# Projective Jet Geometry and Hold-Out Rank of Binary Goppa Codes*

Yongge Wang
Department of Cybersecurity, University of North Carolina at Charlotte
Charlotte, NC 28223, USA
yonwang@charlotte.edu

3 November 2026

**Abstract**

We develop a projective-geometric and algebraic framework for the higher-order polynomial relations among the public generator-matrix columns of a binary Goppa code, the relations exploited by the recent structural distinguishers and key-recovery methods for Classic McEliece. The alternant representation realizes the public columns as rational points on a hidden curve parametrized over $\mathbb{F}_{2^m}$; higher-order vanishing conditions become fat-point conditions, pulled back along the hidden parametrization.

Our first contribution is an exact characterization of a codimension-$t$ ambient space containing all products of $d$ code polynomials: the products satisfy $(\pi^d h)' \equiv 0 \pmod{g^2}$ for the support polynomial $\pi$ and the Goppa polynomial $g$, and the space of degree-$\leq dD$ polynomials with this property has dimension $dD + 1 - t$ once $dD \geq 2t - 1$. Equality with the actual product space is proved in a local regime, verified computationally on every instance tested, and stated as a conjecture. Our second contribution is a Frobenius-descent theorem: a binary form vanishes on the hidden curve if and only if it vanishes on the union of its $m$ Frobenius conjugates, with the $\mathbb{F}_2$-dimension of such forms equal to the $K$-dimension of the ideal of the union; from it we derive Hilbert-function bounds, including a multilinear version attained on all but one of the twenty-seven keys with $t \geq 5$ tested (and missed by one on that key), and the local structure of the pulled-back forms at the roots of $g$.

We then analyze the universal dependencies among Hasse-jet conditions in the multilinear setting of the attacks: a closed formula for the rank at a single column, and an explicit family of $\binom{k}{2s-d-2}$ relations between the conditions at any two columns. At a held-out column we prove an exact codimension formula for the Goppa subcode generated by the gradients of the vanishing forms. Finally, running our experimental pipeline (the jet system of the attacks, without the key-recovery steps) on six independently generated keys of a Classic-McEliece-shaped code shows that the kernel dimension is comparable to that of a random control, while every kernel form exhibits the predicted hidden-curve vanishing, the gradients span the predicted subcode, and the first-order normal data of the kernel is nearly independent of the key. Together these results explain the main components of the nuisance space of the recent attacks and isolate the conjectures that remain for an exact description of the kernel.

*This research was conducted with the assistance of AI models, including Anthropic's Claude Fable 5.1 and OpenAI's GPT-6. These models were used to develop and debug experimental scripts, execute and cross-validate computational results, test the findings, rephrase portions of the manuscript, and assist in reviewing, checking, and refining mathematical proofs. The AETHER IDE https://amalithlab.github.io was also used to coordinate and facilitate collaboration among multiple LLMs throughout the research process.

# Contents


**1 Introduction** **3**
- 1.1 Context: the 2024–2026 structural attacks . . . 4
- 1.2 Conventions . . . 5

**2 Binary Goppa codes and their alternant representation** **5**
- 2.1 Definition and parity checks . . . 5
- 2.2 Code polynomials and the GRS containment . . . 5
- 2.3 Shortening . . . 6

**3 Public projective geometry and jet spaces** **6**
- 3.1 The public point configuration . . . 6
- 3.2 Hasse jets . . . 7
- 3.3 Evaluation matrix and scalar extension . . . 8
- 3.4 Universal relations over $\mathbb{F}_2$ . . . 8
- 3.5 Multilinear forms and systematic form . . . 8

**4 Fat points and restricted polynomial spaces** **9**

**5 The hidden rational curve and clean-coordinate completeness** **9**
- 5.1 The hidden map . . . 9
- 5.2 Pullback of jets and completeness . . . 10

**6 The space of pulled-back forms** **10**

**7 Towards Conjecture 6.4: structure of the product space** **12**
- 7.1 The product spaces as local conditions at the roots of $g$ . . . 12
- 7.2 Derivatives of products of code polynomials . . . 13
- 7.3 A reduction of the $d = 2$ case . . . 14
- 7.4 What the conjecture is really about . . . 14
- 7.5 Results on the strategy: S1 proved, S2 reduced . . . 15
- 7.6 A proof of the conjecture in the local regime . . . 18
- 7.7 What remains . . . 19

**8 Frobenius structure: the conjugate curves** **20**
- 8.1 Compatibility . . . 20
- 8.2 The descent theorem . . . 20
- 8.3 The Hilbert function of the union of conjugate curves . . . 22
- 8.4 Exact rank of the public jet matrix for quadrics . . . 23

**9 Exact hold-out rank** **24**
- 9.1 The Holdout/Hermite distinguisher in this language . . . 24

**10 Jets at the held-out column** **25**

**11 The multilinear setting: universal relations and the bridge to the attacks** **27**
11.1 No universal value relations for multilinear forms . . . . . 27
11.2 The universal relations at one point . . . . . 27
11.3 The multilinear part of the curve ideal . . . . . 30
11.4 The smallest shortened systematic instances . . . . . 31
11.5 A Classic-McEliece-shaped instance: $(m, t, n) = (6, 6, 64)$ . . . . . 32
11.6 Anatomy of the structural kernel . . . . . 35

**12 Dirty-coordinate soundness** **36**

**13 Quadratic relations and the rank-two attack** **37**

**14 The generic baseline: Alexander–Hirschowitz over $\mathbb{F}_2$** **37**

**15 A computational program** **38**
15.1 Experimental methods and reproducibility . . . . . 38

**16 Conclusion** **40**

# 1 Introduction

Let $\Gamma(L, g)$ be a binary Goppa code over $\mathbb{F}_2$, defined by a support $L \subset \mathbb{F}_{2^m}$ and a Goppa polynomial $g$. A public generator matrix hides the support and the polynomial, but its columns form a finite configuration of points over $\mathbb{F}_2$. The algebraic-cryptanalytic approach that has advanced furthest in 2025–2026 searches for low-degree polynomials that vanish, to higher order, on many of these public points, and asks what the vanishing forces at the remaining points.

This paper isolates the geometric and rank-theoretic foundations of that approach. Its goals are:

(1) to formulate higher-order public polynomial relations as projective jets and fat-point conditions, with the characteristic-two bookkeeping (Hasse derivatives, Euler's identity, universal relations over $\mathbb{F}_2$) made explicit;

(2) to use the alternant structure to prove the clean-coordinate completeness bound and to determine exactly the space of pulled-back forms;

(3) to describe the Frobenius structure correctly, as $m$ conjugate curves through the same rational points, and to prove the corresponding descent theorem;

(4) to express the information contributed by held-out columns as an exact restricted-rank invariant, and to connect it to the published attacks;

(5) to describe the universal relations among the Hasse conditions of the attacks in closed form at one column (Theorem 11.4) and to construct, for every pair of columns, $\binom{k_\ell}{2s-d-2}$ explicit relations (Theorem 11.6); together these give a first algebraic description of the main components of the nuisance space that Saarinen and Weis remove numerically, while the complete structure of the cross-point dependencies remains open;

(6) to run the jet system of the attacks, without the key-recovery steps, on keys of Classic-McEliece shape, $(m, t, n) = (6, 6, 64)$ and $(7, 7, 128)$, and identify what its kernel is made of;

(7) to state the open problems whose solution would turn the heuristic cost estimates of the current literature into theorems.

**Main contributions.** Three results carry the paper: (1) the exact ambient space $\mathcal{W}_d$ of the products of code polynomials and its codimension $t$ (Theorem 6.1, with the root characterization of Lemma 7.1 and the local structure of Theorem 7.21); (2) the Frobenius descent theorem and the Hilbert-function bounds for the union of conjugate curves (Theorems 8.4, 8.8, 11.11); (3) the exact single-point rank and the pairwise relations among Hasse conditions (Theorems 11.4 and 11.6), together with the gradient theorem at a held-out column (Theorem 10.1) and the multi-key experiments of Section 11.5. Everything else is a refinement, an experiment, or an open problem, and is labelled as such.

The paper is foundational rather than a claim of a new attack. Several of its statements are reformulations, in projective language, of lemmas that appear in the attack papers; where this is so we say so. Because characteristic two invites several natural-looking but false shortcuts (ordinary instead of Hasse derivatives, Frobenius acting on the support instead of on the parameterization, characteristic-zero genericity over $\mathbb{F}_2$), we flag these explicitly where they arise.

### 1.1 Context: the 2024–2026 structural attacks

The syzygy distinguisher of Randriambololona [13] (Eurocrypt 2025) used graded Betti numbers of point configurations obtained from shortened duals of the public code, and was the first distinguisher for binary Goppa codes that is asymptotically subexponential in $t$. The higher-order vanishing (HOV) distinguisher of Hemmert and Wiemers [14] (ePrint 2025/1661, Crypto 2026) computes homogeneous polynomials that vanish to higher order at the columns of a shortened public *dual*; Hemmert [15] turned it into a key recovery that exploits the decomposition of the scalar extension of the dual into $m$ Frobenius-conjugate GRS blocks, and Saarinen's HOVER [16] replaced its MinRank search by linear algebra.

The Holdout/Hermite distinguisher of Ghoshal, Ishai, Jain and Sun [18] (ePrint 2026/1630, August 2026) works on the *primal* side: it is a classical distinguisher for Goppa–McEliece that runs in time $n^{O(\log n)}$ in the regime $m = \Theta(\log n)$, $t = \Theta(n/\log n)$, $k = \Theta(n)$, with advantage $1 - o(1)$, and it originated from a failed attempt to build doubly efficient PIR from algebraic locally decodable codes. Its mechanism is exactly the one formalized in Sections 3–5 below: the public columns are points of a hidden rational curve, multilinear polynomials of degree $d$ are forced to vanish to order $s$ at all but one column of a shortened, systematic instance, and on a Goppa key the pulled-back polynomial has more forced zeros than its degree allows, so it vanishes at the held-out column as well. Weis [19] sharpened the parameters (from $(d, s) = (8, 7)$ to $(7, 6)$ for the Classic McEliece sets), lowered the cost of one sparse-kernel "run" to $2^{89}$–$2^{98}$ bit operations in a memory-free model, gave two key-recovery routes (one from gradients at held-out columns, one from a single run), proved two facts about binary Goppa codes that we re-derive and extend in Sections 6 and 10, and recovered the TII-253 challenge key. Saarinen [22] costed the GIJS locator route item by item, and solved TII-254 ($(m, t, n) = (8, 12, 223)$, a $[223, 127]$ code) by a two-anchor Holdout/Hermite method [23] that computed two 121-dimensional relation kernels conditioned at distinct public coordinates, an 80-dimensional pair core, a 64-dimensional common nuisance space, and a 16-dimensional remainder identified as an $\mathbb{F}_{2^8}$ projective-line geometry. Vedenev [20] proposed key recovery from derivative spaces of the hidden curve; Apon [21] showed that one step of Vedenev's original route needs more than $2t + 3$ held positions, because the derivative-flag block at each held point restricts to a single evaluation condition on an explicit $(2t + 3)$-dimensional family, and Vedenev's October revision replaced that step. In a different direction, Briaud, Lemoine, Randriambololona and Tillich [24] gave a heuristic, conjecturally subexponential key recovery via rank-two matrices in the code of quadratic relations, at best of order $2^{454}$ for the level-1 Classic McEliece parameters. The German BSI advised against new deployments of Classic McEliece in a note dated 29 September 2026 and announced on

1 October 2026 [25], while stating that no practical attack on its recommended parameter sets exists; the scheme had been standardized by ISO/IEC in June 2026. Fernandez [26] gives a version-aware assessment of all of these results, which we have used to check the statements above.

This paper takes no position on the policy question. It aims to provide exact geometric tools for the structural analysis, and to state precisely which of the ingredients of the current estimates are proved and which are heuristic.

### 1.2 Conventions

Throughout, $K = \mathbb{F}_{2^m}$, $\sigma : K \to K$ is the Frobenius $a \mapsto a^2$, and for $f = \sum c_a X^a \in K[X]$ we write $f^{(2)} = \sum c_a^2 X^a$ for the polynomial with Frobenius-conjugated coefficients, so that $f^{(2)}(\alpha^2) = f(\alpha)^2$. All derivatives are Hasse derivatives. *Vanishing to order $s$* at a point means membership in the $s$-th power of the maximal ideal of the point; we use the multiplicity $s$ throughout, rather than the jet order $s-1$, to avoid off-by-one errors in the counting arguments.

## 2 Binary Goppa codes and their alternant representation

### 2.1 Definition and parity checks

Let $L = (\alpha_1, \dots, \alpha_n) \subset K$ consist of pairwise distinct elements, let $g \in K[X]$ have degree $t$ with $g(\alpha_i) \neq 0$ for all $i$, and let

$$\pi(X) = \prod_{i=1}^{n} (X - \alpha_i), \qquad \pi_S(X) = \prod_{i \in S} (X - \alpha_i) \quad (S \subseteq \{1, \dots, n\}).$$

The binary Goppa code is

$$\Gamma(L, g) = \left\{ c \in \mathbb{F}_2^n : \sum_{i=1}^{n} \frac{c_i}{X - \alpha_i} \equiv 0 \pmod{g(X)} \right\} = \ker_{\mathbb{F}_2}(\widehat{H}_g), \tag{2.1}$$

where $\widehat{H}_g \in \mathbb{F}_2^{mt \times n}$ is the $\mathbb{F}_2$-expansion of $H_g = (\alpha_i^j / g(\alpha_i))_{0 \le j < t,\, 1 \le i \le n}$. Hence $k = \dim_{\mathbb{F}_2} \Gamma(L, g) \ge n - mt$, with equality for the parameters considered here (we always record the actual $k$). Multiplying the congruence in (2.1) by $\pi$ gives the classical identity

$$\sum_{i:\, c_i = 1} \frac{\pi(X)}{X - \alpha_i} = \pi_c'(X)\, \pi_{\bar{c}}(X), \qquad c \in \Gamma(L, g) \iff g \mid \pi_c', \tag{2.2}$$

where $\pi_c = \pi_{\{i : c_i = 1\}}$ and $\pi_{\bar{c}} = \pi / \pi_c$. In characteristic two the derivative of any polynomial has only even-degree terms, so $\pi_c' = s_c^2$ for a unique $s_c \in K[X]$; if $g$ is square-free then $g \mid \pi_c' = s_c^2$ forces $g \mid s_c$, hence $g^2 \mid \pi_c'$, which is the identity $\Gamma(L, g) = \Gamma(L, g^2)$ and the source of the designed distance $2t + 1$. We assume $g$ square-free from now on; in Classic McEliece $g$ is irreducible.

### 2.2 Code polynomials and the GRS containment

By $\Gamma(L, g) = \Gamma(L, g^2)$ the code is the subfield subcode of the dual of $\mathrm{GRS}_{2t}(L, 1/g^2)$, which is the GRS code

$$\mathrm{GRS}_{D+1}(L, v) = \{(v_i f(\alpha_i))_i : \deg f \le D\}, \qquad v_i = \frac{g(\alpha_i)^2}{\pi'(\alpha_i)}, \qquad D = n - 2t - 1. \tag{2.3}$$

Every codeword $c$ therefore has a unique *code polynomial* $f_c \in K[X]_{\le D}$ with $c_i = v_i f_c(\alpha_i)$. Identity (2.2) gives it explicitly.

**Lemma 2.1** (Explicit code polynomial)**.** *For $c \in \Gamma(L, g)$,*

$$f_c = \frac{\pi_c' \, \pi_{\bar{c}}}{g^2} = u_c^2 \, \pi_{\bar{c}}, \qquad u_c = \frac{s_c}{g} \in K[X], \tag{2.4}$$

*and*

$$\pi_c \, \pi_{\bar{c}}' = \pi' + g^2 f_c. \tag{2.5}$$

*Proof.* The Lagrange interpolant of the values $\pi'(\alpha_i)c_i$ at $L$ is $\sum_{c_i=1} \pi/(X - \alpha_i) = \pi_c' \pi_{\bar{c}}$ by (2.2); dividing by $g^2$, which is coprime to $\pi$ and divides $\pi_c'$, gives a polynomial of degree $\leq n - 1 - 2t = D$ taking the values $c_i \pi'(\alpha_i)/g(\alpha_i)^2 = c_i/v_i$ at $L$. This is $f_c$. For (2.5), differentiate $\pi = \pi_c \pi_{\bar{c}}$ and substitute $\pi_c' \pi_{\bar{c}} = g^2 f_c$. □

Let $G \in \mathbb{F}_2^{k \times n}$ be a full-row-rank generator matrix with rows $c^{(1)}, \ldots, c^{(k)}$ and let $f_j = f_{c^{(j)}}$. Since $\mathbb{F}_2$-independent binary vectors remain $K$-independent, the $f_j$ are $K$-linearly independent, and the $i$-th column of $G$ is

$$y_i = v_i \, (f_1(\alpha_i), \ldots, f_k(\alpha_i))^\top \in \mathbb{F}_2^k. \tag{2.6}$$

**Proposition 2.2** (Point at infinity)**.** *The coefficient of $X^D$ in $f_c$ is* $\mathrm{wt}(c) \bmod 2 \in \mathbb{F}_2$. *Consequently the vector of leading coefficients $(\mathrm{lc}_D f_1, \ldots, \mathrm{lc}_D f_k)$ equals the sum $\sum_{i=1}^n y_i$ of all public columns. Whenever some row of $G$ has odd weight, this sum is a nonzero binary vector and $[\sum_i y_i]$ is the image of the point $X = \infty$ under the hidden map of Section 5.*

*Proof.* $\pi_c'$ has degree $\leq \mathrm{wt}(c) - 1$ and its coefficient of $X^{\mathrm{wt}(c)-1}$ is $\mathrm{wt}(c) \bmod 2$; $\pi_{\bar{c}}$ is monic of degree $n - \mathrm{wt}(c)$; so the coefficient of $X^{n-1}$ in $\pi_c' \pi_{\bar{c}}$ is $\mathrm{wt}(c) \bmod 2$, and division by the monic $g^2$ moves it to $X^D$. For a row $c$ of $G$, $\mathrm{wt}(c) \bmod 2 = \sum_i c_i$ is the row's entry in $\sum_i y_i$. □

This is Weis's observation [19, Prop. 2] that the sum of all columns of an instance is one more point of the hidden curve, at which vanishing conditions can also be imposed; the proof above shows that it is literally $\Phi(\infty)$, with no scalar ambiguity, when the multipliers (2.3) are used.

### 2.3 Shortening

The attacks [18, 19] do not work with $G$ but with a shortened code. For $S \subseteq \{1, \ldots, n\}$ with $|S| = \ell$, the shortened code $C_S$ (codewords vanishing on $S$, restricted to the complement) is the binary Goppa code $\Gamma(L \setminus L_S, g)$ of length $n_\ell = n - \ell$ and, generically, dimension $k_\ell = k - \ell$. Its code polynomials are the $f_c/\pi_S$ for $c \in C$ vanishing on $S$ (since $\pi_S \mid \pi_{\bar{c}}$), of degree at most $D_\ell = n_\ell - 2t - 1 = D - \ell$. Everything below applies to the shortened code with $(n, k, D)$ replaced by $(n_\ell, k_\ell, D_\ell)$.

## 3 Public projective geometry and jet spaces

### 3.1 The public point configuration

Each column $y_i$ of $G$ is nonzero (otherwise a coordinate is identically zero in the code). Two equal columns would mean that $e_i + e_{i'}$ lies in the *dual* code, i.e. that $\Gamma(L, g)^\perp$ has minimum distance $\leq 2$; we assume throughout that this does not happen:

(A0) the columns of $G$ are pairwise distinct ($d(\Gamma^\perp) \geq 3$).

(The distinctness of the parity-check columns $(\alpha_i^j/g(\alpha_i))_j$, which follows from $\alpha_i \neq \alpha_{i'}$, is the statement $d(\Gamma) \geq 3$ and is not the same thing.) (A0) held for every key generated in this paper and is implicit in the attack literature, where the systematic non-unit columns are treated as distinct points; a proof from the Goppa structure would be welcome. Under (A0) the columns determine $n$ distinct rational points

$$P_i = [y_i] \in \mathbb{P}^{k-1}(\mathbb{F}_2). \tag{3.1}$$

Let $S = \mathbb{F}_2[Y_1, \ldots, Y_k]$ and $S_d$ its degree-$d$ component, of dimension $\binom{k+d-1}{d}$. We write $S_{d,K} = S_d \otimes_{\mathbb{F}_2} K$.

### 3.2 Hasse jets

For a multi-index $\nu \in \mathbb{N}^k$ the Hasse derivative $\partial^{[\nu]}H$ of $H \in S_d$ is defined by the Taylor expansion $H(Y + Z) = \sum_\nu \partial^{[\nu]}H(Y)\, Z^\nu$. In characteristic two the Hasse derivatives differ from iterated ordinary derivatives and from the naive rule "$\partial Y_i^2 = 0$":

*Remark* 3.1 (Hasse derivatives of low-degree forms over $\mathbb{F}_2$). Three facts about Hasse derivatives are used repeatedly.

(i) For a quadratic form $Q = \sum_{i \leq j} M_{ij} Y_i Y_j$, the first Hasse derivatives are the polar forms

$$\partial^{[e_i]}Q = \sum_{j \neq i} M_{ij} Y_j = B_Q(e_i, \cdot),$$

which are not identically zero. First-order jet vanishing of $Q$ at $P$ (i.e. $Q \in \mathfrak{p}_P^2$) means $Q(P) = 0$ and $P$ lies in the radical of $B_Q$; it is a non-vacuous condition.

(ii) For $Q = Y_1 Y_2$ the ordinary and Hasse derivatives agree, $\partial Q/\partial Y_1 = \partial^{[e_1]}Q = Y_2$; for $Q = Y_1^2$ the ordinary derivative is $2Y_1 = 0$ and so is $\partial^{[e_1]}Q$, but $\partial^{[2e_1]}Q = 1$: the second-order Hasse derivative sees what the ordinary derivative loses. Likewise $\partial^{[e_i]}(Y_i^2 Y_j) = 0$ but $\partial^{[e_j]}(Y_i^2 Y_j) = Y_i^2$. Hence a cubic monomial with a squared variable has a nonzero gradient at most rational points, and the multilinear reduction of a form has a different jet from the form itself.

(iii) (Euler) $\sum_i Y_i\, \partial^{[e_i]}H = d \cdot H$. For odd $d$ this expresses the value of $H$ at a point through its first Hasse derivatives, so the "value condition" at a point is a consequence of the $k$ affine gradient conditions; the number of independent order-$\leq 1$ conditions at a point is $k$ either way.

Fix $P \in \mathbb{P}^{k-1}(\mathbb{F}_2)$ and an affine chart with local coordinates $z_1, \ldots, z_{k-1}$ centered at $P$. A form $H \in S_d$ dehomogenizes to $H_P(z)$, and $H$ vanishes to order $\geq s$ at $P$ if and only if all coefficients of $H_P$ of total degree $< s$ vanish. The number of such coefficients is

$$q_{k,s} = \binom{k+s-2}{k-1}. \tag{3.2}$$

Vanishing to order $s$ is intrinsic: it is membership in $\mathfrak{p}_P^s$, where $\mathfrak{p}_P$ is the homogeneous prime of $P$.

**Definition 3.2** (Jet relation space)**.** For $I \subseteq \{1, \ldots, n\}$ and a multiplicity vector $s = (s_i)_{i \in I}$,

$$\mathcal{K}_{d,s}(G; I) = \{H \in S_d :\ H \in \mathfrak{p}_{P_i}^{s_i} \text{ for every } i \in I\}. \tag{3.3}$$

When all $s_i$ equal $s$ we write $\mathcal{K}_{d,s}(G; I)$ with $s$ a number.

### 3.3 Evaluation matrix and scalar extension

Choosing charts and monomial orderings gives the public jet-evaluation map

$$E_{d,s,I} : S_d \longrightarrow \bigoplus_{i\in I} \mathbb{F}_2^{q_{k,s_i}}, \qquad \mathcal{K}_{d,s}(G; I) = \ker E_{d,s,I}. \tag{3.4}$$

**Proposition 3.3** (Scalar-extension invariance)**.** $\operatorname{rank}_{\mathbb{F}_2} E_{d,s,I} = \operatorname{rank}_K (E_{d,s,I} \otimes K)$.

*Proof.* Rank is the largest size of a nonzero minor, and minors in $\mathbb{F}_2$ stay nonzero in $K$. □

Frobenius phenomena therefore never enter through scalar extension of the public matrix; they enter through the hidden parameterization (Section 8).

### 3.4 Universal relations over $\mathbb{F}_2$

Over a finite field some forms vanish at *every* rational point. The homogeneous ideal of $\mathbb{P}^{k-1}(\mathbb{F}_2)$ is generated by the forms $Y_i^2 Y_j + Y_i Y_j^2$ $(i < j)$, and its degree-three component has dimension $\binom{k}{2}$. Consequently, for any configuration of rational points, the pure point-evaluation part of the cubic jet matrix has a kernel of dimension at least $\binom{k}{2}$, independent of the configuration. These forms do not vanish to order two at rational points (their Hasse gradients are nonzero by Remark 3.1(ii)), so they do not lie in $\mathcal{K}_{3,2}$; but any "generic rank" prediction for the value conditions over $\mathbb{F}_2$ must quotient them out. More generally, the universal relations among Hasse conditions that hold for every configuration are what Weis [19, Prop. 6] calls *local relations* and Saarinen [23] removes as a *nuisance space* (4 617 031 directions in the TII-254 computation). The one-point part of this universal subspace is described exactly in Theorem 11.4; its cross-point part is Open Problem 14.2.

### 3.5 Multilinear forms and systematic form

The attacks use the multilinear subspace $W_d^{\mathrm{ml}} \subset S_d$ (square-free monomials, $\dim = \binom{k}{d}$) and a generator matrix in systematic form. The reason is the following elementary fact.

**Lemma 3.4** (Unit columns)**.** *A multilinear form of degree $d$ vanishes to order $\geq d-1$ at every coordinate point $e_i \in \mathbb{P}^{k-1}$.*

*Proof.* In the chart $Y_i = 1$ at $e_i$, a square-free monomial of degree $d$ containing $Y_i$ becomes a monomial of degree $d-1$ in the other variables, and one not containing $Y_i$ has degree $d$. □

Thus in a systematic instance with $k_\ell$ unit columns and $mt$ non-unit columns, only the non-unit columns need conditions: order $s$ at the constrained columns and none at the held-out one [18, 19]. Over $\mathbb{F}_2$ the Hasse conditions on multilinear forms are further reducible: Hemmert and Wiemers [14] observe, and Weis uses, that the conditions of orders $d-2, d-4, d-8, \ldots$ imply all the others, and the HOV literature works with the square-free partial derivatives of orders $p - 2^\ell$ [17]. We do not reprove these reductions here; they change the row count of the public matrix but not the kernel.

*Caution* 3.5. It is tempting to argue that restricting to multilinear forms loses no jet information over $\mathbb{F}_2$ because "the gradient of $Y_i^2 Y_j$ vanishes". By Remark 3.1(ii) this is false: the restriction does change the jet conditions, and its justification is Lemma 3.4 together with the kernel-dimension heuristic of Section 14, not a conservation of information.

# 4 Fat points and restricted polynomial spaces

For a finite set $I$ of distinct public points and multiplicities $s_i$, let $Z = \sum_{i\in I} s_i P_i$ be the fat-point subscheme with ideal $I(Z) = \bigcap_{i\in I} \mathfrak{p}_{P_i}^{s_i}$.

**Theorem 4.1** (Fat-point identity)**.** *$\mathcal{K}_{d,s}(G;I) = I(Z)_d$ and $\dim \mathcal{K}_{d,s}(G;I) = \binom{k+d-1}{d} - \operatorname{rank} E_{d,s,I}$.*

*Proof.* Local membership in $\mathfrak{p}_{P_i}^{s_i}$ is exactly the vanishing of the Hasse jet of order $< s_i$; intersect over $i$ and apply rank–nullity. □

*Remark* 4.2 (Restricted spaces). For a subspace $W_d \subseteq S_d$ (multilinear, sparse, ...) the relation space is $W_d \cap I(Z)_d = \ker(E_{d,s,I}|_{W_d})$, of dimension $\dim W_d - \operatorname{rank}(E_{d,s,I}|_{W_d})$. The completeness statements below hold for each $H \in W_d$ individually; the dimension statements do not transfer and must be computed in $W_d$.

# 5 The hidden rational curve and clean-coordinate completeness

## 5.1 The hidden map

The representation (2.6) defines a rational map

$$\Phi : \mathbb{P}^1_K \dashrightarrow \mathbb{P}^{k-1}_K, \qquad X \longmapsto [f_1(X) : \cdots : f_k(X)], \qquad \Phi(\alpha_i) = P_i. \tag{5.1}$$

Since $y_i \neq 0$ and $v_i \neq 0$, $\Phi$ is defined at every support point. For $H \in S_{d,K}$ define the *pullback*

$$h_H(X) = H(f_1(X), \ldots, f_k(X)) \in K[X]_{\leq dD}. \tag{5.2}$$

Let $\Gamma = \overline{\Phi(\mathbb{P}^1)} \subset \mathbb{P}^{k-1}_K$ be the image curve and $I(\Gamma) \subset S \otimes K$ its homogeneous ideal, so that $I(\Gamma)_d = \ker(H \mapsto h_H)$. Let $\delta \leq D$ be the degree of $\Phi^*\mathcal{O}(1)$ (the common degree of the $f_j$ after removing a common factor; such a factor is coprime to $\pi$ since no $y_i$ vanishes, but nothing so far excludes a common root outside $L$). The algebraic results of this paper (Theorems 5.2, 6.1, 8.4, 8.8, 10.1) are statements about the polynomials $f_j$ and need no geometric hypothesis beyond (A0); the completeness bound is stated with $\delta$. Where the *image curve* is discussed as a geometric object (its degree, its singularities, the genus of $X_\Gamma$ in Section 8.3) we assume in addition:

(A1) $\gcd(f_1, \ldots, f_k) = 1$, so that $\Phi$ is a morphism; together with (A3), which supplies an $f_j$ of degree exactly $D$ (Proposition 2.2), this gives $\delta = D$;

(A2) $\Phi$ is birational onto its image, the points $\Phi(\beta)$, $\beta \in Z(g)$, are distinct cusps with $\delta$-invariant one, and there are no other singularities;

(A3) $\sum_i y_i \neq 0$, i.e. $\Gamma(L, g)$ contains a codeword of odd weight, so that $\Phi(\infty)$ is a point.

(A1) and (A3) are checked directly on an instance and held for all instances computed here; (A2) is the geometric picture suggested by the ramification in Proposition 7.13(iii) and is used only in the Riemann–Roch interpretation, which is marked as such.

### 5.2 Pullback of jets and completeness

**Lemma 5.1** (Jets pull back to multiplicities)**.** *If $H \in \mathfrak{p}^s_{P_i}$ then $(X-\alpha_i)^s \mid h_H$.*

*Proof.* In a chart at $P_i$ in which some coordinate of $\Phi(\alpha_i)$ is a unit, the local coordinate functions $z_j(X)$ of $\Phi$ vanish at $\alpha_i$, so $H_{P_i}(z(X))$, which has only terms of degree $\geq s$ in $z$, has order $\geq s$ at $\alpha_i$; $h_H$ differs from it by the $d$-th power of a unit at $\alpha_i$. □

**Theorem 5.2** (Clean-coordinate completeness)**.** *Let $H \in S_{d,K}$ satisfy $H \in \mathfrak{p}^{s_i}_{P_i}$ for all $i \in I$. If*

$$\sum_{i\in I} s_i > d\,\delta \quad (\textit{in particular if } \sum_{i\in I} s_i > dD), \tag{5.3}$$

*then $h_H = 0$; hence $H(y_j) = 0$ for every $j \in \{1,\dots,n\}$ and $H(\sum_i y_i) = 0$, and in fact $H$ vanishes on the whole curve $\Gamma$.*

*Proof.* The pairwise coprime factors $(X-\alpha_i)^{s_i}$ divide $h_H$, which is a section of $\Phi^*\mathcal{O}(d)$ of degree $d\delta$ after cancelling the common factor; if $\sum s_i > d\delta$ the section is zero. Then $H(y_j) = v_j^d h_H(\alpha_j) = 0$, and $H(\sum_i y_i)$ is the leading coefficient of $h_H$ by Proposition 2.2. □

*Remark* 5.3 (Attribution and scope). Theorem 5.2 is the core lemma of the Holdout/Hermite distinguisher [18]: on a Goppa key the restriction $\varphi_h(Z) = h(p(Z))$ has degree at most $dD_\ell$ and is forced to vanish to order $s$ at the constrained columns and to order $d-1$ at the unit columns (Lemma 3.4), so once the forced zeros outnumber the degree, $\varphi_h \equiv 0$ and $h(y_\tau) = \lambda_\tau^d \varphi_h(\alpha_\tau) = 0$ at the held-out column [18, 19, 26]. Our contribution is the projective fat-point language, the non-uniform multiplicities, and the inclusion of the point at infinity. The theorem proves completeness for clean coordinates only; it says nothing about dirty ones (Section 12), and it assumes the full jet conditions (3.3).

*Remark* 5.4 (Which $(d,s)$ are usable). For the unshortened code with uniform multiplicity $s$ at all $n$ columns, (5.3) reads $sn > d(n-2t-1)$, i.e. $s \geq d$ or $n < d(2t+1)/(d-s)$. For cubics with double points ($d=3$, $s=2$) this requires $n < 6t+3$, which fails at every Classic McEliece parameter set; for cubics with triple points the $N\binom{k+1}{2} \times \binom{k+2}{3}$ system is heavily overdetermined and its kernel is generically zero. Neither cubic system is usable at cryptographic size. The actual attacks therefore shorten to a small $k_\ell$ (so that the unit columns supply $(d-1)k_\ell$ free forced zeros), use multilinear forms of degree $d=7$ with $s=6$ at the $mt-1$ constrained columns, and need $(d-1)k_\ell + s(mt-1)$ to exceed $dD_\ell - t$, the $t$ being Theorem 6.1 below. With $s = d-1$ this is $k_\ell + mt < 2dt + t + 1$, i.e. $k_\ell < (2d+1-m)t+1$, which explains why $d \geq 6$ is needed for $m=12$ and $d \geq 7$ for $m=13$ (at $m=13$, $d=6$ leaves no room for a single unit column), and why the choice $(d,s)=(7,6)$ of [19] is natural. The kernel then has to be nonzero, which is the separate, heuristic condition of Section 14.

## 6 The space of pulled-back forms

Let $\mathcal{V}_d = \{h_H : H \in S_{d,K}\} \subseteq K[X]_{\leq dD}$ be the space of pulled-back forms, i.e. the $K$-span of the products of $d$ code polynomials. For a generic $k$-dimensional subspace of $K[X]_{\leq D}$ one expects $\dim \mathcal{V}_d = \min\{\binom{k+d-1}{d}, dD+1\}$. Ghoshal et al. observed experimentally that binary Goppa codes fall short of this, and Weis [19, Prop. 1] proved that the products lie in an explicit space of dimension $dD+1-t$, under the name "$t$ zeros to spare". We give an independent derivation that identifies the space exactly, from Lemma 2.1.

**Theorem 6.1** (The pulled-back space)**.** *Let $\Gamma(L, g)$ be a binary Goppa code with square-free $g$, and let $h = f_1 \cdots f_d$ be a product of $d$ code polynomials (not necessarily distinct). Then*

$$(\pi^d h)' \equiv 0 \pmod{g^2}. \tag{6.1}$$

*Equivalently, $(\pi^d h)' = g^2 w^2$ for some $w \in K[X]$. Consequently*

$$\mathcal{V}_d \subseteq \mathcal{W}_d := \{h \in K[X]_{\le dD} : \ g^2 \mid (\pi^d h)'\}, \qquad \dim_K \mathcal{W}_d \ \ge \ dD + 1 - t, \tag{6.2}$$

*with $\dim_K \mathcal{W}_d = dD + 1 - t$ whenever $dD \ge 2t - 1$, for either parity of $d$.*

*Proof.* Write $f_j = u_j^2 \pi_{\bar{c}_j}$ as in (2.4) and $U = \prod_j u_j$, so that $h = U^2 \prod_j \pi_{\bar{c}_j}$. Since $(U^2)' = 0$ in characteristic two,

$$h' = U^2 \sum_{j=1}^{d} \pi'_{\bar{c}_j} \prod_{i \ne j} \pi_{\bar{c}_i} \equiv U^2 \sum_{j=1}^{d} \frac{\pi'}{\pi_{c_j}} \prod_{i \ne j} \pi_{\bar{c}_i} \pmod{g^2},$$

using (2.5) in the form $\pi'_{\bar{c}_j} = (\pi' + g^2 f_j)/\pi_{c_j}$, which is legitimate modulo $g^2$ because $\pi_{c_j}$ is coprime to $g$ (so $\pi_{c_j}$ is invertible in $K[X]/(g^2)$ and the identity $\pi_{c_j}\pi'_{\bar{c}_j} = \pi' + g^2 f_j$ may be divided by it). Since $\pi_{\bar{c}_i}\pi_{c_i} = \pi$, multiplying the $j$-th summand by $\pi_{c_j}/\pi_{c_j}$ gives $\prod_{i \ne j} \pi_{\bar{c}_i} = \pi^{d-1}\pi_{c_j} / \prod_i \pi_{c_i}$, so each summand equals $\pi'\pi^{d-1} / \prod_i \pi_{c_i}$ modulo $g^2$, and

$$h' \equiv d \cdot \frac{U^2 \pi' \pi^{d-1}}{\prod_i \pi_{c_i}} = d \cdot \frac{\pi'}{\pi} h \pmod{g^2}.$$

If $d$ is even this says $h' \equiv 0$, and then $(\pi^d h)' = \pi^d h' \equiv 0$. If $d$ is odd it says $\pi h' + \pi' h \equiv 0$, i.e. $(\pi h)' \equiv 0$, and then $(\pi^d h)' = \pi^{d-1}(\pi h)' \equiv 0$. This proves (6.1); the square form follows because derivatives in characteristic two are squares and $g$ is square-free.

For the dimension, consider the $K$-linear map $\Psi : K[X]_{\le dD} \to K[X]/(g^2)$, $h \mapsto (\pi^d h)' \bmod g^2$, whose kernel is $\mathcal{W}_d$. Its image consists of squares modulo $g^2$. The squaring map $w \mapsto w^2$ on $K[X]/(g^2)$ is additive and $\sigma$-semilinear, so its image is a $K$-subspace; its kernel is $\{w : g \mid w\}$, of dimension $t$, so the image has dimension $2t - t = t$. Hence $\operatorname{rank} \Psi \le t$ and $\dim \mathcal{W}_d \ge dD + 1 - t$. For equality it suffices that $\Psi$ hits every square modulo $g^2$. For even $d$ the image of $\Psi$ is $\{\pi^d h' \bmod g^2\}$; as $h$ ranges over $K[X]_{\le dD}$ the derivative $h'$ ranges over all squares of degree $\le dD - 1$, and a square $w^2$ modulo $g^2$ depends only on $w \bmod g$, so these cover all squares modulo $g^2$ once $(dD - 1)/2 \ge t - 1$; multiplication by the square unit $\pi^d = (\pi^{d/2})^2$ permutes the squares modulo $g^2$. For odd $d$ take $h = \pi w$ with $\deg w \le dD - n$; then $(\pi^d h)' = \pi^{d+1} w'$ and the same argument applies once $dD - n - 1 \ge 2(t - 1)$. The sharper and parity-free condition $dD \ge 2t - 1$ follows from the root characterization of Lemma 7.1 below: $\mathcal{W}_d$ is cut out of $K[X]_{\le dD}$ by the $t$ conditions $h'(\beta) = (d \bmod 2)\lambda_\beta h(\beta)$ at the roots $\beta$ of $g$, and these are independent as soon as Hermite interpolation can prescribe $h(\beta)$ and $h'(\beta)$ freely at the $t$ roots, i.e. as soon as $dD + 1 \ge 2t$. □

**Corollary 6.2.** $\dim_K I(\Gamma)_d = \binom{k+d-1}{d} - \dim_K \mathcal{V}_d \ \ge \ \binom{k+d-1}{d} - (dD + 1 - t)$*, with equality if and only if $\mathcal{V}_d = \mathcal{W}_d$.*

**Experiment 6.3** ($\mathcal{V}_d = \mathcal{W}_d$ on small instances)**.** For random square-free $g$ and random supports, we computed $\dim_K \mathcal{V}_d$ (rank of the coefficient matrix of all $\binom{k+d-1}{d}$ products), $\dim_K \mathcal{W}_d$, and the identity (6.1) on every product. In all cases the identity held, $\dim \mathcal{W}_d = dD + 1 - t$, and $\mathcal{V}_d = \mathcal{W}_d$ for every $d \ge 2$ (for $d = 1$, $\mathcal{V}_1$ is the $k$-dimensional space of code polynomials and $\mathcal{W}_1$ is larger).

| $m$ | $n$ | $t$ | $k$ | $D$ | $d$ | $\binom{k+d-1}{d}$ | $dD+1$ | $\dim \mathcal{V}_d = \dim \mathcal{W}_d$ | $dD+1-t$ |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 16 | 2 | 8 | 11 | 2 | 36 | 23 | 21 | 21 |
| 4 | 16 | 2 | 8 | 11 | 3 | 120 | 34 | 32 | 32 |
| 4 | 16 | 2 | 8 | 11 | 4 | 330 | 45 | 43 | 43 |
| 4 | 14 | 2 | 6 | 9 | 2 | 21 | 19 | 17 | 17 |
| 4 | 14 | 2 | 6 | 9 | 3 | 56 | 28 | 26 | 26 |
| 5 | 32 | 3 | 17 | 25 | 2 | 153 | 51 | 48 | 48 |
| 5 | 32 | 3 | 17 | 25 | 3 | 969 | 76 | 73 | 73 |
| 5 | 28 | 3 | 13 | 21 | 2 | 91 | 43 | 40 | 40 |
| 5 | 28 | 3 | 13 | 21 | 3 | 455 | 64 | 61 | 61 |
| 5 | 32 | 4 | 12 | 23 | 2 | 78 | 47 | 43 | 43 |
| 5 | 32 | 4 | 12 | 23 | 3 | 364 | 70 | 66 | 66 |
| 6 | 64 | 3 | 46 | 57 | 2 | 1081 | 115 | 112 | 112 |
| 6 | 60 | 4 | 36 | 51 | 2 | 666 | 103 | 99 | 99 |

Across keys: for 20 random keys each of $(5, 32, 3)$ $(d = 2, 3)$, $(6, 64, 3)$ $(d = 2)$ and $(6, 64, 6)$ $(d = 2)$, seeds 100–119, $\dim \mathcal{V}_d = dD + 1 - t$ held in all 80 cases.

**Conjecture 6.4.** $\mathcal{V}_d = \mathcal{W}_d$ for every binary Goppa code with square-free $g$ and every $d \geq 2$ such that $\binom{k+d-1}{d} > dD + 1 - t$. (Equality $\binom{k+d-1}{d} = dD + 1 - t$ does not suffice: see Experiment 7.18; a refined form is Conjecture 7.19.)

*Remark* 6.5. Weis states his Proposition 1 as a structural identity ("the product of code polynomials is $\Gamma^2$ times a square" in his normalization) and uses the $t$ saved zeros heuristically; his Theorem 3 shows that, for $\Gamma$ uniform among monic irreducibles of degree $t$, a nonzero polynomial with the forced zeros exists with probability at most $n_\ell q^{-(t-\delta+1)}$ when the zero count falls short by $\delta \leq t$ [19, 26]. Theorem 6.1 gives the exact description of the ambient space as a kernel; Conjecture 6.4 is the statement that the products fill it.

*Caution* 6.6. The generic expectation $\dim \mathcal{V}_d = \min\{\binom{k+d-1}{d}, d\delta + 1\}$, which one might try to justify by the base-point-free pencil trick, is false for binary Goppa codes: the span of the code polynomials is a Galois-structured, not a generic, subspace of $K[X]_{\leq D}$, and the defect is $t$. For the same reason $\sum s_i \leq \dim \mathcal{V}_d$ Hermite functionals need not be independent on $\mathcal{V}_d$: they are independent on all of $K[X]_{\leq dD}$, not on an arbitrary proper subspace. The correct rank statement is Theorem 6.1 together with Weis's probabilistic Theorem 3.

# 7 Towards Conjecture 6.4: structure of the product space

Throughout this section $g$ is square-free, $\pi = \prod_i (X - \alpha_i)$, $D = n - 2t - 1$, $\mathcal{V}_d = \mathcal{V}_1^d$ is the $K$-span of products of $d$ code polynomials, and $\mathcal{W}_d = \{h \in K[X]_{\leq dD} : g^2 \mid (\pi^d h)'\}$. We write $\bar{K}$ for an algebraic closure of $K$ and $Z(g) \subset \bar{K}$ for the $t$ distinct roots of $g$. For $\beta \in Z(g)$ put

$$\lambda_\beta := \frac{\pi'(\beta)}{\pi(\beta)} \in \bar{K}^\times \qquad (\pi(\beta) \neq 0 \text{ since } \gcd(g, \pi) = 1).$$

## 7.1 The product spaces as local conditions at the roots of $g$

**Lemma 7.1** (Root characterization)**.** *For $h \in K[X]_{\leq dD}$,*

$$h \in \mathcal{W}_d \iff h'(\beta) = (d \bmod 2)\, \lambda_\beta\, h(\beta) \quad \text{for every } \beta \in Z(g).$$

*In particular $\mathcal{W}_a \cdot \mathcal{W}_b \subseteq \mathcal{W}_{a+b}$ for all $a, b \geq 1$, and $\dim_K \mathcal{W}_d = dD + 1 - t$ whenever $dD \geq 2t - 1$.*

*Proof.* In characteristic two $(\pi^d h)'$ is a square $s^2$, and since $g$ is square-free, $g^2 \mid s^2 \iff g \mid s \iff s(\beta)=0$ for all $\beta \in Z(g) \iff (\pi^d h)'(\beta)=0$ for all $\beta$. Now $(\pi^d h)' = \pi^{d-1}\big(d\,\pi' h + \pi h'\big)$, and $\pi(\beta)\neq 0$, which gives the displayed condition. If $h_a \in \mathcal{W}_a$, $h_b \in \mathcal{W}_b$ then $(h_a h_b)'(\beta) = h_a'(\beta)h_b(\beta) + h_a(\beta)h_b'(\beta) = \big((a \bmod 2) + (b \bmod 2)\big)\lambda_\beta(h_a h_b)(\beta) = ((a+b) \bmod 2)\lambda_\beta(h_a h_b)(\beta)$, so $h_a h_b \in \mathcal{W}_{a+b}$. Finally the $t$ functionals $h \mapsto h'(\beta) - \varepsilon\lambda_\beta h(\beta)$ ($\beta \in Z(g)$) are linearly independent on $K[X]_{\le M}$ as soon as $M \ge 2t-1$, since Hermite interpolation with multiplicity two at the $t$ distinct points $\beta$ lets one prescribe $h(\beta)=0$ and $h'(\beta)$ arbitrarily. (The functionals take values in $\bar{K}$; the $K$-linear conditions they cut out are the $t$ conditions "$(\pi^d h)' \equiv 0 \bmod g^2$" of Theorem 6.1, whose image lies in the $t$-dimensional space of squares modulo $g^2$.) □

*Remark* 7.2. Lemma 7.1 gives a one-line proof of Theorem 6.1: by (2.5), $\pi_c \pi'_{\bar{c}} = \pi' + g^2 f_c$, so at $\beta \in Z(g)$, $\pi_c(\beta)\pi'_{\bar{c}}(\beta) = \pi'(\beta)$, and $f_c'(\beta) = u_c(\beta)^2 \pi'_{\bar{c}}(\beta) = u_c(\beta)^2 \pi'(\beta)/\pi_c(\beta) = \lambda_\beta\, u_c(\beta)^2 \pi_{\bar{c}}(\beta) = \lambda_\beta f_c(\beta)$. Hence $\mathcal{V}_1 \subseteq \mathcal{W}_1$, and $\mathcal{V}_d \subseteq \mathcal{W}_1^d \subseteq \mathcal{W}_d$ by the Leibniz rule. It also shows that $\mathcal{W}_d$ depends on $(L,g)$ only through $\pi$ and $Z(g)$, not through the code.

**Lemma 7.3** (Complements)**.** *For $M \ge 2t-1$ and $\varepsilon \in \{0,1\}$ let $\mathcal{W}_M^{(\varepsilon)} = \{h \in K[X]_{\le M} : h'(\beta) = \varepsilon\lambda_\beta h(\beta)\ \forall \beta \in Z(g)\}$ (so $\mathcal{W}_d = \mathcal{W}_{dD}^{(d \bmod 2)}$). Then*

$$K[X]_{\le M} = \mathcal{W}_M^{(\varepsilon)} \oplus g\, K[X]_{\le t-1}, \qquad g^2 K[X]_{\le M-2t} \subseteq \mathcal{W}_M^{(\varepsilon)}.$$

*Explicitly, for every $P \in K[X]_{\le M}$ there is a unique $R_P \in K[X]_{\le t-1}$, namely the polynomial with $R_P(\beta) = \big(\varepsilon\lambda_\beta P(\beta) - P'(\beta)\big)/g'(\beta)$, such that $P + gR_P \in \mathcal{W}_M^{(\varepsilon)}$.*

*Proof.* If $gR \in \mathcal{W}_M^{(\varepsilon)}$ with $\deg R < t$, then $(gR)'(\beta) = g'(\beta)R(\beta) = \varepsilon\lambda_\beta(gR)(\beta) = 0$ and $g'(\beta)\neq 0$ ($g$ square-free), so $R$ vanishes on $Z(g)$ and $R=0$. The dimensions add up to $M+1$ by Lemma 7.1. The formula for $R_P$ is the condition $(P+gR)'(\beta) = \varepsilon\lambda_\beta P(\beta)$. The last inclusion is clear. □

## 7.2 Derivatives of products of code polynomials

**Proposition 7.4** (Derivative of a product)**.** *Let $a, b \in \Gamma(L,g)$ and let $a+b$ denote their sum. Write $f_c = u_c^2 \pi_{\bar{c}}$ as in (2.4), and let $\pi_{\overline{a\cup b}}$ be the polynomial of the coordinates outside* $\operatorname{supp}(a) \cup \operatorname{supp}(b)$*. Then*

$$f_a f_b = \rho_{ab}^2\, \pi_{a+b}, \qquad (f_a f_b)' = g^2\, w_{ab}^2, \qquad \rho_{ab} := u_a u_b\, \pi_{\overline{a\cup b}}, \quad w_{ab} := \rho_{ab}\, u_{a+b} = u_a u_b u_{a+b}\, \pi_{\overline{a\cup b}}. \tag{7.1}$$

*Moreover* $\deg w_{ab} \le D - 1 - t$.

*Proof.* $\pi_{\bar{a}}\pi_{\bar{b}} = \pi^2_{\bar{a}\cap\bar{b}}\,\pi_{\bar{a}\triangle\bar{b}}$ and $\bar{a}\triangle\bar{b} = \operatorname{supp}(a+b)$, so $f_a f_b = u_a^2 u_b^2 \pi^2_{\overline{a\cup b}} \pi_{a+b}$. Since squares have zero derivative, $(f_a f_b)' = \rho_{ab}^2 \pi'_{a+b} = \rho_{ab}^2 g^2 u_{a+b}^2$ by (2.2) and (2.4) applied to the codeword $a+b$. For the degree, every coordinate of $\operatorname{supp}(a)\cup\operatorname{supp}(b)$ lies in exactly two of the three supports of $a, b, a+b$, so $\operatorname{wt}(a)+\operatorname{wt}(b)+\operatorname{wt}(a+b) = 2|a\cup b|$; with $\deg u_c \le (\operatorname{wt}(c)-1)/2 - t$ this gives $\deg w_{ab} \le |a\cup b| - \frac{3}{2} - 3t + n - |a\cup b|$, i.e. $\deg w_{ab} \le D - 1 - t$. □

**Proposition 7.5** (Derivatives at the support)**.** *For every $c \in \Gamma(L,g)$ and every $i$,*

$$v_i\, f_c'(\alpha_i) = \sum_{j:\ c_j \neq c_i} \frac{1}{\alpha_i - \alpha_j}. \tag{7.2}$$

*Equivalently, with the Cauchy matrix $T_{ij} = 1/(\alpha_i - \alpha_j)$ ($i \neq j$), $T_{ii} = 0$, and $\tau_i = \sum_{j\neq i} T_{ij}$: $\big(v_i f_c'(\alpha_i)\big)_i = Tc + \tau \star c$, and $\tau = 0$ when $L = \mathbb{F}_{2^m}$.*

*Proof.* If $c_i = 0$ then $\pi_{\bar{c}}(\alpha_i) = 0$, and (2.5) gives $\pi_c(\alpha_i)\pi'_{\bar{c}}(\alpha_i) = \pi'(\alpha_i)$; with $u_c(\alpha_i)^2 = \pi'_c(\alpha_i)/g(\alpha_i)^2$ we get $v_i f'_c(\alpha_i) = \pi'_c(\alpha_i)/\pi_c(\alpha_i) = \sum_{j\in c}(\alpha_i - \alpha_j)^{-1}$. If $c_i = 1$ then $u_c(\alpha_i)^2\pi_{\bar{c}}(\alpha_i) = f_c(\alpha_i) = 1/v_i$, so $v_i f'_c(\alpha_i) = \pi'_{\bar{c}}(\alpha_i)/\pi_{\bar{c}}(\alpha_i) = \sum_{j\notin c}(\alpha_i - \alpha_j)^{-1}$. For $L = \mathbb{F}_{2^m}$, $\tau_i = \sum_{\gamma\neq 0}\gamma^{-1} = \sum_{\gamma\neq 0}\gamma = 0$. □

### 7.3 A reduction of the $d = 2$ case

Since $2D < 2n$, Hermite evaluation of order two at $L$, $h \mapsto (h(\alpha_i), h'(\alpha_i))_i$, is injective on $K[X]_{\le 2D}$. Applying it to $f_a f_b$ and using Proposition 7.5:

**Corollary 7.6.** $\dim_K \mathcal{V}_2 = \dim_K \operatorname{span}\{(a\star b,\ b\star Ta + a\star Tb) : a, b \in \Gamma(L, g)\} \subseteq K^{2n}$, *where* $\star$ *is the coordinatewise product and the* $\tau$*-terms cancel.*

Write $\Omega := \operatorname{span}_K\{w_{ab}\} \subseteq K[X]_{\le D-1-t}$ and $\mathrm{Sq}(U) = \{s^2 : s \in U\}$ (the $K$-span of squares of elements of a $K$-space $U$ is $\mathrm{Sq}(U)$, since every element of $K$ is a square).

**Theorem 7.7** (Reduction)**.** *Assume* $D \ge 2t - 1$. *The following are equivalent:*

*(i)* $\mathcal{V}_2 = \mathcal{W}_2$ *(Conjecture 6.4 for* $d = 2$*);*

*(ii)* (A) $\Omega = K[X]_{\le D-1-t}$, *and* (B) $\mathrm{Sq}(K[X]_{\le D}) \subseteq \mathcal{V}_2$*;*

*(iii)* (A′) $\dim_K \operatorname{span}\{b\star Ta + a\star Tb\} = D - t$, *and* (B).

*Proof.* By Lemma 7.1 with $d = 2$, $\mathcal{W}_2 = \{h : h'(\beta) = 0\ \forall\beta\}$, and $\mathcal{W}_2 = \mathrm{Sq}(K[X]_{\le D}) \oplus X\,g^2\,\mathrm{Sq}(K[X]_{\le D-1-t})$ (even and odd parts; the odd part $h'$ must be a square divisible by $g^2$). The derivative maps $\mathcal{V}_2$ onto $\mathcal{V}'_2 := \operatorname{span}\{(f_a f_b)'\} = g^2\,\mathrm{Sq}(\Omega)$ by Proposition 7.4, with kernel $\mathcal{V}_2 \cap \mathrm{Sq}$, so $\dim \mathcal{V}_2 = \dim \Omega + \dim(\mathcal{V}_2 \cap \mathrm{Sq})$ with $\dim\Omega \le D - t$ and $\dim(\mathcal{V}_2 \cap \mathrm{Sq}) \le D + 1$. Hence $\dim \mathcal{V}_2 = 2D + 1 - t$ iff both bounds are attained, which is (ii). For (iii): $\mathcal{V}'_2$ consists of squares of degree $\le 2D - 1 < 2n$, so it is determined by its values on $L$, and $(f_a f_b)'(\alpha_i) = v_i^{-2}(b\star Ta + a\star Tb)_i$ by Proposition 7.5. □

**Experiment 7.8.** On all instances of Experiment 6.3 (full and non-full supports) the identity (7.1) held for every pair, $\dim\Omega = D - t$, every $X^{2j}$ $(j \le D)$ lay in $\mathcal{V}_2$, and $\dim \mathcal{V}_2 = 2D + 1 - t$; i.e. (A) and (B) both hold. The identity (7.2) was verified on every codeword and coordinate.

### 7.4 What the conjecture is really about

Lemma 7.1 shows that $\mathcal{W}_1$ depends only on $(\pi, Z(g))$. The following observation, verified on every instance tested (table below), locates the difficulty precisely.

**Experiment 7.9.** For the instances listed below: (i) $\mathcal{W}_1 \cdot \mathcal{W}_1 = \mathcal{W}_2$; (ii) a uniformly random $k$-dimensional $K$-subspace $V \subset \mathcal{W}_1$ satisfies $V \cdot V = \mathcal{W}_2$; (iii) $\mathcal{V}_1 \cdot \mathcal{W}_1 = \mathcal{W}_2$ and $\mathcal{V}_1 \cdot g^2 K[X]_{\le D-2t} = g^2 K[X]_{\le 2D-2t}$; (iv) a uniformly random $k$-dimensional subspace of $K[X]_{\le D}$ has product space all of $K[X]_{\le 2D}$.

| $m$ | $n$ | $t$ | $k$ | $D$ | $\dim \mathcal{W}_1$ | $\dim \mathcal{W}_1\mathcal{W}_1$ | random $V \subset \mathcal{W}_1$: $\dim VV$ | $\dim \mathcal{W}_2$ |
|---|---|---|---|---|---|---|---|---|
| 4 | 16 | 2 | 8 | 11 | 10 | 21 | 21 | 21 |
| 5 | 32 | 3 | 17 | 25 | 23 | 48 | 48 | 48 |
| 5 | 32 | 4 | 12 | 23 | 20 | 43 | 43 | 43 |

Thus Conjecture 6.4 at $d = 2$ says that the Goppa space $\mathcal{V}_1$ is in *general position inside* $\mathcal{W}_1$ *with respect to multiplication*: the Goppa structure enters only through $\mathcal{V}_1 \subseteq \mathcal{W}_1$, and $(m-3)t$ further conditions cut $\mathcal{V}_1$ out of $\mathcal{W}_1$ without reducing the product space. Two remarks delimit the possible proofs.

*Remark* 7.10 (Leading terms do not suffice). A natural attempt is to bound $\dim \mathcal{V}_2$ from below by the number of degrees attained by products. It fails: the degree set of $\mathcal{V}_1$ misses $(m-2)t$ degrees in structured low blocks (e.g. $\{0, \dots, 2t-1\} \cup \{4t-2, 4t-1, 4t\}$ for $(m, n, t) = (5, 32, 3)$), its sumset misses several degrees that $\mathcal{W}_2$ attains, and the low-degree elements of $\mathcal{V}_2$ (including the constant 1) arise only from cancellations. The same is true for a random $k$-dimensional $V \subset \mathcal{W}_1$, whose degree set is a top interval.

*Remark* 7.11 (The genuinely generic case is a maximal-rank statement). For a general $k$-dimensional $V \subset K[X]_{\le D}$, $\dim V \cdot V = \min\{\binom{k+1}{2}, 2D+1\}$ is the statement that a general rational curve of degree $D$ in $\mathbb{P}^{k-1}$ lies on the expected number of quadrics (maximal rank in degree two), known in characteristic zero and confirmed here over $\mathbb{F}_{2^m}$ numerically (Experiment 7.9(iv)). Conjecture 6.4 is the analogous maximal-rank statement for the *non-generic* curve $\Gamma$, relative to the ambient space $\mathcal{W}_2$ rather than $H^0(\mathbb{P}^1, \mathcal{O}(2D))$.

### 7.5 Results on the strategy: S1 proved, S2 reduced

We now carry out the first step of the strategy and reduce the second to a single explicit inclusion. Throughout, GRS denotes $\mathrm{GRS}_{D+1}(L, v)$, identified with $K[X]_{\le D}$ via $f \mapsto (v_i f(\alpha_i))_i$, $\sigma$ acts on $K^n$ coordinatewise, and $B_s = \mathrm{span}_K\{(\alpha_i^{2^s j} g(\alpha_i)^{-2^s})_i : j < t\}$ are the Frobenius blocks, so that $C^\perp \otimes K = \bigoplus_{s<m} B_s$ and, by Delsarte, $\mathcal{V}_1 = C \otimes K = \bigcap_{s<m} \sigma^s(\mathrm{GRS})$.

**Theorem 7.12** (S1: products of $\mathcal{W}_1$). *If $n \ge 6t$, then $\mathcal{W}_1 \cdot \mathcal{W}_1 = \mathcal{W}_2$.*

*Proof.* Write $A = K[X]_{\le D-2t}$ and $M = A \cdot \mathcal{W}_1 \subseteq K[X]_{\le 2D-2t}$. By Lemma 7.3, $\mathcal{W}_1 \supseteq g^2 A$, so $M \supseteq g^2 A \cdot A = g^2 K[X]_{\le 2D-4t}$; and $\mathcal{W}_1 \ni w_0 := 1 + gR_1$, which is a unit modulo $g^2$ since $w_0(1 + gR_1) = 1 + g^2 R_1^2$. Since $D - 2t \ge 2t - 1$, the image of $M$ in $K[X]/(g^2)$ contains $X^j \bar{w}_0$ for $0 \le j \le 2t-1$, hence $K[X]_{<2t} \cdot \bar{w}_0 = K[X]/(g^2)$. As $K[X]_{\le 2D-2t} = g^2 K[X]_{\le 2D-4t} \oplus K[X]_{<2t}$, this gives $M = K[X]_{\le 2D-2t}$, and therefore

$$\mathcal{W}_1 \cdot \mathcal{W}_1 \;\supseteq\; g^2 A \cdot \mathcal{W}_1 \;=\; g^2 M \;=\; g^2 K[X]_{\le 2D-2t}.$$

Now $\mathcal{W}_2 \cap g^2 K[X] = g^2 K[X]_{\le 2D-2t}$ (every $g^2 Q$ lies in $\mathcal{W}_2$), and by Lemma 7.3 with $\varepsilon = 0$ the image of $\mathcal{W}_2$ in $K[X]/(g^2)$ has dimension $\dim \mathcal{W}_2 - (2D-2t+1) = t$. The squares $(P + gR_P)^2 = P^2 + g^2 R_P^2$ of elements of $\mathcal{W}_1$ lie in $\mathcal{W}_1 \cdot \mathcal{W}_1$ and reduce modulo $g^2$ to all squares $a^2$, $\deg a < t$ (a square modulo $g^2$ depends only on $a \bmod g$), which form a $t$-dimensional $K$-subspace of $K[X]/(g^2)$ because $K$ is perfect. Hence $\mathcal{W}_1 \cdot \mathcal{W}_1$ contains $\mathcal{W}_2 \cap g^2 K[X]$ and maps onto the image of $\mathcal{W}_2$ modulo $g^2$, so $\mathcal{W}_1 \cdot \mathcal{W}_1 = \mathcal{W}_2$. □

**Proposition 7.13** (What $\mathcal{W}_1$ is). *(i) $\mathcal{W}_1 = \mathrm{GRS} \cap \sigma(\mathrm{GRS})$; equivalently, when $\mathrm{GRS}^\perp = B_0 \oplus B_1$, $\mathcal{W}_1 = \mathrm{GRS} \cap B_2^\perp$. Thus $\mathcal{W}_1$ is the two-fold Delsarte intersection and $\mathcal{V}_1 = \mathcal{W}_1 \cap \bigcap_{s \ge 3} B_s^\perp$: the $(m-3)t$ conditions cutting $\mathcal{V}_1$ out of $\mathcal{W}_1$ are exactly the Frobenius blocks $B_3, \dots, B_{m-1}$.*

*(ii)* (Ramification.) *For every $\beta \in Z(g)$ the vector $(f_1'(\beta), \dots, f_k'(\beta))$ equals $\lambda_\beta(f_1(\beta), \dots, f_k(\beta))$. Hence the differential of the hidden map $\Phi : \mathbb{P}^1 \to \mathbb{P}^{k-1}$ vanishes at the $t$ roots of $g$: $\Phi$ is ramified at $Z(g)$. Under (A1)–(A2) the image curve $\Gamma$ has the $t$ unibranch singular points $\Phi(\beta)$ (over $\bar{K}$) as its only singularities; vanishing of the differential alone does not determine their type, which is why (A2) is an assumption.*

*Proof.* (i) For $f \in K[X]_{\le D}$ and $\beta \in Z(g)$, Lagrange interpolation at $L$ gives $f'(\beta) - \lambda_\beta f(\beta) = -\pi(\beta)\sum_i f(\alpha_i)/(\pi'(\alpha_i)(\beta-\alpha_i)^2)$. Since $g(\beta)=0$, $1/(\beta-\alpha) = -q_\beta(\alpha)/g(\alpha)$ with $q_\beta = g/(X-\beta) \in \bar{K}[X]_{\le t-1}$, so with $c_i = v_i f(\alpha_i)$,

$$f'(\beta) - \lambda_\beta f(\beta) = -\pi(\beta)\sum_i c_i \frac{q_\beta(\alpha_i)^2}{g(\alpha_i)^4}.$$

As $\beta$ runs over $Z(g)$ the $q_\beta$ form a basis of $\bar{K}[X]_{\le t-1}$, and their squares span $\operatorname{span}\{X^{2j} : j < t\}$. So $f \in \mathcal{W}_1$ iff $c \perp \sigma\big(\{X^j/g^2\}_{j<t}\big)$, and since $\{X^j g/g^2\}_{j<t} \oplus \{X^j/g^2\}_{j<t} = \{q/g^2 : \deg q < 2t\} = \mathrm{GRS}^\perp$, this is $c \perp \sigma(\mathrm{GRS}^\perp)$, i.e. $c \in \sigma(\mathrm{GRS})$. The form with $B_2$ follows from $\sigma(B_0 \oplus B_1) = B_1 \oplus B_2$ and $B_1 \subseteq \mathrm{GRS}^\perp$. The last sentence is Delsarte's theorem. (ii) is Lemma 7.1 applied to each $f_j \in \mathcal{V}_1 \subseteq \mathcal{W}_1$. □

*Remark* 7.14 (The cuspidal model). Proposition 7.13(ii) explains $\mathcal{W}_d$ geometrically. Let $\tilde{\Gamma}$ be the curve obtained from $\mathbb{P}^1_{\bar{K}}$ by pinching the tangent direction at each point of $Z(g)$ (a rational curve with $t$ cusps, arithmetic genus $t$). A form of degree $d$ pulled back along $\Phi$ is a section of $\mathcal{O}_{\mathbb{P}^1}(dD)$ whose derivative at each $\beta$ is tied to its value exactly as in Lemma 7.1, i.e. a section of the degree-$dD$ line bundle on $\tilde{\Gamma}$; so $\mathcal{W}_d = H^0(\tilde{\Gamma}, \mathcal{O}(d))$, of dimension $dD + 1 - t$ by Riemann–Roch, $\mathcal{W}_1$ is the *complete* linear system, and Theorem 7.12 is normal generation of a line bundle of degree $D \ge 2t+1$ on a curve of arithmetic genus $t$. The Goppa subsystem $\mathcal{V}_1 \subsetneq \mathcal{W}_1$ (for $m > 3$) is cut out by the higher Frobenius blocks.

**Lemma 7.15.** $\mathcal{V}_1 \cap gK[X] = \mathcal{V}_1 \cap g^2K[X]$. *If* (H1) *the evaluation map* $\mathcal{V}_1 \to \bar{K}^{Z(g)}$, $f \mapsto (f(\beta))_\beta$, *is onto (equivalently* $\mathcal{V}_1 \to K[X]/(g)$ *is onto), then* $\dim(\mathcal{V}_1 \cap g^2K[X]) = k - t$.

*Proof.* If $f \in \mathcal{V}_1 \subseteq \mathcal{W}_1$ and $f(\beta) = 0$ then $f'(\beta) = \lambda_\beta f(\beta) = 0$, so $(X-\beta)^2 \mid f$ for all $\beta$ and $g^2 \mid f$. □

**Theorem 7.16** (Reduction II)**.** *Assume* $n \ge 6t$ *and* (H1)*. Then*

$$\mathcal{V}_2 = \mathcal{W}_2 \iff g^2K[X]_{\le 2D-2t} \subseteq \mathcal{V}_2 \impliedby (\mathcal{V}_1 \cap g^2K[X]) \cdot \mathcal{V}_1 = g^2K[X]_{\le 2D-2t}.$$

*In particular, under these hypotheses Conjecture 6.4 for* $d = 2$*, and with it statements* (A) *and* (B) *of Theorem 7.7, follow from the single inclusion* $(\mathcal{V}_1 \cap g^2K[X]) \cdot \mathcal{V}_1 \supseteq g^2K[X]_{\le 2D-2t}$.

*Proof.* $\mathcal{V}_2 \subseteq \mathcal{W}_2$ always, and $\mathcal{W}_2 \cap g^2K[X]_{\le 2D} = g^2K[X]_{\le 2D-2t}$. If $\mathcal{V}_2$ contains this space, then, since $\mathcal{V}_2 \supseteq \mathrm{Sq}(\mathcal{V}_1)$ and by (H1) the reductions $(f \bmod g)^2$ of these squares exhaust the squares modulo $g^2$, which is the whole image of $\mathcal{W}_2$ in $K[X]/(g^2)$ (proof of Theorem 7.12), we get $\mathcal{V}_2 = \mathcal{W}_2$. The other implications are clear. □

Write $\mathcal{V}_1 = E \oplus g^2U'$ with $U' = (\mathcal{V}_1 \cap g^2K[X])/g^2 \subseteq K[X]_{\le D-2t}$ and $E$ a lift of a basis of $\mathcal{V}_1 \bmod g^2$ ($\dim E = t$ under (H1)). The inclusion in Theorem 7.16 reads $U' \cdot \mathcal{V}_1 = K[X]_{\le 2D-2t}$, and the proof of Theorem 7.12 gives half of it:

**Proposition 7.17.** *Assume* (H2)*:* $U' \to K[X]/(g^2)$ *is onto (equivalently* $\dim(\mathcal{V}_1 \cap g^4K[X]) = k-3t$*). Then* $U' \cdot \mathcal{V}_1 + g^2K[X]_{\le 2D-4t} = K[X]_{\le 2D-2t}$*. Consequently* $U' \cdot \mathcal{V}_1 = K[X]_{\le 2D-2t}$ *if and only if* $U' \cdot \mathcal{V}_1 \supseteq g^2K[X]_{\le 2D-4t}$.

*Proof.* $E \bmod g^2$ contains a unit (an element congruent to 1 modulo $g$), so $(U' \bmod g^2)\cdot(E \bmod g^2) = K[X]/(g^2)$, and $K[X]_{\le 2D-2t} = g^2K[X]_{\le 2D-4t} \oplus K[X]_{<2t}$. □

**Experiment 7.18.** On all instances tested (those of Experiments 6.3 and 7.9, full and non-full supports, $m \le 7$), (H1) and (H2) held, and $U' \cdot E + g^2 U' \cdot U' = U' \cdot \mathcal{V}_1 = K[X]_{\le 2D-2t}$ held whenever $\binom{k+1}{2} > 2D + 1 - t$. The one instance with $\binom{k+1}{2} = 2D + 1 - t$ (namely $(m, n, t) = (4, 13, 2)$, $k = 5$) has $\dim \mathcal{V}_2 = 14 < 15 = \dim \mathcal{W}_2$: the strict inequality in Conjecture 6.4 cannot be relaxed to equality. Also, $U' \cdot U' \subsetneq K[X]_{\le 2D-4t}$ in some instances, so the term $U' \cdot E$ is needed, and the naive downward induction ($U^{(j)} \cdot \mathcal{V}_1 = K[X]_{\le 2D-2jt}$ for $U^{(j)} := (\mathcal{V}_1 \cap g^{2j} K[X])/g^{2j}$) fails at the levels $j$ where $U^{(j)} \bmod g^2$ is not onto; these are governed by the structure below.

**Conjecture 7.19** (refined)**.** If $n \ge 6t$ and $\binom{k+1}{2} > 2D + 1 - t$, then $\mathcal{V}_2 = \mathcal{W}_2$; more generally $\mathcal{V}_d = \mathcal{W}_d$ for $d \ge 2$ when $\binom{k+d-1}{d} > dD + 1 - t$.

**Experiment 7.20** (The $g$-adic filtration of $\mathcal{V}_1$)**.** Let $d_j = \dim(\mathcal{V}_1 \cap g^j K[X])$. On all instances tested the sequence $k, d_1, d_2, \ldots$ decreases by exactly $t$ at each step except that $d_j = d_{j+1}$ for $j = 2^s - 2$, $s = 2, \ldots, m-1$ (and for boundary effects when $d_j < t$). For instance for $(m, n, t) = (7, 128, 3)$, $k = 107$: $d_j = 104, 104, 101, 98, 95, 95, 92, \ldots, 74, 74, 71, \ldots, 29, 29, 29, 26, \ldots$, with repeats at $j = 1, 5, 13, 29$. Equivalently:

$$\dim(\mathcal{V}_1 \cap g^j K[X]) = k - jt + t \cdot \#\{s \in \{2, \ldots, m-1\} : 2^s - 2 \le j\} \qquad \text{(while positive)}.$$

**Theorem 7.21** (Local structure of $\mathcal{V}_1$ at the roots of $g$)**.** *For $\beta \in Z(g)$ put $u_\beta := g^2/\big((X-\beta)^2 \pi\big)$, a unit at $\beta$. For $f \in K[X]_{\le D}$ with $c = (v_i f(\alpha_i))_i$ and $2 \le s \le m-1$,*

$$c \perp B_s \iff D^{[2^s-3]}\big[u_\beta f\big](\beta) = 0 \quad \text{for all } \beta \in Z(g),$$

*where $D^{[j]}$ is the $j$-th Hasse derivative. Consequently*

$$\mathcal{V}_1 = \Big\{ f \in K[X]_{\le D} : \ D^{[2^s-3]}[u_\beta f](\beta) = 0 \ \ \forall \beta \in Z(g), \ 2 \le s \le m-1 \Big\},$$

*and for $f \in \mathcal{V}_1$ and $2 \le s \le m-1$: $g^{2^s-3} \mid f \Longrightarrow g^{2^s-2} \mid f$. In particular*

$$\dim\big(\mathcal{V}_1 \cap g^j K[X]\big) \ \ge \ k - jt + t \cdot \#\{s \in \{2, \ldots, m-1\} : 2^s - 2 \le j\} \qquad (j \ge 1),$$

*with equality exactly when the maps $\mathcal{V}_1 \to \bigoplus_\beta \bar{K}^j$, $f \mapsto (f^{[0]}(\beta), \ldots, f^{[j-1]}(\beta))_\beta$ have the largest image compatible with the constraints above (observed in every instance of Experiment 7.20).*

*Proof.* $B_s$ is spanned over $\bar{K}$ by the vectors $\big((q(\alpha_i)/g(\alpha_i))^{2^s}\big)_i$, $q \in \bar{K}[X]_{\le t-1}$, and the $q_\beta = g/(X-\beta)$ form a basis of $\bar{K}[X]_{\le t-1}$ with $q_\beta(\alpha)/g(\alpha) = 1/(\alpha - \beta)$. So $c \perp B_s$ iff $\sum_i c_i (\alpha_i - \beta)^{-2^s} = 0$ for all $\beta$. Now $c_i = P(\alpha_i)/\pi'(\alpha_i)$ with $P = g^2 f$ of degree $\le n-1$, and partial fractions give $P(Y)/\pi(Y) = \sum_i P(\alpha_i)/(\pi'(\alpha_i)(Y - \alpha_i))$; applying $D_Y^{[M-1]}$ and using $D^{[j]}(Y-\alpha)^{-1} = (Y-\alpha)^{-1-j}$ in characteristic two,

$$\sum_i \frac{P(\alpha_i)}{\pi'(\alpha_i)(Y-\alpha_i)^M} = D_Y^{[M-1]}\Big[\frac{g^2 f}{\pi}\Big](Y).$$

With $M = 2^s$ the condition is $D^{[2^s-1]}[g^2 f/\pi](\beta) = 0$. Write $g^2 f/\pi = (X-\beta)^2 \cdot u_\beta f$. By the Leibniz rule $D^{[N]}[(X-\beta)^2 F] = \sum_{a+b=N} D^{[a]}[(X-\beta)^2]\, D^{[b]} F$, and at $X = \beta$ only $a = 2$ survives ($D^{[0]}$ vanishes there, $D^{[1]}(X-\beta)^2 = 0$ in characteristic two, $D^{[2]}(X-\beta)^2 = 1$). Hence $D^{[2^s-1]}[g^2 f/\pi](\beta) = D^{[2^s-3]}[u_\beta f](\beta)$. Since $\mathrm{GRS}^\perp = B_0 \oplus B_1$ accounts for $f \in K[X]_{\le D}$, Delsarte's theorem gives the description of $\mathcal{V}_1$. If $g^{2^s-3} \mid f$ then $f^{[j]}(\beta) = 0$ for $j < 2^s - 3$, so $0 = D^{[2^s-3]}[u_\beta f](\beta) = u_\beta(\beta) f^{[2^s-3]}(\beta)$ and $(X-\beta)^{2^s-2} \mid f$ for every $\beta$. Each constraint determines one Hasse coefficient per root in terms of the lower ones, which gives the dimension inequality. $\square$

*Remark* 7.22. For $s = 2$ the condition reads $u_\beta(\beta)f'(\beta) + u'_\beta(\beta)f(\beta) = 0$; since $(g^2)' = ((X-\beta)^2)' = 0$, $u'_\beta/u_\beta = \pi'/\pi$ at $\beta$, and we recover the cusp condition $f'(\beta) = \lambda_\beta f(\beta)$ of Lemma 7.1 and Proposition 7.13. The theorem is thus the complete local form, at the roots of $g$, of the Frobenius blocks $B_2, \dots, B_{m-1}$: block $B_s$ pins the Hasse coefficient of order $2^s - 3$. For $f \in \mathcal{V}_1$ nothing else is pinned, and the products $u_\beta f \cdot u_\beta \tilde{f}$ satisfy only the order-one constraint (their order-1 coefficient is $a_0\tilde{a}_1 + a_1\tilde{a}_0 = 0$), consistent with $\mathcal{V}_2 \subseteq \mathcal{W}_2$ and with no further local constraint on $\mathcal{V}_2$.

*Remark* 7.23 (Status of S3 and S4). *S3.* Under (H1) and $n \geq 6t$, statement (B) (and (A)) follows from the inclusion of Theorem 7.16; no separate argument for the squares is needed. *S4.* Coordinatewise Frobenius does not preserve the instance: for $L = \mathbb{F}_{2^m}$ it maps the pair ($\Gamma(L, g)$, annihilator of $\mathcal{V}_2$) to the same pair for the conjugate code $\Gamma(L, g^{(2)})$, since $v_{\tau(i)} = g(\alpha_i^2)^2 = g^{(1/2)}(\alpha_i)^4 \neq v_i^2$ unless $g \in \mathbb{F}_2[X]$. So an $\mathbb{F}_2$-descent of the annihilator is available only for Goppa polynomials with coefficients in $\mathbb{F}_2$ (which Classic McEliece does not use); in general the Frobenius only shows that the problem for $g$ and for $g^{(2)}$ are equivalent.

## 7.6 A proof of the conjecture in the local regime

For $2 \leq r \leq m-1$ let

$$\mathcal{V}^{[r]} := \{ f \in K[X]_{\leq D} : \ D^{[2^s-3]}[u_\beta f](\beta) = 0 \ \ \forall \beta \in Z(g), \ 2 \leq s \leq r \},$$

so that $\mathcal{V}^{[2]} = \mathcal{W}_1$ and $\mathcal{V}^{[m-1]} = \mathcal{V}_1$ (Theorem 7.21). Put $M_r := 2^r - 2$, one more than the largest pinned order.

**Theorem 7.24** (Products in the local regime)**.** *If $D \geq M_r t + 1$, then $\mathcal{V}^{[r]} \cdot \mathcal{V}^{[r]} = \mathcal{W}_2$. In particular Conjecture 6.4 at $d = 2$ holds whenever $D \geq (2^{m-1} - 2)t + 1$.*

*Proof.* Write $\mathcal{V} = \mathcal{V}^{[r]}$, $M = M_r$, and let $\Lambda_{\beta,s} : F \mapsto D^{[2^s-3]}[u_\beta F](\beta)$ on $K[X]_{\leq D}$, so that $\mathcal{V} = \bigcap \ker \Lambda_{\beta,s}$ and $\mathcal{V}^\perp = \operatorname{span}\{\Lambda_{\beta,s}\}$ in $(K[X]_{\leq D})^*$. Each $\Lambda_{\beta,s}$ vanishes on multiples of $(X-\beta)^{2^s-2}$, hence on $g^M K[X]_{\leq D-Mt}$. Since $D \geq Mt - 1$, reduction $K[X]_{\leq D} \to K[X]/(g^M) \cong \bigoplus_\beta \bar{K}[z]/(z^M)$ ($z = X - \beta$) is onto, and $\mathcal{V}$ is the full preimage of $\prod_\beta \mathcal{V}_\beta$ with $\mathcal{V}_\beta = u_\beta^{-1}\mathcal{U}$, $\mathcal{U} = \{F \in \bar{K}[z]/(z^M) : F_a = 0 \text{ for } a \in S\}$, $S = \{2^s - 3 : 2 \leq s \leq r\}$.

*Shift-span.* $\mathcal{V} \supseteq g^M K[X]_{\leq D-Mt}$, and $\mathcal{U} + z\mathcal{U} = \bar{K}[z]/(z^M)$: coefficients of orders $a \notin S$ come from $\mathcal{U}$, and those of orders $a \in S$ from $z\mathcal{U}$ (the coefficient of $z^a$ in $zF$ is $F_{a-1}$, and $a - 1 \notin S$ because no two elements of $S$ differ by one). Hence $\mathcal{V} + X\mathcal{V}$ maps onto $K[X]/(g^M)$ and contains $g^M K[X]_{\leq D+1-Mt}$, so $\mathcal{V} + X\mathcal{V} = K[X]_{\leq D+1}$, and inductively $K[X]_{\leq J} \cdot \mathcal{V} = K[X]_{\leq D+J}$ for all $J \geq 1$.

Now let $\psi \in (K[X]_{\leq 2D})^*$ vanish on $\mathcal{V} \cdot \mathcal{V}$. For fixed $f \in \mathcal{V}$ the functional $F \mapsto \psi(fF)$ on $K[X]_{\leq D}$ vanishes on $\mathcal{V}$, so it lies in $\mathcal{V}^\perp$ and kills $g^M K[X]_{\leq D-Mt}$; thus $\psi$ kills $g^M \cdot \mathcal{V} \cdot K[X]_{\leq D-Mt} = g^M K[X]_{\leq 2D-Mt}$ by shift-span with $J = D - Mt \geq 1$. As $K[X]_{\leq 2D} = g^M K[X]_{\leq 2D-Mt} \oplus K[X]_{<Mt}$, $\psi$ factors through $K[X]/(g^M)$: $\psi(h) = \sum_\beta \sum_{j<M} \mu_{\beta,j} h^{[j]}(\beta)$. The image of $\mathcal{V} \cdot \mathcal{V}$ in $K[X]/(g^M)$ is $\bigoplus_\beta \operatorname{span}(\mathcal{V}_\beta \mathcal{V}_\beta)$, and $\mathcal{V}_\beta \mathcal{V}_\beta = u_\beta^{-2}\mathcal{U}\mathcal{U}$. Finally $\mathcal{U} \cdot \mathcal{U} = \{G : G_1 = 0\}$: the inclusion $\subseteq$ holds since $F_1 = \tilde{F}_1 = 0$; conversely $z^a \in \mathcal{U}$ for every $a \notin S$, and every $N \neq 1$ is a sum $a + b$ with $a, b \notin S$ ($N \notin S$: $a = 0$; $N \in S$, $N \geq 5$: $a = 2$, $b = N - 2 \equiv 3 \pmod 4 \notin S$). Since $u_\beta^2$ is a square unit, $(u_\beta^2 H)_1 = u_\beta(\beta)^2 H_1$, so $\mathcal{V}_\beta \mathcal{V}_\beta = \{H : H_1 = 0\}$ and $\psi$ kills it iff $\mu_{\beta,j} = 0$ for $j \neq 1$. Hence $\psi \in \operatorname{span}\{h \mapsto h'(\beta)\} = \mathcal{W}_2^\perp$, and $\mathcal{V} \cdot \mathcal{V} = \mathcal{W}_2$. □

*Remark* 7.25 (Scope, and what the proof shows). The hypothesis $D \geq (2^{m-1} - 2)t + 1$ reads $n \geq 2^{m-1}t + 2$, which at full support forces $t = 1$; for $t = 2$ and $n = 2^m$ it fails by two. So Theorem 7.24 proves the conjecture for $t = 1$ (checked numerically on six instances with $m \leq 6$), and for the intermediate spaces $\mathcal{V}^{[r]}$ it proves $\mathcal{V}^{[r]}\mathcal{V}^{[r]} = \mathcal{W}_2$ whenever $(2^r - 2)t < D$: at `mceliece348864`

($t = 64$, $D = 3359$) this covers $r \le 5$, i.e. the first four Frobenius blocks, while $\mathcal{V}_1$ has ten. The proof makes the remaining difficulty precise: a block $B_s$ with $(2^s - 2)t > D + 1$ pins a Hasse coefficient of order $2^s - 3$ at each root of $g$, which exceeds the degree budget per root, so the block condition is *not* local at $Z(g)$ within $K[X]_{\le D}$, and the $g$-adic/local induction cannot reach it. Closing (7.3) in the cryptographic regime therefore requires a tool that sees the deep blocks globally, not at the roots of $g$: the natural candidates are the Frobenius-conjugate representations, in which block $B_s$ is the degree bound and the other blocks are again local, or the displacement structure of the annihilator (Section 7.5, S2).

The local-regime theorem is unconditional where it applies: it proves Conjecture 6.4 at $d = 2$ for every binary Goppa code with $t = 1$, and for every code whose Frobenius blocks beyond the local range are absent; in the general case it is the base step of a $g$-adic induction whose remaining inductive step is the inclusion (7.3) below.

### 7.7 What remains

Outside the local regime of Theorem 7.24, under $n \ge 6t$, (H1) and (H2), the whole of Conjecture 6.4 for $d = 2$ has been reduced (Theorem 7.16, Proposition 7.17) to one inclusion:

$$(\mathcal{V}_1 \cap g^2 K[X]) \cdot \mathcal{V}_1 \;\supseteq\; g^4 K[X]_{\le 2D-4t}. \tag{7.3}$$

The analogous statement with $\mathcal{W}_1$ in place of $\mathcal{V}_1$ is Theorem 7.12; the difference is that $\mathcal{W}_1 \cap g^2 K[X] = g^2 K[X]_{\le D-2t}$ is a full polynomial space, whereas $\mathcal{V}_1 \cap g^2 K[X]$ is cut out by the blocks $B_3, \dots, B_{m-1}$ and has the $g$-adic structure of Experiment 7.20. Two routes suggest themselves. First, Theorem 7.21 gives the exact local structure of $\mathcal{V}_1$ at $Z(g)$; (7.3) can be attacked level by level in the $g$-adic filtration, where the obstruction at the paired levels $j = 2^s - 2$ is a single determined Hasse coefficient per root, and a base-point-free-pencil argument on the small spaces $U^{(j)}$ handles the top levels. Second, by Proposition 7.13(i) the inclusion is a statement about $\mathrm{GRS} \cap_\sigma \mathrm{GRS} \cap \cdots \cap \sigma^{m-1}\,\mathrm{GRS}$, and the trace-code technique by which Mora and Tillich [10] control products of $\sum_s \sigma^s(\cdot)$ may have a dual version for products of $\bigcap_s \sigma^s(\cdot)$. Either would complete the proof of Conjecture 7.19 at $d = 2$ and make the "$t$ zeros to spare" of [19] a theorem in the exact form that the attacks use.

A third piece of evidence says what kind of statement the proof must be. For an exponent $e \ge 3$ let $B^{(e)}$ be the "block" $\{(v_i \alpha_i^{ej} g(\alpha_i)^{-e})_i : \; j < t\}$, so that $\mathcal{V}_1 = K[X]_{\le D} \cap \bigcap_{s=2}^{m-1} (B^{(2^s)})^\perp$ and $\mathcal{W}_1 = K[X]_{\le D} \cap (B^{(4)})^\perp$ (the blocks $e = 1, 2$ impose nothing on $K[X]_{\le D}$). For a general exponent set $E \ni 4$ of the same size put $U_E = K[X]_{\le D} \cap \bigcap_{e \in E} (B^{(e)})^\perp$.

**Experiment 7.26** (Which conditions at $Z(g)$ are compatible with products)**.** At $(6, 64, 6)$ ($\dim \mathcal{W}_2 = 97$, $\dim U_E = 28$) the products $U_E \cdot U_E$ span $\mathcal{W}_2$ for the Frobenius set $E = \{4, 8, 16, 32\}$ and for every set $\{4, e, 16, 32\}$ with $e \in \{5, 7, 9, \dots, 20, 22, 23, 24, 33, \dots, 36, 48, 56\}$, for $\{4, 8, 12, 16\}$, $\{4, 30, 31, 32\}$, $\{4, 9, 19, 39\}$ and random sets; they fail exactly for $e = 3$ (rank 79) and $e = 6$ (rank 91), and for sets containing them ($\{4, 5, 6, 7\}$: 91; $\{3, 4, 5, 7\}$: 66; $\{4, 6, 10, 14\}$: 91). The same at $(5, 32, 3)$. The reason is visible for $e = 3$: by the residue theorem the block $B^{(3)}$ is the condition $g \mid f$ (its functionals are $\sum_\beta w_\beta f(\beta) \beta^{3j}$, a Vandermonde system at the roots), so every product is divisible by $g^2$ and cannot fill $\mathcal{W}_2$; $e = 6$ is its Frobenius relative and imposes a divisibility on the odd part. The Frobenius blocks $e = 2^s$, $s \ge 2$, impose single Hasse conditions of order $2^s - 3$ at the roots (Theorem 7.21) and no divisibility.

So Conjecture 6.4 at $d = 2$ is neither a general-position statement (the blocks are far from generic) nor a statement about the Frobenius exponents in particular (almost every exponent works); what has to be proved is that the conditions $D^{[2^s-3]}[u_\beta f](\beta) = 0$ are *product-free*: they force no

common divisibility or common vanishing on the products $ff'$, $f, f' \in \mathcal{V}_1$, beyond $h'(\beta) = 0$. A proof of (7.3) along these lines would show directly that $(\mathcal{V}_1 \cap g^2 K[X]) \cdot \mathcal{V}_1$ has no fixed component at the roots of $g$.

# 8 Frobenius structure: the conjugate curves

*Notation for this and the following sections.* $S_d$ is the space of all $\mathbb{F}_2$-forms of degree $d$ in $Y_1, \dots, Y_k$ and $S_{d,K} = S_d \otimes K$ (Section 3); $W_d \subseteq S_d$ is a chosen $\mathbb{F}_2$-subspace of forms, the attacks using the multilinear one $W_d^{\mathrm{ml}}$ (Section 3.5); $\mathcal{V}_d \subseteq \mathcal{W}_d \subseteq K[X]_{\le dD}$ are the pulled-back polynomial spaces of Section 6; and $\mathcal{J}_d = I(X_\Gamma)_d \subseteq S_{d,K}$ is the degree-$d$ part of the ideal of the union of conjugate curves, defined below.

## 8.1 Compatibility

**Proposition 8.1** (Frobenius compatibility)**.** *For every support point,* $[f_1^{(2)}(\alpha_i^2) : \cdots : f_k^{(2)}(\alpha_i^2)] = [f_1(\alpha_i)^2 : \cdots : f_k(\alpha_i)^2] = P_i$.

*Proof.* $f^{(2)}(\alpha^2) = f(\alpha)^2$, and an $\mathbb{F}_2$-rational projective point is fixed by coordinatewise squaring. □

Let $\Phi^{\sigma^s} = [f_1^{(2^s)} : \cdots : f_k^{(2^s)}]$ and $\Gamma^{\sigma^s} = \overline{\Phi^{\sigma^s}(\mathbb{P}^1)}$, $0 \le s < m$. Proposition 8.1 says that the conjugate curve $\Gamma^{\sigma^s}$ passes through $P_i$ at the parameter $\alpha_i^{2^s}$. Thus all $m$ conjugate curves pass through all $n$ public points (and through $\Phi(\infty)$), at different parameters. This is the geometric content of what Weis calls the *Frobenius ambiguity*, the choice among the $m$ conjugate curves through the binary points [19, 26], and of the $m$ Frobenius-conjugate GRS blocks of the HOV literature [15, 16].

*Caution* 8.2. Proposition 8.1 does *not* say that $\alpha_i^2 \in L$ implies any relation between $P_i$ and the point $P_j = \Phi(\alpha_i^2)$. If $\alpha_j = \alpha_i^2$ then $\Phi(\alpha_j) = P_j$ and $\Phi^\sigma(\alpha_j) = P_i$: two different curves through two different points at the same parameter. In particular $h_H(\alpha_i) = 0$ does not imply $h_H(\alpha_i^2) = 0$ for an $\mathbb{F}_2$-form $H$ (the identity $f^{(2)}(\alpha^2) = f(\alpha)^2$ concerns $f^{(2)}$, not $f$), and the Frobenius orbit structure of the support induces no rank defect by itself. Experiment 8.6 exhibits counterexamples on every instance tested.

## 8.2 The descent theorem

**Lemma 8.3.** *For $H \in S_d$ (coefficients in $\mathbb{F}_2$) and $f \in K[X]^k$, $H(f^{(2)}) = \big(H(f)\big)^{(2)}$. Hence for $H \in S_d$,*

$$h_H = 0 \iff H \in I(\Gamma)_d \iff H \in I(\Gamma^{\sigma^s})_d \text{ for all } s.$$

*Proof.* The map $f \mapsto f^{(2)}$ is a ring automorphism of $K[X]$ fixing $\mathbb{F}_2[X]$; apply it to $H(f) = \sum_\mu c_\mu f^\mu$ with $c_\mu \in \mathbb{F}_2$. □

Let $X_\Gamma = \bigcup_{s=0}^{m-1} \Gamma^{\sigma^s}$, a curve of degree at most $m\delta$ defined over $\mathbb{F}_2$, and let $\mathcal{J}_d = \bigcap_s I(\Gamma^{\sigma^s})_d = I(X_\Gamma)_d \subseteq S_{d,K}$.

**Theorem 8.4** (Descent)**.** *(i) $S_d \cap I(\Gamma)_d = S_d \cap \mathcal{J}_d$, and $\mathcal{J}_d$ is stable under the Galois action of* $\mathrm{Gal}(K/\mathbb{F}_2)$ *on coefficients.*

*(ii)* $\dim_{\mathbb{F}_2}\big(S_d \cap I(\Gamma)_d\big) = \dim_K \mathcal{J}_d$, *and* $\mathcal{J}_d = \big(S_d \cap I(\Gamma)_d\big) \otimes_{\mathbb{F}_2} K$.

*(iii) The* Frobenius defect

$$\Delta_{\mathrm{Frob}}(d) := \dim_K I(\Gamma)_d - \dim_{\mathbb{F}_2}\big(S_d \cap I(\Gamma)_d\big) = \dim_K I(\Gamma)_d - \dim_K \mathcal{J}_d$$

*satisfies* $0 \le \Delta_{\mathrm{Frob}}(d) \le (m-1)\dim_K \mathcal{V}_d \le (m-1)(dD+1-t)$.

*Proof.* (i) is Lemma 8.3; the Galois action permutes the $I(\Gamma^{\sigma^s})_d$. (ii) A Galois-stable $K$-subspace $\mathcal{J}$ of $S_{d,K}$ carries a semilinear action of the cyclic group $\mathrm{Gal}(K/\mathbb{F}_2)$, and by Galois descent for vector spaces (Speiser's lemma, equivalently $H^1(\mathrm{Gal}(K/\mathbb{F}_2), \mathrm{GL}_N(K)) = 1$) the natural map $\mathcal{J}^{\mathrm{Gal}} \otimes_{\mathbb{F}_2} K \to \mathcal{J}$ is an isomorphism; here $\mathcal{J}_d^{\mathrm{Gal}} = S_d \cap \mathcal{J}_d = S_d \cap I(\Gamma)_d$. (iii) $I(\Gamma)_d/\mathcal{J}_d$ injects into $\bigoplus_{s=1}^{m-1} S_{d,K}/I(\Gamma^{\sigma^s})_d$, and $S_{d,K}/I(\Gamma^{\sigma^s})_d \cong \mathcal{V}_d^{\sigma^s}$ via $H \mapsto H(f^{(2^s)})$, because $I(\Gamma^{\sigma^s})_d$ is by definition the kernel of that map and $\mathcal{V}_d^{\sigma^s}$ its image; in particular $\dim_K\big(S_{d,K}/I(\Gamma)_d\big) = \dim_K \mathcal{V}_d$, and each summand has dimension $\dim \mathcal{V}_d$ since $\sigma$ is a bijection. □

**Corollary 8.5** (Rank lower bound in the completeness regime)**.** *If* $(s_i)_{i\in I}$ *satisfies (5.3), then* $\mathcal{K}_{d,s}(G;I) \subseteq S_d \cap I(X_\Gamma)_d$*, hence*

$$\operatorname{rank} E_{d,s,I} \;\ge\; \binom{k+d-1}{d} - \dim_K \mathcal{J}_d \;=\; \dim_K\big(S_{d,K}/I(X_\Gamma)_d\big).$$

In the completeness regime the relation space is thus the space of $\mathbb{F}_2$-forms vanishing on the union of the $m$ conjugate curves and vanishing to the prescribed orders at the public points; the latter are conditions in the normal directions to $X_\Gamma$ at its rational points. Determining $\dim \mathcal{J}_d$ and the normal conditions is the exact-rank problem (Open Problem 14.3).

**Experiment 8.6** (Descent and the Frobenius defect)**.** On the instances of Experiment 6.3 we computed $\dim_K I(\Gamma)_d$, $\dim_K \mathcal{J}_d$ (kernel of the $K$-linear map $H \mapsto (h_H, h_H^{(2)}, \dots, h_H^{(2^{m-1})})$) and $\dim_{\mathbb{F}_2}(S_d \cap I(\Gamma)_d)$ (kernel of the $\mathbb{F}_2$-linear map $H \mapsto h_H$, $K$ expanded in an $\mathbb{F}_2$-basis). The identity of Theorem 8.4(ii) held in every case. We also tested the claim refuted in Caution 8.2: for each $\alpha \in L$ with $\alpha^2 \in L$, we took $\mathbb{F}_2$-forms $H$ with $h_H(\alpha) = 0$ and evaluated $h_H(\alpha^2)$; it was nonzero in 18–36 of the tests on each instance.

| $m$ | $n$ | $t$ | $k$ | $d$ | $\dim_K I(\Gamma)_d$ | $\dim_K \mathcal{J}_d$ | $\dim_{\mathbb{F}_2}(S_d \cap I(\Gamma)_d)$ | $\Delta_{\mathrm{Frob}}$ | $(m-1)(dD+1-t)$ |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 16 | 2 | 8 | 2 | 15 | 3 | 3 | 12 | 63 |
| 4 | 16 | 2 | 8 | 3 | 88 | 43 | 43 | 45 | 96 |
| 4 | 16 | 2 | 8 | 4 | 287 | 209 | 209 | 78 | 129 |
| 4 | 14 | 2 | 6 | 2 | 4 | 0 | 0 | 4 | 51 |
| 4 | 14 | 2 | 6 | 3 | 30 | 2 | 2 | 28 | 78 |
| 5 | 32 | 3 | 17 | 2 | 105 | 45 | 45 | 60 | 192 |
| 5 | 32 | 3 | 17 | 3 | 896 | 736 | 736 | 160 | 292 |
| 5 | 28 | 3 | 13 | 2 | 51 | 7 | 7 | 44 | 160 |
| 5 | 28 | 3 | 13 | 3 | 394 | 266 | 266 | 128 | 244 |
| 5 | 32 | 4 | 12 | 2 | 35 | 0 | 0 | 35 | 172 |
| 5 | 32 | 4 | 12 | 3 | 298 | 166 | 166 | 132 | 264 |
| 6 | 64 | 3 | 46 | 2 | 969 | 734 | 734 | 235 | 560 |

*Remark* 8.7 (What descent does and does not give). Theorem 8.4 determines the space of $\mathbb{F}_2$-forms vanishing on the hidden curve. It does not recover the curve: the public data determine $X_\Gamma$ (over $\mathbb{F}_2$), and key recovery amounts to selecting one of its $m$ components and parameterizing it, which is the "branch selection" step in [19] and the Frobenius-block selection in [15, 16]. One might hope to classify the hidden map by a cocycle in $H^1(\mathrm{Gal}(K/\mathbb{F}_2), \mathrm{PGL}_2(K))$; but $\Phi^\sigma$ is not of the form $\Phi \circ \lambda$ for a Möbius transformation $\lambda$ (the Frobenius on $\mathbb{P}^1$ is not a Möbius transformation, and the conjugate curves are different curves), and in any case $H^1(\mathrm{Gal}(K/\mathbb{F}_2), \mathrm{PGL}_2(K))$ is trivial for finite fields. The descent datum is the $\mathbb{F}_2$-curve $X_\Gamma$ with its $m$ conjugate components, as above.

### 8.3 The Hilbert function of the union of conjugate curves

The data of Experiment 8.6 fit a single formula, which is Riemann–Roch on $X_\Gamma$.

**Theorem 8.8** (Rank of the conjugate evaluation map)**.** *Assume (A3), let $d \geq 2$ and $n \geq 6t+2$ (so that $dD - 2t \geq n$). Let $\rho_d := \operatorname{rank}_K\big(S_{d,K} \to \bigoplus_{s<m} K[X]_{\leq dD},\ H \mapsto (H(f^{(2^s)}))_s\big)$, so that $\dim_K \mathcal{J}_d = \binom{k+d-1}{d} - \rho_d$. Then*

$$\dim_K \mathcal{V}_d \ \leq\ \rho_d \ \leq\ m\,(dD+1-t) - (m-1)(n+1),$$

*and therefore*

$$\dim_K \mathcal{J}_d \ \geq\ \binom{k+d-1}{d} - m(dD+1-t) + (m-1)(n+1),$$

$$\Delta_{\mathrm{Frob}}(d) \ \leq\ m(dD+1-t) - (m-1)(n+1) - \dim_K \mathcal{V}_d.$$

*If $\mathcal{V}_d = \mathcal{W}_d$ (Conjecture 7.19), the last bound reads $\Delta_{\mathrm{Frob}}(d) \leq (m-1)(dD-t-n)$.*

*Proof.* The $s$-th component of the map lands in $\mathcal{V}_d^{(s)} \subseteq \mathcal{W}_d^{(s)}$, the spaces of Theorem 6.1 for the conjugate code $\Gamma(L^{(2^s)}, g^{(2^s)})$, each of dimension $dD+1-t$. For every $H$, $i$ and $s$, $v_i^{2^s d}\, H(f^{(2^s)})(\alpha_i^{2^s}) = H(y_i) = v_i^d H(f)(\alpha_i)$ (Proposition 8.1), and the coefficient of $X^{dD}$ in $H(f^{(2^s)})$ is $H(\sum_i y_i)$ for every $s$ (Proposition 2.2 applied to the conjugate code). So the image lies in the subspace $Z \subseteq \bigoplus_s \mathcal{W}_d^{(s)}$ cut out by these $(m-1)(n+1)$ linear conditions. They are independent: for $s \geq 1$, $\mathcal{W}_d^{(s)} \supseteq (g^{(2^s)})^2 K[X]_{\leq dD-2t}$, and the map $Q \mapsto \big((Q(\alpha_i^{2^s}))_i, [X^{dD-2t}]Q\big)$ on $K[X]_{\leq dD-2t}$ is onto $K^{n+1}$ once $dD - 2t \geq n$ (the $n$ points $\alpha_i^{2^s}$ are distinct), and $(g^{(2^s)})^2$ is monic and nonzero on them. Hence $\dim Z = m(dD+1-t) - (m-1)(n+1)$. The lower bound $\rho_d \geq \dim \mathcal{V}_d$ is the $s=0$ component. (Without (A3) every component has zero leading coefficient, which is $m$ conditions instead of $m-1$ and lowers the bound by one.) □

**Conjecture 8.9** (Hilbert function of $X_\Gamma$)**.** For $d \geq 2$, $n \geq 6t+2$ and $\binom{k+d-1}{d} \geq m(dD+1-t) - (m-1)(n+1)$,

$$\rho_d = m\,(dD+1-t) - (m-1)(n+1), \qquad \Delta_{\mathrm{Frob}}(d) = (m-1)(dD-t-n),$$

$$\dim_K \mathcal{J}_d = \binom{k+d-1}{d} - m(dD+1-t) + (m-1)(n+1).$$

**Experiment 8.10.** Conjecture 8.9 holds in every instance of Experiment 8.6 with full support ($\rho_2 = 33, 108, 347$ and $\rho_3 = 77, 233$, $\rho_4 = 121$ for the respective rows) and in the non-full-support instances with $k \geq 8$; e.g. $(5,28,3)$: $\rho_2 = 84 = 5\cdot 40 - 4\cdot 29$, $\rho_3 = 189$. When $\binom{k+d-1}{d}$ is smaller than the right-hand side, $\rho_d = \binom{k+d-1}{d}$ and $\mathcal{J}_d = 0$, as for $(5,32,4)$ at $d=2$. The only deviation is the $k=6$ instance $(4,14,2)$ at $d=3$, where $\rho_3 = 54$ against the bound $\min\{56,59\}$.

*Remark* 8.11 (Riemann–Roch reading). $X_\Gamma$ (Figure 1) is the union of the $m$ conjugate cuspidal rational curves $\tilde{\Gamma}^{\sigma^s}$ (each of arithmetic genus $t$), glued at the $n+1$ common rational points $P_1, \dots, P_n, \Phi(\infty)$. If the $m$ branches through each common point have independent tangents, the gluing is seminormal, $\delta = m-1$ per point, and $p_a(X_\Gamma) = mt + (m-1)(n+1) - (m-1) = mt + (m-1)n$. A line bundle of degree $mdD$ on $X_\Gamma$ then has $h^0 = mdD + 1 - p_a = m(dD+1-t) - (m-1)(n+1)$, which is the conjectured $\rho_d$: Conjecture 8.9 says that forms of degree $d$ restrict onto all of $H^0(X_\Gamma, \mathcal{O}(d))$. The branches at $P_i$ have tangent directions $\Phi^{\sigma^s\prime}(\alpha_i^{2^s})$, $s < m$; independence of these $m$ vectors and of $y_i$ is also what the count of normal conditions below requires.

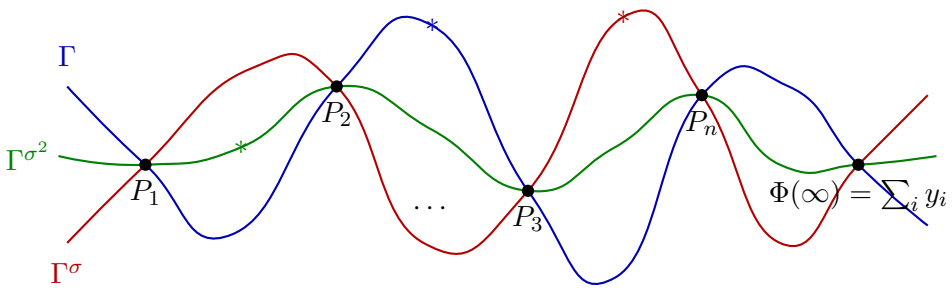



Figure 1: The curve $X_\Gamma$ for $m = 3$: the $m$ Frobenius-conjugate images of $\mathbb{P}^1$ all pass through the $n$ public points $P_i$ (at the parameters $\alpha_i^{2^s}$) and through $\Phi(\infty)$, with distinct tangents; each component is ramified at the $t$ roots of its Goppa polynomial ($*$, one shown per curve). An $\mathbb{F}_2$-form vanishes on one component iff it vanishes on all (Lemma 8.3); Riemann–Roch on this seminormal union gives $h^0(\mathcal{O}(d)) = m(dD + 1 - t) - (m - 1)(n + 1)$.

*Remark* 8.12 (What a proof of Conjecture 8.9 must supply). Riemann–Roch on $X_\Gamma$ computes $h^0(X_\Gamma, \mathcal{O}(d))$; it does not show that the restriction map $S_{d,K} \to H^0(X_\Gamma, \mathcal{O}(d))$ is onto, and that surjectivity is the entire content of the conjecture. Its $s = 0$ component is the statement $\mathcal{V}_d = \mathcal{W}_d$ of Conjecture 7.19: the image of $S_{d,K}$ in the $s = 0$ factor is $\mathcal{V}_d$, while $H^0$ of the $s = 0$ branch is $\mathcal{W}_d$. So "projective normality" or "normal generation" of $X_\Gamma$ in degree $d$ is not an input that can be quoted from the theory of curves of this type; it is Conjecture 7.19 together with the independence of the $(m - 1)(n + 1)$ gluing conditions proved in Theorem 8.8. In particular, the local methods of Theorem 7.24 cannot close it at cryptographic $t$ (Section 7.7), and Conjecture 8.9 should be regarded as exactly as hard as Conjecture 7.19, no harder and no easier. Riemann–Roch on $X_\Gamma$ moreover controls only the order-0 restriction; the dimension of the attack kernel (Section 11.6) depends on which first- and second-order conormal data of $\Gamma$ are realized by forms of $\mathbb{P}^{k-1}$, a question about the projection $\mathbb{P}^D \dashrightarrow \mathbb{P}^{k-1}$ that no statement about $H^0(X_\Gamma, \mathcal{O}(d))$ answers.

### 8.4 Exact rank of the public jet matrix for quadrics

For $d = 2$ the completeness condition $\sum_i s_i > 2D$ holds with $s_i = 2$ on a set $I$ and $s_i = 1$ elsewhere as soon as $|I| > 2D - n = n - 4t - 2$, and then $\mathcal{K}_{2,s}(G) = \{H \in S_2 \cap \mathcal{J}_2 : \nabla H(y_i) = 0 \ \forall i \in I\}$ (Corollary 8.5; over $\mathbb{F}_2$ the Hasse gradient of a quadric is its polar form $B_H(\cdot, y_i)$).

**Proposition 8.13** (Normal conditions for quadrics)**.** *Let $H \in S_{2,K} \cap \mathcal{J}_2$ and $i \in I$. Among the $k$ gradient conditions $B_H(e_l, y_i) = 0$ at $y_i$, the condition in the direction $y_i$ is automatic (Euler), and so are the $m$ conditions in the tangent directions $\Phi^{\sigma^s\prime}(\alpha_i^{2^s})$ of the branches; if these $m+1$ vectors are $K$-independent, the gradient conditions at $y_i$ have rank at most $k - 1 - m$ on $S_{2,K} \cap \mathcal{J}_2$. Moreover for $i \neq j$ the condition at $y_i$ in the direction $y_j$ coincides with the condition at $y_j$ in the direction $y_i$, since $B_H(y_j, y_i) = B_H(y_i, y_j)$.*

**Experiment 8.14** (Exact kernel for quadrics)**.** On the instances $(5, 32, 3)$ and $(6, 64, 3)$, for $I$ of sizes $1, 2, 5, 10, 18, 19, 25$ the rank of the gradient conditions on $S_2 \cap \mathcal{J}_2$ was exactly $|I|(k - 1 - m) - \binom{|I|}{2}$ whenever this is below $\dim \mathcal{J}_2$, and $\dim \mathcal{J}_2$ otherwise. Hence, in the completeness regime and until saturation,

$$\dim \mathcal{K}_{2,s}(G) = \binom{k+1}{2} - m(2D + 1 - t) + (m - 1)(n + 1) - |I|(k - 1 - m) + \binom{|I|}{2},$$

and $\operatorname{rank} E_{2,s} = \binom{k+1}{2} - \dim \mathcal{K}_{2,s}(G)$; e.g. $734 - 25 \cdot 39 + 300 = 59$ for $(6, 64, 3)$ with $|I| = 25$, as computed. The pair dependencies of Proposition 8.13 are the only ones.

Taken together, Theorem 8.8, Conjecture 8.9 and Experiment 8.14 give, for quadrics, the exact rank of the public jet matrix in the completeness regime as a closed formula in $(m, t, n, k, |I|)$; at cryptographic sizes the formula gives $\dim \mathcal{K}_{2,s} = 0$, which is why the attacks need higher degree, multilinear forms and shortening. Open Problem 14.3 is thereby reduced to proving Conjecture 8.9 and extending the count of normal conditions to higher $d$ and $s$.

# 9 Exact hold-out rank

Let $I$ be the retained coordinates and $J$ the held-out ones, $I \sqcup J = \{1, \dots, n\}$, with multiplicities $s_I$ and $s_J$. Write $E_I = E_{d,s_I,I}$, $E_J = E_{d,s_J,J}$, $\mathcal{K}_I = \ker E_I$ and $\mathcal{K}_{\text{all}} = \ker \binom{E_I}{E_J}$.

**Theorem 9.1** (Exact hold-out increment)**.** $\Delta(J) := \dim \mathcal{K}_I - \dim \mathcal{K}_{\text{all}} = \operatorname{rank}(E_J|_{\mathcal{K}_I})$, *and* $0 \le \Delta(J) \le \sum_{j \in J} q_{k,s_j}$.

*Proof.* Rank–nullity for $E_J|_{\mathcal{K}_I}$, whose kernel is $\mathcal{K}_{\text{all}}$. □

**Proposition 9.2** (Cutoff)**.** *If $(s_i)_{i \in I}$ satisfies (5.3), then every $H \in \mathcal{K}_I$ vanishes at every column, so the value conditions at $J$ contribute nothing to $\Delta(J)$; if moreover $s_j = 1$ for all $j \in J$, then $\Delta(J) = 0$. For uniform $s$ and $|J| = h$ this holds when $s(n - h) > dD$, i.e. for*

$$h \;<\; n - \frac{dD}{s} \qquad (\text{for } s = d\colon\ h \le 2t).$$

*Proof.* Theorem 5.2. □

*Remark* 9.3 (The direction of the cutoff; what Apon proved). The retained set must be *large* for completeness, so $\Delta(J) = 0$ for *small* $h$ (at most $2t$ when $s = d$), and $\Delta(J)$ can be positive only beyond that. Note that $\Delta(J)$ for higher-order held-out conditions cannot be bounded by a count in $\mathcal{V}_d$: the higher-order jets of $H$ at a held-out point do not factor through $h_H$ (only the value does). We also note that the "waterfall" of Apon [21] is a different statement: for a Goppa polynomial of degree $t$ he constructs an explicit $(2t + 3)$-dimensional family modulo the true solution on which the whole derivative-flag block at a held support point restricts to a single evaluation condition, so that a specific reconstruction step of [20] needs $c_{\text{need}} > 2t + 3$ held positions. Neither of Weis's routes contains that step [26].

## 9.1 The Holdout/Hermite distinguisher in this language

The GIJS/Weis instance is a shortened systematic code with $k_\ell$ unit columns and $mt$ non-unit columns. In our notation: $W_d = W_d^{\text{ml}}$ (multilinear forms of degree $d$ in $k_\ell$ variables), $I =$ the $mt - 1$ constrained non-unit columns with $s_i = s$, the unit columns carry $s = d - 1$ automatically (Lemma 3.4), $J = \{\tau\}$ with $s_\tau = 1$, and the test is whether $E_J|_{W_d \cap \mathcal{K}_I} = 0$. On a Goppa key this holds by Theorem 5.2 once $(d - 1)k_\ell + s(mt - 1) > dD_\ell - t$ (Theorem 6.1 supplying the $-t$); on a random matrix a nonzero multilinear $h$ has $h(y_\tau) \neq 0$ with probability at least $2^{-d}$ [18, 19]. Saarinen's two-anchor method [23] computes $W_d \cap \mathcal{K}_{I_1}$ and $W_d \cap \mathcal{K}_{I_2}$ for two anchors and works in their intersection modulo the universal nuisance space of Section 3.4; in the TII-254 computation the two kernels had dimension 121, their intersection 106, the certified pair core 80, the nuisance space 64, and the signal 16 dimensions carrying an $\mathbb{F}_{2^8}$ projective line.

## 10 Jets at the held-out column

The value of $H \in \mathcal{K}_I$ at a held-out column is forced; its higher jets there are not, and they carry key information. Weis proves the first-order statement [19, Thm. 4]: if $h_H = 0$ and only the value condition is imposed at $y_\tau$, then the first-order Hasse derivative vector $\nabla H(y_\tau)$, read as a message of the instance, encodes a codeword of the explicit subcode

$$C'_\tau = \Gamma\big(L \setminus \{\alpha_\tau\},\ g \cdot (X - \alpha_\tau)\big),$$

defined by $m+1$ binary linear conditions; the gradients of about $k_\ell + 16$ random kernel vectors span $C'_\tau$ (his Assumption 2), and knowing $C'_\tau$ for every column gives the support by linear algebra. Geometrically: $H$ vanishes on $\Gamma$, so $\nabla H$ at a point of $\Gamma$ is normal to the tangent line, and the tangent at $P_\tau$ is determined by $f'(\alpha_\tau)$, i.e. by the derivative of the hidden parameterization. In our framework this is a two-line theorem.

**Theorem 10.1** (Gradients at a held-out column)**.** *Let $H \in S_d$ be an $\mathbb{F}_2$-form with $h_H = 0$ (for instance any $H \in \mathcal{K}_{d,s}(G; I)$ in the completeness regime), let $\tau$ be any coordinate, and let $w = \nabla H(y_\tau) \in \mathbb{F}_2^k$ be the Hasse gradient. Then the codeword $c = wG \in C$ has code polynomial $f_c = \langle w, f\rangle = \sum_l w_l f_l$, and*

$$f_c(\alpha_\tau) = 0, \qquad f'_c(\alpha_\tau) = 0, \qquad \text{and, by the parity of (2.4)}, \qquad (X - \alpha_\tau)^3 \mid f_c.$$

*Equivalently $c_\tau = 0$ and $\sum_{j \in \operatorname{supp} c} (\alpha_\tau - \alpha_j)^{-1} = 0$, i.e. $c|_{L\setminus\{\alpha_\tau\}} \in \Gamma(L \setminus \{\alpha_\tau\}, g \cdot (X - \alpha_\tau))$. Writing $C'_\tau := \{c \in C : (X-\alpha_\tau)^2 \mid f_c\}$ and $r_\tau := \operatorname{rank}_{\mathbb{F}_2}\big(w \mapsto \langle w, f'(\alpha_\tau)\rangle \;\big|\; \langle w, y_\tau\rangle = 0\big) \le m$, this subcode has $\mathbb{F}_2$-codimension exactly $1 + r_\tau$ in $C$; $r_\tau \ge 1$ when $f'(\alpha_\tau) \notin K \cdot f(\alpha_\tau)$, and $r_\tau = m$, i.e. codimension $m+1$, exactly when $\dim \Gamma(L \setminus \{\alpha_\tau\}, g(X-\alpha_\tau)) = \dim \Gamma(L \setminus \{\alpha_\tau\}, g) - m$, the generic dimension of that Goppa code (the same kind of condition as $\dim \Gamma(L, g) = n - mt$). Also $\Gamma(L \setminus \{\alpha_\tau\}, g(X-\alpha_\tau)) = \Gamma(L \setminus \{\alpha_\tau\}, g(X-\alpha_\tau)^2)$.*

*Proof.* Since $wG = \sum_l w_l c^{(l)}$, its code polynomial is $\sum_l w_l f_l = \langle w, f\rangle$. Put $\tilde{y} = f(\alpha_\tau) = y_\tau / v_\tau$; by homogeneity $\nabla H(\tilde{y}) = v_\tau^{1-d} \nabla H(y_\tau)$. Euler's identity gives $\langle \nabla H(\tilde{y}), \tilde{y}\rangle = dH(\tilde{y}) = 0$, so $f_c(\alpha_\tau) = 0$. Differentiating $H(f(X)) \equiv 0$ gives $\langle \nabla H(f(X)), f'(X)\rangle \equiv 0$, so $f'_c(\alpha_\tau) = \langle \nabla H(\tilde{y}), f'(\alpha_\tau)\rangle = 0$. By Proposition 7.5, $v_\tau f'_c(\alpha_\tau) = \sum_{j : c_j \neq c_\tau} (\alpha_\tau - \alpha_j)^{-1}$, which with $c_\tau = 0$ is the stated sum; it is the condition $\sum_{j \neq \tau} c_j/(X - \alpha_j) \equiv 0 \bmod (X - \alpha_\tau)$, and together with the Goppa condition modulo $g$ this is membership in $\Gamma(L \setminus \{\alpha_\tau\}, g(X-\alpha_\tau))$. Finally $f_c = u_c^2 \pi_{\bar{c}}$ with $\alpha_\tau$ a simple root of $\pi_{\bar{c}}$, so $(X-\alpha_\tau)^2 \mid f_c$ forces $(X-\alpha_\tau) \mid u_c$ and $(X-\alpha_\tau)^3 \mid f_c$; the same parity argument gives the equality of the two Goppa codes. For the codimension: $c = wG$ lies in $C'_\tau$ iff $\langle w, y_\tau\rangle = 0$ (one binary condition) and $\langle w, f'(\alpha_\tau)\rangle = 0$, an $\mathbb{F}_2$-linear map into $K$ whose rank on $\{w : \langle w, y_\tau\rangle = 0\}$ is $r_\tau \le m$; so the codimension is $1 + r_\tau$. Since $C'_\tau \cong \Gamma(L \setminus \{\alpha_\tau\}, g(X-\alpha_\tau))$ sits inside $\Gamma(L \setminus \{\alpha_\tau\}, g)$ (the words with $c_\tau = 0$) with the $m$ binary expansions of one extra $K$-linear parity check, $r_\tau = m$ iff those $m$ checks are independent, which is the stated dimension condition. □

In every instance computed here $r_\tau = m$ (e.g. the gradient span of dimension $k_\ell - m - 1 = 21$ in Experiment 11.15); a proof that the binary Goppa structure forces $r_\tau = m$ would be the same as a proof that Goppa codes have their generic dimension, which is not available.

This is Weis's Theorem 4 [19] in the language of the hidden curve.

**Experiment 10.2.** For $d = 2$ and the instances $(5, 32, 3)$ and $(6, 64, 3)$, the span of $\{\nabla H(y_\tau)G : H \in S_2 \cap I(\Gamma)_2\}$ over the whole space of $\mathbb{F}_2$-quadrics containing $\Gamma$ was exactly $C'_\tau$, of dimension $k - m - 1$ (11 and 39); every gradient codeword had $(X - \alpha_\tau)^3 \mid f_c$. Weis's Assumption 2 (that

the gradients of $\approx k_\ell + 16$ random kernel vectors span $C'_\tau$) is thus the statement that random elements of a space whose gradients span $C'_\tau$ exactly already do so for a sample of size slightly above $\dim C'_\tau$. The rank $r_\tau$ of Theorem 10.1 was $m$ at every non-unit column of 20 random keys of $(6, 64, 6)$ shortened to $k_\ell = 20$ (seeds 100–119), i.e. $C'_\tau$ had codimension $m+1$ in all 720 cases.

**Proposition 10.3** (Jet consistency along the curve)**.** *Let $H \in S_{d,K}$ with $h_H = 0$ and $\tilde{y} = f(\alpha_\tau)$, and expand $f(Y) = \tilde{y} + \sum_{a\geq 1} f^{[a]}(\alpha_\tau)\,(Y-\alpha_\tau)^a$ with the Hasse coefficients $f^{[a]} = D^{[a]}f \in K^k$. By the multivariate Hasse–Taylor formula $H(\tilde{y}+Z) = \sum_\nu \partial^{[\nu]}H(\tilde{y})Z^\nu$, substituting $Z = \sum_{a\geq 1} f^{[a]}(\alpha_\tau)T^a$ and extracting the coefficient of $T^N$ gives, for every $N \geq 1$,*

$$\sum_{\nu} \partial^{[\nu]}H(\tilde{y})\ [T^N] \prod_{j=1}^{k} \Big(\sum_{a\geq 1} f_j^{[a]}(\alpha_\tau)T^a\Big)^{\nu_j} = 0, \tag{10.1}$$

*a relation between the Hasse jet of $H$ at $P_\tau$ and the osculating data $f^{[1]}, f^{[2]}, \ldots$ of the hidden curve. In particular, if $H \in \mathfrak{p}^r_{P_\tau}$ (all $\partial^{[\nu]}H(\tilde{y})$ with $|\nu| < r$ vanish), then the $N = r$ instance of (10.1) reads $\sum_{|\nu|=r} \partial^{[\nu]}H(\tilde{y})\, f'(\alpha_\tau)^\nu = 0$: the degree-$r$ part of $H$ at $\tilde{y}$ vanishes on the tangent vector. Moreover the two-variable polynomial $\Psi(X,Y) := \langle \nabla H(f(X)), f(Y)\rangle$ is divisible by $(Y-X)^2$, so the $K$-vector $w_1 := \partial_X \nabla H(f(X))|_{X=\alpha_\tau}$ also defines a $K$-point $\langle w_1, f\rangle$ of $C'_\tau \otimes K$.*

*Proof.* (10.1) is the coefficient of $(Y-\alpha_\tau)^N$ in $0 = H(f(Y))$, computed with the Hasse–Taylor expansion, which is valid in every characteristic (the Hasse derivatives $\partial^{[\nu]}$ are defined by that expansion, so no polarization of the homogeneous parts of $H$ is needed—in characteristic two such polarizations need not exist, e.g. $Y_1^2$ has zero polar form). For $\Psi$: $\Psi(X,X) = dH(f(X)) = 0$ and $\partial_Y \Psi(X,Y)|_{Y=X} = \langle \nabla H(f(X)), f'(X)\rangle = 0$, so $\Psi$ vanishes to order two along the diagonal; then $\partial_X[(Y-X)^2\Theta] = (Y-X)^2\partial_X\Theta$ in characteristic two. □

*Remark* 10.4 (What higher jets do and do not encode). One might hope that the order-$r$ jet of $H$ at $y_\tau$ "encodes codewords of $\Gamma(L \setminus \{\alpha_\tau\}, g(X-\alpha_\tau)^r)$", deeper subcodes for higher $r$. This is not the right statement. The order-$r$ jet is the collection of Hasse derivatives $\partial^{[\nu]}H(y_\tau)$, $|\nu| = r$; the only canonical way to read it as messages is to contract the degree-$r$ part of $H$ at $y_\tau$ with $r-1$ public vectors $u \in \mathbb{F}_2^k$ (i.e. to take $\nabla$ of $H$ restricted to the line $y_\tau + \varepsilon u$), and the resulting codewords satisfy only $c_\tau = 0$ (homogeneity). Deeper vanishing of $f_c$ at $\alpha_\tau$ requires contracting with the *secret* osculating vectors $f'(\alpha_\tau), f^{[2]}(\alpha_\tau), \ldots$, which is exactly what Proposition 10.3 expresses: the higher jets encode the osculating flag of the hidden curve at $P_\tau$ through the relations above, not deeper Goppa subcodes. This is the mechanism of the derivative-flag route of [20], and the dependence of these relations across held positions is the content of Apon's bound [21]. The deterministic key leakage per held-out column from first-order data is therefore exactly Theorem 10.1: $m+1$ binary conditions, the subcode $C'_\tau$; the second order adds the relation $T_2(f', f') + \langle \nabla H, f^{[2]}\rangle = 0$ between the public jet and the secret flag.

A proof of the deeper relations' rank would quantify how much key information several held-out columns yield jointly, which is the question the hold-out increment of Section 9 measures only in aggregate. The single-run route of [19] reads instead the *column kernel* (relations among the conditions): the Hermite interpolation relation expressing the value condition at a held-out column through the others has coefficients depending on the nodes, and Weis's Theorem 5 (proving a conjecture of [18] along a reduction of [22]) shows that two column-kernel vectors $V_1, V_2$ satisfy $V_2 = \alpha_j^r V_1$ on a locator block of each column, $r$ a power of two, exposing a Frobenius conjugate of every support element as a ratio. In the language of this paper the column kernel is the cokernel side of Theorem 4.1; its structure modulo the universal relations is Open Problem 14.2.

# 11 The multilinear setting: universal relations and the bridge to the attacks

The theorems of Sections 6–9 are stated for the full space $S_d$, while the attacks use the multilinear space $W_d^{\mathrm{ml}}$ in a shortened systematic instance (Section 3.5). This section records what transfers, what the universal (configuration-independent) relations are in the multilinear setting, and what the smallest instances look like.

## 11.1 No universal value relations for multilinear forms

**Proposition 11.1.** *A nonzero multilinear form over $\mathbb{F}_2$ does not vanish on all of $\mathbb{F}_2^k$. Hence $I\big(\mathbb{P}^{k-1}(\mathbb{F}_2)\big)_d \cap W_d^{\mathrm{ml}} = 0$ for every $d$: the universal relations of Section 3.4 (such as $Y_i^2 Y_j + Y_i Y_j^2$) have no multilinear analogue.*

*Proof.* The map from multilinear polynomials to Boolean functions on $\mathbb{F}_2^k$ is the algebraic normal form, which is a bijection (both spaces have dimension $2^k$, and the map is injective by Möbius inversion). □

So in the multilinear setting the universal nuisance space of Saarinen [23] and the local relations of Weis [19] live entirely among the *Hasse conditions*, not among the forms: they are linear dependencies among the functionals $h \mapsto \partial^{[S]} h(y)$ that hold for every configuration.

## 11.2 The universal relations at one point

For a multilinear form $h = \sum_{|T|=d} h_T Y^T$ and $S \subseteq \{1, \dots, k\}$, $\partial^{[S]} Y^T = Y^{T \setminus S}$ if $S \subseteq T$ and 0 otherwise (Hasse derivatives of order $\geq 2$ in one variable vanish on multilinear forms). At a binary point $y$ with support $Y$,

$$\partial^{[S]} h(y) = \sum_{T \supseteq S,\ T \setminus S \subseteq Y} h_T. \tag{11.1}$$

Let $r_2(w; \leq i, j)$ denote the $\mathbb{F}_2$-rank of the inclusion matrix whose rows are the subsets of size $\leq i$ and whose columns the $j$-subsets of a $w$-set (entry 1 for inclusion).

**Proposition 11.2** (Single-point rank)**.** *Let $1 \leq s \leq d$ and let $y \in \mathbb{F}_2^k$ have weight $w$. The functionals $h \mapsto \partial^{[S]} h(y)$, $|S| \leq s-1$, on $W_d^{\mathrm{ml}}$ span a space of dimension*

$$R(k, w, d, s) = \sum_{a=0}^{\min(s-1,\, k-w)} \binom{k-w}{a} r_2\big(w;\ \leq s-1-a,\ d-a\big), \tag{11.2}$$

*with the convention $r_2(w; \leq i, j) = 0$ when $j > w$ (there are no $j$-subsets of a $w$-set); the range of $a$ stops at $s-1$ because a block with $a > s-1$ has no rows, so for $a \leq s-1$ the row bound $s-1-a$ is nonnegative. Hence a single constrained column contributes exactly $R(k, w, d, s)$ independent conditions out of the nominal $\sum_{j<s} \binom{k}{j}$, and the difference is the universal nuisance at that point.*

*Proof.* In (11.1), $T \setminus S \subseteq Y$ and $S \subseteq T$ force $T \setminus Y = S \setminus Y =: A$. Group the monomials $T$ by $A = T \setminus Y$ (a subset of the complement of $Y$, $|A| = a \leq d$) and the conditions $S$ by $S \setminus Y = A$. Conditions and monomials with different $A$ involve disjoint sets of coefficients, so the span decomposes as a direct sum over $A$. For fixed $A$, writing $T = A \sqcup T'$ and $S = A \sqcup S'$ with $S' \subseteq T' \subseteq Y$, $|T'| = d - a$, $|S'| \leq s-1-a$, the functional is $h \mapsto \sum_{T' \supseteq S'} h_{A \sqcup T'}$: the inclusion matrix of $S'$ against $(d-a)$-subsets of $Y$. There are $\binom{k-w}{a}$ choices of $A$. □

The stacked inclusion ranks in (11.2) have a closed form. For a single row size Wilson's theorem [6] gives $\operatorname{rank}_2 W_{i,j}(w) = \sum_{l\le i,\ \binom{j-l}{i-l}\ \text{odd}} \left[\binom{w}{l} - \binom{w}{l-1}\right]$ in the range $i \le \min(j, w-j)$, which is in general smaller than $\binom{w}{i}$; stacking all row sizes $\le i$ restores the characteristic-zero rank, by an elementary argument.

**Lemma 11.3** (Stacked inclusion matrices over $\mathbb{F}_2$)**.** *Let $0 \le j \le w$ and $i \ge 0$, and let $W_{\le i,j}(w)$ be the $0/1$ matrix whose rows are the subsets $B$ of a $w$-set with $|B| \le i$, whose columns are its $j$-subsets $T$, and whose $(B,T)$ entry is $[B \subseteq T]$. Then*

$$r_2(w; \le i, j) = \operatorname{rank}_{\mathbb{F}_2} W_{\le i,j}(w) = \binom{w}{\min(i,\ j,\ w-j)}.$$

*Proof.* Write $i_0 = \min(i, j, w-j)$. *Upper bound.* For an integer matrix the rank of its reduction modulo 2 is at most its rank over $\mathbb{Q}$, since a minor that is nonzero modulo 2 is a nonzero integer. Over $\mathbb{Q}$, for $l \le i' \le j$ one has $W_{l,i'}W_{i',j} = \binom{j-l}{i'-l} W_{l,j}$ with a nonzero scalar, so the row space of $W_{l,j}$ lies in that of $W_{i',j}$. Taking $i' = i_0$ shows that the $\mathbb{Q}$-row space of $W_{\le i_0,j}$ is that of $W_{i_0,j}$, of dimension $\le \binom{w}{i_0}$; and if $i > i_0$ then either $i_0 = j$, in which case rows of size $> j$ are zero, or $i_0 = w-j$, in which case $\binom{w}{i_0} = \binom{w}{j}$ is the number of columns. In all cases $\operatorname{rank}_{\mathbb{F}_2} \le \operatorname{rank}_{\mathbb{Q}} \le \binom{w}{i_0}$.

*Lower bound.* Since $W_{\le i,j} \supseteq W_{\le i_0,j}$ as a set of rows, it suffices to treat $i \le \min(j, w-j)$ and prove $\operatorname{rank}_{\mathbb{F}_2} W_{\le i,j}(w) \ge \binom{w}{i}$, by induction on $w$. Two reductions first. (a) Complementation: over $\mathbb{F}_2$, $[B \cap T' = \emptyset] = \sum_{B' \subseteq B} [B' \subseteq T']$ and conversely $[B \subseteq T'] = \sum_{B' \subseteq B} [B' \cap T' = \emptyset]$, so the row spaces of $([B \subseteq T])_{|B| \le i}$ and $([B \cap T = \emptyset])_{|B| \le i}$ coincide, and $T \mapsto T^c$ identifies the latter with $W_{\le i, w-j}(w)$: $r_2(w; \le i, j) = r_2(w; \le i, w-j)$. Hence we may assume $j \le w-j$. If $i = j = w/2$, the rows of size $i$ form the identity matrix and we are done; otherwise $i < w-j$ (since $i = w-j$ together with $i \le j \le w-j$ would force $i = j = w/2$), so $i \le \min(j, w-1-j)$. (b) If $i = 0$ the rank is $1 = \binom{w}{0}$.

Now fix an element $\omega$ of the $w$-set and order the columns as $T \not\ni \omega$ (the $j$-subsets of the remaining $(w-1)$-set) followed by $T = T' \sqcup \{\omega\}$ (the $(j-1)$-subsets $T'$). A row $B \not\ni \omega$ reads $\big(W_{|B|,j}(w-1)_B \,\big|\, W_{|B|,j-1}(w-1)_B\big)$ and a row $B = B' \sqcup \{\omega\}$, $|B'| \le i-1$, reads $\big(0 \,\big|\, W_{|B'|,j-1}(w-1)_{B'}\big)$. For a block matrix $M = (L \,|\, R)$, $\operatorname{rank} M \ge \operatorname{rank} L + \dim\{\text{rows in the row space of } M \text{ of the form } (0\,|\,*)\}$. Here the left block $L$ has the row space of $W_{\le i,j}(w-1)$, and the rows $(0\,|\,*)$ include all of $(0\,|\,W_{\le i-1,j-1}(w-1))$. Both are in the induction range ($i \le \min(j, (w-1)-j)$ and $i-1 \le \min(j-1, (w-1)-(j-1))$), so

$$\operatorname{rank}_{\mathbb{F}_2} W_{\le i,j}(w) \ \ge\ \binom{w-1}{i} + \binom{w-1}{i-1} = \binom{w}{i}. \qquad \square$$

**Theorem 11.4** (Closed form of the single-point rank)**.** *Let $1 \le s \le d$ (as throughout: a multiplicity larger than the degree makes the form vanish). For a binary point of weight $w$,*

$$R(k,w,d,s) = \sum_{a=0}^{s-1} \binom{k-w}{a} \binom{w}{\min(s-1-a,\ d-a,\ w-d+a)},$$

*with the convention $\binom{x}{y} = 0$ for $y < 0$ or $y > x$; thus a term is $0$ when $d-a > w$ (the minimum is then negative) and when $a > k-w$. In particular:*

(i) *if $w \ge d+s-1$ then $R(k,w,d,s) = \binom{k}{s-1}$: the conditions of all orders $< s$ at the point have exactly the rank of the nominal order-$(s-1)$ conditions, and the universal nuisance space at the point has dimension $\sum_{l \le s-2} \binom{k}{l}$, independent of the point;*

(ii) *the universal nuisance exceeds* $\sum_{l\le s-2}\binom{k}{l}$ *only at columns of weight* $w < d+s-1$, *by* $\sum_a \binom{k-w}{a}\big[\binom{w}{s-1-a} - \binom{w}{\min(s-1-a,\,w-d+a)}\big]$.

*Proof.* Insert Lemma 11.3 into (11.2); the range of $a$ is cut at $s-1$ because the block $A$ has no rows when $|A| = a > s-1$. For (i), if $w \ge d+s-1$ then $s-1-a \le w-d+a$ for all $a \ge 0$ and $s-1-a \le d-a$ because $s \le d$, so the minimum is $s-1-a$ and Vandermonde's identity gives $\sum_a \binom{k-w}{a}\binom{w}{s-1-a} = \binom{k}{s-1}$. (ii) is the difference of the two expressions. □

The theorem was checked against the inclusion-matrix computation of (11.2) for all $6 \le k \le 14$, $2 \le d \le 6$, $1 \le s < d$ and all weights $w$ (no discrepancy), and (11.2) itself against direct computation for $(k,d,s) \in \{(9,4,3),(10,5,4),(11,6,5),(12,4,3),(12,5,4)\}$; e.g. for $(k,d,s) = (12,5,4)$ and $w = 7 < d+s-1 = 8$ there are 299 nominal conditions of order $< 4$ at the point, of which 206 are independent, while at $w = 10$ exactly $\binom{12}{3} = 220$ are.

**Experiment 11.5** (Order reduction and cross-point nuisance)**.** (i) For random configurations of $N$ binary points, the rank of all conditions of order $< s$ equalled the rank of the conditions of orders $d-2, d-4, d-8, \dots$ only ($\{0,2\}$ for $(4,3)$; $\{1,3\}$ for $(5,4)$), in every case tested: this is the reduction of Hemmert–Wiemers and Weis quoted in Section 3.5, confirmed on the nose. (ii) Across several points the nuisance exceeds the sum of the single-point values: the cross-point part is the subject of the next experiment.

**Theorem 11.6** (Pairwise universal relations)**.** *Let* $H$ *be a multilinear form of degree* $d$, *let* $y_1, y_2 \in \mathbb{F}_2^k$ *have supports* $Y_1, Y_2$, *let* $1 \le s \le d$, *and suppose* $r := 2s-d-2 \ge 0$ *(so that* $s-1-r = d+1-s \ge 1$ *and the sums below are nonempty). Then for every* $R \subseteq \{1,\dots,k\}$ *with* $|R| = r$,

$$\sum_{\substack{S\subseteq Y_2\setminus R\\ |S|=s-1-r}} \partial^{[R\cup S]}H(y_1) \;=\; \sum_{\substack{S\subseteq Y_1\setminus R\\ |S|=s-1-r}} \partial^{[R\cup S]}H(y_2).$$

*Both sides are combinations of Hasse conditions of order* $s-1$ *at the two points, so this is a relation among the rows of the jet matrix that holds for every pair of columns. For* $r < 0$ *there is no relation of this shape, and for* $(d,s) = (5,3),(4,2),(3,2),(6,3)$ *none was found.*

*Proof.* Both sides are linear in $H$, so it suffices to take $H = Y^T$, $|T| = d$. By (11.1), $\partial^{[R\cup S]}Y^T(y_1) = [R\cup S \subseteq T]\,[T\setminus(R\cup S) \subseteq Y_1]$, and $T\setminus(R\cup S) \subseteq Y_1$ means $(T\setminus Y_1)\setminus R \subseteq S$. Hence the left side equals $[R\subseteq T]$ times the number of $S$ with $(T\setminus Y_1)\setminus R \subseteq S \subseteq (T\cap Y_2)\setminus R$ and $|S| = s-1-r$. Such $S$ exist only if $T \subseteq Y_1\cup Y_2\cup R$; then, writing $a = |(T\setminus R)\setminus Y_2|$, $b = |(T\setminus R)\setminus Y_1|$ and $c = |(T\setminus R)\cap Y_1\cap Y_2|$, so that $a+b+c = d-r$, the number is $\binom{c}{s-1-r-b}$. By symmetry the right side is $[R\subseteq T]\,[T\subseteq Y_1\cup Y_2\cup R]\binom{c}{s-1-r-a}$. Since $(s-1-r-b)+(s-1-r-a) = 2s-2-2r-(d-r-c) = c$ by the choice of $r$, the two binomial coefficients are equal (both are 0 when one lower index is negative, because the other then exceeds $c$). □

**Experiment 11.7** (Pairwise universal nuisance)**.** The relations of Theorem 11.6 were checked to lie in the row spaces of both points in every instance tested, as they must. Their number is $\binom{k}{2s-d-2}$, and for two points of weight $\ge d+s-1$ whose supports are in general position (neither contained in the other, overlap not extreme) they are independent modulo the single-point relations and span all cross-point relations: the rank of all conditions of order $< s$ at the pair fell short of $2\binom{k}{s-1}$ by exactly $\binom{k}{2s-d-2}$ in every such case (1 for $(d,s) = (4,3)$ at $k = 12,13,14,16,18$ and for $(6,4)$ at $k = 14,16$; $k$ for $(5,4)$ at $k = 14,18,20$; $\binom{k}{2}$ for $(6,5)$ at $k = 16,20$). Pairs with a large overlap, or of weight below $d+s-1$, have more relations (104 instead of 20 for $(5,4)$, $k = 20$, weights 10,

overlap 8; 132 instead of 91 for $(6,5)$, $k=14$, weights $9<10$), and then the $\binom{k}{2s-d-2}$ relations of the theorem can even become dependent. For $N$ points in general position the deficiency was $\binom{N}{2}\binom{k}{2s-d-2}$ for small $N$ ($N\le 5$ at $(5,4)$, $k=20$; $N\le 4$ at $(4,3)$, $k=16$) and grew faster beyond, so relations involving three or more points exist as well.

**Conjecture 11.8** (Pairwise nuisance)**.** For two points of weight $\ge d+s-1$ in general position the relations of Theorem 11.6 span the cross-point relations, so that the conditions of order $<s$ at the pair have rank exactly $2\binom{k}{s-1}-\binom{k}{2s-d-2}$.

*Remark* 11.9 (What this says about the attacks). In a shortened systematic Goppa matrix the non-unit columns have weight close to $k_\ell/2$, far above $d+s-1$ (at TII-253, $(m,t,n)=(8,9,214)$, in Weis's parameters $k_\ell=46$, $(d,s)=(6,5)$ [19]: weights near 23 against $d+s-1=10$). By Theorem 11.4(i) each constrained column then imposes exactly $\binom{k_\ell}{s-1}$ independent conditions on its own. If all $mt=72$ non-unit columns are constrained, the single-point counts add up to $72\binom{46}{4}=11\,749\,320$ against $\binom{46}{6}=9\,366\,819$ multilinear monomials, so a kernel exists only through cross-point dependencies, and Theorem 11.6 supplies $\binom{72}{2}\binom{46}{2}=2\,645\,460$ of them from pairs alone (independent by Conjecture 11.8), before the higher-order terms—already more than the excess $2\,382\,501$. So the kernel at TII-253 is predicted to be dominated by the universal nuisance, of a size comparable to the $4\,617\,031$ nuisance directions Saarinen reports at TII-254 [23], with the structural signal a small remainder; this is precisely what Experiment 11.15 exhibits in miniature at $(6,6,64)$ (kernels of 428–650 against random controls of 448–642 over six keys). Theorems 11.4 and 11.6 thus settle the one-point part and the existence half of the pairwise part of Open Problem 14.2; the structural part of the kernel is $\mathcal{J}_d^{\mathrm{ml}}$ cut by the normal conditions, Section 8.3.

### 11.3 The multilinear part of the curve ideal

Since $W_d^{\mathrm{ml}}$ is defined over $\mathbb{F}_2$ and is Galois-stable, the descent theorem applies verbatim:

$$\dim_{\mathbb{F}_2}\bigl(W_d^{\mathrm{ml}}\cap I(\Gamma)_d\bigr)=\dim_K\bigl(W_{d,K}^{\mathrm{ml}}\cap\mathcal{J}_d\bigr)=:\dim\mathcal{J}_d^{\mathrm{ml}},$$

and in the completeness regime the kernel of the multilinear jet matrix is $W_d^{\mathrm{ml}}\cap\mathcal{J}_d$ cut by the normal conditions at the constrained columns, exactly as in Corollary 8.5.

**Experiment 11.10.** For $d=2$, every $\mathbb{F}_2$-quadric vanishing on $\Gamma$ was multilinear: $\dim\mathcal{J}_2^{\mathrm{ml}}=\dim\mathcal{J}_2$ ($45=45$ for $(5,32,3)$, $734=734$ for $(6,64,3)$). For $d=3$ the multilinear part is proper: $\dim\mathcal{J}_3^{\mathrm{ml}}=549$ against $\dim\mathcal{J}_3=736$ for $(5,32,3)$, with 680 of the 969 cubic monomials multilinear.

In systematic coordinates the multilinear part has its own Hilbert function, and it is the one that governs the attacks. Let the code be in systematic form with $k_\ell$ unit columns and $mt$ non-unit columns, let $U$ be the set of support elements of the unit columns, and $\pi_U=\prod_{a\in U}(X-\alpha_a)$.

**Theorem 11.11** (Rank of the multilinear evaluation map)**.** *Let $d\ge 2$ and $dD-2t-(d-1)k_\ell\ge mt$. Let $\rho_d^{\mathrm{ml}}$ be the $\mathbb{F}_2$-rank of $W_d^{\mathrm{ml}}\to\bigoplus_{s<m}K[X]_{\le dD}$, $H\mapsto(H(f^{(2^s)}))_s$ (equivalently, by Theorem 8.4, $\dim_K$ of the span of the conjugate evaluations of the multilinear monomials), so that $\dim\mathcal{J}_d^{\mathrm{ml}}=\binom{k_\ell}{d}-\rho_d^{\mathrm{ml}}$. Then*

$$\rho_d^{\mathrm{ml}}\ \le\ m\bigl(dD+1-t-(d-1)k_\ell\bigr)-(m-1)(mt+1),$$

*and the $s=0$ component $\mathcal{V}_d^{\mathrm{ml}}:=\operatorname{span}_K\{f_T:|T|=d\}$ has dimension at most $dD+1-t-(d-1)k_\ell$.*

*Proof.* A multilinear $H$ vanishes to order $d-1$ at every unit point $e_a$ (Lemma 3.4), and $\Phi^{\sigma^s}(\alpha_a^{2^s})=e_a$, so $H(f^{(2^s)})$ vanishes to order $\ge d-1$ at $\alpha_a^{2^s}$ for $a\in U$. These $(d-1)k_\ell$ conditions have full rank on

$\mathcal{W}_d^{(s)}$, because already on its subspace $(g^{(2^s)})^2 K[X]_{\le dD-2t}$ they are Hermite interpolation conditions at $k_\ell$ distinct points on polynomials of degree range $\ge (d-1)k_\ell - 1$, with $(g^{(2^s)})^2$ nonvanishing there. So each component lies in a space of dimension $dD + 1 - t - (d-1)k_\ell$. The $(m-1)(mt+1)$ gluing conditions of Theorem 8.8 at the $mt$ non-unit points and at $\infty$ remain independent on these subspaces: $Q \mapsto Q\,(g^{(2^s)})^2(\pi_U^{(2^s)})^{d-1}$, $Q \in K[X]_{\le dD-2t-(d-1)k_\ell}$, lands in the subspace and realizes arbitrary values at the $mt$ points and an arbitrary leading coefficient once $dD - 2t - (d-1)k_\ell \ge mt$. □

**Conjecture 11.12** (Multilinear Hilbert function)**.** Fix $d \ge 2$ and let $(m_i, n_i, t_i, k_{\ell,i})_{i\ge 1}$ be any sequence of parameters with $t_i \to \infty$, $n_i \le 2^{m_i}$, and $dD_i - 2t_i - (d-1)k_{\ell,i} \ge m_i t_i$ (the hypothesis of Theorem 11.11, $D_i = k_{\ell,i} + m_i t_i - 2t_i - 1$). For the $i$-th parameter set draw $g$ uniformly from the monic irreducible polynomials of degree $t_i$ over $\mathbb{F}_{2^{m_i}}$, the support uniformly among the $n_i$-subsets of $\mathbb{F}_{2^{m_i}}$ (the experiments use the first $n_i$ field elements, which for $n_i = 2^{m_i}$ is the same thing), and the $k - k_{\ell,i}$ shortened positions uniformly among subsets of that size. Then the probability that equality holds tends to 1 as $i \to \infty$; we do not claim uniformity in the choice of sequence. Observed frequencies: 7 of 7 instances in Experiment 11.13, 20 of 20 at $d = 4$ and 19 of 20 at $d = 5$ over the keys of seeds 100–119 at $(6, 64, 6)$, $k_\ell = 20$, and 6 of 6 at $k_\ell = 28$:

$$\dim \mathcal{J}_d^{\mathrm{ml}} = \binom{k_\ell}{d} - m\big(dD+1-t-(d-1)k_\ell\big) + (m-1)(mt+1), \qquad \dim \mathcal{V}_d^{\mathrm{ml}} = dD+1-t-(d-1)k_\ell.$$

**Experiment 11.13.** Computing the conjugate evaluation of every multilinear monomial from the code polynomials of the systematic rows (Lemma 2.1):

| $(m,t,n)$ | $k_\ell$ | $d$ | $D_\ell$ | $\binom{k_\ell}{d}$ | $\rho_d^{\mathrm{ml}}$ | bound of Theorem 11.11 |
|---|---|---|---|---|---|---|
| $(6,6,64)$ | 16 | 4 | 39 | 1 820 | 433 | 433 |
| $(6,6,64)$ | 20 | 4 | 43 | 4 845 | 457 | 457 |
| $(6,6,64)$ | 20 | 5 | 43 | 15 504 | 595 | 595 |
| $(6,6,64)$ | 24 | 5 | 47 | 42 504 | 619 | 619 |
| $(6,6,64)$ | 28 | 5 | 51 | 98 280 | 643 | 643 |
| $(5,3,32)$ | 12 | 4 | 20 | 495 | 146 | 146 |
| $(6,3,64)$ | 15 | 5 | 26 | 3 003 | 290 | 313 |
| $(7,7,128)$ | 31 | 5 | 65 | 169 911 | 1 065 | 1 065 |
| $(7,7,128)$ | 32 | 5 | 66 | 201 376 | 1 072 | 1 072 |

The bound is attained in every instance of the table with $t \ge 5$, including the two $(7,7,128)$ instances with 169 911 and 201 376 monomials. Over 20 random keys of $(6, 64, 6)$ shortened to $k_\ell = 20$ (seeds 100–119) it was attained in all 20 cases for $d = 4$ and in 19 of 20 for $d = 5$: the key with seed 101 ($g = X^6 + 29X^5 + 48X^4 + 50X^3 + 34X^2 + 42X + 14$) has $\rho_5^{\mathrm{ml}} = 594$ against 595 and $\dim \mathcal{V}_5^{\mathrm{ml}} = 129$ against 130, while its $\mathcal{V}_d^{\mathrm{ml}}$ for $d \le 4$ and the $\mathcal{V}_2$ of its unshortened code are of full dimension. So equality fails for some keys, by one, and the conjecture is stated as a probability over keys, with the observed frequencies recorded; the $t = 3$ instance falls short by 23 (its $\mathcal{V}_5^{\mathrm{ml}}$ has dimension 61 against 68), the same small-$t$ deviation as in Experiment 6.3 and in Weis's experiments. For $d = 2$ the conjecture is equivalent to Conjecture 8.9, the $k_\ell$ squares $Y_a^2$ accounting exactly for the difference $\rho_2 - \rho_2^{\mathrm{ml}} = k_\ell$, which is Experiment 11.10.

## 11.4 The smallest shortened systematic instances

**Experiment 11.14.** We built the GIJS/Weis matrix (Section 9) for the smallest codes available here: shorten $\Gamma(L, g)$ to $k_\ell$ unit columns, keep the $mt - 1$ constrained non-unit columns and one held-out column, impose the reduced-order Hasse conditions on multilinear forms of degree $d$, and compute the kernel over $\mathbb{F}_2$; the same was done for a random systematic matrix of the same shape.

| $(m,t)$ | $k_\ell$ | $(d,s)$ | forced | $dD_\ell - t$ | monomials | rows | rank (Goppa) | ker (Goppa) | ker (random) |
|---|---|---|---|---|---|---|---|---|---|
| $(6,3)$ | 9 | $(4,3)$ | 78 | 77 | 126 | 629 | 126 | 0 | 0 |
| $(6,3)$ | 15 | $(5,4)$ | 128 | 127 | 3003 | 7990 | 3003 | 0 | 0 |
| $(5,3)$ | 12 | $(4,3)$ | 78 | 77 | 495 | 938 | 495 | 0 | 3 |

In each case the completeness condition $(d-1)k_\ell + s(mt-1) > dD_\ell - t$ holds with one zero to spare, but the kernel is trivial: at these sizes the single-point conditions of Theorem 11.4 outnumber the monomials several times over. (In the third row the Goppa matrix even has full rank where the random one does not, a small-parameter deviation of the kind Weis reports for $t \le 5$.) The first nontrivial instance is the subject of the next subsection.

### 11.5 A Classic-McEliece-shaped instance: $(m,t,n) = (6,6,64)$

The toy parameters at which Weis reports experiments with the attack are $(m,t,n) = (6,6,64)$ at degree 5 [19]. We built the complete GIJS/Weis system for binary Goppa codes with these parameters: $k = n - mt = 28$, shortened to $k_\ell$ unit columns, the $mt = 36$ non-unit columns of the systematic form split into $N$ constrained columns ($N$ denotes the number of constrained columns in this and the next subsection) (reduced-order Hasse conditions of multiplicity $s$ on multilinear forms of degree $d$) and $36 - N$ held-out columns; the kernel was computed over $\mathbb{F}_2$ with a bit-sliced eliminator (98 280 unknowns and 115 640 conditions at $k_\ell = 28$, about 25 minutes on one core). For every kernel we recorded: whether the kernel forms vanish at the held-out columns (the distinguisher of Section 9); whether they vanish *identically on the hidden curve*, i.e. lie in $\mathcal{J}_d^{\mathrm{ml}}$, computed from the explicit code polynomials of Lemma 2.1; the gradient test of Theorem 10.1; and the same quantities for a random systematic matrix of the same shape. By Theorem 11.4 every column here imposes exactly $\binom{k_\ell}{s-1}$ conditions on its own (all weights exceed $d+s-1$), so the column $N\binom{k_\ell}{s-1}$ of the table is the rank the system would have without cross-point dependencies.

**Experiment 11.15** ($(6,6,64)$, $(d,s) = (5,4)$, one held-out column)**.**

| $k_\ell$ | $N$ | complete | monomials | $N\binom{k_\ell}{3}$ | rank | kernel (Goppa / random) | in $\mathcal{J}_5$ |
|---|---|---|---|---|---|---|---|
| 18 | 35 | yes | 8 568 | 28 560 | 8 568 | 0 / 0 | – |
| 20 | 35 | yes | 15 504 | 39 900 | 15 504 | 0 / 0 | – |
| 24 | 35 | yes | 42 504 | 70 840 | 42 504 | 0 / 0 | – |
| 28 | 35 | yes | 98 280 | 114 660 | 97 630 | **650** / 642 | **650** |
| 20 | 16 | no | 15 504 | 18 240 | 15 227 | 277 / 324 | 0 |
| 20 | 17 | no | 15 504 | 19 380 | 15 500 | 4 / 1 | 0 |
| 20 | 18 | no | 15 504 | 20 520 | 15 503 | 1 / 0 | 0 |
| 20 | 19 | no | 15 504 | 21 660 | 15 504 | 0 / 0 | – |

Seed 11 throughout the table. Below completeness, twenty further keys at $k_\ell = 20$, $N = 17$ (seeds 100–119) gave the following kernel dimensions (Goppa / random), with 0 forms in $\mathcal{J}_5$ in every case:

| seed | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 |
|---|---|---|---|---|---|---|---|---|---|---|
| Goppa | 0 | 53 | 0 | 31 | 106 | 0 | 0 | 0 | 0 | 0 |
| random | 0 | 0 | 3 | 0 | 1 | 6 | 0 | 0 | 0 | 27 |
| seed | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 |
| Goppa | 154 | 31 | 2 | 0 | 0 | 2 | 6 | 2 | 249 | 0 |
| random | 0 | 15 | 2 | 0 | 0 | 48 | 1 | 3 | 1 | 2 |

Four more keys at $N = 17$ (seeds 12–15) gave $5, 2, 34, 11$ against $43, 0, 3, 1$, and $N = 18$ gave 0 throughout. At $k_\ell = 16$, $(d,s) = (4,3)$, $N = 16$: Goppa kernel 154 against random 109, 136 of the 154 forms vanishing at all 20 held-out columns, 0 in $\mathcal{J}_4$.

A second code, $(m, t, n) = (7, 7, 128)$ ($k = 79$), was shortened to $k_\ell = 31$ with all $N = mt = 49$ non-unit columns constrained (complete, $320 > 318$, again only through the $-t$): 169 911 monomials, 221 774 conditions, $N\binom{31}{3} = 220\,255$, rank 169 911, *kernel* 0 (three hours, 3.4 GB). At $k_\ell = 32$, the largest complete value ($324 > 323$): 201 376 monomials, 244 608 conditions, rank 201 376, *kernel* 0 again (five hours, 4.8 GB).

For the $k_\ell = 28$ kernel of seed 11: all 650 forms vanish at the held-out column (as Theorem 5.2 requires) and identically on the hidden curve ($\dim \mathcal{J}_5^{\mathrm{ml}} = 97\,637$, Experiment 11.13), so the Hasse conditions at the 35 columns have rank 96 987 on $\mathcal{J}_5^{\mathrm{ml}}$, i.e. 2 771 per column against the single-point value $\binom{28}{3} = 3\,276$. The gradients at the held-out column span a space of dimension exactly $21 = k_\ell - m - 1$, and all 650 gradient codewords lie in $C'_\tau \cong \Gamma(L \setminus \{\alpha_\tau\}, g(X - \alpha_\tau))$ (Theorem 10.1), which they generate (an observation, not part of the theorem). The random matrix of the same shape has rank 97 638 and a kernel of dimension 642, of which 347 forms vanish at a random further column.

Five further keys with the same parameters (seeds 12–16), each with its own random control, the same measurements, and the filtration of Experiment 11.17:

| seed | kernel | random | in $\mathcal{J}_5$ | $\rho_5^{\mathrm{ml}}$ | grad. rank | $r_\tau$ | $\mathrm{gr}_1$ | $\mathrm{gr}_2$ | $\delta_1$ |
|---|---|---|---|---|---|---|---|---|---|
| 11 | 650 | 642 | 650 | 643 | 21 | 6 | 134 | 516 | 38 |
| 12 | 539 | 601 | 539 | 643 | 21 | 6 | 134 | 405 | 38 |
| 13 | 428 | 448 | 428 | 643 | 21 | 6 | 134 | 294 | 38 |
| 14 | 498 | 518 | 498 | 643 | 21 | 6 | 134 | 364 | 38 |
| 15 | 555 | 503 | 555 | 643 | 21 | 6 | 128 | 427 | 44 |
| 16 | 456 | 476 | 456 | 643 | 21 | 6 | 134 | 322 | 38 |

($N_1 = 172$ for every key, as it must be: it depends only on the dimension counts.) The gradient span was compared with the dimension of $\Gamma(L \setminus \{\alpha_\tau\}, g(X - \alpha_\tau))$ computed independently from its parity-check matrix: 21 for all six keys, equal to the span.

The table shows the GIJS mechanism at its smallest instance, and shows what it is made of. Shortening below $k_\ell = 28$ leaves the completeness regime with a *trivial* kernel: the single-point conditions outnumber the monomials by a factor 1.7 to 3.3 and no amount of cross-point dependency makes up the gap. At $(7, 7, 128)$ the kernel is trivial at both complete values $k_\ell = 31, 32$ (factors 1.30 and 1.21). At $(6, 6, 64)$, $k_\ell = 28$, the factor is 1.17: the universal cross-point nuisance (Experiment 11.7: $\binom{35}{2} \cdot 28 = 16\,660$ pairwise relations, plus higher-order ones) exceeds the excess 16 380, and a kernel appears *for the random matrix as well*. Across six keys the Goppa kernel ranges over 428–650 and the random control over 448–642, with no consistent sign to the difference: at this size the kernel is, to within the key-to-key fluctuation, the universal nuisance, and the difference between a Goppa key and a random matrix is not visible in the dimension alone. (This is Saarinen's picture at TII-254—4 617 031 nuisance directions against a 16-dimensional signal [23]—with the signal here below the noise.) What the geometry adds is that *all* kernel forms, nuisance included, vanish on the hidden curve (six keys, 3 126 forms, no exception), exactly as Theorem 5.2 predicts (completeness holds with three zeros to spare, $252 > 249$, and only because of the $-t$ of Theorem 6.1: with the naive bound $dD_\ell$ the instance would not be certified); that the kernel is the structural object $\mathcal{J}_5^{\mathrm{ml}} \cap$ (normal conditions) of Corollary 8.5, with $\dim \mathcal{J}_5^{\mathrm{ml}}$ given on the nose by Conjecture 11.12; and that the gradients at the held-out column lie in $C'_\tau$ with $r_\tau = m$ (Theorem 10.1) and, as an observation beyond the theorem, span all of it on every key, so that one held-out column leaks the $(m+1)$-codimensional subcode $\Gamma(L \setminus \{\alpha_\tau\}, g(X - \alpha_\tau))$ in full. (The theorem gives the containment and the codimension; that the span is the whole subcode is what the experiments show, on six keys here and in Experiment 10.2.)

*Hold-out protocol.* The random control is a systematic matrix of the same shape: $k_\ell$ unit columns and $mt$ non-unit columns drawn uniformly and independently from $\mathbb{F}_2^{k_\ell} \setminus \{0\}$ (seed seed$+7$), of which $N = mt - 1$ are constrained with the same $(d, s)$ and orders and one is held out; the dimensions of unknowns and rows are therefore identical to the Goppa instance, the ranks are not matched by construction. The statistic is per form: for each kernel form we evaluate it at the held-out column and count the forms that vanish there. On the Goppa keys this is 100% on every key (3 126 of 3 126 forms over six keys; Theorem 5.2 makes it a certainty, not a statistic); on the random controls it was 347 of 642 forms (54%) for the seed-11 control at its held-out column, and for the five other controls, evaluated at 10 random columns each, the fraction ranged over 0.445–0.558 with per-control means 0.493–0.507. A random nonzero multilinear form vanishes at a random point with probability about $1/2$, so this is the expected behaviour of a structureless kernel. The distinguisher of Section 9 therefore separates the two cases perfectly on these instances although the kernel dimensions alone do not. Below completeness the kernels that appear in a window of one or two values of $N$ are small, fluctuate from key to key in both directions relative to a random matrix, and contain *no* form vanishing on the curve; their forms do vanish at most held-out columns, but so do the forms of the random kernel at random columns, because a multilinear form vanishing to order $s$ at $y$ vanishes on the whole Hamming ball of radius $s-1$ about $y$ (Taylor's formula with Hasse derivatives). Hold-out vanishing of a few forms is therefore not by itself evidence of structure; membership in $\mathcal{J}_d$ is.

*Heuristic* 11.16 (Window-intersection criterion for relation kernels). Theorem 11.6, Conjecture 11.8 and Experiment 11.15 together give a quantitative prediction of *when* the attack system has a kernel. With $N$ constrained columns of weight $\geq d+s-1$, the conditions have nominal rank $N\binom{k_\ell}{s-1}$ (Theorem 11.4), the monomials number $\binom{k_\ell}{d}$, and the pairwise universal relations of Theorem 11.6 number $\binom{N}{2}\binom{k_\ell}{2s-d-2}$; relations among three or more points add a smaller correction. So a (nuisance) kernel is predicted exactly when

$$\binom{N}{2}\binom{k_\ell}{2s-d-2} \gtrsim N\binom{k_\ell}{s-1} - \binom{k_\ell}{d},$$

and the structural signal rides on top of it. This is a *lower* bound on $k_\ell$; completeness, $(d-1)k_\ell + sN > dD_\ell - t$, is an *upper* bound. The attack needs the two windows to intersect:

| instance | $(d,s)$, $N$ | $k_\ell$ | excess $N\binom{k_\ell}{s-1} - \binom{k_\ell}{d}$ | pairwise | observed kernel |
|---|---|---|---|---|---|
| $(6,6,64)$ | $(5,4)$, 35 | 27 | 21 645 | 16 065 | – |
| | | 28 | 16 380 | 16 660 | 650 (random 642) |
| $(7,7,128)$ | $(5,4)$, 49 | 31 | 50 344 | 36 456 | 0 |
| | | 32 | 41 664 | 37 632 | 0 |
| | | 33 | 30 008 | 38 808 | not complete |
| TII-253 | $(6,5)$, 72 | 45 | 2 582 580 | 2 530 440 | – |
| | | 46 | 2 382 501 | 2 645 460 | Weis's run [19] |

At $(6,6,64)$ the pairwise count predicts a kernel of at least 280 at $k_\ell = 28$ and none at 27; the observed random kernel is 642, the difference being the higher-order universal relations. At $(7,7,128)$ with $(d,s) = (5,4)$ the two windows do *not* intersect: the first $k_\ell$ with a kernel is 33, one beyond completeness, which is why the two runs above found nothing and why larger codes need larger $(d,s)$. At TII-253, $k_\ell = 46$ is precisely the smallest complete value at which the pairwise relations exceed the excess, i.e. the smallest system with a kernel—which is the value Weis used. The criterion is only as good as the pairwise count, which Experiment 11.7 established for points in general position with $N$ small relative to $k_\ell$; in the three instances above $N/k_\ell$ is between 1.2 and 1.6, and they behave as predicted. It fails when $N \gg k_\ell$: at $(7,9,128)$ and $(7,10,128)$ with $k_\ell = 16$ and

$N = 63, 69$, and at $(7, 11, 128)$ with $k_\ell = 17$ and $N = 76$, the pairwise count predicts kernels of 336, 3 264 and 2 958, and all three systems had full rank—the pair relations are no longer independent when every element lies in most of the supports. The same happened at $(d, s) = (6, 5)$ and $(7, 6)$ on the small-$t$ instances within reach of a dense eliminator ($(6, 5, 64)$ with $k_\ell = 15, 16$, $(7, 4, 128)$ with $k_\ell = 15$; $N/k_\ell \approx 1.8$–$1.9$, and column weights near $k_\ell/2$, below $d + s - 1$, so that neither count applies): all complete, all of full rank. At the Classic McEliece parameter sets with $(d, s) = (7, 6)$, completeness (Remark 5.4) caps $k_\ell$ at $3t$ for $m = 12$ and at $2t$ for $m = 13$, so $N/k_\ell \geq 4$ there, outside the range in which the criterion has been validated. For $t \geq 6$ the smallest $(6, 5)$ instance in which the two windows intersect at the attack's ratio $N/k_\ell$ is $(7, 7, 128)$ at $k_\ell \approx 40$, i.e. $4 \cdot 10^6$ unknowns, a sparse-solver job.

### 11.6 Anatomy of the structural kernel

The number that would turn the kernel-dimension heuristic into a formula is the rank of the Hasse conditions on $\mathcal{J}_d^{\mathrm{ml}}$, i.e. the 650 of Experiment 11.15. Over $K$ the kernel is a statement about a single curve: by Theorem 8.4 its $\mathbb{F}_2$-dimension equals the $K$-dimension of

$$\mathcal{K}_K = \{H \in S_{5,K} : \ \mathrm{mult}_{e_a} H \geq 4 \text{ (28 unit points)}, \ \mathrm{mult}_{P_i} H \geq 4 \text{ (35 constrained points)}\},$$

forms of degree 5 with multiplicity 4 at 63 smooth points of the rational curve $\Gamma$ of degree 51 in $\mathbb{P}^{27}$ (multilinearity is the multiplicity-4 condition at the coordinate points, Lemma 3.4). This is a fat-points-on-a-curve problem, and the classical way to count such forms (Catalisano–Ellia–Gimigliano for rational normal curves) is to filter by the order of vanishing along the curve and count, at each order $j$, the admissible $j$-th normal data: the tuples $(h_S)_{|S|=j}$, $h_S = (\partial^{[S]} H)(f)$, in the polynomial spaces that Theorems 6.1 and 11.11 describe, with the forced zeros and the relations that come from $H \in I(\Gamma)^{(j)}$ (Euler's identity and the Hasse chain rule along $\Gamma$). We computed both sides.

**Experiment 11.17** (Filtration of the kernel by vanishing order)**.** For the $k_\ell = 28$ kernel of Experiment 11.15:

(i) Over $\mathbb{F}_2$, filtered by vanishing order along the union $X_\Gamma$ of the conjugate curves: $650 = 576 + 74 + 0$; i.e. 74 forms have all first derivatives in $\mathcal{J}_4$, none has all second derivatives in $\mathcal{J}_3$.

(ii) Over $K$, filtered by vanishing order along the single curve $\Gamma$: $650 = 134 + 516 + 0$; i.e. a 516-dimensional subspace vanishes to order exactly 2 along $\Gamma$, and no form vanishes to order 3.

(iii) The first-order data $(h_a)_{a \leq 28}$ of the kernel span a 134-dimensional space. The polynomial model allows more: each $h_a$ lies in $\mathcal{W}_4$ with zeros of order 3 at the other unit points, order 5 at $\alpha_a$ (the tangent-cone condition at $e_a$) and order 3 at the constrained points—a space of dimension $199 - 81 - 5 - 105 = 8$—and the $28 \cdot 8 = 224$ coefficients are subject to Euler's relation $\sum_a f_a h_a = 0$ (rank 33) and the tangent relation $\sum_a f'_a h_a = 0$ (rank 52, which implies Euler's). So the admissible first-order data form a space $N_1$ of dimension $224 - 52 = 172$, of which only 134 are realized by forms of degree 5.

(iv) At second order the gap is much larger: the admissible data $(h_{ab})$ have dimension at least $6804 - 3140 = 3664$ after all the pointwise relations we could derive, against the 516 realized.

(v) Across the six keys of Experiment 11.15 the first-order part is much more stable than the second: $\mathrm{gr}_1 = 134$ for five keys and 128 for one (seed 15), so $\delta_1 = 38$ or 44, while $\mathrm{gr}_2 = 294, 322, 364, 405, 427, 516$. No key had a third-order part. Six keys are too few to call 134 an invariant of the parameter set; what they support is a stable first-order profile with an occasional defect-one exception and a strongly key-dependent second-order part.

The lesson is that the kernel dimension is *not* an expected-dimension count of admissible normal data, not even at the first normal level: the restriction map from forms of degree 5 vanishing on $\Gamma$ to first-order normal data (sections of the conormal sheaf of $\Gamma$ twisted by $\mathcal{O}(5)$, with the prescribed zeros) is not onto, missing 38 dimensions here. We call

$$\delta_r := \dim N_r - \dim \operatorname{im}\big(\mathcal{K}_K \cap I(\Gamma)^{(r)} \to N_r\big)$$

the *$r$-th conormal projection defect* of $\Gamma$ (with the multiplicity data understood); here $\delta_1 = 38$, and $\delta_2 = \dim N_2 - 516 \geq 3148$ if the pointwise relations of (iv) are all the relations on second-order data. This is the precise sense in which "normal generation" fails for $\Gamma$: $\Gamma$ is the projection of the rational normal curve of degree $D$ from a $\mathbb{P}^{D-k}$ (the annihilator of $\mathcal{V}_1 \subset K[X]_{\leq D}$), it is not linearly normal, and the forms of $\mathbb{P}^{k-1}$ see only part of the normal data that the parameter line would allow. At order 0 the same phenomenon is exactly the question $\mathcal{V}_d = \mathcal{W}_d$, where the answer appears to be "everything is seen" (Conjecture 6.4); at order $\geq 1$ it is not. A formula for the structural kernel must therefore account for the projection, and the numbers 134 (stable across keys) and $\mathrm{gr}_2$ (key-dependent) and the defect 38 are the targets. Why the first-order normal data should be nearly independent of the key while the second-order data is not is a question in its own right: a natural hypothesis is that the first graded piece is governed by the tangent data common to all keys with these parameters, and the second by the projection centre of the particular key, i.e. by $\mathcal{V}_1$ itself; characterizing the first-order image independently of the key, and identifying the key-dependent obstruction at second order, is the next theorem to aim for. What we can say exactly is the shape of the answer: the kernel is the sum of a first-order part, bounded above by the admissible conormal data, and a second-order part that vanishes on the first infinitesimal neighbourhood of $\Gamma$; no higher order contributes. This is the remaining content of Open Problem 14.3.

## 12 Dirty-coordinate soundness

Completeness says sufficiently many clean retained coordinates force every relation to vanish on every point of $\Gamma$ (and of $X_\Gamma$). A distinguisher also needs the converse: a coordinate not on the curve should fail the test. This is a different statement, and it is where "random" enters.

**Proposition 12.1.** *Let $\mathcal{K} \subseteq W_d^{\mathrm{ml}}$ be a nonzero space of multilinear forms vanishing on $\Gamma$, and let $y^* \in \mathbb{F}_2^k \setminus \{0\}$.*

*(i) (Random column) If $y^*$ is uniform, then for each fixed nonzero $H \in \mathcal{K}$, $\Pr[H(y^*) = 0] \leq 1 - 2^{-d}$ [18]; for $\Delta$ $\mathbb{F}_2$-independent forms whose evaluations at $y^*$ are independent, the acceptance probability is at most $(1 - 2^{-d})^{\Delta}$.*

*(ii) (Curve point off the support) If $y^*$ is a scalar multiple of $\Phi(\beta)$ or of $\Phi^{\sigma^s}(\beta)$ for some $\beta \in \mathbb{P}^1(K)$, then $H(y^*) = 0$ for every $H \in \mathcal{K}$: the test cannot distinguish support points from other points of $X_\Gamma$.*

*(iii) (Low-weight error) If $y^* = y_j + e$ with $y_j$ a column and $e \neq 0$, then $H(y^*) = \sum_{|\nu| \geq 1} \partial^{[\nu]} H(y_j) e^\nu$, and rejection requires some Hasse derivative of $H$ at $y_j$ to be nonzero in a direction visible to $e$; if the order-$\prec$ $s$ jet of $H$ vanishes at $y_j$ this needs $\mathrm{wt}(e) \geq s$ in general.*

*Proof.* (i) is the multilinear Schwartz–Zippel bound used in [18]. (ii) is Theorem 5.2 and Lemma 8.3. (iii) is Taylor expansion. □

Item (ii) is the fundamental limitation of hold-out tests: they are *curve membership* tests, not *support membership* tests. Distinguishing the support inside $X_\Gamma$ needs the subfield condition $y \in \mathbb{F}_2^k$ together with the Frobenius structure of Section 8; this is exactly what the finishing steps of [19, 23, 20] do.

*Remark* 12.2. The Chevalley–Warning theorem gives a parity constraint on the number of common zeros of the forms in $\mathcal{K}$ when $\Delta d < k$; it does not improve the bound in (i) usefully, and (i) is the operative statement.

## 13 Quadratic relations and the rank-two attack

Let $\mathcal{S}_2(G)$ be the space of quadratic forms vanishing at all columns of $G$, that is, $\mathcal{S}_2(G) = \mathcal{K}_{2,1}(G; \{1, \dots, n\})$, the kernel of the catalecticant map from symmetric coefficient arrays to $\mathbb{F}_2^n$. The attack of [24] searches for rank-two elements in a related code of quadratic relations.

**Proposition 13.1.** *(i) $\mathcal{K}_{2,2}(G; I) \subsetneq \mathcal{K}_{2,1}(G; I)$ in general: a quadratic form vanishes to order two at $P$ if and only if $Q(P) = 0$ and $P$ lies in the radical of its polar form (Remark 3.1(i)).*

*(ii) No nonzero linear form vanishes to order two at any point: $\mathcal{K}_{1,2}(G; I) = 0$.*

*(iii) If $\ell_1, \ell_2$ are linear forms vanishing at all $P_i$, $i \in I$, then $\ell_1\ell_2 \in \mathcal{K}_{2,1}(G; I)$ has rank $\leq 2$; but when the columns indexed by $I$ span $\mathbb{F}_2^k$ there are no such $\ell$, so the rank-two elements of $\mathcal{S}_2(G)$ do not arise this way for Goppa codes of cryptographic size.*

*Caution* 13.2. It is tempting to conclude that $\mathcal{K}_{2,s} = \mathcal{K}_{2,1}$ for all $s$ over $\mathbb{F}_2$ because "the gradient of a quadratic form vanishes identically in characteristic two", so that the rank-two attack would coincide with the degree-two jet attack at every order. The premise uses the ordinary derivative of $Y_i^2$, not the Hasse derivative; (i) above is the correct statement. The relation between the rank-two locus of [24] (whose notion of rank in characteristic two should be taken from that paper and from [11]) and the jet filtration on $\mathcal{K}_{2,\bullet}$ is open.

**Open problem 13.3.** Does the rank-two locus of [24] correspond to a distinguished subvariety of the secant variety of $X_\Gamma$, and does the jet filtration $\mathcal{K}_{2,2} \subset \mathcal{K}_{2,1}$ isolate part of it?

## 14 The generic baseline: Alexander–Hirschowitz over $\mathbb{F}_2$

For general points over an algebraically closed field of characteristic zero, the Alexander–Hirschowitz theorem [4, 5] says that $N$ double points in $\mathbb{P}^n$ impose independent conditions on forms of degree $d$ (the evaluation map has maximal rank) except in the cases

$$d = 2,\ 2 \leq N \leq n; \qquad (n, d, N) \in \{(2, 4, 5), (3, 4, 9), (4, 3, 7), (4, 4, 14)\}.$$

Two issues arise over $\mathbb{F}_2$.

(a) *Characteristic.* The theorem is known to hold in characteristic $p > d$ (for $d > 2$), and the case $p = d$ is open [5]; characteristic 2 with $d \geq 3$ is not covered by the literature. A Schwartz–Zippel reduction to the algebraically closed case does not work either: over $\mathbb{F}_2$ a polynomial of degree $\geq 2$ vanishes at a uniformly random point with probability that need not be small. The statement remains open.

(b) *Rationality.* "General points" is not meaningful over $\mathbb{F}_2$; the correct question is about uniformly random rational points, and the universal relations of Section 3.4 show that even pure point evaluation never attains the characteristic-zero generic rank in degree $\geq 3$.

What the attacks actually use is a Fröberg-type heuristic for the dimension $\kappa$ of the kernel of the multilinear jet matrix after the characteristic-two order reduction; that the kernel has dimension at least $\kappa$ is Weis's Assumption 1, and it is the only assumption that enters his parameter choice [19, 26]. The generic baseline is therefore not a theorem but the main open problem of the subject:

**Open problem 14.1** (Kernel dimension)**.** For multilinear forms of degree $d$ in $k_\ell$ variables over $\mathbb{F}_2$ and Hasse conditions of order $< s$ at $N$ uniformly random rational points (after the order reduction), determine the rank of the jet matrix, or prove the Fröberg-type prediction with high probability. The classical results on fat points supported on rational normal curves [7, 8] are the nearest tools, but they are in characteristic zero and for full homogeneous forms.

**Open problem 14.2** (Universal relations)**.** Proposition 11.1 shows there are no universal relations among multilinear *forms*, and Theorem 11.4 gives the universal relations among the Hasse conditions at *one* point in closed form (none beyond the trivial $\sum_{l\leq s-2}\binom{k}{l}$ at a column of weight $\geq d+s-1$), and Theorem 11.6 gives $\binom{k}{2s-d-2}$ explicit relations for every pair of columns. What remains is to prove that these span the pairwise relations in general position (Conjecture 11.8) and to describe the relations among three or more points: the dependencies among the conditions at several points that hold for every configuration (Experiment 11.5(ii)), equivalently a closed form for the rank of the reduced-order multilinear jet matrix of a random configuration. This is the universal nuisance space of [23] and the local relations of [19]; subtracting it from the row count is what turns the Fröberg-type heuristic for the kernel dimension into a statement about the structural kernel alone.

**Open problem 14.3** (Exact rank)**.** Prove Conjecture 8.9 ($\rho_d = m(dD+1-t)-(m-1)(n+1)$, i.e. degree-$d$ forms restrict onto $H^0(X_\Gamma, \mathcal{O}(d))$), and extend the count of normal conditions of Proposition 8.13 and Experiment 8.14 from quadrics to degree $d$ and multiplicity $s$, so that $\operatorname{rank} E_{d,s,I}$ in the completeness regime becomes a closed formula in $(m,t,n,k,d,s,|I|)$. Theorem 8.8 gives the inequality in one direction. Experiment 11.17 turns the second half into a precise geometric question: determine the conormal projection defects $\delta_r$ of the projected curve $\Gamma = \Phi(\mathbb{P}^1) \subset \mathbb{P}^{k-1}$ in terms of the linear series $\mathcal{V}_1 \subset K[X]_{\leq D}$ that defines the projection $\mathbb{P}^D \dashrightarrow \mathbb{P}^{k-1}$, for $r = 1, 2$ (no higher order contributes). Then $\dim\ker E_{d,s,I} = \sum_{r\geq 1}(\dim N_r - \delta_r)$ with $N_r$ the admissible $r$-th order data of the polynomial model, and the kernel-dimension heuristic of the attacks becomes a formula. At $(6,6,64)$: $N_1 = 172$, $\delta_1 = 38$, and $\dim N_2 - \delta_2 = 516$.

# 15 A computational program

## 15.1 Experimental methods and reproducibility

Every instance in this paper is generated deterministically from $(m, n, t, \text{seed})$ by the same procedure, so that it can be regenerated exactly.

- *Field.* $K = \mathbb{F}_{2^m} = \mathbb{F}_2[x]/(p_m)$ with $p_3 = x^3+x+1$, $p_4 = x^4+x+1$, $p_5 = x^5+x^2+1$, $p_6 = x^6+x+1$, $p_7 = x^7+x+1$, $p_8 = x^8+x^4+x^3+x^2+1$; elements are written as integers via the polynomial basis ($x \mapsto 2$), and arithmetic uses log/antilog tables.
- *Support.* $L$ is the first $n$ elements of $K$ in this integer order, $L = \{0, 1, \ldots, n-1\}$ (the full field for $n = 2^m$).

- *Goppa polynomial.* Python's `random` module is seeded with the seed; monic polynomials of degree $t$ with uniformly random lower coefficients are drawn until one is irreducible over $K$. The main instances are: $(6, 64, 6)$, seed 11: $g = X^6 + 60X^5 + 23X^4 + 24X^3 + 57X^2 + 59X + 57$, $k = 28$; $(7, 128, 7)$, seed 11: $g = X^7 + 47X^6 + 53X^5 + 5X^4 + 77X^3 + 69X^2 + 54$, $k = 79$; $(5, 32, 3)$, seed 3: $g = X^3 + 23X^2 + 8X + 15$, $k = 17$.

- *Code and systematic form.* $\Gamma(L, g)$ is the $\mathbb{F}_2$-kernel of the $mt \times n$ binary expansion of the parity-check matrix $(\alpha_i^j / g(\alpha_i))_{j<t,\, i\le n}$; shortening to $k_\ell$ unit columns picks the $k - k_\ell$ shortened positions with the same module seeded with seed+1, and the systematic form is obtained by row reduction with the first available pivot. The constrained/held-out split is a seeded shuffle of the non-unit columns.

- *Linear algebra.* Ranks over $\mathbb{F}_2$ were computed with a dense bit-sliced Gaussian elimination in C (64-bit words, pivot on the lowest set bit) and, for every instance of at most $10^4$ unknowns, independently with a pure-Python eliminator; ranks over $K$ were computed with the `galois` package using the same field polynomials and cross-checked against a pure-Python implementation on the smaller instances. All counts of monomials, rows and single-point ranks are exact integers.

- *Code.* The scripts (field arithmetic, code generation, jet-matrix construction, the eliminators, and the scripts producing every table) will be made available with the paper.

Keys are counted as follows. The $(6, 6, 64)$, $k_\ell = 28$ system was run on six keys (seeds 11–16): six independent Goppa polynomials, no shortening ($k_\ell = k$), each with its own random choice of held-out column and its own random control. The $k_\ell = 20$ statistics use twenty keys (seeds 100–119): twenty independent Goppa polynomials, each with one independent random choice of the eight shortened positions, i.e. twenty independent (polynomial, shortening) pairs, not several shortenings of one key. The $\mathcal{V}_d = \mathcal{W}_d$ checks use twenty independent Goppa polynomials per parameter set and no shortening. The $(7, 7, 128)$ systems were run on one key each ($k_\ell = 31$ and 32 are two shortenings of the same seed-11 polynomial). All parameter sets use the full-field support $n = 2^m$ except $(4, 13, 2)$, $(4, 14, 2)$, $(5, 28, 3)$ and $(6, 60, 4)$ in Sections 6 and 8. Structural statements (membership of the kernel in $\mathcal{J}_d$, the gradient subcode, $r_\tau$, $\rho_d^{\mathrm{ml}}$) held on every key; the kernel dimension and its second-order part vary from key to key and are reported per key.

**Use of AI tools.** This research was assisted by AI models (large language models of Anthropic's Claude family). They were used to write and debug the scripts for the experiments (field arithmetic, code generation, construction of the jet matrices, the $\mathbb{F}_2$ eliminator and the post-processing of kernels), to run the computations and cross-check them against independent implementations, to test the stated formulas and conjectures against the computed data, and to assist in drafting, checking and editing proofs and text. The author has verified every mathematical statement, proof and reported number, and takes full responsibility for the content of the paper. No AI tool is an author of this work.

The experiments of Sections 6, 8 and 11.5 were run with a self-contained Python implementation (arithmetic in $\mathbb{F}_{2^m}$ for $m \le 8$, random irreducible $g$, binary kernel of the expanded parity-check matrix, Lagrange interpolation of code polynomials with the multipliers (2.3), exact rank over $K$ and over $\mathbb{F}_2$) together with a small bit-sliced $\mathbb{F}_2$ eliminator in C for the jet matrices of Section 11.5 (the $115\,640 \times 98\,280$ system took 25 minutes on one core and 1.2 GB). The next experiments, in order of value, are:

(1) **TII-253.** Reproduce Weis's run at $(m, t, n) = (8, 9, 214)$, $k_\ell = 46$, $(d, s) = (6, 5)$ with the three measurements of Experiment 11.15 (membership of the kernel in $\mathcal{J}_6^{\mathrm{ml}}$, the random control, the gradient subcode); the system has $9.4 \cdot 10^6$ unknowns and needs a sparse or block-Wiedemann eliminator, not the dense one used here. Remark 11.9 predicts a kernel dominated by universal nuisance with a small structural signal, and Conjecture 11.12 predicts $\dim \mathcal{J}_6^{\mathrm{ml}}$.

(2) **Conjecture 6.4** and **Conjecture 11.12** at larger $(m, t, d)$ and for shortened codes; a counterexample would be as informative as a proof.

(3) **A second structural kernel.** By Heuristic 11.16 no complete $(5, 4)$ instance of $(7, 7, 128)$ has a kernel; the next instance with one is $(7, 7, 128)$ at $(d, s) = (6, 5)$ (window $k_\ell < 48$, of the order of $10^7$ unknowns) or a code with $(m, t) = (6, 7)$ or $(7, 6)$ chosen so that the two windows intersect at a dense-eliminator size. Either would give a second value of $\delta_1$.

(4) **The kernel dimension itself.** Predict the $650 = 134 + 516$ of Experiments 11.15 and 11.17: determine which first-order normal data of $\Gamma$ are realized by forms of $\mathbb{P}^{k-1}$ (the defect $172 - 134 = 38$), and the second-order part; this is where the projection of the rational normal curve enters. Prove Conjecture 11.8, that the relations of Theorem 11.6 span the cross-point relations of a pair, and find the three-point relations of Experiment 11.7.

(5) **Hold-out increment.** For a fixed instance, record $h \mapsto \Delta(J)$ with $s_j = 1$ and with $s_j > 1$ on the held-out set, and verify Proposition 9.2.

(6) The jet-consistency relations of Proposition 10.3 at orders 2 and 3, and the rank of the system they impose across several held-out columns (Apon's bound).

(7) **Dirty coordinates.** Measure acceptance rates under the three models of Proposition 12.1.

All experiments should use reproducible seeds, record the actual $k$, and verify ranks by a second implementation. Small instances falsify conjectures; they do not establish asymptotic soundness.

# 16 Conclusion

The projective viewpoint turns public polynomial relations into fat-point conditions on a configuration of rational points. The alternant embedding supplies a hidden rational parameterization, and pulling relations back to the parameter line converts jet conditions into root multiplicities, which is the completeness mechanism of the Holdout/Hermite distinguisher. Two exact results describe where the binary structure enters. The pulled-back forms lie in the explicit space $\mathcal{W}_d = \{h : g^2 \mid (\pi^d h)'\}$ of codimension $t$, and appear to fill it; this is the structural reason for the $t$ saved zeros that lower the cost of every current attack. And the $\mathbb{F}_2$-rational forms vanishing on the hidden curve are exactly those vanishing on the union of its $m$ Frobenius conjugates, with the dimension given by Galois descent; this is the structural reason for the Frobenius ambiguity that every key-recovery route has to resolve.

On a real key of Classic-McEliece shape, $(m, t, n) = (6, 6, 64)$, the picture is now concrete (Section 11.5): over six keys the kernel of the GIJS/Weis system has dimension 428–650, every one of its forms vanishes on the hidden curve as completeness predicts, its gradients at the held-out column generate the $(m+1)$-codimensional subcode of Theorem 10.1, and the random controls have kernels of 448–642. The attack's kernel is, to within the key-to-key fluctuation, the universal nuisance space, whose one-point part is now known in closed form (Theorem 11.4) and whose pairwise part

is $\binom{k\ell}{2s-d-2}$ explicit relations per pair (Theorem 11.6); the structural signal is $\mathcal{J}_d^{\mathrm{ml}}$, whose dimension is Conjecture 11.12, cut by the normal conditions.

What remains open is the exact size of that signal: the rank of the normal conditions on $\mathcal{J}_d^{\mathrm{ml}}$, which all current cost estimates take from a Fröberg-type heuristic, and a proof of the nuisance counts. A proof of either would convert the conditional estimates of [19, 22] into theorems about the concrete security of binary Goppa McEliece; a counterexample would change the estimates. Both are stated precisely enough in Sections 3.5 and 14 to be attacked.

## References


[1] R. J. McEliece, *A public-key cryptosystem based on algebraic coding theory*, DSN Progress Report 42–44, JPL, 1978, pp. 114–116.

[2] V. D. Goppa, *A new class of linear error-correcting codes*, Problems of Information Transmission 6 (1970), 207–212.

[3] V. M. Sidelnikov and S. O. Shestakov, *On insecurity of cryptosystems based on generalized Reed–Solomon codes*, Discrete Math. Appl. 2 (1992), 439–444.

[4] J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Algebraic Geom. 4 (1995), 201–222.

[5] M. C. Brambilla and G. Ottaviani, *On the Alexander–Hirschowitz theorem*, J. Pure Appl. Algebra 212 (2008), 1229–1251.

[6] R. M. Wilson, *A diagonal form for the incidence matrices of $t$-subsets vs. $k$-subsets*, European J. Combin. 11 (1990), 609–615.

[7] M. V. Catalisano and A. Gimigliano, *On the Hilbert function of fat points on a rational normal cubic*, J. Algebra 183 (1996), 245–265.

[8] M. V. Catalisano, Ph. Ellia and A. Gimigliano, *Fat points on rational normal curves*, J. Algebra 216 (1999), 600–619.

[9] J.-C. Faugère, V. Gauthier-Umaña, A. Otmani, L. Perret and J.-P. Tillich, *A distinguisher for high-rate McEliece cryptosystems*, IEEE Trans. Inform. Theory 59 (2013), 6830–6844.

[10] R. Mora and J.-P. Tillich, *On the dimension and structure of the square of the dual of a Goppa code*, Des. Codes Cryptogr. 91 (2023), 1351–1372.

[11] A. Couvreur, R. Mora and J.-P. Tillich, *A new approach based on quadratic forms to attack the McEliece cryptosystem*, Asiacrypt 2023.

[12] E. Kirshanova and A. May, *Breaking Goppa-based McEliece with hints*, Information and Computation 293 (2023), 105045.

[13] H. Randriambololona, *The syzygy distinguisher*, Eurocrypt 2025. ePrint 2024/1193.

[14] T. Hemmert and A. Wiemers, *Distinguishing Goppa codes using higher-order vanishing*, Crypto 2026, Part IV, pp. 191–223. ePrint 2025/1661.

[15] T. Hemmert, *Key recovery for the McEliece cryptosystem using higher-order vanishing*, ePrint 2026/1339.

[16] M.-J. O. Saarinen, *HOVER: Higher-Order Vanishing Endomorphism Recovery*, ePrint 2026/1778; NordSec 2026.

[17] S. Kim and I. Kim, *Zeta-transform evaluation for higher-order vanishing key recovery*, arXiv:2609.26132, 2026.

[18] A. Ghoshal, Y. Ishai, A. Jain and N. Sun, *Quasipolynomial cryptanalysis of the McEliece cryptosystem (or: PIR meets McEliece)*, ePrint 2026/1630.

[19] S. A. Weis, *Improving GIJS key recovery for Classic McEliece*, ePrint 2026/1984 (version 0.5, October 2026).

[20] K. Vedenev, *Extending distinguishing to key recovery for subfield subcodes of GRS codes*, ePrint 2026/1747 (revised October 2026).

[21] D. Apon, *An algebraic-geometry lower bound against the ePrint:2026/1747 McEliece key-recovery attack*, ePrint 2026/1810.

[22] M.-J. O. Saarinen, *Bit operation cost of hold-out key-recovery attacks against Classic McEliece*, ePrint 2026/1786.

[23] M.-J. O. Saarinen, *Two-anchor Holdout/Hermite: solving the TII-254 McEliece key recovery challenge*, ePrint 2026/1986.

[24] P. Briaud, A. Lemoine, H. Randriambololona and J.-P. Tillich, *A heuristic subexponential attack on the McEliece cryptosystem*, ePrint 2026/1232.

[25] Bundesamt für Sicherheit in der Informationstechnik, *Notes on recent developments concerning Classic McEliece*, 29 September 2026 (announced 1 October 2026).

[26] R. Fernandez, *Structural cryptanalysis of Classic McEliece: assessment of claims, costs and evidence*, arXiv:2610.31196, October 2026.

[27] D. J. Bernstein et al., *Classic McEliece: conservative code-based cryptography*, https://classic.mceliece.org.