large, then the polynomials $L_{\tilde{\boldsymbol{a}}_i}^{[2^\ell]}$ for $\ell \geq 1$ and $i = 1, \ldots, \tilde{n}$ form a semi-regular sequence in $\tilde{R}_k$ with high probability.

**Heuristic Assumption 2.** Let $p \geq 2$ and suppose $\mathrm{char}(\mathbb{F}_q) > p$ and $\tilde{n} > k$. Let $\tilde{\boldsymbol{A}} \in \mathbb{F}_q^{k \times \tilde{n}}$ be a randomly selected matrix with columns $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_{\tilde{n}}$. If $k$ is sufficiently large, then the polynomials $L_{\tilde{\boldsymbol{a}}_i}^{[2]}$ for $i = 1, \ldots, \tilde{n}$ form a $(p+1)$-regular sequence in $\tilde{R}_k$ with high probability.

Now, assuming Heuristic Assumption 1 and Heuristic Assumption 2 hold, we are able to predict (at least with high probability for sufficiently large parameter sets) the dimension of the kernel of our map $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ for randomly chosen $\boldsymbol{A}$. As already indicated above, these formulas have been confirmed in computations for concrete parameters, see also Remark 7.

**Corollary 1.** *Suppose Heuristic Assumption 1 and Heuristic Assumption 2 hold and $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ is randomly selected. Let $\tilde{n} \coloneqq n - k > k$. If $k$ is sufficiently large, then the following holds with high probability: If $\mathrm{char}(\mathbb{F}_q) = 2$ or $\mathrm{char}(\mathbb{F}_q) > p$, then*

$$\begin{aligned}\dim_{\mathbb{F}_q}\left(\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)\right) &= \textit{degree } p \textit{ coefficient of } \left[(1+t)^k \cdot \left(1-t^2\right)^{\tilde{n}}\right]^+ \\ &= \begin{cases} \sum_{i=0}^{\lfloor p/2 \rfloor} (-1)^i \cdot \binom{\tilde{n}}{i} \cdot \binom{k}{p-2i} & \textit{if } p < d_{reg} \\ 0 \quad \textit{if } p \geq d_{reg} \end{cases}\end{aligned}$$

*Proof.* For $\mathrm{char}(\mathbb{F}_q) > p$, this follows immediately from Equation 4, Lemma 5 and Proposition 6.

Suppose $\mathrm{char}(\mathbb{F}_q) = 2$, let $d_{reg}$ be the degree of regularity of the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ and $D \coloneqq \lfloor \log_2(d_{reg}) \rfloor$. Lemma 5 and Proposition 6 yield that

$$\dim_{\mathbb{F}_q}\left(\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)\right) = \text{degree } p \text{ coefficient of } \left[\frac{(1+t)^k}{\prod_{\ell=1}^{D}(1+t^{2^\ell})^{\tilde{n}}}\right]^+$$

Note that $\frac{1}{\prod_{\ell=1}^{D}(1+t^{2^\ell})} = 1 - t^2 + o(t^{d_{reg}})$, so the result follows. □

We see that under the above heuristic assumptions, we obtain the same formula in the cases $\mathrm{char}(\mathbb{F}_q) = 2$ and $\mathrm{char}(\mathbb{F}_q) > p$. This is not a coincidence. Consider the generators $L_{\tilde{\boldsymbol{a}}_i}^{[\ell]}$ of the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ where $\ell \geq 2$. The essential difference between different characteristics of the ground field is (or at least seems to be) that by Lemma 3, depending on the characteristic, $L_{\boldsymbol{c}}^{[i]} \cdot L_{\boldsymbol{c}}^{[j]}$ is either 0 or a non-zero multiple of $L_{\boldsymbol{c}}^{[i+j]}$. However, since we include all $L_{\boldsymbol{c}}^{[\ell]}$ in the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ anyway, this should not make a difference for the vector space dimension of $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ in degree $p$. Note that this suggests that the formula from Corollary 1 should actually hold for *all* characteristics, even though we cannot deduce it

**Table 1.** Results of computing the dimension of $\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)$ for randomly selected $k \times n$ matrices $\boldsymbol{A}$ over $\mathbb{F}_2$ with $p = 4$ and $k = 25$. For different values of $n$, the table displays the expected dimension according to Heuristic 1, the observed average dimension over 100 experiments as well as the maximum observed deviation from the expected dimension. The observed dimension was always lower bounded by the dimension according to Heuristic 1.

| $n$ | 35 | 45 | 55 | 65 | 69 | 70 | 71 |
|---|---|---|---|---|---|---|---|
| Dim. acc. to Heuristic 1 | 9695 | 6840 | 4085 | 1430 | 396 | 140 | 0 |
| Average dimension | 9695.25 | 6841.72 | 4090.32 | 1444.14 | 415.34 | 160.69 | 0 |
| Maximum deviation | 5 | 47 | 52 | 62 | 72 | 72 | 0 |

**Table 2.** Results for ground field $\mathbb{F}_2$, $p = 4$, $k = 35$, with 100 experiments.

| $n$ | 50 | 70 | 90 | 110 | 129 | 130 | 131 |
|---|---|---|---|---|---|---|---|
| Dim. acc. to Heuristic 1 | 43540 | 32130 | 21120 | 10510 | 801 | 300 | 0 |
| Average dimension | 43540 | 32130 | 21120.2 | 10510.6 | 801.95 | 300.95 | 0 |
| Maximum deviation | 0 | 5 | 5 | 10 | 10 | 10 | 0 |

**Table 3.** Results for ground field $\mathbb{F}_2$, $p = 5$, $k = 35$, with 10 experiments.

| $n$ | 45 | 55 | 65 | 75 | 85 | 93 | 94 |
|---|---|---|---|---|---|---|---|
| Dim. acc. to Heuristic 1 | 260757 | 200382 | 143507 | 90132 | 40257 | 2877 | 0 |
| Average dimension | 260759.8 | 200387.6 | 143512.6 | 90141.8 | 40280.8 | 2924.6 | 0 |
| Maximum deviation | 14 | 14 | 14 | 28 | 56 | 84 | 0 |

**Table 4.** Results for ground field $\mathbb{F}_7$, $p = 6$, $k = 20$, with 100 experiments.

| $n$ | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|
| Dim. acc. to Heuristic 1 | 20516 | 16425 | 12520 | 8800 | 5264 | 1911 | 0 |
| Average dimension | 20516 | 16425 | 12520 | 8800 | 5264.42 | 1911.42 | 0 |
| Maximum deviation | 0 | 0 | 0 | 0 | 42 | 42 | 0 |

from semi-regularity assumptions in the case that $2 < \mathrm{char}(\mathbb{F}_q) < p$ (see Example 1). Indeed, this is supported by experimental evidence. We summarise this in the following heuristic which we shall use to predict the behaviour of our distinguisher in the random case:

**Heuristic 1.** Suppose $\boldsymbol{A} \in \mathbb{F}_q^{k \times n}$ is randomly selected. Let $\tilde{n} \coloneqq n - k > k$. If $k$ is sufficiently large, then with high probability,

$$\dim_{\mathbb{F}_q}\left(\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)\right) = \text{degree } p \text{ coefficient of } \left[(1+t)^k\left(1-t^2\right)^{\tilde{n}}\right]^+$$

*Remark 7.* Note that we are being vague about what constitutes 'sufficiently large' parameters. We have run a series of experiments to provide evidence for the validity of Heuristic 1 and to illustrate its limits, see Table 1 to Table 4. In each experiment, for a fixed value of $q$ and $p$ and $k$, we generated a number of random $k \times n$ matrices $\boldsymbol{A}$ over $\mathbb{F}_q$ for different values of $n$ and computed the

dimension of $\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)$. We chose $q = 2$ and $q = 7$ to illustrate the cases of the binary and a larger field.

Let us first consider the case $q = 2$ and $p = 4$. Table 1 shows that for $k = 25$, there are significant deviations from the prediction of Heuristic 1. However, Table 2 illustrates that if we increase $k$ to 35, the experiments adhere to Heuristic 1 with few exceptions. Note that even for $k = 25$, in all runs of our experiment, the smallest $n$ for which the kernel becomes trivial matches precisely the theoretical prediction. This is actually the important property which we will need for our distinguisher. Thus we see that even if $k$ is not large enough for the experimental results to precisely match Heuristic 1, our distinguisher often still works.

Table 3 shows that, as expected, the results for $p = 5$ and $k = 35$ match the heuristic less closely than for $p = 4$ and the same value of $k$. However, the point at which the kernel becomes trivial again matches precisely the theoretical prediction. Note that we have run this experiment only 10 times (as opposed to 100 times for the other cases) because the computational complexity quickly increases with increasing $p$.

Finally, Table 4 shows that for $q = 7$ and $p = 6$, a value of $k = 20$ already suffices for our experiments to match Heuristic 1 almost perfectly. This is not surprising since in larger fields, degenerate behaviour is less likely than in small fields.

## 5 Higher-order vanishing for alternant and Goppa codes

Suppose $\boldsymbol{A} \in \mathbb{F}_q^{mr \times n}$ is a generator matrix of the $\mathbb{F}_q$-code $\mathrm{Alt}_r(\boldsymbol{x}, \boldsymbol{y})^\perp$ where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_{q^m}^n$. We assume throughout that the alternant code is proper, i.e. that $\boldsymbol{A}$ has full rank $mr$. Our aim in this section is to prove that (for values of $p$ lower than some bound depending on the code parameters) the kernel of $\varphi_{\boldsymbol{A}}^{(p)}$ is non-trivial in this case. In subsection 5.3, we will prove slightly improved bounds in the special case of binary Goppa codes.

Since $\varphi_{\boldsymbol{A}}^{(p)}$ is $\mathbb{F}_q$-linear, we may extend scalars to $\mathbb{F}_{q^m}$. This preserves the dimension of its kernel. By Proposition 1, over the extension field $\mathbb{F}_{q^m}$, the parity check matrix $\boldsymbol{A}$ is row-equivalent to the matrix

$$\boldsymbol{A}_{\mathbb{F}_{q^m}} = \begin{pmatrix} \boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}} \\ (\boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}})^{(q)} \\ (\boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}})^{(q^2)} \\ \cdots \\ (\boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}})^{(q^{m-1})} \end{pmatrix} \text{ with } \boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1x_1 & y_2x_2 & \cdots & y_nx_n \\ y_1x_1^2 & y_2x_2^2 & \cdots & y_nx_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ y_1x_1^{r-1} & y_2x_2^{r-1} & \cdots & y_nx_n^{r-1} \end{pmatrix}$$

where $(\boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}})^{(q^i)}$ denotes the entry-wise $q^i$-th power of the matrix $\boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}}$. By Proposition 3, the dimension of the kernel of the $\mathbb{F}_q$-linear map $\varphi_{\boldsymbol{A}}^{(p)}$ and the $\mathbb{F}_{q^m}$-linear map $\varphi_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}^{(p)}$ are equal.

### 5.1 The kernel in the alternant case

In this section, we prove lower bounds on the dimension of the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ by explicitly constructing elements in the kernel as follows:

It will be convenient to label the indeterminates in $R_{rm}$ as

$$X_0^{(0)}, \ldots, X_{r-1}^{(0)}, X_0^{(1)}, \ldots, X_{r-1}^{(1)}, \ldots, X_0^{(m-1)}, \ldots, X_{r-1}^{(m-1)} \tag{5}$$

where each superscript corresponds to one sub-matrix $(\boldsymbol{G}_{\boldsymbol{x},\boldsymbol{y}})^{(q^i)}$ of $\boldsymbol{A}_{\mathbb{F}_{q^m}}$. Let $e \coloneqq \left\lfloor \log_q \left( \frac{r-1}{p-1} \right) \right\rfloor$. For $0 \le u \le e$ and $0 \le v < m$, we define the $p \times (r-(p-1)q^u)$ matrix

$$\boldsymbol{B}_{u,v} \coloneqq \begin{pmatrix} X_0^{(v)} & X_1^{(v)} & \ldots & X_{r-1-(p-1)q^u}^{(v)} \\ X_{q^u}^{(v)} & X_{1+q^u}^{(v)} & \ldots & X_{r-1-(p-2)q^u}^{(v)} \\ \ldots & \ldots & \ldots & \ldots \\ X_{(p-1)q^u}^{(v)} & X_{1+(p-1)q^u}^{(v)} & \ldots & X_{r-1}^{(v)} \end{pmatrix}$$

and the $p \times n_{col}$ block matrix

$$\boldsymbol{\Phi}^{(p)} \coloneqq (\boldsymbol{B}_{0,e} \mid \boldsymbol{B}_{1,e-1} \mid \ldots \mid \boldsymbol{B}_{e,0})$$

where $n_{col} \coloneqq (e+1)r - (p-1) \cdot \frac{q^{e+1}-1}{q-1}$.

For $p = 2$, this is the same matrix $\Phi$ considered in [23, section 3].

**Lemma 6.** *The $p \times p$ minors of $\boldsymbol{\Phi}^{(p)}$ lie in the kernel of $\varphi^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$. They are linearly independent over $\mathbb{F}_{q^m}$.*

*Proof.* Note that $\boldsymbol{\Phi}^{(p)}$ is designed such that its rank is 1 if we substitute any column of $\boldsymbol{A}_{\mathbb{F}_{q^m}}$ for the indeterminates $X_i^{(j)}$. For, substituting $X_i^{(j)} \mapsto (y_t \cdot (x_t)^i)^{q^j}$ for any fixed $1 \le t \le n$, we obtain each row from the previous one by multiplying by $(x_t)^{q^e}$. Thus, for all $2 \le \ell \le p$, any $\ell \times \ell$ minor of $\boldsymbol{\Phi}^{(p)}$ vanishes when evaluated at $X_i^{(j)} \mapsto (y_t \cdot (x_t)^i)^{q^j}$.

By the rule for differentiating a determinant, any partial derivative of a $p \times p$ minor of $\boldsymbol{\Phi}^{(p)}$ of order $p - \ell$ for $2 \le \ell \le p$ is a linear combination of $\ell \times \ell$ minors of $\boldsymbol{\Phi}^{(p)}$, so it vanishes when evaluated at $X_i^{(j)} \mapsto (y_t \cdot (x_t)^i)^{q^j}$ (for any $t$). Hence any such $p \times p$ minor is in the kernel of $\varphi^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$.

We prove linear independence by induction on $p$ and first introduce some notation. Given a set $S \subset \{1, \ldots, n_{col}\}$ with $|S| = p$, we denote by $\boldsymbol{\Phi}^{(p)}_S$ the $p \times p$ submatrix formed by the columns of $\boldsymbol{\Phi}^{(p)}$ with indices in $S$.

The entries in any fixed row of $\boldsymbol{\Phi}^{(p)}$ are distinct indeterminates of the polynomial ring and therefore linearly independent. Now suppose that $p \ge 2$ and suppose for contradiction that the $p \times p$ minors are not linearly independent. Let a non-trivial linear combination be given by

$$\sum_{i=1}^{\ell} \lambda_i \cdot \det\left( \boldsymbol{\Phi}^{(p)}_{S_i} \right) = 0 \tag{6}$$

where $\ell \geq 1$, $\lambda_i \in \mathbb{F}_{q^m} \setminus \{0\}$ for all $1 \leq i \leq \ell$ and the $S_i \subset \{1, \ldots, n_{col}\}$ are distinct subsets of size $p$. Let $a \in \{1, \ldots, n_{col}\}$ be the largest column index contained in some $S_i$, and let $S_{i_1}, \ldots, S_{i_t}$ be those sets that contain $a$. By maximality of $a$, the indeterminate (let us call it $X_b^{(c)}$) in the bottom rightmost entry of the matrices $\boldsymbol{\Phi}^{(p)}_{S_{i_d}}$ for $1 \leq d \leq t$ only occurs in this entry and nowhere else in the matrices $\boldsymbol{\Phi}^{(p)}_{S_i}$ for $1 \leq i \leq \ell$. Hence using Laplace expansion of the determinants in (6) by the last row and comparing coefficients of $X_b^{(c)}$, we deduce that

$$\sum_{d=1}^{t} \lambda_{i_d} \cdot \det\left(\hat{\boldsymbol{\Phi}}^{(p)}_{S_{i_d} \setminus \{a\}}\right) = 0$$

where $\hat{\boldsymbol{\Phi}}^{(p)}$ denotes the matrix $\boldsymbol{\Phi}^{(p)}$ with the last row deleted. Since the determinants in this sum are maximal minors of $\boldsymbol{\Phi}^{(p-1)}$, by induction, we deduce that $\lambda_{i_d} = 0$ for $1 \leq d \leq t$, contradicting the assumption that $\lambda_i \neq 0$ for all $i$. This concludes our proof. □

The preceding lemma implies that the kernel of $\varphi^{(p)}_{\boldsymbol{A}}$ has dimension at least $\binom{n_{col}}{p}$. In order to deduce from this a lower bound on the dimension of the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}}$, we need a lemma about the behaviour of $\boldsymbol{\Phi}^{(p)}$ under a change of basis. To this end, let $\boldsymbol{S} \in \mathbb{F}_{q^m}^{mr \times mr}$ be an invertible matrix such that $\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}$ is in reduced row echelon form. We define the change of variables

$$\alpha_{\boldsymbol{S}^{-1}} \colon \mathbb{F}_{q^m}\left[X_i^{(j)}\right]_{i=0,\ldots,r-1}^{j=0,\ldots,m-1} \to \mathbb{F}_{q^m}\left[Z_1, \ldots, Z_{mr}\right], \quad f(\boldsymbol{X}) \mapsto f(\boldsymbol{Z} \cdot (\boldsymbol{S}^{-1})^T)$$

where $\boldsymbol{X} = (X_0^{(0)}, X_1^{(0)}, \ldots, X_{r-1}^{(0)}, X_0^{(1)}, \ldots, X_{r-1}^{(m-1)})$ and $\boldsymbol{Z} = (Z_1, \ldots, Z_{mr})$ are row vectors. We denote by $\tilde{\boldsymbol{\Phi}}^{(p)}$ the matrix obtained from $\boldsymbol{\Phi}^{(p)}$ by applying $\alpha_{\boldsymbol{S}^{-1}}$ to each entry. In the following, we let $\mathbb{F}_{q^m}\left[Z_1, \ldots, Z_{mr}\right]$ be the domain of $\varphi^{(p)}_{\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ to distinguish it from the domain of $\varphi^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ (which we took to be $\mathbb{F}_{q^m}[X_i^{(j)}]$). This is just a naming convention for the indeterminates.

**Lemma 7.** *The $p \times p$ minors of $\tilde{\boldsymbol{\Phi}}^{(p)}$ are in the kernel of $\varphi^{(p)}_{\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ and are linearly independent. Furthermore, in any $p \times p$ minor, each indeterminate $Z_j$ occurs to degree at most 1.*

*Proof.* By definition of $\alpha_{\boldsymbol{S}^{-1}}$, substituting any column of $\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}$ for $\boldsymbol{Z}$ in the entries of $\tilde{\boldsymbol{\Phi}}^{(p)}$ yields the same matrix as substituting the corresponding column of $\boldsymbol{A}_{\mathbb{F}_{q^m}}$ for $\boldsymbol{X}$ in $\boldsymbol{\Phi}^{(p)}$. Hence Lemma 6 implies that the $p \times p$ minors of $\tilde{\boldsymbol{\Phi}}^{(p)}$ are linearly independent and in the kernel of $\varphi^{(p)}_{\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}}$.

Now fix a $p \times p$ submatrix $\boldsymbol{M}$ of $\tilde{\boldsymbol{\Phi}}^{(p)}$. We want to show that in $\det(\boldsymbol{M})$, each indeterminate occurs to degree at most 1. The entries of $\boldsymbol{M}$ are homogeneous linear polynomials in the indeterminates. Consider a fixed but arbitrary indeterminate $Z_j$. The standard basis vector $\boldsymbol{e}_j$ occurs as a column in $\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}$,

so $\tilde{\boldsymbol{\Phi}}^{(p)}$ and therefore $\boldsymbol{M}$ has rank at most 1 when evaluated at $\boldsymbol{e}_j$. Hence the matrix formed by the coefficients of $Z_j$ in $\boldsymbol{M}$ has rank at most 1. This means that using elementary column operations on $\boldsymbol{M}$, one can transform it such that $Z_j$ occurs only in a single column. The determinant of such a matrix is clearly linear in $Z_j$. Since elementary column operations do not alter the determinant, this implies that the same holds for $\boldsymbol{M}$. □

**Proposition 7.** *The dimension of the kernel of* $\tilde{\varphi}^{(p)}_{\boldsymbol{A}}$ *is at least* $\binom{n_{col}}{p}$.

*Proof.* By Lemma 7, the $p \times p$ minors of $\tilde{\boldsymbol{\Phi}}^{(p)}$ lie in the kernel of $\varphi^{(p)}_{\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ and are linearly independent. Since each indeterminate in such a minor occurs to degree at most 1, these minors lie in the domain of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ and therefore in its kernel. Thus, the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ has dimension at least $\binom{n_{col}}{p}$. Since the dimension of the kernel is invariant under field extension, this proves the claim. □

Note that $\boldsymbol{\Phi}^{(p)}$ must have at least $p$ columns for our construction to yield an element in the kernel. Furthermore, the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}}$ is generally larger than the lower bound from Proposition 7. For our purposes, Proposition 7 in combination with the considerations on shortening (see the next section) will be sufficient.

### 5.2 The effect of shortening

It turns out that shortening the dual of alternant and Goppa codes in not too many positions (we make this more precise below) preserves enough of the code structure that the kernel of $\tilde{\varphi}^{(p)}$ is still non-trivial. The advantage of shortening is that it decreases the dimension of our code, which in turn significantly decreases the complexity of computing the dimension of the kernel of $\tilde{\varphi}^{(p)}$. We briefly recall the definition of a shortened code.

**Definition 9.** *Let* $C \subset \mathbb{F}_q^n$ *be a linear code and* $i \in \{1, \ldots, n\}$. *The* code shortened in position $i$ *is given by* $\{\boldsymbol{c} \in C \mid c_i = 0\}$.

We may of course puncture such a shortened code in its $i$th position essentially without losing any information since the $i$th coordinate of any codeword in the shortened code is 0 anyway. Now we analyse the dimension of the kernel of $\tilde{\varphi}^{(p)}$ for a shortened dual of an alternant code. Note that the following result is very similar to [23, Theorem 1].

**Proposition 8.** *Let* $\boldsymbol{A}_{s_*}$ *be a generator matrix of the code* $\mathrm{Alt}_r(\boldsymbol{x}, \boldsymbol{y})^{\perp}$ *shortened in* $s_* \coloneqq n_{col} - p$ *places. Then the kernel of* $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{s_*}}$ *is non-trivial.*

*Proof.* Since shortening commutes with field extensions, it suffices to prove this for the extended code with generator matrix $\boldsymbol{A}_{\mathbb{F}_{q^m}}$ defined at the beginning of section 5. As before, let $\boldsymbol{S} \in \mathbb{F}_{q^m}^{mr \times mr}$ be an invertible matrix such that $\boldsymbol{S}\boldsymbol{A}_{\mathbb{F}_{q^m}}$ is in reduced row echelon form. We may assume that we shorten in the first $s_*$ positions; if not, we can reorder the columns of $\boldsymbol{A}_{\mathbb{F}_{q^m}}$ such that this is the case.

A generator matrix of the shortened code is given by the matrix $\boldsymbol{B}$ obtained by deleting the rows in the echelonised matrix $\boldsymbol{SA}_{\mathbb{F}_{q^m}}$ that have non-zero entries in the first $s_*$ columns. Suppose these are rows $1, 2, \ldots, s'$ where $s' \leq s_*$. This implies that (up to renaming indeterminates) $\tilde{\varphi}^{(p)}_{\boldsymbol{B}}$ is the restriction of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$ to the subalgebra $\mathbb{F}_{q^m}[Z_{s'+1}, \ldots, Z_{mr}] \subset \mathbb{F}_{q^m}[Z_1, \ldots, Z_{mr}]$.

Since $\boldsymbol{SA}_{\mathbb{F}_{q^m}}$ contains the standard unit vectors as columns and $\tilde{\boldsymbol{\Phi}}^{(p)}$ has rank 1 when evaluated at the columns of $\boldsymbol{SA}_{\mathbb{F}_{q^m}}$, the following holds: For any indeterminate $Z_i$, the matrix formed by the coefficients of $Z_i$ in $\tilde{\boldsymbol{\Phi}}^{(p)}$ has rank 1. In particular, this holds for $Z_1, \ldots, Z_{s'}$. Therefore, using elementary column operations, we may transform $\tilde{\boldsymbol{\Phi}}^{(p)}$ to a matrix $\tilde{\boldsymbol{\Phi}}'^{(p)}$ such that its first $n_{col} - s'$ columns only contain the indeterminates $Z_{s'+1}, \ldots, Z_{mr}$. Note that $n_{col} - s' \geq n_{col} - s_* \geq p$, so the $p \times p$ minor obtained from the first $p$ columns of $\tilde{\boldsymbol{\Phi}}'^{(p)}$ contains only those indeterminates. Since elementary column operations of $\tilde{\boldsymbol{\Phi}}^{(p)}$ leave the space of maximal minors invariant, Lemma 7 implies that this minor is a non-trivial element in the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{\mathbb{F}_{q^m}}}$. It is also a non-trivial element in the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{B}}$ since it only involves the indeterminates $Z_{s'+1}, \ldots, Z_{mr}$. □

Our computations suggest that if we shorten the dual of a random alternant code in more than $s_*$ positions, then the kernel becomes trivial with high probability. See Table 5 for some concrete computations. Thus, the expression for $s_*$ seems to be tight in the sense that it is almost surely the *maximum* number of places that we can shorten without obtaining a trivial kernel. Of course, it may still occur that we can shorten in more places. For example, we will see that in the case of Goppa codes, the above expression for $s_*$ can be improved.

**Table 5.** Dimension of the kernel of $\tilde{\varphi}^{(3)}_{\boldsymbol{A}_s}$ for different numbers $s$ of shortened positions, where $\boldsymbol{A}_s$ is the generator matrix of the shortened dual of an alternant code with (in the left table) $q = 2$, $m = 10$, $r = 7$ and $n = 1024$, and (in the right table) $q = 5$, $m = 4$, $r = 9$ and $n = 625$. Indeed, we have $s_* = 5$ on the left and $s_* = 4$ on the right.

| $s$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $\dim\left(\ker\left(\tilde{\varphi}^{(3)}\right)\right)$ | 950 | 600 | 350 | 150 | 40 | 10 | 0 |

| $s$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $\dim\left(\ker\left(\tilde{\varphi}^{(3)}\right)\right)$ | 224 | 108 | 40 | 16 | 4 | 0 |

### 5.3 The special case of Goppa codes

For simplicity, we restrict our discussion to the case $q = 2$. It turns out that in the case of Goppa codes, we can shorten more than for generic alternant codes while still obtaining a non-trivial kernel:

**Proposition 9.** *Let $\boldsymbol{x} \in \mathbb{F}_{2^m}^n$ have pairwise distinct entries and $C = \mathrm{Gop}(\boldsymbol{x}, g)^\perp$ be the dual of a proper binary Goppa code with square-free Goppa polynomial $g$*

*of degree $t$. Let $\hat{e} \coloneqq \left\lfloor \log_4\left(\frac{2t-1}{p-1}\right)\right\rfloor$ and $\boldsymbol{A}_{\hat{s}_*}$ be a generator matrix of the code obtained by shortening $C$ in*

$$\hat{s}_* \coloneqq 2(\hat{e}+1)t - (p-1)\cdot\frac{4^{\hat{e}+1}-1}{3} - p$$

*positions. Then the kernel of $\tilde{\varphi}^{(p)}_{\boldsymbol{A}_{\hat{s}_*}}$ is non-trivial.*

*Proof.* We construct a matrix similar to $\boldsymbol{\Phi}^{(p)}$ in the alternant case, but with some adaptions. By Proposition 2, we have $\bigoplus_{i=0}^{\lfloor m/2\rfloor-1}\mathrm{GRS}_{2t}(\boldsymbol{x}^{4^i},\boldsymbol{y}^{2\cdot 4^i}) \subset \mathrm{span}_{\mathbb{F}_{q^m}}(C)$. A generator matrix of this subcode of $\mathrm{span}_{\mathbb{F}_{q^m}}(C)$ is given by

$$\boldsymbol{A}'_{\mathbb{F}_{q^m}} = \begin{pmatrix} \boldsymbol{G}'_{\boldsymbol{x},\boldsymbol{y}} \\ (\boldsymbol{G}'_{\boldsymbol{x},\boldsymbol{y}})^{(4)} \\ (\boldsymbol{G}'_{\boldsymbol{x},\boldsymbol{y}})^{(4^2)} \\ \cdots \\ (\boldsymbol{G}'_{\boldsymbol{x},\boldsymbol{y}})^{(4^{\lfloor m/2\rfloor-1})} \end{pmatrix} \text{ with } \boldsymbol{G}'_{\boldsymbol{x},\boldsymbol{y}} = \begin{pmatrix} y_1^2 & y_2^2 & \cdots & y_n^2 \\ y_1^2x_1 & y_2^2x_2 & \cdots & y_n^2x_n \\ y_1^2x_1^2 & y_2^2x_2^2 & \cdots & y_n^2x_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ y_1^2x_1^{2t-1} & y_2^2x_2^{2t-1} & \cdots & y_n^2x_n^{2t-1} \end{pmatrix}$$

It is sufficient to prove the claim for this subcode. We rename indeterminates as

$$X_0^{(0)},\ldots,X_{2t-1}^{(0)},X_0^{(1)},\ldots,X_{2t-1}^{(1)},\ldots,X_0^{(\lfloor m/2\rfloor-1)},\ldots,X_{2t-1}^{(\lfloor m/2\rfloor-1)}$$

in order to correspond to the blocks of $(\boldsymbol{G}'_{\boldsymbol{x},\boldsymbol{y}})^{(4^j)}$ in $\boldsymbol{A}'_{\mathbb{F}_{q^m}}$. For $0 \le u \le \hat{e}$ and $0 \le v < \lfloor m/2\rfloor$, we define the $p \times (2t-(p-1)4^u)$ matrix

$$\boldsymbol{B}'_{u,v} \coloneqq \begin{pmatrix} X_0^{(v)} & X_1^{(v)} & \ldots & X_{2t-1-(p-1)4^u}^{(v)} \\ X_{4^u}^{(v)} & X_{1+4^u}^{(v)} & \ldots & X_{2t-1-(p-2)4^u}^{(v)} \\ \cdots & \cdots & \cdots & \cdots \\ X_{(p-1)4^u}^{(v)} & X_{1+(p-1)4^u}^{(v)} & \cdots & X_{2t-1}^{(v)} \end{pmatrix}$$

and the block matrix $\boldsymbol{\Phi}^{(p)}_{Gop} \coloneqq (\boldsymbol{B}_{0,\hat{e}} \mid \boldsymbol{B}_{1,\hat{e}-1} \mid \ldots \mid \boldsymbol{B}_{\hat{e},0})$. Note that this matrix has $p$ rows and $2(\hat{e}+1)t-(p-1)\cdot\frac{4^{\hat{e}+1}-1}{3}$ columns. It is constructed such that if we substitute any column of $\boldsymbol{A}'_{\mathbb{F}_{q^m}}$ in $\boldsymbol{\Phi}^{(p)}_{Gop}$ via $X_i^{(j)} \mapsto (y_\ell^2x_\ell^i)^{4^j}$, we obtain a matrix where each row is obtained from the previous one by multiplication by $x_\ell^{4^{\hat{e}}}$. Now the exact same argument leading to Proposition 8 proves our assertion. □

*Remark 8.* Experiments show that this expression for $\hat{s}_*$ is not tight, i.e. that one can usually shorten more while still maintaining a non-trivial kernel. For example, for binary Goppa codes with $m = 10$, $r = 7$ and $n = 1024$, the formula in Proposition 9 implies we can shorten in 15 positions. In experiments, we find that we can in fact shorten in 17 positions and still obtain a non-trivial kernel (of dimension 10).

## 6 The distinguisher

Suppose now that for fixed $m, r, n \in \mathbb{N}$, a $\mathbb{F}_q$-code $C$ is randomly sampled from one of

(i) the set of proper alternant (or binary Goppa) codes over $\mathbb{F}_q$ with field extension degree $m$, degree $r$ and length $n$
(ii) the set of linear $[n, k]$-codes over $\mathbb{F}_q$ where $k = mr$

and one is to decide whether $C$ comes from (i) or (ii). Our distinguisher proceeds as follows. Initially, we set $p = 2$. Then:

1. Let $s$ be the integer defined in Proposition 8 (if we want to distinguish alternant codes) or the integer defined in Proposition 9 (in the case we want to distinguish binary Goppa codes).
   If $s < 0$, then abort. Otherwise, shorten the code $C^\perp$ in $s$ positions and let $\boldsymbol{A} \in \mathbb{F}_q^{k' \times n'}$ be a generator matrix of this shortened code where $k' = k - s$ and $n' = n - s$.
2. Determine the smallest $\hat{n}$ such that the degree $p$ coefficient of
$$\left[(1+t)^{k'} \cdot \left(1 - t^2\right)^{\hat{n}}\right]^+ \tag{7}$$
   is zero. If $\hat{n} > n' - k'$, set $p = p + 1$ and return to step 1.
3. Determine whether or not the kernel of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ is trivial. One way is to determine
$$\beta^{(p)} := \dim_{\mathbb{F}_q}\left(\ker\left(\tilde{\varphi}_{\boldsymbol{A}}^{(p)}\right)\right)$$
   by computing a matrix representation of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ and applying Gaussian elimination. Alternatively, as we shall see below, one can use Wiedemann's algorithm [26] to determine whether $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ has full rank. Another way is to compute a Gröbner basis of the ideal $\mathfrak{L}_{\tilde{\boldsymbol{A}}}$ (defined in Definition 7) with respect to some monomial order and to check whether all monomials of degree $p$ in $\tilde{R}_{k'}$ occur as leading monomials or not. If the kernel of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ is trivial, we output that $C$ is random, otherwise we output that $C$ is alternant/Goppa.

In step 1, we shorten in just enough positions that we still expect $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ to have a non-trivial kernel in the alternant or Goppa case. Note that in the case of binary Goppa codes, by Remark 8, we can usually shorten a bit more than the amount predicted by Proposition 9. Thus, we are overly strict in the Goppa case by aborting in case the expression for $s$ is negative. In practice, we could try to increase the amount of shortening a bit more in the Goppa case before aborting. Assuming Heuristic 1 holds, step 2 ensures that the kernel of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ is indeed trivial in the random case with high probability.

Step 3 dominates the complexity of our distinguisher. We can employ Wiedemann's algorithm [26] to determine whether $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ has full rank. Given a $n_0 \times n_1$ matrix over $\mathbb{F}_q$ with $n_1 < n_0$ such that at most $\omega$ field operations are required to

apply the matrix to a vector (which can be achieved for instance if the matrix is sparse with at most $\omega$ non-zero entries), Wiedemann's algorithm runs in time

$$\mathcal{O}(n_1 \cdot (\omega + n_0 \cdot \log(n_0)) \tag{8}$$

It either finds a non-trivial element in the kernel of the matrix, or it outputs that it has full rank with an error probability bounded away from 1. By re-running it sufficiently many times, we can determine with high probability whether the matrix has full rank or not.

We thus obtain the following complexity estimate:

**Proposition 10.** *Suppose we wish to distinguish binary Goppa codes from random binary codes, so $q = 2$. Let $k, n, p \in \mathbb{N}$. Let $s$ be chosen as in Proposition 9 and let $\boldsymbol{A} \in \mathbb{F}_2^{(k-s)\times(n-s)}$ be as above. Assume that $p$ has been chosen such that the degree $p$ coefficient in (7) is zero, and let $\ell \coloneqq \lfloor \log_2(p) \rfloor$. Using Wiedemann's algorithm in step 3, the distinguisher's runtime is*

$$\mathcal{O}\left(N_1 \cdot \left(N_1(n-k)\sum_{i=1}^{\ell}\binom{p}{2^i} + N_0 \cdot \log(N_0)\right)\right)$$

*where $N_0 \coloneqq (n-k) \cdot \sum_{i=1}^{\ell}\binom{k-s}{p-2^i}$ and $N_1 \coloneqq \binom{k-s}{p}$.*

*Proof.* Since $q = 2$, by Remark 4 it is sufficient to compute the derivatives of order $p-2$, $p-4$, ... and $p-2^\ell$ in the definition of $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$. Let $\mathcal{M}$ be the matrix representing $\tilde{\varphi}_{\boldsymbol{A}}^{(p)}$ with respect to the monomials of degree $p$ with each indeterminate of degree at most 1 as a basis of the domain and the standard basis as a basis of the range. Since step 3 dominates the complexity, it suffices to estimate the complexity of Wiedemann's algorithm to determine whether $\mathcal{M}$ has full rank. $\mathcal{M}$ has $N_0$ rows and $N_1$ columns. For a fixed monomial of degree $p$, at most $\sum_{i=1}^{\ell}\binom{p}{2^i}$ of the considered partial derivatives are non-zero. Hence the matrix $\mathcal{M}$ contains at most $N_1 \cdot (n-k) \cdot \sum_{i=1}^{\ell}\binom{p}{2^i}$ non-zero entries. The claim now follows from (8). □

Given any concrete parameter set, it is easy to determine whether our distinguisher is able to successfully distinguish or not, and which amount of shortening and which value of $p$ are required. This allows us to determine that our distinguisher does indeed work e.g. for Classic McEliece parameter sets, and to estimate its complexity in this case in subsection 6.2. Giving a closed expression in terms of the parameters $n$ and $k$ for the size of $p$ required seems to be difficult though. For, it is hard to provide a closed formula in terms of $n$ and $k$ that determines the lowest degree where the coefficient in $(1+t)^k \cdot \left(1-t^2\right)^{\tilde{n}}$ is negative. In other words, it is hard to determine a closed formula for the degree of regularity of a semi-regular sequence. For any given concrete parameters, this is easy to compute of course. [4] provides asymptotic formulas and [14] determines bounds in some cases.

**Table 6.** Results for $q = 2$, $m = 10$, $r = 10$ at $p = 3$, shortening in 27 positions.

| $n$ | $\geq 953$ | 952 | 951 | 950 | 949 | 948 | 947 | 946 |
|---|---|---|---|---|---|---|---|---|
| $\beta^{(3)}$ (Goppa) | 350 | 350 | 350 | 350 | 350 | 350 | 365 | 438 |
| $\beta^{(3)}$ (random) | 0 | 1 | 73 | 146 | 219 | 292 | 365 | 438 |

**Table 7.** Results for $q = 2$, $m = 12$, $r = 17$ at $p = 3$, shortening in 64 positions.

| $n$ | $\geq 3401$ | 3400 | 3399 | 3398 | 3397 |
|---|---|---|---|---|---|
| $\beta^{(3)}$ (Goppa) | 240 | 240 | 280 | 420 | 560 |
| $\beta^{(3)}$ (random) | 0 | 140 | 280 | 420 | 560 |

**Table 8.** Results for $q = 2$, $m = 8$, $r = 6$ at $p = 4$, shortening in 5 positions.

| $n$ | $\geq 197$ | 196 | 195 | 194 |
|---|---|---|---|---|
| $\beta^{(4)}$ (Goppa) | 1120 | 1120 | 1400 | 2157 |
| $\beta^{(4)}$ (random) | 0 | 644 | 1400 | 2157 |

**Table 9.** Results for $q = 3$, $m = 5$, $r = 9$ at $p = 4$, shortening in 2 positions.

| $n$ | $\geq 194$ | 193 | 192 | 191 |
|---|---|---|---|---|
| $\beta^{(4)}$ (alternant) | 5 | 644 | 1400 | 2157 |
| $\beta^{(4)}$ (random) | 0 | 644 | 1400 | 2157 |

### 6.1 Some example computations

We present some experimental results for our distinguisher for different parameter sets below. We implemented our distinguishing algorithm in SageMath and ran our experiments on a machine equipped with two AMD EPYC 9745 128-Core Processors and 4TiB of RAM. Our implementation parallelises the computation of the representing matrix. We conducted each experiment several times. The values of $\beta^{(p)}$ that we obtained in below experiments in the case of randomly chosen codes almost always matched the prediction of Heuristic 1. There were only very slight deviations. In the random case in Table 6, we sometimes obtained dimension 2 for $n = 952$ instead of 1 as indicated in the table, while the prediction from Heuristic 1 would be 0. Similarly, in Table 7, we sometimes obtained $\beta^{(3)} = 1$ instead of $\beta^{(3)} = 0$ for $n = 3401$ in the random case. We observed no more significant deviations. This provides further evidence for the validity of Heuristic 1.

*Example 2.* We consider the same parameters as in [23, Example 2] which seeks to distinguish binary Goppa codes with parameters $m = 10$, $n = 1024$ and $r = 10$ from random codes. For these parameters, our distinguisher works at degree $p = 3$. Proposition 9 predicts that we can shorten the code in 27 positions. Using this, we experimentally find that $\beta^{(3)} = 350$ in the Goppa case, while $\beta^{(3)} = 0$ in the random case. See Table 6 for more detailed experimental results. Our computation required less than 1GB of RAM, about 4.2 CPU hours and a few minutes of elapsed real time.

*Example 3.* We consider binary Goppa codes with parameters $m = 12$, $r = 17$. This is significantly larger than Example 2. Shortening in 64 positions, we obtain the results displayed in Table 7 for different values of $n$. Thus, we see that we can distinguish at degree $p = 3$ for $n \geq 3400$. The formula in Proposition 9 predicts that we can shorten in 57 positions for these parameter sets and $p = 3$, showing that Proposition 9 is not quite tight. The computation took less than 32GB of RAM, about 231 CPU hours and less than 4 hours of elapsed real time.

*Example 4.* We apply the distinguisher at degree $p = 4$ to the parameter set $m = 8$, $r = 6$ to distinguish binary Goppa codes and random codes. The experimental results can be found in Table 8, indicating that we can distinguish for $n \geq 196$. The computation took less than 3GB of RAM, about 5 CPU hours and less than 10 minutes of elapsed real time.

*Example 5.* In Table 9 we provide an example for distinguishing alternant codes in characteristic 3 by applying the distinguisher at $p = 4$ to the parameter set $m = 5$, $r = 9$. The values of $\beta^{(4)}$ we obtain in the random case for various $n$ precisely match the expression in Heuristic 1, confirming our conjecture that it holds in all characteristics. The computation took less than 90GB of RAM, about 8 CPU hours and less than 2.5 hours of elapsed real time.

### 6.2 Performance for Classic McEliece parameters

Consider a binary Goppa code $\mathrm{Gop}(\boldsymbol{x}, g)$ with field extension degree $m$, length $n$ and Goppa degree $\deg(g) = r$ with irreducible Goppa polynomial $g$. For the Classic McEliece parameters, Table 10 presents the choice of $p$ and the required shortening in order to distinguish Goppa codes with these parameters from random linear codes. These estimates hold under the assumption that Heuristic 1 is valid. By Remark 8, it is likely we can actually shorten in more positions than indicated in the table, so the parameters can probably be made a bit smaller. We also estimate the complexity of our distinguisher. This is, of course, only a very rough estimate since it only computes the asymptotically dominant term given in Proposition 10 and does not take into account constant factors or lower-order terms. This caveat should be kept in mind.

**Table 10.** Required degree $p$ and number of shortened positions $\hat{s}_*$ for distinguishing Goppa codes with indicated Classic McEliece parameters from random linear codes. $N$ and $M$ denote the number of rows and columns, respectively, of the matrix representing $\tilde{\varphi}^{(p)}$. $\kappa_W$ is the term inside $\mathcal{O}(\cdot)$ in Proposition 10.

| $(n, r, m)$ | $(3488, 64, 12)$ | $(4608, 96, 13)$ | $(6688, 128, 13)$ | $(6960, 119, 13)$ | $(8192, 128, 13)$ |
|---|---|---|---|---|---|
| $\hat{s}_*$ | 123 | 91 | 175 | 217 | 271 |
| $p$ | 23 | 51 | 57 | 44 | 41 |
| $\log_2(N)$ | 146.3 | 304.8 | 352.4 | 282.3 | 270.5 |
| $\log_2(M)$ | 139.6 | 297.5 | 344.7 | 274.8 | 262.8 |
| $\kappa_W$ | $2^{310}$ | $2^{650}$ | $2^{748}$ | $2^{601}$ | $2^{575}$ |

### 6.3 Relation to the syzygy distinguisher

Let $C$ be a $[n,k]$-code over $\mathbb{F}_q$ with generator matrix $\boldsymbol{B}$ whose columns we denote by $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_n$ and rows by $\boldsymbol{c}_1,\ldots,\boldsymbol{c}_k$. [23, Proposition 14] allows one to compare the syzygy distinguisher with the approach presented here. [23, Proposition 14 and Lemma 14] describes the syzygy module as $\ker \partial|_T$ where, in our notation,

$$\partial : \left(\bigwedge^{p-1} R_k^{(1)}\right) \otimes C \longrightarrow \left(\bigwedge^{p-2} R_k^{(1)}\right) \otimes \mathbb{F}_q^n$$

and $T$ is the subspace spanned by basis vectors

$$\boldsymbol{e}_{i_1,\ldots,i_{p-1};l} \coloneqq (X_{i_1} \wedge \cdots \wedge X_{i_{p-1}}) \otimes \boldsymbol{c}_l$$

with $1 \leq i_1 < \cdots < i_{p-1} \leq k$ and $1 \leq l \leq i_{p-1}$.

$\partial|_T$ is given explicitly by the images of the basis vectors

$$\boldsymbol{e}_{i_1,\ldots,i_{p-1};l} \mapsto \sum_{j=1}^{p-1} (-1)^j X_{i_1} \wedge \cdots \hat{X_{i_j}} \cdots \wedge X_{i_{p-1}} \otimes [\mathrm{ev}_{\boldsymbol{b}_1}(X_{i_j} X_l), \ldots, \mathrm{ev}_{\boldsymbol{b}_n}(X_{i_j} X_l)]$$

where notation $\hat{X_{i_j}}$ means that $X_{i_j}$ is omitted and $\mathrm{ev}_{\boldsymbol{c}}$ denotes evaluation of a polynomial at a vector $\boldsymbol{c}$. We now assume that $\boldsymbol{B}$ is in reduced row-echelon form and that $\boldsymbol{b}_s \in \mathbb{F}_q^k$ is equal to the $s$-th standard basis vector for $s \leq k$. Let $T' \subset T$ be spanned by all basis vectors $\boldsymbol{e}_{i_1,\ldots,i_{p-1};l}$ in $T$ such that $l$ is different from all $i_1,\ldots,i_{p-1}$. Then one can see that $\ker \partial|_T$ is equal to the kernel of the map

$$\partial'' \colon T' \to \left(\bigwedge^{p-2} R_k^{(1)}\right) \otimes \mathbb{F}_q^{n-k}$$

defined by

$$\begin{aligned}\boldsymbol{e}_{i_1,\ldots,i_{p-1};l} \mapsto \sum_{j=1}^{p-1} ((-1)^j X_{i_1} \wedge \cdots \hat{X_{i_j}} \cdots \wedge X_{i_{p-1}} \\ \otimes [\mathrm{ev}_{\boldsymbol{b}_{k+1}}(X_{i_j} X_l), \ldots, \mathrm{ev}_{\boldsymbol{b}_n}(X_{i_j} X_l)])\end{aligned}$$

If we modify $\bar{\varphi}_{\boldsymbol{B}}^{(p)}$ (as defined in Definition 6) to consider derivatives of order equal to $p-2$ only, this can be understood as the map defined on monomials $\mu$ as

$$\begin{aligned} U_k^{(p)} &\to \left(\bigwedge^{p-2} R_k^{(1)}\right) \otimes \mathbb{F}_q^{n-k}, \\ \mu &\mapsto \sum_{\substack{\nu \text{ with } \nu \mid \mu, \\ \deg(\nu) = p-2}} \nu \otimes \left[\mathrm{ev}_{\boldsymbol{b}_{k+1}}\left(\frac{\mu}{\nu}\right), \ldots, \mathrm{ev}_{\boldsymbol{b}_n}\left(\frac{\mu}{\nu}\right)\right] \end{aligned}$$

This map is very similar to the map $\partial''$ by considering $\nu = X_{i_1} \wedge \cdots \hat{X_{i_j}} \cdots \wedge X_{i_{p-1}}$ and $\mu = \nu X_l X_{i_j}$. However, there are differences. For each $\mu$ the image is a sum over $\binom{p}{2}$ terms, whereas $\partial''$ is a sum over $p-1$ terms. The number of basis elements in $\partial''$ of the form $\boldsymbol{e}_{i_1,\ldots,i_{p-1};l}$ is equal to $(p-1)\binom{k}{p}$. In the end, it does not seem likely that the distinguisher presented here can be reduced to the distinguisher from [23]. But both distinguishers consider similar elements in the space $\left(\bigwedge^{p-2} R_k^{(1)}\right) \otimes \mathbb{F}_q^{n-k}$.

**Acknowledgments.** We would like to thank Jan Geuenich as well as the anonymous reviewers for thoroughly reading a previous version of this paper and providing very helpful comments and suggestions.

## References

1. Classic McEliece: conservative code-based cryptography: cryptosystem specification. https://classic.mceliece.org/mceliece-spec-20221023.pdf (2022), [Online; accessed 29-January-2026]
2. Classic McEliece: conservative code-based cryptography: guide for security reviewers. https://classic.mceliece.org/mceliece-security-20221023.pdf (2022), [Online; accessed 29-January-2026]
3. Alagic, G., Bros, M., Ciadoux, P., Cooper, D., Dang, Q., Dang, T., Kelsey, J., Lichtinger, J., Liu, Y.K., Miller, C., et al.: Status report on the fourth round of the NIST post-quantum cryptography standardization process. US Department of Commerce, National Institute of Standards and Technology (2025)
4. Bardet, M., Faugere, J.C., Salvy, B., Yang, B.Y.: Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In: Proc. of MEGA. vol. 5, pp. 2–2 (2005)
5. Bardet, M., Mora, R., Tillich, J.: Polynomial time key-recovery attack on high rate random alternant codes. IEEE Trans. Inf. Theory **70**(6), 4492–4511 (2024), https://doi.org/10.1109/TIT.2023.3334592
6. Becker, A., Joux, A., May, A., Meurer, A.: Decoding random binary linear codes in 2 (n/20): How 1 + 1 = 0 improves information set decoding. In: Pointcheval, D., Johansson, T. (eds.) Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, April 15-19, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7237, pp. 520–536. Springer (2012), https://doi.org/10.1007/978-3-642-29011-4_31
7. Berlekamp, E.R.: Algebraic coding theory (revised edition). Aegean Park Press (1984)
8. Berlekamp, E.R., McEliece, R.J., van Tilborg, H.C.A.: On the inherent intractability of certain coding problems. IEEE Trans. Inf. Theory **24**(3), 384–386 (1978), https://doi.org/10.1109/TIT.1978.1055873
9. Cascudo, I., Cramer, R., Mirandola, D., Zémor, G.: Squares of random linear codes. IEEE Trans. Inf. Theory **61**(3), 1159–1173 (2015), https://doi.org/10.1109/TIT.2015.2393251
10. Couvreur, A., Mora, R., Tillich, J.: A new approach based on quadratic forms to attack the McEliece cryptosystem. In: Guo, J., Steinfeld, R. (eds.) Advances in

Cryptology - ASIACRYPT 2023 - 29th International Conference on the Theory and Application of Cryptology and Information Security, Guangzhou, China, December 4-8, 2023, Proceedings, Part IV. Lecture Notes in Computer Science, vol. 14441, pp. 3–38. Springer (2023), https://doi.org/10.1007/978-981-99-8730-6_1
11. Emsalem, J., Iarrobino, A.: Inverse system of a symbolic power, I. Journal of Algebra **174**(3), 1080–1090 (1995). https://doi.org/https://doi.org/10.1006/jabr.1995.1168
12. Faugère, J., Gauthier-Umaña, V., Otmani, A., Perret, L., Tillich, J.: A distinguisher for high rate McEliece cryptosystems. In: 2011 IEEE Information Theory Workshop, ITW 2011, Paraty, Brazil, October 16-20, 2011. pp. 282–286. IEEE (2011), https://doi.org/10.1109/ITW.2011.6089437
13. Hodges, T.J., Iyer, H.: Semi-regularity of pairs of boolean polynomials. Finite Fields Their Appl. **79**, 101992 (2022), https://doi.org/10.1016/j.ffa.2022.101992
14. Kousidis, S.: Krawtchouk polynomials and quadratic semi-regular sequences. In: Davenport, J.H., Wang, D., Kauers, M., Bradford, R.J. (eds.) Proceedings of the 2019 on International Symposium on Symbolic and Algebraic Computation, ISSAC 2019, Beijing, China, July 15-18, 2019. pp. 243–250. ACM (2019), https://doi.org/10.1145/3326229.3326230
15. Kruglov, V.: Growth of the ideal generated by a quadratic multivariate function. University of Cincinnati (2010)
16. MacWilliams, F.J., Sloane, N.J.A.: The theory of error-correcting codes, vol. 16. Elsevier (1992), seventh impression
17. Massey, J.L., von Seeman, N., Schöller, P.: Hasse derivatives and repeated-root cyclic codes. In: IEEE Symposium of Info. Theory, Ann. Arbor, USA (1986)
18. May, A., Meurer, A., Thomae, E.: Decoding random linear codes in O(2 0.054n)). In: Lee, D.H., Wang, X. (eds.) Advances in Cryptology - ASIACRYPT 2011 - 17th International Conference on the Theory and Application of Cryptology and Information Security, Seoul, South Korea, December 4-8, 2011. Proceedings. Lecture Notes in Computer Science, vol. 7073, pp. 107–124. Springer (2011), https://doi.org/10.1007/978-3-642-25385-0_6
19. McEliece, R.J.: A public-key cryptosystem based on algebraic coding theory. The Deep Space Network Progress Report **42-44**, 114–116 (1978)
20. Mora, R., Tillich, J.: On the dimension and structure of the square of the dual of a Goppa code. Des. Codes Cryptogr. **91**(4), 1351–1372 (2023), https://doi.org/10.1007/s10623-022-01153-w
21. Pellikaan, R., Márquez-Corbella, I.: Error-correcting pairs for a public-key cryptosystem. In: Journal of Physics: Conference Series. vol. 855, p. 012032. IOP Publishing (2017)
22. Prange, E.: The use of information sets in decoding cyclic codes. IRE Trans. Inf. Theory **8**(5), 5–9 (1962), https://doi.org/10.1109/TIT.1962.1057777
23. Randriambololona, H.: The syzygy distinguisher. IACR Cryptol. ePrint Arch. p. 1193 (16/10/2025), https://eprint.iacr.org/2024/1193
24. Semaev, I., Tenti, A.: Probabilistic analysis on Macaulay matrices over finite fields and complexity of constructing Gröbner bases. Journal of Algebra **565**, 651–674 (2021), https://doi.org/10.1016/j.jalgebra.2020.08.035
25. Stern, J.: A method for finding codewords of small weight. In: Cohen, G.D., Wolfmann, J. (eds.) Coding Theory and Applications, 3rd International Colloquium, Toulon, France, November 2-4, 1988, Proceedings. Lecture Notes in Computer Science, vol. 388, pp. 106–113. Springer (1988), https://doi.org/10.1007/BFb0019850
26. Wiedemann, D.H.: Solving sparse linear equations over finite fields. IEEE Trans. Inf. Theory **32**(1), 54–62 (1986), https://doi.org/10.1109/TIT.1986.1057137