Hidden Curves and Short Duals: Cubic Jets, Dual Leaks, and the Limits of Masked-Code Encryption*

Yongge Wang
Department of Cybersecurity, UNC Charlotte
Charlotte, NC, USA

November 3, 2026

Abstract

Masked-code encryption hides a decodable code behind a transformation meant to make it look random. We study the two masks proposed as compact alternatives to Goppa codes—Wang's RLCE and the $(U \mid U+V)$ construction of UVW-KEM—over the decodable hidden codes they have been or could be applied to, and find that, among the families examined, none yields a compact trapdoor.

Over Reed–Solomon and algebraic-geometry codes the mask hides a *curve*, and curves are not hidden from jets. We explain the October 2026 key recovery of UVW-KEM (NGCC) and show it is not specific to UVW: for *hidden-pairing codes* we bound from below the dependencies among order-r jets of degree-d forms at the columns of any shortening—for cubics $2k_1 - k_2 - k_l + 2$ (UVW) and $2(n-k) - w - k_l + 2$ (RLCE), attained in every regime tested, and we prove that the structured part of the dependency space carries the pairing and the secret ratios in closed form. Every proposed UVW parameter set has the predicted cubic deficiency (key recovery has been demonstrated at UVW-KEM-128). RLCE's long-key regime $w \geq n-k$, out of reach of the square-code attack for eight years, falls at degree three: on a full-size Level-1 parameter set we recover an equivalent key from the public key alone and decrypt fresh ciphertexts. HermitianRLCE is distinguished the same way. A forcing law, a master count and a redundancy law make the *jet degree* an invariant of hidden-pairing codes: within that model, each unit of structural depth costs a unit of redundancy.

Without a curve the masks fail differently: both *preserve the dual* of the hidden code, at most doubling the weight of every dual word. polarRLCE at rate $\leq 2/3$ leaks dual words found in 47–94 bits; at higher rates no certified decoding radius supports 143-bit security up to length 16384, and at 32768 the key is 58 MiB, 234× Classic McEliece. LDPC codes leak their sparse checks, MDPC codes are better left unmasked, and UVW carries a matched-pair subspace for every hidden code. Reed–Muller-heavy codes under RLCE resist every attack here at 4 MiB keys, but only if a decoder at 3.4× the bounded-distance radius exists.

Keywords: Code-based cryptography · masked codes · Reed-Solomon codes · polar codes · LDPC codes · higher-order vanishing · jets · dual leak · UVW-KEM · RLCE.

*This research was conducted with the assistance of Anthropic's Claude Fable 5.1 AI model and the AETHER IDE (`https://amalithlab.github.io`). The scripts used for cryptographic analysis are available at `https://github.com/yonggewang/RLCE/tree/master/scripts`.

1 Introduction

Masking a generalized Reed-Solomon (GRS) code so that its public generator looks random has been attempted many times, because an $[n, k]$ GRS code corrects $\lfloor (n-k)/2 \rfloor$ errors with the smallest possible redundancy and gives McEliece-type keys far smaller than binary Goppa codes. Every masking proposed before 2018 fell to a *structural* attack that ignores the decoding problem: Sidelnikov-Shestakov [26] for plain GRS codes, and the *square-code* distinguisher [14, 9] for Wieschebrink's random columns [32], the BBCRS masking [1], and the short-key parameters of Wang's RLCE [27, 28], recovered by Couvreur, Lequesne and Tillich (CLT) [13]. The square-code attack exploits that the Schur square of a k-dimensional GRS code has dimension $2k-1$ instead of the generic $\binom{k+1}{2}$, and works on *shortenings* of the public code until the GRS part is exposed. Two families survived it. RLCE with $w \geq n-k$ inserted random columns, which CLT proved out of reach of their method, and the $(U{+}V, U{+}W)$ construction of Yang and Zhang [33], which masks the RS code *additively* so that no coordinate projection of the public code is GRS at all. The latter is the basis of UVW-KEM, submitted to the first round of China's Next-generation Commercial Cryptographic Algorithms program (NGCC) in September 2026 [23].

Higher-order vanishing. In 2026 a different algebraic tool reached the binary Goppa codes of Classic McEliece: instead of asking which *quadratic* forms vanish on the columns of a (shortened) public generator, one asks which forms of higher degree vanish there *to higher order*, i.e. together with their derivatives. The syzygy distinguisher [24], the quasi-polynomial distinguisher and key recovery of Ghoshal, Ishai, Jain and Sun [15], its improvement by Weis [30], and the higher-order vanishing attacks of Hemmert and Wiemers [21, 19] and Saarinen [25] all belong to this line, which led the German BSI to advise against Classic McEliece for new developments [6]. For Goppa codes the mechanism is the subfield subcode structure and its Frobenius action. Three different objects are being counted in this literature, and we keep them apart throughout. The syzygy distinguisher counts *relations among the quadratic generators* of a shortened dual code (graded Betti numbers); the Goppa attacks count forms of degree $d = \Theta(\log n)$ vanishing to high order at the columns, and read the secret off a *Frobenius orbit* of hidden GRS blocks that appears after scalar extension to $\mathbb{F}_{2^m}$ (HOVER extracts it from the endomorphism algebra of the vanishing tensors); this paper counts forms of degree *three* whose first derivatives vanish at the columns, over the public field and with no subfield, and reads the secret off *incidences*—public columns whose span meets a hidden curve. The first mechanism is combinatorial-algebraic, the second needs a subfield subcode, the third needs a curve of low degree and a decoder that unmasks locally; none of the three implies another, and a code may be immune to two of them and fall to the third (Sect. 9.6).

The UVW break. On 26 October 2026 the NGCC audit site ngcc.dev published finding kem-38-8 on UVW-KEM [22]: following a forum post of L. Wang that applied the method of [15, 30] to UVW, Saarinen counted the *cubic forms whose gradient vanishes at every column* of the public code shortened to dimension $k_l \approx \sqrt{6m}$ and found $k_2 - k_l + 2$ more than for random columns (800 against 636 at UVW-KEM-128); he then read the secret pairing and scaling ratios off the row dependencies of the jet system, subtracted the scaled twins to obtain a $[430, 215]$ code with square dimension 429, ran Sidelnikov-Shestakov, and decapsulated fresh ciphertexts. UVW-KEM-128 is broken in practice. The report is operational: it states the count as measured, verifies it with controls, and gives the recipe. It does not say *why* the count is $k_2 - k_l + 2$, why the dependency blocks have their shape, what happens at other parameters, or whether other masked-GRS schemes are affected. Those are the questions of this paper.

Beyond curves. If the hidden curve is the weakness, the obvious repair is to hide a decodable code that has none. This was the motivation of polarRLCE [43], which places a binary polar code under the RLCE mask, and it applies equally to LDPC and MDPC codes and to the UVW mask. We therefore study the whole matrix {GRS, AG, polar, LDPC/MDPC} × {RLCE, UVW}. The jet does not reach the curve-free codes, but the masks fail on them anyway, because both act on the primal code and leave the dual of the hidden code in place.

Our results. We show that the attack is a consequence of one geometric fact shared by all masked-GRS schemes whose decoder undoes the mask coordinate-pair by coordinate-pair, and that the curve-free alternatives fail through their duals.

1. **Hidden-pairing codes (Sect. 3).** We define a public code to have a *hidden pairing* when its columns are partitioned into secret pairs (and possibly singletons) such that the span of each pair contains a point of a hidden rational normal curve—a column of a secret GRS code—and show that UVW1, UVW2, RLCE and the $(U \,|\, U{+}V)$ constructions with a Reed-Solomon V are instances. We condense the structural anatomy of these codes (the decoder's fold, the hidden dual subcode, the incidence test and the reduction from the pairing to an equivalent key), which is the back end of every attack below.

2. **The jet dependency theorem (Sect. 4).** For the public code shortened at $l = k - k_l$ positions, let n_0, n_1, n_2 be the numbers of pairs with $0, 1, 2$ shortened coordinates and s_0, s_1 the numbers of surviving and shortened singleton curve points. The system of order-r jets of degree-d forms at the surviving columns has at least

$$(2r-d)^+ n_0 + r\,(n_1 + s_0) + d\,(n_2 + s_1) - d(k_2 - 1) - 1$$

   row dependencies supported on its structured rows; for the cubic gradient system, where random points impose independent conditions, these are dependencies beyond the generic rank. The proof is two lines: a form vanishing to order r at both ends of a line vanishes, with its first $2r-d-1$ derivatives, along the whole line (Lemma 4.2), so each pair contributes Hermite data at its curve point, and Hermite data at more points than the composed polynomial has coefficients are dependent. For cubics and gradients this is $2k_1 - k_2 - k_l + 2$ for UVW and $2(n-k) - w - k_l + 2$ for RLCE; at the NGCC parameters it is the reported $k_2 - k_l + 2$. The bound is attained in every non-degenerate instance we generated (Sect. 8).

3. **The recovery stage, explained (Sect. 5).** The proof of the dependency theorem is constructive, and we prove that the dependencies it produces form a subspace K_s of the left kernel of exact dimension $T - \dim \mathcal{V}$, under a non-degeneracy condition on the shortening that we state (Lemma 4.3). On K_s the dependency block of a surviving twin has a closed form: $\lambda_i = \delta_i^{-2}\big(x_i/(3\delta_i) - x_j/\delta_j\big)$ for UVW2, whence the blocks are rank one, twins share a coefficient vector, and the secret ratio is $\delta_j/\delta_i = -3a/b$ when $\lambda_i = a\,x_i + b\,x_j$; for RLCE the block direction gives the hidden GRS column itself (Theorem 5.1). The read-off needs the left kernel to equal K_s (*exactness*); we show that a candidate pair must pass a consistency test with $k_l - 2$ conditions, so a wrong kernel can hide pairs but not produce a false one, that the final key is verified by a GRS test, and we measure where exactness fails: beyond the Veronese threshold of Remark 9.5 the read-off recovers nothing, and recovery there is open; that threshold needs more surviving curve-point columns than $\approx 1.4\,k_2$, against 0.33–0.69 of k_2 at every proposed parameter set (Proposition 5.3). We verified the formulas on every instance, and we carry the recovery through end to end on RLCE (item 4).

4. **Consequences for parameters (Sect. 6).** Every UVW parameter set—the six of [33] and the three NGCC sets—is distinguished at degree 3 by a linear system of 1.8×10^4 to 1.9×10^5 unknowns; the unbalanced family $k_1 > k_2$ proposed for smaller keys is the *most* exposed, and in fact already falls to the square code (item 5). Key recovery with decapsulation of fresh ciphertexts has been demonstrated at UVW-KEM-128 [22]; at the larger sets the distinguisher is proved and the recovery stage is the same computation at larger size, but has not been run, and we say so. No split $k_1{:}k_2$ escapes the hierarchy: order-$(d-1)$ jets of degree d reach any $k_2 < m(1-1/d)$, and the Guruswami-Sudan radius ties a large k_2/m to a small error weight; the degree, and with it the exponent of the attack, grows with k_2/m, so this is a family of polynomial attacks indexed by the rate, not one uniform polynomial attack, and we separate the regimes accordingly. For RLCE the long-key regime $w \geq n-k$ falls at degree 3: the 2017 NIST sets ID 0, ID 2 and ID 4 give predicted excesses 116, 183 and 478, confirmed at full size on ID 0 and on a rate-0.87 Level-1 set over $\mathbb{F}_{2^{10}}$; on the latter we run the complete attack from the public key alone—jet rounds to recover most pairs with their ratios, a square-code stage on the folded code for the rest (the count says when CLT's regime is reached), fold, Sidelnikov-Shestakov—and decrypt a fresh ciphertext of weight $t = 57$ in about half an hour on two cores, including a run in which the attack process reads only a public key and ciphertext written by a separate key-generation program (Sect. 8.1). $w = c(n-k)$ random columns only raise the required degree to about $c+2$.

5. **The jet degree, an invariant (Sect. 9).** We define the *jet degree* of a code and prove the *forcing law* (Theorem 9.2): the order-r jets at the j columns of a tuple force the hidden point in their span iff $jr > (j-1)d$, with $(jr-(j-1)d)^+$ Hermite orders forced—the columns are a projective frame, the fat-point ideal is monomial, and the question is a count of exponent vectors, in every characteristic. The dependency theorem for any widths follows, the *master count* $\Delta_{d,r} = \sum_\ell (j'_\ell r - (j'_\ell - 1)d)^+ - d(k_2-1) - 1$ (Theorem 9.3), which on the hierarchy $r = d-1$ collapses to $\Delta_{d,d-1} \geq d(\rho+1) - R - k_l - 1$ (ρ, R the hidden and public redundancies), with equality once d exceeds the mask width. Its degree-two row is the square-code attack: it reproduces the CLT threshold $w < n-k$ for RLCE and predicts that the three unbalanced sets of [33] fall to the plain shortened square code, confirmed at full size (excess 14, 28, 45, as predicted). A width-3 masking is blind at degrees 2 and 3 and falls at $(4, 3)$ by exactly the count; a UVW key with a tiny U-part escapes degree 3 and falls at 4: "$\delta = j+1$" is a statement about balanced rates, and the count says the rest.

6. **The redundancy law (Theorem 9.8).** A hidden-pairing code blind at degree d has $R \geq d(\rho+1) - k_l - 1$; a hierarchy jet degree D costs public redundancy $(D-1)\rho$ up to the shortening slack, tightly (UVW, RLCE: $D = 3$, $R = 2\rho$). This is a lower bound within the hidden-pairing model—decoders that unmask by a linear map on a few public coordinates per hidden coordinate—not an impossibility theorem for every conceivable masking; within that model there is no compact, jet-resistant GRS trapdoor. We then pass the masked-GRS literature through the count (Table 11): Berger-Loidreau subcodes, Wieschebrink's random columns, BBCRS, RLCE, UVW, three $(U \,|\, U{+}V)$ variants, twisted RS, subspace subcodes, Goppa. One by-product: the Reed-Solomon $(U \,|\, U{+}V)$ codes of Márquez-Corbella and Tillich [8] lie on a rational normal scroll, so every shortening has square dimension $\leq 3(k_l - 1)$ and the construction has jet degree 2 for all parameters. Another: the count extends verbatim from rational to algebraic-geometry curves (the composed space becomes $L(d\alpha P)$, dimension $d\alpha + 1 - g$), so HermitianRLCE [29] is distinguished at degree three with provable excess $2(n-k) - w - k_l + 2 - 2g$, positive at all proposed parameters and confirmed in characteristic

two (Sect. 9.3).

7. **Why square-code analyses saw nothing (Sect. 7).** We prove, for the record, that Schur-square dimension tests of pair-constrained subcodes certify a defect only when all but at most one of the hidden pairs are already correct, the statement on which the square-code literature and the first structural audit of UVW-KEM relied. It is true and irrelevant: jets count *relations* among the derivatives of the ambient shortened code, not dimensions of subcodes cut out by guesses.

8. **Dual leaks and the limits of masking (Sect. 10).** Both masks preserve the dual of the hidden code: a dual word $\mathbf{h}$ becomes a public dual word of weight $\mathrm{wt}(\mathbf{h}) + Y \le 2\,\mathrm{wt}(\mathbf{h})$ under RLCE for every w, and $(\mathbf{h}, -\mathbf{h})$ of weight exactly $2\,\mathrm{wt}(\mathbf{h})$ under UVW for every k_1 (Lemma 10.1). A polar code is short on one side, $d \cdot d^{\perp} \le 2n$ (Proposition 10.4). polarRLCE leaks its affine minimum-weight words at small w, a leak the Schur-square analysis of [43] does not see, and closes it only by inserting enough columns; at rate $\le 2/3$ it then leaks dual words of weight $\le 2d^{\perp}$, $d^{\perp} \le 64$, found in 47–94 bits at parameters meeting their ISD targets (Proposition 10.6). At higher rates the dual is heavy but the decoder is weak: with the successive-cancellation failure probability bounded by the Tal-Vardy degrading construction, no information set up to $n = 16384$ supports a 2^{143} dual-word cost and a 2^{143} message-recovery cost at once, and at $n = 32768$ the cheapest that does has a 58 MiB key, 234× Classic McEliece (Proposition 10.7, Table 15). Resisting the dual leak forces the information set to contain a Reed-Muller code (Corollary 10.5); RM$(6, 13)$ and RM$(7, 13)$ under RLCE resist every attack here at 3.7–4.0 MiB keys, 15× Classic McEliece, if a decoder at 3.4–3.8 times the bounded-distance radius exists (Table 16). LDPC codes leak their sparse checks (found in 35–63 bits); an MDPC code under RLCE is dominated by the same code unmasked; and UVW with any hidden code carries a matched-pair subspace of codimension k_1 (Theorem 10.8). Table 18 classifies every instantiation; a compact masked trapdoor would need a decodable code with a heavy primal, a heavy dual, and no curve or subfield structure.

Table 1 classifies the claims. Everything measured is reproducible from the scripts described in Sect. 8.2.

Relation to the Goppa attacks. The mechanism here is not the one that breaks Goppa codes: no subfield, no Frobenius, no catalecticant. The GRS code lives over the public field and what the jets see is the *twin line*, two public columns whose span contains a curve point; that is why the attack transfers to RLCE and works in characteristic 2 and odd characteristic alike.

Disclosure. The UVW attack is public [22]; this paper adds its theory. The RLCE result (distinguisher and end-to-end key recovery), the square-code result on the unbalanced UVW sets, the scroll result on $(U \,|\, U{+}V)$-RS codes, and everything in Sect. 10 are new. Concurrently and independently, Lieb, Mazumder and Schaller [39] attack GRS codes under a weight-two *secant* mask (each public column a combination of two curve points) with a Schur-cube distinguisher at very high rate; their mask is outside the hidden-pairing model and their regime is disjoint from ours (Sect. 9.6).

2 Preliminaries

q is a prime power and $\mathbb{F}_q$ the field with q elements; vectors are rows unless they are columns of a generator matrix. An $[N, k]$ code $\mathcal{C} \subseteq \mathbb{F}_q^N$ has a generator $G \in \mathbb{F}_q^{k \times N}$ of full row rank; we

Table 1: Status of each claim in this paper.

Claim	Where	Status
UVW1, UVW2, RLCE, $(U\|U{+}V)$ with RS V have hidden pairings	Prop. 3.2	Proved
Jet dependency lower bound	Thm. 4.4	Proved
Structured dependencies independent; $\dim K_s = T - \dim \mathcal{V}$	Lemma 4.3	Proved under a stated non-degeneracy condition
Generic baseline: line relations $\binom{n'}{2}(2r{-}d{-}1)^+$	Lemma 2.2	Proved (lower bound); exact in tests
Block structure, ratio and curve-direction formulas on K_s	Thm. 5.1	Proved; verified
Recovery sound without exactness (no false pairs; verified output)	Prop. 5.3	Proved (soundness); completeness per round heuristic
UVW: all nine parameter sets distinguished at degree 3	Table 2	Proved count; toy scale here; full size at NGCC-128 in [22]
UVW key recovery	Sect. 5, [22]	Demonstrated at UVW-KEM-128 [22] and on toys here; not run at 256/512
No $k_1{:}k_2$ split escapes the hierarchy	Cor. 4.8, Sect. 6	Proved (bound); degree, hence exponent, grows with k_2/m
RLCE long keys distinguished at degree 3	Cor. 4.6, Table 3	Proved count; measured at full size (ID 0, R-1)
RLCE long keys: end-to-end key recovery and decryption	Sect. 8.1, Table 8	Demonstrated at full size (R-1), two keys, one in strict public-key-file mode; 25–35 min on two cores
Jet-to-square-code handover at $w' < n - k - k_l + 1$	Sect. 8.1	Proved (count); measured (R-1)
Schur-filtration certificates need all but ≤ 1 pair	Thm. 7.1	Proved
Forcing law: tuple forces hidden point iff $jr > (j-1)d$	Thm. 9.2	Proved $\forall j, d, r$, any characteristic; verified 210 triples
Master count, hierarchy closed form	Thm. 9.3, Cor. 9.4	Proved (lower bound); exact in all non-degenerate tests
Square code breaks unbalanced UVW sets (128b/192b/256b)	Sect. 9.4	Proved count; measured at full size
Width-3 mask: blind at $2, 3$, falls at $(4, 3)$	Sect. 9.4	Proved count; measured (toy)
Redundancy law $R \geq d(\rho + 1) - k_l - 1$	Thm. 9.8	Proved (hidden-pairing model); tight
$(U\|U{+}V)$-RS lies on a scroll, $\dim \mathcal{C}_{(Z)}{}^{\star 2} \leq 3(k_l - 1)$	Sect. 9.6	Proved; measured (equality)
Landscape of masked GRS codes	Table 11	Count where it applies; cited otherwise
AG master count: composed space $L(d\alpha P)$, dim $d\alpha{+}1{-}g$	Thm. 9.6	Proved (lower bound)
HermitianRLCE distinguished at degree 3	Cor. 9.7, Table 9	Proved (bound > 0); measured (char 2)
Dual preservation under RLCE and UVW	Lemma 10.1	Proved; leak law measured
Polar codes: $d \cdot d^\perp \leq 2n$; min-weight counts	Prop. 10.4	Proved; enumeration (94 codes)
polarRLCE primal leak at small w; closed by inserted-column weight	Table 12	Proved (dimension, weight); measured
polarRLCE at rate $\leq 2/3$ falls to its dual	Prop. 10.6	Proved (i); computed (ii), (iii)
Dual leak forces $\mathcal{C}(I) \supseteq \mathrm{RM}(r_0, m)$; width-$b$ masks leak at $\leq b\,\mathrm{wt}(\mathbf{h})$	Cor. 10.5, Rem. 10.2	Proved
SC failure bounds of polar codes on BSC(p)	Table 14	Computed (rigorous upper bound, degrading construction)
polarRLCE excluded at every rate for $n \leq 16384$; ≥ 58 MiB at $n = 32768$	Prop. 10.7, Table 15	Computed (Prange filter; Stern re-pricing of survivors)
RM-heavy RLCE: dual ≥ 191 bits; needs a decoder at $3.4\times$ BD radius	Table 16	Computed (Stern family); open (decoder)
LDPC leaks under RLCE and UVW; MDPC dominated	Sect. 10, Table 17	Measured; computed
UVW matched-pair subspace, any hidden code	Thm. 10.8	Proved; measured
Landscape of masked codes	Table 18	Classification of known families

write $\mathbf{g}_1, \ldots, \mathbf{g}_N \in \mathbb{F}_q^k$ for its columns. For $Z \subseteq [N]$ the *shortening* $\mathcal{C}_{(Z)}$ is the code of codewords vanishing on Z, with Z deleted. If the columns $\{\mathbf{g}_c\}_{c \in Z}$ are linearly independent, $\mathcal{C}_{(Z)}$ has dimension $k - |Z|$, and a generator of it is LG restricted to $[N] \setminus Z$, where the rows of $L \in \mathbb{F}_q^{(k-|Z|)\times k}$ span $R_Z := \{\mathbf{r} : \mathbf{r} \cdot \mathbf{g}_c = 0 \ \forall c \in Z\}$. We call $\Pi_Z : \mathbb{F}_q^k \to \mathbb{F}_q^{k-|Z|}$, $\mathbf{g} \mapsto L\mathbf{g}$, the *shortening projection*; the columns of the shortened generator are $\mathbf{x}_c = \Pi_Z \mathbf{g}_c$, and $\mathbf{x}_c = \mathbf{0}$ for $c \in Z$.

Schur products. $\mathbf{x} \star \mathbf{y} = (x_1 y_1, \ldots, x_N y_N)$ and $\mathcal{C}^{\star 2} = \mathrm{span}\{\mathbf{x} \star \mathbf{y} : \mathbf{x}, \mathbf{y} \in \mathcal{C}\}$, of dimension $\min(N, \binom{k+1}{2})$ for a random code with probability $1 - o(1)$ [7] and $\min(N, 2k - 1)$ for a GRS code. The gap is the lever of every square-code attack.

GRS codes and the rational normal curve. For distinct $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_q$ and nonzero $\beta_1, \ldots, \beta_m$ the $[m, k_2]$ GRS code has generator with columns $\beta_\ell\, \gamma(\alpha_\ell)$, where

$$\gamma(\alpha) = (1, \alpha, \alpha^2, \ldots, \alpha^{k_2 - 1})^T \in \mathbb{F}_q^{k_2}$$

parametrizes the *rational normal curve* of degree $k_2 - 1$. The dual of an $[m, k_2]$ GRS code is a GRS code of dimension $m - k_2$ on the same support; Sidelnikov-Shestakov [26] recovers $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ from any generator in $O(m^3)$ operations; and the Guruswami-Sudan algorithm [16] list-decodes up to $m - \sqrt{mk_2}$ errors, which is the radius UVW uses.

Forms, jets and the Euler identity. $\mathrm{Sym}^d(V^*)$ denotes the space of homogeneous polynomials (forms) of degree d in k_l variables, of dimension $N_d := \binom{k_l + d - 1}{d}$. For $f \in \mathrm{Sym}^d$ and $s \geq 0$, $D^s f(\mathbf{x})$ is the symmetric s-linear form of s-th partial derivatives at $\mathbf{x}$; f *vanishes to order* r at $\mathbf{x}$ if $D^s f(\mathbf{x}) = 0$ for all $s < r$. Since the coordinates of $D^s f$ are forms of degree $d - s$, Euler's identity $\sum_t x_t\, \partial_t (D^s f) = (d - s)\, D^s f$ shows that, whenever $d, d - 1, \ldots, d - r + 2$ are nonzero in $\mathbb{F}_q$, vanishing of all derivatives of order *exactly* $r - 1$ at $\mathbf{x}$ already forces vanishing to order r. For $(d, r) = (3, 2)$ the hypothesis is $3 \neq 0$, i.e. q not a power of 3; it holds for every field in this paper, including characteristic 2.

Definition 2.1 (jet system of a point set)**.** For points $\mathbf{x}_1, \dots, \mathbf{x}_{n'} \in \mathbb{F}_q^{k_l}$ and integers $d \geq r \geq 1$, the (d, r) *jet system* $\mathsf{J}_{d,r}$ is the linear map $\mathrm{Sym}^d \to \bigoplus_{c=1}^{n'} \mathrm{Sym}^{r-1}$, $f \mapsto \big(D^{r-1} f(\mathbf{x}_c)\big)_c$, written as a matrix with $n'\binom{k_l+r-2}{r-1}$ rows indexed by (point, derivative multi-index) and N_d columns indexed by monomials. Its kernel is the space of degree-d forms vanishing to order r at all points; its *row dependencies* are the vectors $\Lambda = (\lambda_c)_c$, $\lambda_c \in \mathrm{Sym}^{r-1}(V)$, with $\sum_c \langle \lambda_c, D^{r-1} f(\mathbf{x}_c) \rangle = 0$ for all f. For $r = 2$, $\lambda_c \in \mathbb{F}_q^{k_l}$ is a direction and the dependency reads $\sum_c D_{\lambda_c} f(\mathbf{x}_c) = 0$: a vanishing combination of directional derivatives.

The number of row dependencies is rows $-$ rank $\mathsf{J}_{d,r}$ and the kernel dimension is $N_d - \operatorname{rank} \mathsf{J}_{d,r}$; the two differ by the constant rows $- N_d$. A *distinguisher* compares either with its value on random points.

Lemma 2.2 (generic baseline)**.** *For any n' points in general position, $\mathsf{J}_{d,r}$ has at least $\binom{n'}{2}(2r-d-1)^+$ row dependencies: for each pair of points, the restriction of a degree-d form to the line through them is a binary form with $d+1$ coefficients, on which order-r vanishing at both points imposes $2r$ conditions. In particular the baseline is zero whenever $2r \leq d+1$, which includes the cubic gradient system $(3, 2)$ and the quartic gradient system $(4, 2)$, and it is $\binom{n'}{2}$ for $(4, 3)$.*

Proof. Parametrize the line as $\mathbf{x}_a + t(\mathbf{x}_b - \mathbf{x}_a)$. The $2r$ functionals "derivatives of order $< r$ at $t = 0$ and $t = 1$ of the restricted binary form" are linear in the $d+1$ coefficients of a univariate polynomial of degree $\leq d$, hence satisfy at least $2r - d - 1$ relations; each relation is a combination of rows of $\mathsf{J}_{d,r}$ supported on the two points. Relations for different pairs are independent as long as $2n' \leq \binom{k_l+r-2}{r-1}$ and the points are in general position: the relation of the pair (a, b) restricted to the rows of a is a form of degree $r-1$ in $\mathbf{x}_a, \mathbf{x}_b$ with a nonzero $\mathbf{x}_b^{\otimes(r-1)}$ term, and these are linearly independent across b. □

In all our experiments the baseline of Lemma 2.2 is exact on random codes when the system is underdetermined ($N_d \geq$ rows): the rank is rows $- \binom{n'}{2}(2r - d - 1)^+$. For cubics with double points in characteristic zero this is the Alexander-Hirschowitz theorem [5], whose exceptional cases do not include degree 3 in five or more variables; over finite fields we verify it on random codes at every size we use. We take it as the reference value and call *excess* the number of row dependencies beyond it.

Hermite interpolation. Let $\mathcal{P}_D$ be the polynomials of degree $\leq D$ in one variable. For distinct nodes $a_1, \dots, a_u$ with multiplicities $m_1, \dots, m_u$, the functionals $F \mapsto F^{(s)}(a_i)$, $0 \leq s < m_i$, are linearly independent on $\mathcal{P}_D$ as long as $\sum m_i \leq D + 1$, and span its dual when $\sum m_i = D + 1$; this holds in any characteristic for Hasse derivatives ($F^{(s)}(a)$ read as the coefficient of $(X - a)^s$ in the expansion of F at a, so that $F^{(s)}(a) = 0$ for $s < m$ is equivalent to $(X - a)^m \mid F$), and for ordinary derivatives whenever $m_i \leq \operatorname{char} \mathbb{F}_q$, which covers every system in this paper. Consequently, if T such functionals are given with $T > \dim \mathcal{V}$ for a subspace $\mathcal{V} \subseteq \mathcal{P}_D$, their restrictions to $\mathcal{V}$ satisfy at least $T - \dim \mathcal{V}$ independent linear relations.

3 Hidden-Pairing Codes

3.1 The schemes

UVW. Let $n = 2m$. Yang and Zhang [33] take a random $[m, k_1]$ code U with generator G_U, a random $[m, k_2]$ code W with generator G_W, a secret $[m, k_2]$ RS code with generator G_{RS} (columns

Table 2: UVW parameter sets [33, 23] ($n = 2m$, $k = m$) and the cubic-jet system that distinguishes each (k_l the smallest shortening dimension with $N_3 \geq n'k_l$, $n' = m + k_l$; excess the bound $2k_1 - k_2 - k_l + 2$ of Corollary 4.5).

Set	q	n	k_1	k_2	w	k_l	rows $n'k_l$	cols N_3	excess
128a	401	800	200	200	116	51	23 001	23 426	151
128b	337	672	188	148	112	47	18 001	18 424	183
192a	601	1200	300	300	172	62	41 044	41 664	240
192b	499	996	279	219	167	57	31 635	32 509	284
256a	797	1592	398	398	230	71	61 557	62 196	329
256b	677	1352	379	297	226	66	48 972	50 116	397
NGCC-128	433	860	215	215	116	53	25 599	26 235	164
NGCC-256	857	1708	427	427	232	74	68 672	70 300	355
NGCC-512	1709	3412	853	853	463	103	186 327	187 460	752

$\gamma(\alpha_\ell)$, $\ell \leq m$), $G_V = G_{\mathrm{RS}} + G_W$, an invertible S, and a permutation P:

$$G_{\mathrm{sk}} = \begin{pmatrix} G_U & G_U \\ G_V & G_W \end{pmatrix}, \qquad G_{\mathrm{pk}} = S\, G_{\mathrm{sk}}\, P, \qquad k = k_1 + k_2. \tag{1}$$

A codeword is $(\mathbf{u} + \mathbf{v} \,\|\, \mathbf{u} + \mathbf{w})$; the decoder subtracts the halves, $\mathbf{c}_{11} - \mathbf{c}_{12} = \mathbf{r}_2 G_{\mathrm{RS}} + (\mathbf{e}_1 - \mathbf{e}_2)$, and list-decodes the RS code. *UVW2* replaces P by a monomial matrix $\Pi\Delta$, so public column c is $\delta_c S G_{\mathrm{sk}}[:, \sigma(c)]$ with a secret nonzero δ_c; this is the variant of the NGCC candidate UVW-KEM, whose specification fixes a public RS support $\alpha_\ell = \ell$, uniform errors of weight w and balanced $k_1 = k_2 = m/2$ [23]. Table 2 lists the six parameter sets of [33] (a balanced and an unbalanced family, $k_1 > k_2$) and the three NGCC sets; all have $k = m$.

RLCE. Wang's RLCE [27, 28] starts from an $[n, k]$ GRS code over $\mathbb{F}_q$, $q = 2^\mu$, with generator G_s, appends w random columns $\mathbf{r}_0, \ldots, \mathbf{r}_{w-1}$, each placed next to one of the last w GRS columns, and mixes every such pair by a secret $A_i \in \mathrm{GL}_2(\mathbb{F}_q)$ with all four entries nonzero:

$$\begin{aligned} G_1 &= [\mathbf{g}_0, \ldots, \mathbf{g}_{n-w-1},\ \mathbf{g}_{n-w}, \mathbf{r}_0,\ \ldots,\ \mathbf{g}_{n-1}, \mathbf{r}_{w-1}], \\ G_{\mathrm{pk}} &= S\, G_1\, \mathrm{diag}(I_{n-w}, A_0, \ldots, A_{w-1})\, P_2, \end{aligned} \tag{2}$$

of length $N = n + w$. The decoder applies A_i^{-1} to each pair, keeps the GRS coordinate, and uniquely decodes $t = \lfloor (n-k)/2 \rfloor$ errors. CLT [13] recover the key in polynomial time when $w < n - k$ and prove that the square-code method does not reach $w \geq n - k$; the *long-key* sets of the 2017 NIST submission have $w = n - k$ (Table 3).

Reed-Solomon $(U \,|\, U{+}V)$ codes. Márquez-Corbella and Tillich [8] consider $\{(\mathbf{u} \,\|\, \mathbf{u} + \mathbf{v})\}$, up to diagonal scalings of the halves, with U and V both RS codes; with U random instead it is (1) with $G_W = 0$. Both carry the pairing of the V curve below; the RS-U version carries much more (Sect. 9.6).

Table 3: RLCE long-key parameter sets ($w = n - k$, $N = n + w$, $t = (n - k)/2$): the three 2017 NIST sets out of reach of CLT [13] and a rate-0.87 set at the NIST Level-1 ISD target; k_l as in Table 2 with $n' = N - k + k_l$, excess the bound $(n - k) - k_l + 2$ of Corollary 4.6, measured at full size on ID 0 and R-1 (Sect. 8).

Set	q	n	k	t	w	N	k_l	rows	cols N_3	excess
ID 0 (128)	2^{10}	630	470	80	160	790	46	16 836	17 296	116
ID 2 (192)	2^{10}	1000	764	118	236	1236	55	28 985	29 260	183
ID 4 (256)	2^{11}	1360	800	280	560	1920	84	101 136	102 340	478
R-1 (rate 0.87)	2^{10}	846	732	57	114	960	39	10 413	10 660	77

3.2 The common structure

Definition 3.1 (hidden pairing)**.** Let $\mathcal{C} \subseteq \mathbb{F}_q^N$ have generator G with columns $\mathbf{g}_1, \ldots, \mathbf{g}_N \in \mathbb{F}_q^k$. A *hidden pairing of rank* k_2 on $\mathcal{C}$ consists of an injective linear map $M : \mathbb{F}_q^{k_2} \to \mathbb{F}_q^k$, so that $\Gamma(\alpha) := M\gamma(\alpha)$ is a rational normal curve of degree $k_2 - 1$ in $W := M(\mathbb{F}_q^{k_2})$; distinct $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_q$; and a partition of a subset of $[N]$ into *twin pairs* $\{i_\ell, j_\ell\}$ and *singletons* $\{c_\ell\}$, one per $\ell \le m$, such that

$$\begin{aligned} \Gamma(\alpha_\ell) &\in \operatorname{span}(\mathbf{g}_{i_\ell}, \mathbf{g}_{j_\ell}) \setminus \left(\mathbb{F}_q\mathbf{g}_{i_\ell} \cup \mathbb{F}_q\mathbf{g}_{j_\ell}\right) && \text{for twin pairs,} \\ \mathbf{g}_{c_\ell} &\in \mathbb{F}_q^\times \Gamma(\alpha_\ell) && \text{for singletons.} \end{aligned}$$

The curve point of a twin pair is *genuinely mixed*: it lies in the plane of the pair but on neither of its two lines. Columns outside the partition (*free* columns) are allowed but occur in none of our schemes.

Proposition 3.2 (instances)**.** *(i) UVW1 has a hidden pairing of rank k_2 with m twin pairs $\{c, \tau(c)\}$ (the positions carrying the same column of G_U) and no singletons: $\mathbf{g}_i - \mathbf{g}_j = S(\mathbf{0}; \gamma(\alpha_\ell))$. (ii) UVW2 likewise, with $\delta_i^{-1}\mathbf{g}_i - \delta_j^{-1}\mathbf{g}_j = S(\mathbf{0}; \gamma(\alpha_\ell))$. (iii) RLCE has a hidden pairing of rank k with w twin pairs (the mixed pairs) and $n - w$ singletons (the unmixed GRS columns): for the pair mixed by $A_i = (a_{uv})$, $S\mathbf{g}_{n-w+i} = (a_{11}\mathbf{x}_i - a_{10}\mathbf{x}_j)/\det A_i$ where $\mathbf{x}_i, \mathbf{x}_j$ are the two public columns. (iv) $(U \,|\, U{+}V)$ with an RS code V and random U is (i) with $G_W = 0$; with RS U and V the pair $\{i, i'\}$ spans a plane containing $(\mathbf{0}; \gamma_V(\alpha_i))$ as well, so the V curve is hidden-paired.*

Proof. (i),(ii),(iv): the secret columns of position ℓ and $m{+}\ell$ are $(G_U[:, \ell]; G_V[:, \ell])$ and $(G_U[:, \ell]; G_W[:, \ell])$, whose difference is $(\mathbf{0}; G_{\mathrm{RS}}[:, \ell]) = (\mathbf{0}; \gamma(\alpha_\ell))$ by $G_V = G_{\mathrm{RS}} + G_W$; $M = S(\mathbf{0}; \cdot)$. The point is genuinely mixed because $(\mathbf{0}; \gamma(\alpha_\ell))$ is not proportional to a column with nonzero G_U-part. (iii): the public pair is $(\mathbf{x}_i\ \mathbf{x}_j) = S(\mathbf{g}\ \mathbf{r})A_i$, so $(\mathbf{g}\ \mathbf{r}) = S^{-1}(\mathbf{x}_i\ \mathbf{x}_j)A_i^{-1}$ and $S\mathbf{g} = (a_{11}\mathbf{x}_i - a_{10}\mathbf{x}_j)/\det A_i$, genuinely mixed because $a_{10}, a_{11} \neq 0$; here $M = S$, the multipliers β_ℓ being absorbed in the scalar freedom of Definition 3.1, and $k_2 = k$. □

3.3 What the pairing hides: the anatomy

We condense the structural facts that make the pairing the whole secret; they are the back end of the attack and are elementary.

Lemma 3.3 (fold)**.** *Given the twin pairs of a hidden-pairing code and, for each pair, the point $\Gamma(\alpha_\ell)$ expressed in the pair's plane (for UVW1 the difference, for UVW2 the scaled difference, for*

RLCE the combination of Proposition 3.2(iii)), the matrix F whose columns are these points and the singleton columns has $\operatorname{rowsp} F = \operatorname{rowsp}(M\,[\gamma(\alpha_\ell)]_\ell)$*, a GRS code with permuted support, of dimension k_2 and square dimension* $\min(m, 2k_2 - 1)$*.*

Proof. $F = M\,[\beta_\ell \gamma(\alpha_\ell)]_\ell$ up to column scalings, and M has full column rank. □

Theorem 3.4 (incidence; UVW)**.** *Let $W = S(\{\mathbf{0}\} \times \mathbb{F}_q^{k_2})$. For distinct public columns i, j of a UVW1 key, $\mathbf{g}_i - \mathbf{g}_j \in W$ if (i, j) is a twin pair, and with probability q^{-k_1} over G_U otherwise; with probability $\geq 1 - \binom{n}{2} q^{-k_1}$ the incident pairs are exactly the m twin pairs. For UVW2 the same holds for proportionality modulo W with false-positive rate at most $q^{-(k_1-1)}$.*

Proof. $\mathbf{g}_i - \mathbf{g}_j \in W$ iff the two columns have the same G_U-part, two independent uniform vectors of $\mathbb{F}_q^{k_1}$ unless (i, j) is a twin pair; union bound. □

Corollary 3.5 (reduction)**.** *Given the twin pairs (and the ratios, for UVW2 and RLCE), an equivalent trapdoor correcting the same number of errors as the legitimate decoder follows in $O(N^3)$ field operations: form F (Lemma 3.3), run Sidelnikov-Shestakov on* $\operatorname{rowsp} F$ *to obtain a support and multipliers, and decode by folding the ciphertext along the pairs.*

Proposition 3.6 (hidden dual subcode)**.** *For UVW1 with pairing τ, $E := \mathcal{C}_{\mathrm{pk}}^{\perp} \cap \{\mathbf{y} : y_i + y_{\tau(i)} = 0\ \forall i\}$ is isomorphic to* $\mathrm{RS}^{\perp}$ *via $\mathbf{y} \mapsto \mathbf{z}_1$ (the first half in secret order), has dimension $e = m - k_2$ and square dimension* $\min(m, 2e - 1)$*, and the map $\mathbf{y} \mapsto (y_i + y_{\tau(i)})_i$ on $\mathcal{C}_{\mathrm{pk}}^{\perp}$ has kernel E and image $U^{\perp}$, of rank $m - k_1$. For RLCE the analogous subcode is cut out of $\mathcal{C}_{\mathrm{pk}}^{\perp}$ by one proportionality constraint per mixed pair and is isomorphic to* $\mathrm{GRS}^{\perp}$*, of dimension $n - k$.*

Proof. For UVW1, $\mathbf{y} \in \mathcal{C}^{\perp}$ iff $G_U(\mathbf{z}_1 + \mathbf{z}_2)^T = \mathbf{0}$ and $G_{\mathrm{RS}}\mathbf{z}_1^T + G_W(\mathbf{z}_1 + \mathbf{z}_2)^T = \mathbf{0}$; on the anti-repetition subspace $\mathbf{z}_1 + \mathbf{z}_2 = \mathbf{0}$ this reduces to $\mathbf{z}_1 \in \mathrm{RS}^{\perp}$, and the halves of a Schur product of two such vectors coincide, so $E^{\star 2}$ is the doubling of $(\mathrm{RS}^{\perp})^{\star 2}$. The rank statement follows from $\dim \mathcal{C}^{\perp} - \dim E = (n-k) - (m-k_2) = m - k_1 = \dim U^{\perp}$. For RLCE, $\mathbf{y} \in \mathcal{C}^{\perp}$ with the random-column coordinates of $(AP_2)^{-T}$-transformed $\mathbf{y}$ equal to zero reduces to $G_s \mathbf{h}^T = \mathbf{0}$ (cf. [13]). □

Proposition 3.6 is the structure the square-code attacks of [13] exploit when they can reach it, and the structure that Sect. 7 shows they cannot reach incrementally. The jets of Sect. 4 reach the *primal* curve Γ directly.

4 The Jet Dependency Theorem

Throughout, $\mathcal{C}$ is a hidden-pairing code of rank k_2 (Definition 3.1) and $Z \subseteq [N]$ a shortening set of size $l = k - k_l$ whose columns are independent, with shortening projection $\Pi = \Pi_Z$ and surviving columns $\mathbf{x}_c = \Pi \mathbf{g}_c$, $c \notin Z$. Let n_0, n_1, n_2 be the numbers of twin pairs with $0, 1, 2$ coordinates in Z and s_0, s_1 the numbers of singletons outside and inside Z, so that $n_0 + n_1 + n_2 + s_0 + s_1 = m$ and $n_1 + 2n_2 + s_1 = l$ when there are no free columns. The *projected curve* is

$$p(\alpha) := \Pi\,\Gamma(\alpha) = \big(p_1(\alpha), \ldots, p_{k_l}(\alpha)\big), \qquad \deg p_t \leq k_2 - 1,$$

and for a form $f \in \mathrm{Sym}^d$ the *composed polynomial* is $F_f(\alpha) := f(p(\alpha)) \in \mathcal{P}_{d(k_2-1)}$.

Lemma 4.1 (shortened structure)**.** *(a) For a twin pair with both coordinates surviving, $p(\alpha_\ell) \in \operatorname{span}(\mathbf{x}_i, \mathbf{x}_j)$. (b) For a twin pair with exactly one surviving coordinate i, $\mathbf{x}_i \in \mathbb{F}_q^{\times} p(\alpha_\ell)$ whenever $p(\alpha_\ell) \neq \mathbf{0}$. (c) For a twin pair with both coordinates in Z, and for a singleton in Z, $p(\alpha_\ell) = \mathbf{0}$. (d) For a surviving singleton, $\mathbf{x}_c \in \mathbb{F}_q^{\times} p(\alpha_\ell)$.*

Proof. Apply the linear map Π to the relations of Definition 3.1, using $\mathbf{x}_c = \mathbf{0}$ for $c \in Z$. In (b), $\Gamma(\alpha_\ell) = a\mathbf{g}_i + b\mathbf{g}_j$ with $a \neq 0$ (genuinely mixed), so $p(\alpha_\ell) = a\mathbf{x}_i$. □

Lemma 4.2 (twin line)**.** *Let $f \in \mathrm{Sym}^d$ vanish to order r at two linearly independent points $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{F}_q^{k_l}$, with $q > d+1$. Then $D^s f$ vanishes identically on the plane $\mathrm{span}(\mathbf{x}_i, \mathbf{x}_j)$ for every $s < 2r - d$. Consequently, for a twin pair with both coordinates surviving, the functionals*

$$f \longmapsto F_f^{(s)}(\alpha_\ell), \qquad 0 \le s < 2r - d,$$

lie in the row space of $\mathsf{J}_{d,r}$ restricted to the rows of $\mathbf{x}_i$ and $\mathbf{x}_j$. For a surviving column $\mathbf{x}_c \in \mathbb{F}_q^\times p(\alpha_\ell)$ (cases (b) and (d) of Lemma 4.1) the functionals $f \mapsto F_f^{(s)}(\alpha_\ell)$, $0 \le s < r$, lie in the row space of the rows of $\mathbf{x}_c$.

Proof. Fix $s < r$ and a coordinate ϕ of the tensor $D^s f$; it is a form of degree $d - s$. Along the affine line $t \mapsto \mathbf{x}_i + t(\mathbf{x}_j - \mathbf{x}_i)$ it is a polynomial $h(t)$ of degree $\le d - s$ whose derivatives up to order $r - s - 1$ vanish at $t = 0$ and $t = 1$ (they are coordinates of $D^{s+1}f, \dots, D^{r-1}f$ at the two points), so $t^{r-s}(t-1)^{r-s} \mid h$, and $h = 0$ if $2(r-s) > d - s$, i.e. $s < 2r - d$. The homogeneous form ϕ then vanishes on all points $a\mathbf{x}_i + b\mathbf{x}_j$ with $a + b \neq 0$, which are $q^2 - q$ points of the plane, while a nonzero binary form of degree $d - s$ vanishes on at most $(d-s)(q-1) + 1 < q^2 - q$ of them for $q > d + 1$; so ϕ vanishes on the plane. By the chain rule, $F_f^{(s)}(\alpha)$ is a linear combination of the values $D^u f(p(\alpha))[p^{(\cdot)}(\alpha), \dots]$, $u \le s$, of derivative tensors at the curve point contracted with derivatives of p; for a twin pair the curve point lies in the plane (Lemma 4.1(a)), where all these tensors vanish whenever the row conditions of the pair hold. A functional that vanishes on the common kernel of a set of functionals lies in their span. For a column proportional to $p(\alpha_\ell)$ the same holds for $u < r$ directly, since $D^u f(\lambda\mathbf{x}) = \lambda^{d-u} D^u f(\mathbf{x})$. □

Before counting, we make precise how the structured functionals sit inside the row space, since the count of row dependencies rests on it. Write $W := \mathbb{F}_q^{\mathrm{rows}}$ for the space of row combinations of $\mathsf{J}_{d,r}$ and $J : W \to (\mathrm{Sym}^d)^*$ for the linear map sending a row combination to the functional it defines; the left kernel of $\mathsf{J}_{d,r}$ is $\ker J$. A *node* is a pair or singleton ℓ with at least one surviving coordinate; its *order* is $\omega_\ell = 2r - d$ for a pair with both coordinates surviving and $\omega_\ell = r$ for a surviving singleton or a half-shortened pair (a pair with $2r - d \le 0$ and both coordinates surviving is not a node: it contributes nothing), and its *structured functionals* are $\varphi_{\ell,s} : f \mapsto F_f^{(s)}(\alpha_\ell)$, $0 \le s < \omega_\ell$, with Hasse derivatives. Lemma 4.2 provides, for each $\varphi_{\ell,s}$, a row combination $a_{\ell,s} \in W$ supported on the rows of the node's surviving columns with $J(a_{\ell,s}) = \varphi_{\ell,s}$; we fix one such choice. Let $T = \sum_\ell \omega_\ell$ and let $\Phi : \mathbb{F}_q^T \to W$, $e_{\ell,s} \mapsto a_{\ell,s}$.

Lemma 4.3 (independence of the structured functionals)**.** *Assume the* non-degeneracy condition

for every node ℓ, the vectors $p(\alpha_\ell), p^{(1)}(\alpha_\ell), \dots, p^{(\omega_\ell - 1)}(\alpha_\ell) \in \mathbb{F}_q^{k_l}$ are linearly independent, (ND)

where $p^{(i)}$ is the i-th Hasse derivative of the projected curve. Then:

(i) for each node ℓ the functionals $\varphi_{\ell,s}$, $0 \le s < \omega_\ell$, are linearly independent on Sym^d;

(ii) the vectors $a_{\ell,s}$, over all nodes and all s, are linearly independent in W, i.e. Φ is injective;

(iii) let $\mathrm{Rel} := \{c \in \mathbb{F}_q^T : \sum_{\ell,s} c_{\ell,s}\varphi_{\ell,s} = 0\}$ be the relation space. Then $K_s := \Phi(\mathrm{Rel})$ is a subspace of the left kernel of $\mathsf{J}_{d,r}$ of dimension exactly $\dim \mathrm{Rel} = T - \dim \mathrm{span}\{\varphi_{\ell,s}\} \ \ge \ T - \dim \mathcal{V}$, where $\mathcal{V} = \{F_f : f \in \mathrm{Sym}^d\}$.

Proof. (i) By (ND) there are linear forms L_0, L_1 on $\mathbb{F}_q^{k_l}$ with $L_0(p^{(i)}(\alpha_\ell)) = \delta_{i0}$ and $L_1(p^{(i)}(\alpha_\ell)) = \delta_{i1}$ for $0 \le i < \omega_\ell$ (L_1 is needed only if $\omega_\ell \ge 2$). Put $f_s := L_0^{d-s} L_1^s \in \mathrm{Sym}^d$ for $0 \le s < \omega_\ell$; this uses $\omega_\ell \le d$, which holds since $\omega_\ell \le r \le d$. By the Hasse Taylor expansion, valid in every characteristic, $p(\alpha_\ell + \varepsilon) = \sum_{i \ge 0} \varepsilon^i p^{(i)}(\alpha_\ell)$, so $L_0(p(\alpha_\ell + \varepsilon)) = 1 + O(\varepsilon^{\omega_\ell})$ and $L_1(p(\alpha_\ell + \varepsilon)) = \varepsilon + O(\varepsilon^{\omega_\ell})$, whence $F_{f_s}(\alpha_\ell + \varepsilon) = \varepsilon^s + O(\varepsilon^{\omega_\ell})$. Thus $\varphi_{\ell,s'}(f_s) = \delta_{ss'}$ for $s, s' < \omega_\ell$: the matrix $(\varphi_{\ell,s'}(f_s))$ is the identity, and the $\varphi_{\ell,s}$ are independent. (ii) The supports of the $a_{\ell,s}$ for distinct nodes are disjoint, because distinct nodes have disjoint column sets. If $\sum_{\ell,s} c_{\ell,s} a_{\ell,s} = 0$ in W, restricting to the rows of one node gives $\sum_s c_{\ell,s} a_{\ell,s} = 0$, and applying J gives $\sum_s c_{\ell,s} \varphi_{\ell,s} = 0$, so $c_{\ell,s} = 0$ by (i). (iii) For $c \in \mathrm{Rel}$, $J(\Phi(c)) = \sum c_{\ell,s} \varphi_{\ell,s} = 0$, so $\Phi(\mathrm{Rel}) \subseteq \ker J$; Φ is injective by (ii), so $\dim K_s = \dim \mathrm{Rel}$. Finally each $\varphi_{\ell,s}$ factors through the linear map $f \mapsto F_f \in \mathcal{V}$, so $\dim \mathrm{span}\{\varphi_{\ell,s}\} \le \dim \mathcal{V}$. □

Condition (ND) says that the projected curve has no singular or inflectional behaviour of order ω_ℓ at a surviving node. It fails only on a proper closed subset of the projections Π for each node (the projected curve has degree $k_2 - 1 \ge \omega_\ell$ and $k_l \ge \omega_\ell$), so it holds for a generic shortening; we have never seen it fail, and every measurement in Sect. 8 is consistent with it. It is the only genericity hypothesis in the theorem below beyond the two stated in its first sentence. The field-size hypothesis $q > d + 1$ of Lemma 4.2 is carried along for the self-contained proof of this section; the forcing law of Sect. 9.1 (Theorem 9.2(ii) with $j = 2$) gives the same conclusion in every field, so the counts of this section hold without it. Every field in this paper satisfies it in any case.

Theorem 4.4 (jet dependencies of a hidden pairing)**.** *With the notation above, assume $q > d + 1$, that the two columns of every twin pair with both coordinates surviving are linearly independent, and that $p(\alpha_\ell) \ne \mathbf{0}$ for every pair or singleton with a surviving coordinate, and that* (ND) *holds. Then the structured functionals $\varphi_{\ell,s}$ satisfy at least*

$$\Delta_{d,r} := (2r - d)^+ n_0 + r\,(n_1 + s_0) + d\,(n_2 + s_1) - \big(d(k_2 - 1) + 1\big) \tag{3}$$

independent linear relations, each of which is a row dependency of $\mathsf{J}_{d,r}$ supported on the rows of the surviving paired and singleton columns. For $2r \le d + 1$, where random points impose independent conditions (Lemma 2.2), these are $\Delta_{d,r}$ dependencies beyond the generic rank whenever $\Delta_{d,r} > 0$. When there are no free columns and $2r \ge d$, with $l = n_1 + 2n_2 + s_1$ the number of shortened positions and $s = s_0 + s_1$ the number of singletons,

$$\Delta_{d,r} = (2r - d)\,m + (d - r)\,(l + s) - d(k_2 - 1) - 1 : \tag{4}$$

for UVW ($s = 0$) $\Delta_{d,r} = (2r - d)m + (d - r)l - d(k_2 - 1) - 1$, and for RLCE ($m = n$ curve points, $s = n - w$, $k_2 = k$) $\Delta_{d,r} = (2r - d)n + (d - r)(l + n - w) - d(k - 1) - 1$.

Proof. Let $\mathcal{V} := \{F_f : f \in \mathrm{Sym}^d\} \subseteq \mathcal{P}_{d(k_2-1)}$. By Lemma 4.1(c), every α_ℓ with ℓ a fully shortened pair or a shortened singleton is a root of p, hence a root of multiplicity $\ge d$ of every F_f, so $\dim \mathcal{V} \le d(k_2 - 1) + 1 - d(n_2 + s_1)$. By Lemma 4.2, the row space of $\mathsf{J}_{d,r}$ contains the $T := (2r - d)^+ n_0 + r(n_1 + s_0)$ Hermite functionals $F \mapsto F^{(s)}(\alpha_\ell)$ listed there, at $n_0 + n_1 + s_0$ distinct nodes. By Hermite interpolation (Sect. 2) their restrictions to $\mathcal{V}$ satisfy at least $T - \dim \mathcal{V} \ge \Delta_{d,r}$ independent relations, i.e. $\dim \mathrm{Rel} \ge \Delta_{d,r}$. By Lemma 4.3(iii) the map Φ carries Rel injectively into the left kernel of $\mathsf{J}_{d,r}$, so the left kernel contains the subspace $K_s = \Phi(\mathrm{Rel})$ of dimension $\ge \Delta_{d,r}$, each of whose vectors is supported on the rows of the surviving paired and singleton columns. When $2r \le d + 1$ there are no line relations and the generic rank is the number of rows, so each dimension of K_s is a dependency beyond the generic rank. Formula (4) is bookkeeping: with $n_0 + n_1 + n_2 + s_0 + s_1 = m$ and $n_1 + 2n_2 + s_1 = l$, the structured count $(2r - d)n_0 + r(n_1 + s_0) + d(n_2 + s_1)$ equals $(2r - d)m + (d - r)(n_1 + 2n_2 + s_1) + (d - r)(s_0 + s_1)$. □

Corollary 4.5 (UVW)**.** *For UVW1 or UVW2 with $k = m$ shortened at $l = m - k_l$ positions,*

$$\Delta_{3,2} = m + l - 3k_2 + 2 = 2k_1 - k_2 - k_l + 2.$$

In the balanced family $k_1 = k_2$ this is $k_2 - k_l + 2$, the count of [22] ($215 - 53 + 2 = 164$ at UVW-KEM-128); in the unbalanced family $k_1 > k_2$ it is larger; it is positive iff $k_l < 2k_1 - k_2 + 2$, and independent of how the l shortened positions fall on the pairs.

Corollary 4.6 (RLCE)**.** *For RLCE shortened at $l = k - k_l$ positions,*

$$\Delta_{3,2} = 2n - w + l - 3k + 2 = 2(n-k) - w - k_l + 2,$$

which for the long-key regime $w = n - k$ equals $(n-k) - k_l + 2$ and is positive for every $k_l < n - k + 2$. For $w = c(n-k)$ it is $(2-c)(n-k) - k_l + 2$.

Corollary 4.7 (window)**.** *The excess of Theorem 4.4 is visible as a rank deficiency whenever the system is underdetermined, $N_d \geq n' \binom{k_l + r - 2}{r-1}$. For $(3,2)$ and $n' = N - l$ this holds from $k_l \approx \sqrt{6(N-k)}$ on, which is* 51–103 *for the UVW sets and* 39–84 *for the RLCE sets (Tables 2 and 3); the attack chooses the smallest such k_l, where the excess is largest.*

Corollary 4.8 (the jet hierarchy; no split escapes)**.** *For UVW with $k = m$ and the order-$(d-1)$ degree-d jet,*

$$\Delta_{d,d-1} = (d-1)m - d\,k_2 - k_l + d - 1,$$

positive for every $k_2 < m(1 - 1/d)$ up to the k_l term: the cubic gradient jet reaches $k_2 < \frac{2}{3}m$ (all proposed sets), the quartic order-3 jet $k_2 < \frac{3}{4}m$, and so on, at a cost polynomial in m with exponent growing like $d \approx m/(m - k_2)$. (For $d \geq 4$ the generic baseline of Lemma 2.2 is nonzero; that the structured dependencies appear on top of it is verified for $(4,3)$ in Sect. 8.)

Proof. Put $r = d - 1$ in (4) with $l = m - k_l$ and no singletons. □

Equality. Theorem 4.4 is a lower bound. In every non-degenerate instance we generated the excess *equals* $\Delta_{d,r}$: for $(3,2)$ across eleven splits, three fields, UVW1 and UVW2, three sizes and RLCE in characteristics 2 and 257; for $(4,2)$; and for $(4,3)$, where the 992 measured dependencies are the baseline $\binom{45}{2} = 990$ plus $\Delta_{4,3} = 2$ (Sect. 8). When too many surviving columns are themselves projected curve points the deficiency is larger still (Remark 9.5), which only helps the attacker.

5 From Dependencies to the Key

The proof of Theorem 4.4 is constructive: it expresses each structured functional as an explicit combination of jet rows. Making the combination explicit for the cubic gradient system explains the block structure that [22] observed and used.

Theorem 5.1 (block structure)**.** *Let $(d,r) = (3,2)$ and $3 \neq 0$ in $\mathbb{F}_q$. For a twin pair with both coordinates surviving, write $p(\alpha_\ell) = s\,\mathbf{x}_i + t\,\mathbf{x}_j$ with $s, t \neq 0$. Then for every cubic f,*

$$f(p(\alpha_\ell)) = \langle \lambda_i, \nabla f(\mathbf{x}_i) \rangle + \langle \lambda_j, \nabla f(\mathbf{x}_j) \rangle, \quad \lambda_i = \tfrac{s^3}{3}\,\mathbf{x}_i + s^2 t\,\mathbf{x}_j, \quad \lambda_j = st^2\,\mathbf{x}_i + \tfrac{t^3}{3}\,\mathbf{x}_j. \tag{5}$$

Consequently, let K_s be the structured kernel of Lemma 4.3(iii) and $K \in \mathbb{F}_q^{\Delta_s \times \mathrm{rows}}$ a matrix whose rows form a basis of K_s ($\Delta_s = \dim K_s \geq \Delta_{3,2}$). Then the $\Delta_s \times k_l$ block of column i is $\mathbf{c}_\ell \otimes \lambda_i$ and that of column j is $\mathbf{c}_\ell \otimes \lambda_j$ for one coefficient vector $\mathbf{c}_\ell \in \mathbb{F}_q^{\Delta_s}$ *per pair: both blocks have rank one*

and share their column space. The block of a surviving column proportional to a curve point (a singleton, or the survivor of a half-shortened pair) is $\mathbf{c}_\ell \otimes \lambda^{(0)} + \mathbf{c}'_\ell \otimes \lambda^{(1)}$ *and has rank two. If the left kernel of* $\mathsf{J}_{3,2}$ *equals* K_s *—*exactness*, the case* excess $= \Delta_{3,2}$ *observed in every non-degenerate instance of Sect. 8—the same holds for any basis of the full left kernel. Finally, writing a rank-one block direction as* $\lambda_i = a\,\mathbf{x}_i + b\,\mathbf{x}_j$*,*

$$t/s = b/(3a): \qquad \textit{UVW2:}\ \ \delta_j/\delta_i = -3a/b; \qquad \textit{RLCE:}\ \ S\mathbf{g}_\ell \ \propto\ \mathbf{x}_i + \tfrac{b}{3a}\,\mathbf{x}_j,$$

and the same ratio holds for the unshortened public columns $\mathbf{g}_i, \mathbf{g}_j$*.*

Proof. Let $h(a,b) = f(a\mathbf{x}_i + b\mathbf{x}_j) = c_0a^3 + c_1a^2b + c_2ab^2 + c_3b^3$. Differentiating, $\partial_a h(1,0) = \langle \mathbf{x}_i, \nabla f(\mathbf{x}_i)\rangle = 3c_0$, $\partial_b h(1,0) = \langle \mathbf{x}_j, \nabla f(\mathbf{x}_i)\rangle = c_1$, $\partial_a h(0,1) = \langle \mathbf{x}_i, \nabla f(\mathbf{x}_j)\rangle = c_2$ and $\partial_b h(0,1) = \langle \mathbf{x}_j, \nabla f(\mathbf{x}_j)\rangle = 3c_3$, and $f(p(\alpha_\ell)) = h(s,t)$ gives (5). A dependency that is a relation $\sum_\ell c_\ell\, F(\alpha_\ell) + \cdots = 0$ among structured functionals has, on the rows of column i, the vector $c_\ell\lambda_i$, since the gradient rows at a nonzero point are independent functionals and the representation is unique; stacking a basis of dependencies gives the block $\mathbf{c}_\ell \otimes \lambda_i$. For UVW2, $p(\alpha_\ell) = \delta_i^{-1}\mathbf{x}_i - \delta_j^{-1}\mathbf{x}_j$, so $s = \delta_i^{-1}$, $t = -\delta_j^{-1}$ and $b/a = 3t/s = -3\delta_i/\delta_j$. For RLCE the curve point is $s\mathbf{x}_i + t\mathbf{x}_j$ with $t/s = b/(3a)$. The ratio transfers to the public columns because the relation $\Gamma(\alpha_\ell) = s'\mathbf{g}_i + t'\mathbf{g}_j$ projects under Π to $p(\alpha_\ell) = s'\mathbf{x}_i + t'\mathbf{x}_j$ with $\mathbf{x}_i, \mathbf{x}_j$ independent. □

Remark 5.2 (characteristic and normalization)*.* The cubic gradient system and Theorem 5.1 need $3 \neq 0$: in characteristic 3 the Euler identity no longer converts "gradient vanishes" into "multiplicity two" and the factor $\frac{1}{3}$ is undefined; one uses instead the Hasse jets of order two directly (Definition 2.1 with Hasse derivatives), for which Theorem 4.4 and the forcing law hold verbatim and the block formulas change their constants. In characteristic 2 (every RLCE field, HermitianRLCE) the factor $\frac{1}{3}$ is 1 and $-1 = 1$, so $t/s = b/a$ and $S\mathbf{g}_\ell \propto \mathbf{x}_i + (b/a)\mathbf{x}_j$. The normalization $p(\alpha_\ell) = s\mathbf{x}_i + t\mathbf{x}_j$ with $s, t \neq 0$ is the genuinely-mixed condition of Definition 3.1 projected by Π; it can fail only if Π maps $\Gamma(\alpha_\ell)$ onto the line of one public column, a proper closed condition on Z that never occurred in our runs. The forcing law (Theorem 9.2) is a statement about Hasse jets in every characteristic; the read-off formulas of this section are its $(3,2)$ specialization and inherit the restriction $3 \neq 0$.

The hypothesis of exactness in Theorem 5.1 is what the algorithm below would seem to depend on: if the left kernel is larger than K_s, a basis of it mixes structured and unstructured dependencies and the blocks need not be rank one. The algorithm does not in fact depend on it, in the following sense.

Proposition 5.3 (what the read-off needs, and what protects it)**.** *Let* K *be any basis of the full left kernel of* $\mathsf{J}_{3,2}$*, and let the read-off accept a pair* $\{i,j\}$ *of surviving columns only if (a) both blocks have rank one and share their coefficient direction, and (b) the block directions satisfy* $\lambda_i, \lambda_j \in \mathrm{span}(\mathbf{x}_i, \mathbf{x}_j)$ *on all* k_l *coordinates, with* $\lambda_i = a\mathbf{x}_i + b\mathbf{x}_j$*,* $a, b \neq 0$*, and* λ_j *the matching vector of* (5) *for* $t/s = b/(3a)$*.*

(i) (Completeness under exactness.) If the left kernel is K_s*, every twin pair with both coordinates surviving is accepted, with the correct ratio (Theorem 5.1).*

(ii) (No false pairs, heuristic.) Condition (b) imposes $k_l - 2$ *independent linear conditions on each of two vectors of* $\mathbb{F}_q^{k_l}$*. If the blocks of two non-twin columns were uniform subject to rank one, a false acceptance would occur with probability* $q^{-2(k_l-2)}$ *per candidate, i.e. never at any parameter of this paper; we never observed one.*

(iii) (Verified output.) The output of Algorithm 1 is accepted only if the fold matrix F has rank k_2 and Sidelnikov-Shestakov returns a support and multipliers with $\operatorname{rowsp} F = \mathrm{GRS}(\boldsymbol{\alpha}, \boldsymbol{\beta})$*, checked by a rank computation on the stacked generators. Suppose the final fold has at most $f \le (n - k_2 - 3)/3$ wrong pairs, so that it has at least $k_2 + 3$ genuine columns ($f \le 37$ at R-1; $f = 0$ in every run of this paper). Then by Lemma 5.4 the check passes only if* every *column of F is a point of the hidden curve M(curve), at pairwise distinct parameters, and under the key-generation distribution stated after the lemma a false pair or a wrong ratio produces such a column with probability at most $q^{-(k_2-3)}$; so, conditional on $f \le (n - k_2 - 3)/3$, the algorithm outputs a wrong key with probability at most $\binom{N}{2} q^{-(k_2-3)}$. Whether the read-off can accept more than $(n - k_2 - 3)/3$ false pairs is governed by (ii), which is heuristic; we do not claim an unconditional bound.*

(iv) (Where exactness fails.) Beyond the Veronese threshold of Remark 9.5—more surviving curve-point columns (singletons and half-shortened survivors) than about $1.4\,k_2$—the left kernel is strictly larger than K_s, the extra dependencies reach the twin rows (the rows of a pair express, besides $f \mapsto f(p(\alpha_\ell))$, the derivative along the twin line at the curve point, which the curve-point jets can involve), the twin blocks have rank 3*, and the read-off accepts nothing: on RLCE* $(60, 30, 10)$ *over* $\mathbb{F}_{2^{10}}$*, with* 50 *singletons against $k = 30$, the kernel is* 89 *against $\Delta = 31$, the* 18 *twin columns have block rank* 3*, and* 0 *of* 9 *(seed* 1*) and* 0 *of* 7 *(seed* 2*) recoverable pairs are accepted. Recovery in that regime is open (Sect. 11). At every proposed parameter set the ratio of surviving curve-point columns to k_2 is* 0.33 *(R-1, measured),* 0.63 *(ID 0),* 0.61 *(NGCC-128),* 0.69 *(NGCC-512), and exactness held in every full-size run.*

Proof. (i) is Theorem 5.1; (ii) is the stated count under the stated heuristic; (iv) is measured (`rlce_recover.py`). (iii) is Lemma 5.4 and the remark after it. □

The verification rests on a classical fact about rational normal curves, which we state in the form used.

Lemma 5.4 (GRS verification)**.** *Let $F \in \mathbb{F}_q^{k_2 \times n}$ have at least $k_2 + 3$* genuine *columns, i.e. columns $M\beta_c\gamma(\alpha_c)$ with pairwise distinct α_c, and suppose that* $\operatorname{rowsp} F = \mathrm{GRS}(\boldsymbol{\alpha}', \boldsymbol{\beta}')$ *for some pairwise distinct $\alpha'_1, \ldots, \alpha'_n \in \mathbb{F}_q$ and nonzero β'_c. Then every column of F is a point of M(curve), and the parameters of the n columns are pairwise distinct.*

Proof. A generator of $\mathrm{GRS}(\boldsymbol{\alpha}', \boldsymbol{\beta}')$ is $V'D'$ with $V' = [\gamma(\alpha'_c)]_c$ and $D' = \operatorname{diag}(\boldsymbol{\beta}')$, so $F = TV'D'$ for an invertible $T \in \mathbb{F}_q^{k_2 \times k_2}$, and the projectivity $\tau := M^{-1}T$ of $\mathbb{P}^{k_2-1}$ maps the $k_2 + 3$ points $\gamma(\alpha'_c)$ of the rational normal curve Γ_0 at the genuine columns to the points $\gamma(\alpha_c)$ of Γ_0. Any k_2 points of Γ_0 are linearly independent, so these $k_2 + 3$ image points are in general position, and $\tau(\Gamma_0)$ is a rational normal curve through them; through $k_2 + 3$ points of $\mathbb{P}^{k_2-1}$ in general position passes exactly one rational normal curve (Castelnuovo's lemma, [20, Thm. 1.18]). We apply it over the algebraic closure $\overline{\mathbb{F}}_q$: both Γ_0 and $\tau(\Gamma_0)$ are rational normal curves defined over $\mathbb{F}_q$ through the same $k_2 + 3$ $\mathbb{F}_q$-rational points in general position, so they coincide as curves over $\overline{\mathbb{F}}_q$, hence their $\mathbb{F}_q$-points coincide, and $\tau(\Gamma_0) = \Gamma_0$. Hence $T\gamma(\alpha'_c) = M \cdot (\text{a point of } \Gamma_0)$ for every column c, genuine or not, and the parameters are distinct because τ is injective on Γ_0 and the α'_c are distinct. □

The hypothesis of $k_2 + 3$ genuine columns holds whenever at most $(n - k_2 - 3)/3$ pairs are wrong (a wrong pair spoils at most three columns of F; at R-1, 37 wrong pairs leave $846 - 111 = 735 = k_2 + 3$ genuine ones), and in particular in every run of Algorithm 1, whose pairs are all correct. The lemma leaves one way for a wrong fold to pass: a column $\mathbf{g}_i + \rho\mathbf{g}_j$ that lies on the hidden curve at a parameter outside the genuine support. The probability statements that follow are over the

Algorithm 1 JetRecover: public key only

Choose k_l as in Corollary 4.7; set $\mathcal{T} \leftarrow \emptyset$ (known pairs with ratios), $\sigma_c \leftarrow 0$ (survival counts), $\mathcal{R} \leftarrow \emptyset$ (columns that have shown a rank-one block).
repeat
  Choose Z with $|Z| = k - k_l$: first the coordinates of pairs in $\mathcal{T}$, then columns not in $\mathcal{R}$ in decreasing order of σ_c (probable singletons), then unseen columns at random.
  Shorten, form $\mathsf{J}_{3,2}$ at the surviving columns, compute a basis K of its left kernel; increment σ_c for the survivors.
  For each surviving column c with a rank-one block, record its coefficient direction $\mathbf{u}_c \in \mathbb{P}^{\dim K - 1}$ and λ_c; add c to $\mathcal{R}$.
  Group columns by $\mathbf{u}_c$; for each group of size two, solve $\lambda_i = a\mathbf{x}_i + b\mathbf{x}_j$ on all k_l coordinates; if consistent with $a, b \neq 0$ (Proposition 5.3(b)), accept the pair with $t/s = b/(3a)$ and add it to $\mathcal{T}$.
until $|\mathcal{T}|$ = number of twin pairs, or $w - |\mathcal{T}| < n - k - k'_l + 1$ for a square-code window k'_l (RLCE)
(Square-code stage, RLCE.) Fold the pairs of $\mathcal{T}$ into their curve points; repeat: shorten the folded code to dimension k'_l, compute the left kernel of the Veronese matrix of the survivors, read the off-curve columns (rows in no kernel vector), pair them by the planes that meet the curve and record the ratio; until all pairs are known (Sect. 8.1).
Singletons (RLCE) are the columns in no pair.
Form F from the recovered curve points (Lemma 3.3); run Sidelnikov-Shestakov on $\operatorname{rowsp} F$; verify $\operatorname{rowsp} F = \mathrm{GRS}(\boldsymbol{\alpha}, \boldsymbol{\beta})$; output the pairing, ratios and GRS description (Corollary 3.5).

key-generation distribution of the schemes: for RLCE each inserted column $\mathbf{r}_\ell$ is uniform in $\mathbb{F}_q^k$ and independent of the GRS code and of the mixing matrices A_ℓ (whose entries are nonzero); for UVW each column of G_U is uniform in $\mathbb{F}_q^{k_1}$, so that a public pair column is $S(\mathbf{u}_\ell; \cdot)$ with $\mathbf{u}_\ell$ uniform, which plays the role of $\mathbf{r}_\ell$ below. For a true pair with a wrong ratio this needs a second curve point in the plane $\operatorname{span}(\mathbf{g}_i, \mathbf{g}_j) = \operatorname{span}(\mathbf{h}_\ell, \mathbf{r}_\ell)$, i.e. the random column $\mathbf{r}_\ell$ in the span of two curve points, which has probability $\leq q \cdot q^{-(k_2-2)}$ over $\mathbf{r}_\ell$; for two non-twin columns $\mathbf{g}_i, \mathbf{g}_j$ (from different pairs, or a pair column and a singleton) it needs the plane of two columns, at least one of which is uniform in a plane $\operatorname{span}(\mathbf{h}, \mathbf{r})$ with $\mathbf{r}$ uniform, to meet the q points of the curve, again probability $\leq q^{-(k_2-3)}$. A union bound over the $\binom{N}{2}$ candidate pairs gives the bound in Proposition 5.3(iii), which is below 2^{-1800} at every parameter of this paper. (For UVW the same argument applies with $\mathbf{r}_\ell$ replaced by the G_U-column of the pair.)

We verified Theorem 5.1 literally on UVW2 and RLCE toy keys: block ranks, shared coefficient vectors, pairs, ratios and GRS directions all as stated, with no spurious pair (Sect. 8), and we ran Algorithm 1 end to end at full size on RLCE (Sect. 8.1). Whether exactness holds for every non-degenerate hidden-pairing code below the Veronese threshold is open (Sect. 11); what we can say is that it held in every instance below the threshold that we ran, at toy and at full size, and that its failure is visible to the attacker (the kernel exceeds Δ), who then knows the read-off will not work.

Cost and number of rounds. Each round is one dense left-kernel computation on an $n'k_l \times N_3$ matrix with $k_l = \Theta(\sqrt{N-k})$, i.e. $O((N-k)^{1.5\,\omega})$ field operations, at most $O(N^{4.5})$; Sidelnikov-Shestakov is $O(m^3)$. A round recovers exactly the pairs whose two coordinates both survive the shortening. The number of rounds therefore depends on the *rate*: a shortening keeps $n' = N - k + k_l$ columns, so at UVW's rate 1/2 about half the columns survive, a quarter of the unknown pairs are found per round, and three rounds sufficed at UVW-KEM-128 [22]; at RLCE's rate 0.87 (R-1)

only $n'/N \approx 0.28$ of the columns survive, and a random shortening exposes $\approx (n'/N)^2 w \approx 9$ of the 114 pairs per round. The adaptive choice of Z in Algorithm 1—known pairs first, then the columns that have survived most often without ever showing a rank-one block (a paired column survives σ_c rounds without its twin with probability about $(1 - n'/N)^{\sigma_c}$, so these are singletons with growing confidence, and a mistake only delays one pair by a round)—raises this substantially: the survival-model simulation (`sim_rounds.py`) gives 36 ± 7 rounds for R-1 against 64 ± 12 for random shortenings, and 7 against 24 for ID 0, where $N - k$ is larger. Better still, the jet stage need not finish the job: once the known pairs are folded into their curve points the remaining code has $w' = w - |\mathcal{T}|$ pairs, and by Corollary 9.4 its shortened square code is deficient as soon as $w' < n - k - k_l + 1$—the attacker has walked the key into CLT's short-key regime [13], where each round costs seconds. For R-1 this happens after two to four jet rounds. The measured runs, with both stages, are in Sect. 8.1. Algorithm 1 is the procedure of [22] with Theorem 5.1 supplying the step "read the pair and the ratio off the block", Proposition 5.3 supplying the acceptance test, and Corollary 4.6 extending it to RLCE, where the recovered quantity is the hidden GRS column of each mixed pair.

6 Consequences for Parameters

6.1 UVW

Table 2 evaluates Corollaries 4.5 and 4.7 at all nine proposed sets. Three observations.

Every set is distinguished at degree three by a dense linear system of at most 1.9×10^5 unknowns; the NGCC-128 line reproduces the $25\,599 \times 26\,235$ system and the excess 164 of [22], where the three rounds of Algorithm 1 took about sixteen seconds each.

The unbalanced family is the more exposed. The designers propose $k_1 > k_2$ (sets 128b, 192b, 256b) to shrink keys. By Corollary 4.5 the excess $2k_1 - k_2 - k_l + 2$ grows with k_1: 183 at 128b against 151 at 128a. The designers' own structural argument—that the low-weight $(\mathbf{u}\|\mathbf{u})$ codewords are hard to find when $k_1 \le k_2$—is correct and beside the point; the jets never look for codewords.

No split escapes the hierarchy—but the hierarchy is not one attack. The cubic excess is positive iff $k_l < 2k_1 - k_2 + 2$, i.e. (with $k_1 = m - k_2$) iff $k_2 < \frac{2m - k_l + 2}{3}$, about $k_2 < 0.66\,m$ at the proposed sizes. Beyond that, Corollary 4.8 applies with $d = 4, 5, \ldots$, and the only remaining lever, pushing k_2/m towards 1, is forbidden by the Guruswami-Sudan radius $w \le m - \sqrt{mk_2}$: the error weight drops from $0.29\,m$ at $k_2/m = 0.5$ to $0.18\,m$ at 0.67 and $0.11\,m$ at 0.8, so the block length and the key m^2 must grow to keep the decoding security, while the jet degree needed grows only to $d = \lceil m/(m - k_2) \rceil + O(1)$. Three regimes should be kept apart.

- *Fixed degree, polynomial time.* For any fixed d the $(d, d-1)$ test is a rank computation on $\binom{k_l + d - 1}{d}$ unknowns with $k_l = \Theta(\sqrt{d(d-1)m})$, i.e. $O(m^{\omega d/2})$ field operations: polynomial in m. This covers $k_2/m < 1 - 1/d$. All nine proposed sets are in the $d = 3$ regime.

- *Degree growing with the rate.* A family of parameters with $k_2/m \to 1$ needs $d \approx m/(m - k_2) \to \infty$, and the exponent grows with d. "No split escapes" is therefore a statement about a family of polynomial attacks indexed by the rate, not one uniform polynomial-time attack; asymptotically, a family with $m - k_2 = O(1)$ would escape every fixed degree, at the price of a decoding radius $\sqrt{mk_2}$ short of m, i.e. $O(1)$ correctable errors, which is no cryptosystem.

- *Not demonstrated.* A split with $k_2 > \frac{2}{3}m$ is not broken *in practice* by the computations of this paper, and we say so plainly: the $(4, 3)$ system at $m = 860$ (any split with $k_2 < \frac{3}{4}m$) has $k_l \approx \sqrt{12\,n'} \approx 107$ and $\binom{110}{4} \approx 5.8 \times 10^6$ unknowns, beyond dense elimination and within

reach only of the sparse and structured methods of [15, 30], and $(5, 4)$ has about 5×10^8; we have run neither. Such a split has no security argument against the next jet, and a key size that defeats the purpose of the construction (Theorem 9.8 makes this exact).

6.2 RLCE

Table 3 evaluates Corollary 4.6 at the three long-key sets of the 2017 NIST submission, the regime that CLT [13] proved out of reach of the square-code method, and at a rate-0.87 long-key set at the Level-1 ISD target. All four are distinguished at degree three; the Level-1 sets by a rank computation of about 10^4–1.7×10^4 unknowns, which we carried out at full size for both: ID 0 gives excess 116 and R-1 excess 77, each exactly the bound (Sect. 8). On R-1 we then ran the whole of Algorithm 1 from the public key alone and decrypted a fresh ciphertext (Sect. 8.1): the long-key regime is broken in practice at that size, not only distinguished. The mechanism is the one of Theorem 4.4 and nothing else. One might expect the Goppa attacks of 2026 not to transfer to RLCE, which has no subfield subcode and no Frobenius action, and for the Goppa mechanism that expectation is correct; what transfers is the twin line, which needs neither.

Increasing the number of random columns helps less than one might hope. For $w = c(n-k)$, Corollary 4.6 gives a cubic excess $(2-c)(n-k)-k_l+2$, so the cubic jet reaches $c < 2$ at the proposed sizes, and the order-$(d-1)$ degree-d jet, by (4) with $r = d-1$, has excess $(d-1-c)(n-k)+(d-1)-k_l$, positive for $c < d-1$. The required degree is $d \approx c+2$, the attack is polynomial for each fixed c with an exponent that grows with c (the same three regimes as for UVW above), and the public key $k(N-k) = k(n-k)(1+c)$ grows linearly in c. At $c = 3$ the computation is a quintic order-4 jet system with about 2×10^8 unknowns at the ID 0 size, which we have not run; the honest summary is that long keys with $c \geq 2$ are not broken by the computations of this paper, and have no security argument against the next one.

6.3 Other constructions and a design principle

$(U \,|\, U{+}V)$ with a Reed-Solomon V. With U random this is UVW1 with $G_W = 0$: the same pairing, the same counts, and in addition an exposed GRS subcode on half the support that makes it weaker than UVW1 (Sect. 8). With U *also* Reed-Solomon, which is the proposal of Márquez-Corbella and Tillich [8], the code is far more exposed: its columns lie on a rational normal scroll and its shortened squares have dimension at most $3(k_l - 1)$, so the plain square code distinguishes it (Sect. 9.6).

The principle. Whenever the legitimate decoder removes a mask by a linear map on a few public coordinates per secret coordinate and lands on a GRS codeword, the public code has a hidden pairing and Theorem 4.4 (and its general form, Theorem 9.3) applies with the dimension of the GRS code as k_2. The square-code attacks needed the GRS part to be exposed by *shortening alone*; the jets need only that each tuple's span meets the curve. Hiding a Goppa code instead of a GRS code removes the curve and inherits Goppa's jet degree and key size (Remark 9.10).

7 Why Square-Code Analyses Saw Nothing

The structural security of masked-GRS schemes has so far been argued through dimension counts of Schur powers of subcodes cut out by guessed pairs: this is the CLT method [13] and the method of the first structural audit of UVW-KEM. We record what such counts can and cannot do, in the UVW notation of Proposition 3.6; the RLCE case is identical with $e = n - k$.

A *configuration* is a set of j disjoint candidate pairs, c of them true and $t = j - c$ false; it cuts out $X \subseteq \mathcal{C}^{\perp}$ with $\dim X = (n-k) - j$ (constraints assumed independent). Write $E_X = X \cap E$, $e_X =$

$\dim E_X$, $b_X = \dim X - e_X$; when the false constraints are independent on E, $e_X = e - t$ and $b_X = J - c$ with $J = m - k_1$. The configuration is *accepted* if $\dim X^{\star 2} < \beta(X) := \min\bigl(n - j, \binom{\dim X + 1}{2}\bigr)$.

Theorem 7.1 (no partial signal)**.** *(1)* $\dim X^{\star 2} \le n - j$ *and* $\dim X^{\star 2} \le S(X) := \min\bigl(m, 2e - 1, \binom{e_X+1}{2}\bigr) + e_X b_X + \binom{b_X+1}{2}$*. (2) If* $S(X) < \beta(X)$ *then* $(b_X - 1)e_X + \binom{b_X}{2} < k - \min(m, 2e - 1)$*; for* $k = m$ *and* $k_1 \ge k_2$ *this forces* $b_X \le b^*$ *with* $b^* = 1$ *if* $k_1 = k_2$ *and* $b^* = 0$ *if* $k_1 > k_2$*: a square certificate requires all but at most one of the* J *independent true pairs, however many false pairs accompany them. (3) (Measured) acceptance in fact extends to about* $0.6\sqrt{e}$ *missing pairs, always with many false pairs; uniformly random codes under the same constraints are never accepted.*

Proof. (1) A constraint $y_a = -y_b$ makes every Schur product satisfy $v_a = v_b$, one codimension per pair; and $X^{\star 2} \subseteq {E_X}^{\star 2} + E_X \star X'' + {X''}^{\star 2}$ for $X = E_X \oplus X''$, with ${E_X}^{\star 2} \subseteq E^{\star 2}$ of dimension $\min(m, 2e-1)$ by Proposition 3.6. (2) If the cap does not bind, $S(X) = \binom{\dim X + 1}{2} \ge \beta(X)$; if it binds, $S(X) < n - j = k + e_X + b_X$ rearranges to the inequality, whose right side is $m - \min(m, 2e - 1) = \max(0, 2k_2 - m + 1) \in \{0, 1\}$ for $k = m$, $k_1 \ge k_2$, while the left side is $\ge e_X + 1 \ge 2$ for $b_X \ge 2$. (3) Our scans (Sect. 8); the independent audit analysis accompanying [22] reports the same threshold. □

The same bookkeeping shows that higher Schur powers $X^{\star d}$ certify a defect at no larger b_X than the square does, and counting configurations with at least $J - \lceil\sqrt{e}\rceil$ true pairs gives a Schur-filtration adversary, who learns only whether each configuration is accepted, a query lower bound above 2^{889} at every proposed UVW set. These statements are correct, and beside the point. A Schur-filtration adversary measures the dimension of products of codewords of a subcode cut out by a *guess*, which by Theorem 7.1 must already be almost entirely correct before the dimension moves. The jet adversary of Theorem 4.4 measures the rank of a system of *relations*—which forms of degree d vanish, to order r, at the columns of the *ambient* shortened code—with no guess at all: the hidden curve imposes Hermite conditions on every such form through every pair, whether or not the adversary knows which columns are paired, and the conditions are dependent as soon as they outnumber the coefficients of the composed polynomial. Dimension counts of subcodes are blind to this because the dependencies live among the *derivatives* of forms at points, an object no Schur power of a code encodes. The natural question "does the kernel of $\mathrm{Sym}^2(\mathcal{C}^\perp) \to \mathcal{C}^{\perp\star 2}$ carry a signal?" was the right instinct; the kernel that matters is one degree and one derivative up.

8 Experiments

All experiments use our own implementations of the schemes, exact linear algebra over prime fields, the `galois` package for $\mathbb{F}_{2^{10}}$, and two small C programs for the ranks (dense elimination with lazy reduction over $\mathbb{F}_p$, log tables over $\mathbb{F}_{2^\mu}$). *Excess* is the number of row dependencies beyond the generic baseline of Lemma 2.2; Δ is the bound of Theorem 4.4 for the actual shortening pattern. Controls are a uniformly random code of the same length and dimension and, where indicated, the construction with the RS block replaced by a random block ("twin control"), which isolates the curve as the cause.

UVW. Table 4 scans the split $k_1{:}k_2$ at fixed size and shortening: the excess follows $2k_1 - k_2 - k_l + 2$ exactly from $(30, 20)$ down to $(23, 27)$ and vanishes where the bound is non-positive, for UVW1 and UVW2 and across three fields. Varying k_l at $(100, 25, 25)$ over $k_l = 19, \dots, 26$ and at $n = 200$ gives the excess $k_2 - k_l + 2$ throughout the window, including an exactly square system at $n = 200$, and none below it, where the system is overdetermined and the structural dependencies are absorbed by the generic ones, as Corollary 4.7 predicts. Table 5 confirms the higher-order statements of

Table 4: Cubic gradient jets $(3, 2)$ on UVW toy keys, $n = 100$ ($m = 50$, $k = 50$), $k_l = 19$ (1311×1330, generic kernel 19), one key per entry. Excess equals $\Delta_{3,2} = 2k_1 - k_2 - k_l + 2$ in every non-degenerate instance and is zero where $\Delta \le 0$; the $(35, 15)$ key is beyond the Veronese threshold of Remark 9.5 ($k_2 = 15$ against 22 surviving curve-point columns) and shows a larger deficiency. Random-code and twin controls: 0.

(k_1, k_2)	(35,15)	(30,20)	(29,21)	(28,22)	(27,23)	(26,24)	(25,25)	(24,26)	(23,27)	(22,28)	(21,29)	(20,30)
$\Delta_{3,2}$	38	23	20	17	14	11	8	5	2	−1	−2	−3
excess, UVW1 $\mathbb{F}_{101}$	136	23	20	17	14	11	8	5	2	0	0	0
excess, UVW2	—	—	—	17	—	—	8†	—	2	—	—	—

† over $\mathbb{F}_{401}$; UVW1 at $(25, 25)$ also gives 8 over $\mathbb{F}_{211}$ and on three further seeds over $\mathbb{F}_{101}$. For $(20, 30)$ the excess is 0 at every $k_l \in \{19, \dots, 26\}$.

Table 5: Higher degree and order. Left: quartic gradient jets $(4, 2)$ ($\Delta_{4,2} = 2l - 4k_2 + 3$, baseline 0). Right: quartic order-3 jets $(4, 3)$ on a $k_1 < k_2$ key that every cubic jet misses ($n = 48$, $k_1 = 11$, $k_2 = 13$, $k_l = 21$, $n' = 45$; $10\,395 \times 10\,626$): the baseline is the $\binom{45}{2} = 990$ line relations of Lemma 2.2, attained by both controls, and the key shows $990 + \Delta_{4,3}$ with $\Delta_{4,3} = 2m + l - 4k_2 + 3 = 2$.

$(4, 2)$, $n = 100$	$(30, 20)$, k_l=11	$(30, 20)$, k_l=12	$(25, 25)$, k_l=13
$\Delta_{4,2}$	1	−1	−23
excess	1	0	0

$(4, 3)$, $(48, 11, 13)$	key	random	twin control
row dependencies	992	990	990
excess over baseline	2	0	0
$(3, 2)$ excess, $k_l = 14..21$	0	0	0

Theorem 4.4: the quartic gradient system shows the single predicted dependency at $(30, 20)$ and none where the bound is negative; and the quartic order-3 system detects a $k_1 < k_2$ key invisible to every cubic jet, with excess exactly $\Delta_{4,3} = 2$ over a baseline of 990 that both controls attain. UVW1 with $G_W = 0$ (the $(U \,|\, U{+}V)$ construction with a random U) at the same $(100, 25, 25, 19)$ parameters gave 8, 163 and 326 over three seeds against the pairing count 8: its subcode $\{(\mathbf{0} \| \mathbf{v})\}$ is a GRS code on half the support, which the shortening cuts down to a few dimensions with many columns on its curve, the situation of Remark 9.5.

RLCE. Table 6: at the long-key boundary $w = n - k$ the count is $11 = \Delta$ in both characteristics and at rate 0.86, and both controls give zero, so the curve is the cause. At full size, the rate-0.87 Level-1 set R-1 gives excess 77, the bound exactly, in one $10\,413 \times 10\,660$ rank computation with the random control at 0, and the 2017 NIST set ID 0, the Level-1 long-key set that CLT could not reach, gives excess 116, again the bound, from a $16\,836 \times 17\,296$ system in sixteen minutes. The recovery of Sect. 5 runs at this scale as well, and we carry it through to decryption in Sect. 8.1. The short-key set $(60, 30, 10)$, in CLT's regime, has 50 singleton columns against $k = 30$, beyond the Veronese threshold, and shows 89 against the bound 31; there the read-off recovers nothing (Proposition 5.3(iv)).

Recovery. Table 7 verifies Theorem 5.1 literally: block ranks, shared coefficient vectors, pairs, UVW2 ratios and RLCE curve directions are all exactly as predicted, with no spurious pairs.

8.1 End-to-end key recovery of long-key RLCE at full size

Everything above establishes the distinguisher at full size and the read-off on toys and on single rounds. To close the gap between "distinguished" and "broken" we ran Algorithm 1 in full on the rate-0.87 Level-1 set R-1 of Table 3 ($n = 846$, $k = 732$, $w = 114$, $N = 960$, $t = 57$, over $\mathbb{F}_{2^{10}}$), using

Table 6: RLCE: cubic gradient jets. Toy keys over $\mathbb{F}_{257}$ and $\mathbb{F}_{2^{10}}$; the full-size sets R-1 and ID 0 of Table 3 over $\mathbb{F}_{2^{10}}$ (R-1: key generation 14 s, shortening and jet system 14 s, rank 239 s; ID 0: 7 s, 10 s, 932 s; two cores). $\Delta_{3,2} = 2(n-k) - w - k_l + 2$.

(n,k,w)	field	k_l	rows/cols	$\Delta_{3,2}$	excess	random	hidden random	note
(60,30,30)	$\mathbb{F}_{257}$	21	1701/1771	11	11	0	0	$w = n-k$
(60,30,30)	$\mathbb{F}_{2^{10}}$	21	1701/1771	11	11	0	0	$w = n-k$
(210,180,30)	$\mathbb{F}_{2^{10}}$	21	1701/1771	11	11	0	—	rate 0.86, $w = n-k$
(60,30,10)	$\mathbb{F}_{257}$	21	1281/1771	31	89	0	—	$w < n-k$ (CLT regime; beyond threshold)
(846,732,114)	$\mathbb{F}_{2^{10}}$	39	10 413/10 660	77	**77**	0	0	R-1, full size
(630,470,160)	$\mathbb{F}_{2^{10}}$	46	16 836/17 296	116	**116**	—	—	NIST ID 0, full size

Table 7: Recovery from the dependency blocks (Theorem 5.1). UVW2 key $(q,n,k_1,k_2,k_l) = (101,100,25,25,19)$, $\Delta = 8$ dependencies; RLCE keys $(n,k,w,k_l) = (60,30,30,21)$ over $\mathbb{F}_{257}$, two seeds, $\Delta = 11$.

	UVW2	RLCE (seeds 1, 7)
surviving twin columns with rank-one block	46/46	44/44, 46/46
half-shortened columns with rank-two block	23/23	8/8, 6/6
singleton (unmixed GRS) columns with rank-two block	—	29/29, 29/29
twin pairs sharing a coefficient vector	23/23	22/22, 23/23
pairs read off from coefficient vectors (correct / spurious)	23 / 0	22 / 0, 23 / 0
ratios $\delta_j/\delta_i = -3a/b$ correct	23/23	—
hidden GRS direction $t/s = b/(3a)$ correct	—	22/22, 23/23

the public key only (`rlce_e2e.py`; the secret key is read by the script solely to score intermediate rounds, and none of its values enters the computation). Table 8 records the run.

The jet stage. The kernel dimension was $77 = \Delta_{3,2}$ in every round, so exactness (Theorem 5.1) held throughout, every pair with both coordinates surviving was read off with its correct ratio, and no false pair was accepted. The number of pairs exposed per round is what the rate predicts (Sect. 5): a random shortening of R-1 keeps 267 of 960 columns and exposed 12 pairs in round one. The adaptive ordering then found 89 pairs in 13 rounds, but slowed to one pair per round at the end, for a reason worth recording: the deterministic rule "shorten the columns that have survived most often" can lock the two columns of an unknown pair into alternating cohorts that never survive together. Randomizing the order (survival key $\sigma_c/2$ plus Gumbel noise) removes the lock but the tail remains slow, because a pair is exposed only when both of its columns are among the 28% of survivors.

The square-code stage. The tail is better finished by a different tool, and the count says which. Fold the known pairs into their curve points (Lemma 3.3): the folded code is again a hidden-pairing code, of length $N' = N - |\mathcal{T}|$ with $w' = w - |\mathcal{T}|$ pairs and $n - w'$ singletons, and by the master count its degree-two row is $\Delta_{2,1} = (n-k) - w' - k_l + 1$: as soon as $w' < n - k - k_l + 1$ the *shortened square code* of the folded code is deficient—this is exactly the regime in which CLT [13] recover short keys, entered here not by the designer's choice of w but by the attacker's progress. A square-code round costs seconds: shorten the folded code at a random Z to dimension k_l, form the Veronese images $\nu_2(\mathbf{x}_c)$ of the n' surviving columns, and compute the left kernel of the $n' \times \binom{k_l+1}{2}$ matrix

Table 8: End-to-end recovery of a long-key RLCE key at full size (R-1: $n = 846$, $k = 732$, $w = 114$, $N = 960$, $t = 57$, $\mathbb{F}_{2^{10}}$; key seed 1; scripts `rlce_e2e.py` and `rlce_endgame_sq.py`, two cores). Jet stage: each round is one $10\,413 \times 10\,660$ left-kernel computation at $k_l = 39$; "known" is the number of pairs in $\mathcal{T}$ at the start of the round, "exposed" the number of true pairs with both coordinates surviving (read from the secret key for scoring only), "new" the pairs accepted by the read-off. Rounds 1–15 use the deterministic survival ordering of Algorithm 1, rounds 16–17 a randomized one (Sect. 8.1). The kernel dimension was $77 = \Delta_{3,2}$ in every round, every exposed pair was accepted with the correct ratio, and no false pair was ever accepted. Square-code stage: shortened squares of the folded code at $k_l = 25$, ≈ 3 s per round.

stage / round	known	kernel	exposed	new	false	time (s)
jet 1	0	77	12	12	0	302
jet 2	12	77	12	12	0	489
jet 3	24	77	11	11	0	303
jet 4	35	77	4	4	0	308
jet 5	39	77	6	6	0	300
jet 6	45	77	10	10	0	297
jet 7	55	77	6	6	0	299
jet 8	61	77	4	4	0	309
jet 9	65	77	6	6	0	305
jet 10	71	77	5	5	0	306
jet 11	76	77	5	5	0	316
jet 12	81	77	8	8	0	306
jet 13	89	77	6	6	0	301
jet 14	95	77	1	1	0	297
jet 15	96	77	1	1	0	300
jet 16	97	77	3	3	0	296
jet 17	100	77	1	1	0	297
square code, 166 rounds	$101 \to 114$	—	13	13	0	584

Totals: 17 jet rounds (88 min) and 166 square-code rounds (10 min); fold matrix of rank 732; Sidelnikov-Shestakov (34 s): 846 distinct support points, nonzero multipliers, $\operatorname{rowsp} F = \mathrm{GRS}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ verified; fresh ciphertext of error weight 57 decrypted, message recovered, $\mathrm{wt}(\mathbf{c} - \mathbf{m}G_{\mathrm{pk}}) = 57$.

they form. The curve-point columns (singletons, half-shortened survivors, folded pairs) span a space of dimension $2 \deg Q + 1 < n'$, where the shortened code is $\{Z_s(x)Q(x)\}$ with Z_s the vanishing polynomial of the shortened curve points, so they are dependent; a surviving column of an unknown pair whose twin also survives is a point off the curve, and is detected as a row that occurs in no kernel vector. Two off-curve columns a, b are twins iff the plane $\operatorname{span}(\mathbf{x}_a, \mathbf{x}_b)$ meets the curve, i.e. iff $\nu_2(\mathbf{x}_a + \rho\mathbf{x}_b)$ lies in the span of the curve-point images for some ρ, a quadratic condition in ρ solved by trying the $q - 1$ values, and ρ is the ratio, which transfers to the public columns as in Theorem 5.1. Starting from the 101 pairs of the jet stage ($w' = 13$, $k_l = 25$, $\Delta_{2,1} = 77$, Veronese rank 51–75 against $n' \approx 145$) the square-code stage recovered the remaining 13 pairs in 166 rounds and ten minutes, every exposed pair accepted with the correct ratio, no false pair. The handover threshold $w' < n - k - k_l + 1$ is met much earlier: with $k_l = 21$ (the smallest k_l with $\binom{k_l+1}{2} \geq n'$) it is $w' \leq 93$, i.e. after two jet rounds; we report that experiment below.

The back end. The fold matrix has rank 732, Sidelnikov-Shestakov over $\mathbb{F}_{2^{10}}$ returns a support of 846 distinct points and nonzero multipliers with $\operatorname{rowsp} F = \mathrm{GRS}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ (verified by the rank of the

stacked generators), and a fresh ciphertext $\mathbf{c} = \mathbf{m}G_{\mathrm{pk}} + \mathbf{e}$ with $\mathrm{wt}(\mathbf{e}) = 57$ is decrypted by folding $\mathbf{c}$ along the recovered pairs, Berlekamp-Welch in the recovered GRS code, and inversion of the fold; the recovered message equals $\mathbf{m}$ and $\mathbf{c} - \mathbf{m}G_{\mathrm{pk}}$ has weight 57. The whole attack used the public key only (in this run the script reads the secret key solely to score the rounds; the strict-mode run below removes even that), ran on two cores in under two hours of wall-clock time as tabulated, and in about half an hour with the early handover below; it is reproducible from the seeds in Table 8. The same pipeline applies to the 2017 NIST set ID 0 ($N - k = 320$, so a shortening keeps 46% of the columns and the jet stage is predicted to need 7 rounds of sixteen minutes); we have not run it.

Early handover. To test the threshold we restarted the square-code stage from the 24 pairs of jet rounds 1–2 alone (the state file preserves the order of recovery), with $k_l = 21$, where $\Delta_{2,1} = 114 - 90 - 21 + 1 = 4$. It recovered the remaining 90 pairs in 208 rounds and 617 seconds—52 of them in the first 105 seconds, the rest in the coupon-collector tail—again with every exposed pair accepted and no false pair, and the back end decrypted the fresh ciphertext as before. The complete attack on R-1 therefore takes the jet rounds needed to pass 21 pairs—two to four, since a round exposes 9 ± 3 pairs—ten minutes of square-code rounds and one minute of back end: 25–35 minutes on two cores. This is the division of labour the master count prescribes: the jet is needed exactly while $w' \geq n - k - k_l + 1$, i.e. to recover the first $w - (n - k) + k_l \approx 21$ pairs of a long key, and the square code, blind at the start by CLT's theorem, does the rest once those are folded away. For a long key with $w = c(n - k)$, $c > 1$, the jet stage must bring w' below $n - k - k_l$, i.e. recover $(c - 1)(n - k) + k_l$ pairs, before the handover.

Strict mode: an independently generated key. To remove any doubt about what the attack reads, the scripts have a mode in which a separate program (`rlce_keygen_export.py`) generates a key pair and a fresh ciphertext and writes the public key and ciphertext to files, with the secret key and the plaintext under a directory the attack never opens; the attack scripts then read those two files, print no scoring, and write the decrypted message to a file that a separate command compares with the plaintext afterwards. On a new R-1 key (key seed 2) this run took four jet rounds (exposing 7, 6, 8 and 5 pairs, kernel 77 each time: two rounds give $13 < 21$ pairs, too few for the handover, the third reaches the threshold with 21, and a fourth was run for margin, to 26), 134 square-code rounds (442 s) and 48 s of back end; the recovered message equals the plaintext, compared by a separate command, and $\mathbf{c} - \mathbf{m}G_{\mathrm{pk}}$ has weight 57. Both runs, with their logs and state files, are in the scripts repository.

8.2 Reproducibility

The scripts, released with the paper: key generation for every construction (`jet.py`, `jet_mask.py` for hidden pairings of any width profile and the $(U \,|\, U{+}V)$ variants, `rlce_full.py`), the (d, r) jet systems with their predicted counts (`jet_general.py`, `jet_mask.py`), the rank programs `rankp.c`, `rank2m.c`, `nullspace2m.c`, the recovery scripts (`jet_structure.py`, `jet_rlce_structure.py`, `rlce_recover.py`, and `rlce_e2e.py` with `rlce_endgame_sq.py` for the end-to-end run of Sect. 8.1, `sim_rounds.py` for the round-count model), and the checks `forcing_law.py` (Theorem 9.2 on the (j, d, r) grid), `rnc_jets.py` (Remark 9.5) and `grs_rank1.py`. For Sect. 10: `polar_reliab.py` (underflow-free polar construction), `minweight_count.py` (minimum-weight counts, checked by enumeration), `polar_invariants.py` and `family_invariants.py` (exact constructions of polar-RLCE, LDPC-RLCE and UVW over polar and LDPC codes, primal and dual leaks, matched-pair subspace), `nogo_analysis.py` and `hirate_polarrlce.py` (dual-leak costs, parameter searches), `talvardy.py` (degrading construction, Table 14), `polar_closure.py`, `polar_closure_stern.py`

and `polar_closure_stern2.py` (Proposition 10.7, Table 15), `window_search.py` (the illusory window) and `rm_rlce2.py` (Table 16). Every table entry records its seed; all runs used two CPU cores.

9 The Jet Degree of a Masked Code

The cubic jet broke UVW and RLCE because their hidden curve is reached at degree three. This section explains what sets that degree, proves that it is governed by one count, and derives from the count the law that ties structural depth to redundancy. Sections 4–6 are the width-two case of what follows.

Definition 9.1 (jet degree)**.** Let $\mathcal{C} \subseteq \mathbb{F}_q^N$. Its *jet degree* $\delta(\mathcal{C})$ is the least d for which there exist r with $1 \le r \le d$ and a shortening Z such that the (d, r) jet system of $\mathcal{C}_{(Z)}$ has rank strictly below its value on a uniformly random code of the same length and dimension; equivalently, the least degree at which a higher-order vanishing test separates $\mathcal{C}$ from random. The *hierarchy jet degree* $\delta_H(\mathcal{C})$ is the same with $r = d - 1$ only. We set $\delta = \infty$ if no such test exists.

For a uniformly random code $\delta = \infty$ with probability $1 - o(1)$. The square-code distinguisher is the case $(d, r) = (2, 1)$; the cubic jet of this paper is $(3, 2)$; the Goppa attacks of [15, 30, 21] are higher d. Smaller δ means a cheaper attack: a (d, r) test at shortening dimension k_l is a rank computation on a system with $\binom{k_l+d-1}{d}$ columns, polynomial in N for fixed d and quasi-polynomial for $d = \Theta(\log N)$.

9.1 The forcing law

Every scheme in Sect. 3 hides each Reed-Solomon coordinate in the span of a few public columns: one for an unmasked GRS coordinate, two for a mixed pair. Call this number the *mask width* j. The decisive local question is which derivatives of a form at the hidden point inside the span of a tuple are *forced* by the jets at the tuple's columns. For a statement valid in every characteristic we use Hasse derivatives: for a multi-index $\beta \in \mathbb{N}^{k_l}$, $D^{(\beta)}\mathbf{x}^\alpha = \binom{\alpha}{\beta}\mathbf{x}^{\alpha-\beta}$ with $\binom{\alpha}{\beta} = \prod_t \binom{\alpha_t}{\beta_t}$, extended linearly; the *order-r jet* of f at $\mathbf{x}$ is $(D^{(\beta)}f(\mathbf{x}))_{|\beta|<r}$, and f vanishes to order $\ge r$ at $\mathbf{x}$ iff this jet is zero. This is the multiplicity of f at $\mathbf{x}$. Whenever $d, d-1, \dots, d-r+2$ are nonzero in $\mathbb{F}_q$, the ordinary derivatives of order $r-1$ of Definition 2.1 span the same functionals (Euler, Sect. 2), so nothing changes in any system computed in this paper; the general statements below are for Hasse jets. For a tuple of linearly independent columns $\mathbf{x}_1, \dots, \mathbf{x}_j \in \mathbb{F}_q^{k_l}$ with span Λ, a point $Q = \sum_a c_a \mathbf{x}_a$ is *genuinely mixed* if all $c_a \ne 0$, and $K_{d,r}$ denotes the space of degree-d forms vanishing to order $\ge r$ at every $\mathbf{x}_a$.

Theorem 9.2 (forcing law)**.** *Let $\mathbf{x}_1, \dots, \mathbf{x}_j$ be linearly independent, Q genuinely mixed in their span Λ, and $y_1, \dots, y_j$ the coordinates on Λ dual to the basis $(\mathbf{x}_a)$.*

(i) (Monomial description.) *The restriction of $K_{d,r}$ to Λ is the span of the monomials y^α with $|\alpha| = d$ and $\alpha_a \le d - r$ for every a, and each such monomial, read as a form in the variables $y_1, \dots, y_j$ alone, lies in $K_{d,r}$.*

(ii) (Forced order.) *Put $\omega := jr - (j-1)d$. For every β with $|\beta| < \omega$ the functional $f \mapsto D^{(\beta)}f(Q)$ vanishes on $K_{d,r}$ and therefore lies in the span of the order-r jet functionals at $\mathbf{x}_1, \dots, \mathbf{x}_j$. The evaluation $f \mapsto f(Q)$ lies in that span iff $\omega > 0$, i.e. iff $jr > (j-1)d$.*

(iii) (Hierarchy and extremes.) *For $r = d-1$, $\omega = d-j$: the hidden point is forced at degree d iff $d \geq j+1$, and at $d = j$ the obstruction is one form, $K_{j,j-1}|_\Lambda = \mathbb{F}_q\, y_1 \cdots y_j$, the product of the j hyperplanes of Λ spanned by $j-1$ of the columns (the* simplex form*), which is nonzero at every genuinely mixed point. For $r = d$ every functional is forced, but $\mathsf{J}_{d,d}$ has rank N_d on every point set spanning $\mathbb{F}_q^{k_l}$, so no (d,d) test distinguishes anything.*

Proof. (i) Change coordinates on $\mathbb{F}_q^{k_l}$ so that $\mathbf{x}_a = \mathbf{e}_a$ for $a \leq j$. For a monomial, $D^{(\beta)}\mathbf{x}^\alpha(\mathbf{e}_a) = \binom{\alpha}{\beta}$ if $\alpha_b = \beta_b$ for all $b \neq a$ (and $\alpha_a \geq \beta_a$), and 0 otherwise. Take β with $\beta_a = 0$ and $|\beta| \leq r-1$: the only monomial of degree d with $\alpha_b = \beta_b$ for $b \neq a$ is $\alpha(\beta) = \beta + (d - |\beta|)\mathbf{e}_a$, with $\binom{\alpha(\beta)}{\beta} = 1$, so the condition $D^{(\beta)}f(\mathbf{e}_a) = 0$ reads $c_{\alpha(\beta)} = 0$. As β ranges over these multi-indices, $\alpha(\beta)$ ranges over all exponents with $\sum_{b\neq a}\alpha_b \leq r-1$; the conditions with $\beta_a > 0$ involve only coefficients already forced to vanish. Hence $K_{d,r}$ is spanned by the monomials with $\sum_{b\neq a}\alpha_b \geq r$ for every $a \leq j$: a monomial subspace. Monomials involving a variable x_b with $b > j$ restrict to 0 on $\Lambda = \{x_b = 0,\ b > j\}$; the others restrict to y^α with $d - \alpha_a \geq r$, i.e. $\alpha_a \leq d - r$, for every a. Conversely such a y^α, read in $x_1, \ldots, x_j$, satisfies the condition at every $\mathbf{e}_a$. (ii) Let $f \in K_{d,r}$ and $|\beta| = s$. In the monomial description $D^{(\beta)}$ sends $\mathbf{x}^\alpha$ to a multiple of $\mathbf{x}^{\alpha-\beta}$ with $\sum_{b\neq a}(\alpha_b - \beta_b) \geq \sum_{b\neq a}\alpha_b - s$, so $g := D^{(\beta)}f$ has degree $d-s$ and vanishes to order $\geq r-s$ at every $\mathbf{x}_a$; if $s < \omega$ then $s < r$, and (i) applied to g puts $g|_\Lambda$ in the span of the y^α with $|\alpha| = d-s$ and $\alpha_a \leq (d-s)-(r-s) = d-r$. Such monomials exist iff $j(d-r) \geq d-s$, i.e. iff $s \geq \omega$; so for $s < \omega$, $g|_\Lambda = 0$ and in particular $g(Q) = 0$. A functional vanishing on the common kernel of a family of functionals lies in their span (finite dimension), which gives the membership. If $\omega \leq 0$ there is a monomial y^α with $|\alpha| = d$, all $\alpha_a \leq d-r$; it lies in $K_{d,r}$ by (i) and $y^\alpha(Q) = \prod_a c_a^{\alpha_a} \neq 0$, so $f \mapsto f(Q)$ does not vanish on $K_{d,r}$. (iii) $r = d-1$ gives $\omega = jd - j - (j-1)d = d - j$. At $d = j$ the admissible exponents are $|\alpha| = j$, $\alpha_a \leq 1$: only $y_1 \cdots y_j$, and y_a is the hyperplane of Λ through the columns other than $\mathbf{x}_a$. At $r = d$ a form with multiplicity d at $\mathbf{x}$ is, in coordinates with $\mathbf{x} = \mathbf{e}_1$, a form in $x_2, \ldots, x_{k_l}$ alone; if the points span $\mathbb{F}_q^{k_l}$ no nonzero form is free of all of them, so the kernel of $\mathsf{J}_{d,d}$ is 0 for the structured and the random code alike. □

The j columns of a tuple span $\mathbb{P}^{j-1}$, hence *are* a projective frame, the fat-point ideal is monomial, and the independence of j points of multiplicity j in $\mathbb{P}^{j-1}$ on degree-$(j{+}1)$ forms is a count of exponent vectors; Alexander-Hirschowitz, which concerns more points than can be normalized, is needed only for the random baseline (Lemma 2.2). We verified Theorem 9.2, kernel dimension and forcing alike, on all 210 triples (j, d, r) with $j \leq 6$, $d \leq 8$, $1 \leq r \leq d$ over $\mathbb{F}_{101}$ (`forcing_law.py`, no mismatch). For a tuple of which j' columns survive a shortening the forced order reads d for $j' = 0$ (a d-fold zero of the composed polynomial), r for $j' = 1$ (the column is the curve point), $2r - d$ for $j' = 2$ (the twin line), $3r - 2d$ for $j' = 3$: the one expression $(j'r - (j'-1)d)^+$ covers every case of Theorem 4.4.

9.2 The master count

Definition 3.1 extends verbatim to tuples of any width: a *hidden pairing of width profile* (j_ℓ) consists of the injective M, distinct $\alpha_1, \ldots, \alpha_n$, and pairwise disjoint tuples $T_\ell \subseteq [N]$ with $|T_\ell| = j_\ell$ such that $\Gamma(\alpha_\ell)$ is a genuinely mixed combination of $\{\mathbf{g}_c : c \in T_\ell\}$; n is the length of the hidden GRS code, $\rho := n - k_2$ its redundancy, and $R := N - k$ the public redundancy. Columns in no tuple are *free*. After shortening at Z, $\Pi_Z\Gamma(\alpha_\ell)$ is the same combination of the surviving projected columns of T_ℓ, the shortened ones having become $\mathbf{0}$: the projected point is genuinely mixed in the span of the $j'_\ell := |T_\ell \setminus Z|$ survivors. We assume, as holds for a generic shortening with $k_l \geq \max j_\ell$, that these survivors are linearly independent.

Theorem 9.3 (master count)**.** *Let $\mathcal{C}$ have a hidden pairing of width profile (j_ℓ) and rank k_2, shortened at Z to dimension k_l with survivor counts (j'_ℓ). Then $\mathsf{J}_{d,r}$ of $\mathcal{C}_{(Z)}$ has at least*

$$\Delta_{d,r} := \sum_{\ell=1}^{n} \omega_\ell - d(k_2-1) - 1, \qquad \omega_\ell = \begin{cases} \left(j'_\ell\, r - (j'_\ell - 1)\, d\right)^+ & j'_\ell \geq 1, \\ d & j'_\ell = 0, \end{cases}$$

linearly independent row dependencies supported on structured rows, provided $d \cdot \#\{\ell : j'_\ell = 0\} \leq d(k_2-1)+1$ (as in every case considered). For width profiles in $\{1,2\}$ this is Theorem 4.4.

Proof. The proof of Theorem 4.4 goes through with Theorem 9.2(ii) in place of Lemma 4.2. Let $Q_\ell = \Pi_Z \Gamma(\alpha_\ell)$. For a tuple with $j'_\ell \geq 1$ survivors, the functionals $f \mapsto D^{(\beta)} f(Q_\ell)$, $|\beta| < \omega_\ell$, are combinations of the tuple's own rows, and so are the Hermite functionals $f \mapsto F_f^{(s)}(\alpha_\ell)$, $s < \omega_\ell$, of the composed polynomial $F_f(\alpha) = f(\Pi_Z \Gamma(\alpha))$ (chain rule for Hasse derivatives); for $j'_\ell = 0$ the point $\Pi_Z \Gamma(\alpha_\ell)$ is $\mathbf{0}$ and F_f has a d-fold zero at α_ℓ. The composed polynomials live in a space of dimension at most $d(k_2-1)+1-d\cdot\#\{\ell : j'_\ell = 0\}$, on which the $\sum_{j'_\ell \geq 1} \omega_\ell$ Hermite functionals satisfy at least $\Delta_{d,r}$ independent relations (Sect. 2); each relation is a dependency among rows, and relations are independent because the row combinations of distinct tuples have disjoint supports. □

Corollary 9.4 (hierarchy count)**.** *For $r = d-1$, $\omega_\ell = (d - j'_\ell)^+$ and*

$$\Delta_{d,d-1} = \sum_{\ell} (d - j'_\ell)^+ - d(k_2-1) - 1 \;\geq\; d(\rho+1) - R - k_l + f' - 1,$$

where f' is the number of surviving free columns, with equality whenever $d \geq \max_\ell j_\ell$. For a pure width-j masking without free columns,

$$\Delta_{j,j-1} = l - j(k_2-1) - 1, \qquad \Delta_{j+1,j} = n + l - (j+1)(k_2-1) - 1,$$

and for $d \leq j$ only tuples with fewer than d survivors contribute.

Proof. $\sum_\ell (d - j'_\ell)^+ \geq \sum_\ell (d - j'_\ell) = dn - \sum_\ell j'_\ell$, and $\sum_\ell j'_\ell = n' - f' = N - l - f' = R + k_l - f'$; equality holds iff no tuple has more than d survivors. For pure width j and $d = j$ each tuple contributes its number of shortened columns, which sum to l; at $d = j+1$ each contributes one more. □

For UVW ($N = 2m$, $k = k_1 + k_2$, $n = m$) the corollary reads $\Delta_{2,1} = k_1 - k_2 - k_l + 1$, $\Delta_{3,2} = m + k_1 - 2k_2 - k_l + 2$ (which is $2k_1 - k_2 - k_l + 2$ when $k = m$) and $\Delta_{4,3} = 2m + k_1 - 3k_2 - k_l + 3$; for RLCE ($n - w$ singletons, w pairs) it reads $\Delta_{d,d-1} = (d-1)(n-k) + (d-1) - w - k_l$, whose $d = 3$ case is Corollary 4.6 and whose $d = 2$ case, $n - k - w - k_l + 1$, is positive exactly in the regime $w < n - k$ in which CLT recover short keys [13] (up to the slack k_l; measured on RLCE-like toys with $(n,k) = (60,30)$: excess 11 at $w = 10$, $k_l = 10$, and 0 at $w = 20$, $k_l = 11$, as predicted): the square-code attack and the cubic jet are the $d = 2$ and $d = 3$ rows of one count.

Remark 9.5 (exactness and its boundary). $\Delta_{d,r}$ is a lower bound. It is attained when the composed map $f \mapsto F_f$ is onto the polynomials of the stated degree, which holds for a generic shortening, *and* when the surviving columns that are themselves projected curve points—survivors of half-shortened pairs, RLCE singletons—are not too many: the order-2 jets of cubics at m points of a rational normal curve of degree $k_2 - 1$ are independent only up to $m \approx 1.4\, k_2$ and span at most $2(k_2-1)^2 + 2$ dimensions however many points are taken (measured for $k_2 \in \{10, 12, 15, 18, 20, 25, 30\}$, `rnc_jets.py`). Beyond that *Veronese threshold* the jets at the curve-point columns carry dependencies of their own and the deficiency *exceeds* the count. Every full-size instance of this paper has $k_2 \geq 215$ (UVW) or $k \geq 470$ (RLCE) against at most a few hundred such columns and met the count exactly; the toy instances with $k_2 \leq 15$, or with a GRS subcode of small dimension after shortening, cross the threshold and show the larger deficiency (Tables 4, 6; Sect. 9.6). This characterizes the "degenerate" instances.

9.3 Algebraic-geometry codes: HermitianRLCE

The master count used the hidden code only through one fact: the composed functions $F_f(\alpha) = f(\Pi_Z \Gamma(\alpha))$ lie in a space of bounded dimension, namely the polynomials of degree $\le d(k_2 - 1)$, because the hidden curve Γ is *rational* of degree $k_2 - 1$. Nothing else about rationality is used; the twin line of Lemma 4.2 is a statement about lines in projective space and is blind to the genus of the curve a column sits on. We make this explicit and apply it to the one masked-GRS variant built on a higher-genus curve.

Let $\mathcal{X}/\mathbb{F}_q$ be a smooth projective curve of genus g, P a rational point, and $V \subseteq L(\alpha P)$ a k-dimensional space of functions with $\alpha < n$; the *algebraic-geometry (AG) code* $\mathcal{A} = \mathcal{C}(D, \alpha P)$ evaluates V at n rational points $D = P_1 + \cdots + P_n$, and its columns $\Gamma(P_\ell) = (\mathrm{ev}_{P_\ell} f)_{f \in V} \in \mathbb{F}_q^k$ trace the image of $\mathcal{X}$ under the map $\phi_{\alpha P} : \mathcal{X} \to \mathbb{P}^{k-1}$, a curve of degree α and genus g. For $2g - 2 < \alpha$, $k = \alpha + 1 - g$ by Riemann-Roch. A *hidden pairing on an AG code* is Definition 3.1 with Γ this curve in place of the rational normal curve.

Theorem 9.6 (AG master count)**.** *Let $\mathcal{C}$ have a hidden pairing of width profile (j_ℓ) whose hidden code is an AG code $\mathcal{C}(D, \alpha P)$ of genus g with $2g - 2 < d\alpha$. Shortened at Z to dimension k_l, $\mathsf{J}_{d,r}$ of $\mathcal{C}_{(Z)}$ has at least*

$$\Delta^{\mathrm{AG}}_{d,r} := \sum_{\ell=1}^{n} \omega_\ell - (d\alpha + 1 - g) \quad \textit{row dependencies on structured rows,}$$

with ω_ℓ as in Theorem 9.3. Equivalently, writing $\alpha = k + g - 1$, $\Delta^{\mathrm{AG}}_{d,r} = \Delta_{d,r} - (d-1)g$ relative to the genus-0 count at the same (n, k, w, k_l).

Proof. The proof of Theorem 9.3 is reproduced with one substitution. For a degree-d form f and a column $\Gamma(P_\ell)$, the composed object $F_f = f \circ \phi_{\alpha P}$ is a product of d functions of $L(\alpha P)$, hence an element of $L(d\alpha P)$, of dimension $d\alpha + 1 - g$ for $d\alpha > 2g - 2$. The twin line (Lemma 4.2) still places, for each tuple with $j'_\ell \ge 1$ survivors, the Hermite functionals $f \mapsto F_f^{(s)}(P_\ell)$, $s < \omega_\ell$ (derivatives along the curve at P_ℓ), in the row span; across all tuples these are $\sum \omega_\ell$ functionals on a space of dimension $\le d\alpha + 1 - g$, so they satisfy at least $\Delta^{\mathrm{AG}}_{d,r}$ independent relations, each a dependency among rows of disjoint support. The identity $d\alpha + 1 - g = d(k-1) + 1 + (d-1)g$ follows from $\alpha = k + g - 1$. □

The bound is again a lower bound; as in Remark 9.5 it is exceeded when the curve-point columns are many, which for an AG curve of degree $\alpha > k - 1$ they always are, so the measured excess runs well above Δ^{AG} (Table 10). The point is that the genus only enlarges the composed space by $(d-1)g$, a quantity independent of n, while $\sum \omega_\ell$ grows with n: no fixed genus closes the gap.

HermitianRLCE. Replace the GRS code of RLCE (2) by the one-point Hermitian code over $\mathbb{F}_{q^2}$ on the curve $y^q + y = x^{q+1}$, of genus $g = q(q-1)/2$ and length $n = q^3$, keeping the transform (insert w random columns, mix w pairs by secret $A_i \in \mathrm{GL}_2(\mathbb{F}_{q^2})$, scramble and permute); this is the scheme of [29]. By Proposition 3.2(iii) with the Hermitian curve in place of the rational one it is a hidden-pairing code of rank k with $n - w$ singletons and w pairs, so Theorem 9.6 applies with $\alpha = k + g - 1$.

Corollary 9.7 (HermitianRLCE at degree three)**.** *A HermitianRLCE key has cubic-jet excess at least*

$$\Delta^{\mathrm{AG}}_{3,2} = 2(n-k) - w - k_l + 2 - 2g.$$

Table 9: HermitianRLCE parameters [29] ($q = 16$, $\mathbb{F}_{2^8}$, genus $g = 120$, $n = q^3$ truncated to the length shown, $\alpha = k + g - 1$) and the provable cubic-jet excess $\Delta^{\mathrm{AG}}_{3,2} = 2(n-k) - w - k_l + 2 - 2g$; $k_l \approx \sqrt{6(N-k)}$ is the window of Corollary 4.7.

Level	n	k	w	α	k_l	$\Delta^{\mathrm{AG}}_{3,2}$ (lower bound)
128	635	280	22	399	48	≥ 402
192	870	421	33	540	54	≥ 573
256	1090	531	45	650	61	≥ 774

Table 10: HermitianRLCE cubic-jet $(3, 2)$ measurements in characteristic two (the cubic jet needs $3 \neq 0$, so $q \in \{4, 8\}$; script `herm_jet2.py`, `herm_big.py`, ranks by `rank2m`). "excess" is the row dependencies beyond the generic baseline (zero for $(3, 2)$); Δ^{AG} is the lower bound of Theorem 9.6. Both controls—a random code of the same size, and the *same transform applied to a random hidden code*—give 0 in every case, so the excess is the Hermitian curve.

field	g	n	k	w	k_l	excess	Δ^{AG}
$\mathbb{F}_{16}$	6	64	23	8	22	242	42
$\mathbb{F}_{16}$	6	64	23	4	22	318	46
$\mathbb{F}_{16}$	6	64	23	12	22	166	38
$\mathbb{F}_{64}$	28	200	88	16	36	1452	118
$\mathbb{F}_{64}$	28	200	66	16	34	2196	164

This is positive at every parameter set of [29] (Table 9), so HermitianRLCE is distinguished from random at degree three, and the recovery of Sect. 5 applies with the hidden Hermitian code recovered in place of the GRS code (the fold of Lemma 3.3 yields an AG code, structurally recovered by [12] rather than Sidelnikov-Shestakov).

At the $\mathbb{F}_{64}$ row with $k/n = 0.44$—the rate of HermitianRLCE-128—the excess is 1452 on a length-200 instance and grows with n, while both controls are exactly 0: the signal is not a small-field artifact and does not shrink at cryptographic rate. The square code, by contrast, is defeated for $w \geq n - k$ exactly as for GRS-RLCE ($\dim \mathcal{A}^{\star 2} = 2k - 1 + g$ for a Hermitian code, still $< \binom{k+1}{2}$ but reached only below the long-key boundary); it is the degree-three jet, through the curve, that breaks the long-key regime here as it does for RLCE. HermitianRLCE therefore buys the key-size advantage of a higher-genus trapdoor without buying any jet resistance: the genus adds $(d-1)g$ to the composed dimension, a constant, and the mask width stays two.

9.4 Jet degree is a rate statement

Corollary 9.4 says what the width does and does not determine. Below degree $j+1$ a width-j tuple is invisible unless the shortening has cut it down to fewer than d columns; at degree $j+1$ every tuple contributes. So the hierarchy jet degree of a pure width-j masking is $j+1$ exactly when the *shortening budget* $l = k - k_l$ is too small to expose enough tuples at lower degree and the tuples are numerous enough at degree $j+1$:

$$\delta_H = j + 1 \quad \text{iff} \quad \sum_{\ell:\, j'_\ell < d} (d - j'_\ell) \leq d(k_2 - 1) + 1 \ \text{ for } 2 \leq d \leq j,$$

and $n + l > (j+1)(k_2 - 1) + 1$. The first condition holds whenever $l \leq 2k_2 - 1$, and for $j = 2$ it is exactly $k_1 - k_2 < k_l$. Both conditions hold for the *balanced* designs, $k \approx jk_2$ and $n \approx 2k_2$, which is

where UVW ($j = 2$) and the natural width-j generalizations live. Outside that regime the degree moves:

- UVW with $k_1 - k_2 \geq k_l$ has $\delta_H = 2$: the half-shortened pairs alone expose more than $2k_2 - 1$ curve points and the *plain shortened square code* distinguishes the key. This is the case for the three unbalanced sets 128b, 192b, 256b of [33] ($k_1 - k_2 = 40, 60, 82$ against $k_l = 27, 33, 38$): the $(2, 1)$ system at full size has excess 14, 28 and 45, the predicted $k_1 - k_2 - k_l + 1$ exactly, for UVW1 and UVW2 alike, while the balanced 128a set and random controls give 0 (systems 363×378, 531×561, 714×741 at $k_l = 27, 33, 38$, script `jet_mask.py`). The security analysis of [33] is a weight-distribution argument and does not consider shortened squares.
- UVW with $k_1 \leq k_l - 2$ (a tiny U-part) has $\Delta_{3,2} < 0$: the cubic jet is blind and the key falls at degree 4 instead ($\Delta_{4,3} = k_1 + k_2 - k_l + 3$ for $m = 2k_2$). Measured at $k_1 = k_2 = 10$, $m = 20$, $k_l = 20$: no excess at $(2, 1)$, $(3, 2)$; at $(4, 3)$ the random control has exactly its $\binom{40}{2} = 780$ line relations and the key 783.
- A width-3 masking ($k_1 = 16$, $k_2 = 3$, 12 tuples, $N = 36$, $k_l = 19$; and $k_1 = 22$, $k_2 = 3$, 16 tuples, $N = 48$, $l = 2$) is blind at degrees 2 and 3 and distinguished at $(4, 3)$ by exactly the count on top of the baseline attained by the random control ($630 + 3$ and $1035 + 9$): the first full-code demonstration of a jet degree above three.

"$\delta = j + 1$" is the balanced-rate specialization of one count.

9.5 The redundancy law

The hierarchy count has a consequence that does not mention widths at all.

Theorem 9.8 (redundancy law)**.** *Let $\mathcal{C}$ have a hidden pairing whose tuples cover all N columns, hiding a GRS code of redundancy $\rho = n - k_2$, and let $R = N - k$. For every $d \geq 2$ and every shortening to dimension k_l satisfying the independence assumption of Sect. 9.2,*

$$\Delta_{d,d-1} \;\geq\; d(\rho + 1) - R - k_l - 1,$$

with equality whenever d is at least the largest width. Hence if some degree-d hierarchy system of $\mathcal{C}$ carries no structured dependency, then $R \geq d(\rho + 1) - k_l - 1$; and if $\delta_H(\mathcal{C}) \geq D$, then

$$R \;\geq\; (D - 1)(\rho + 1) - k_l - 1.$$

Proof. The inequality is Corollary 9.4 with $f' = 0$; the equality case is the one stated there. If the degree-d system has no structured dependency then $\Delta_{d,d-1} \leq 0$ by Theorem 9.3, and the inequality gives the bound on R. If $\delta_H \geq D$, the degree-$(D-1)$ system has no excess over the generic baseline; for $D - 1 \leq 3$ the baseline is zero and this means no structured dependency, while for larger D one uses in addition that structured dependencies are not absorbed by the pairwise line relations (verified at $(4, 3)$, Sect. 8; open in general, Sect. 11). □

Corollary 9.9 (design principle)**.** *A hidden-pairing trapdoor of hierarchy jet degree D pays public redundancy at least $(D-1)$ times the redundancy $\rho = 2t$ of the GRS code it hides, up to the attacker's shortening slack k_l, which at the window $\binom{k_l+d-1}{d} \approx n'\binom{k_l+d-3}{d-2}$ is $k_l \approx \sqrt{d(d-1)R}$, a lower-order term. The bound is tight: balanced width-$(D-1)$ designs have $R = (D-1)\rho$ and $\delta_H = D$, as UVW and RLCE's long keys do at $D = 3$.*

The theorem is about the hidden-pairing model, decoders that undo the mask by a linear map on a few public coordinates per hidden coordinate. Within it there is no compact, jet-resistant GRS trapdoor: every unit of structural depth is paid for in redundancy, whatever the widths, the mixing coefficients or the field, and the k_1:k_2 discussion of Sect. 6 becomes one line—UVW with $k = m$ has $R = m$, so it is blind at degree 3 only if $m \geq 3(m - k_2 + 1) - k_l - 1$, i.e. $k_2 \geq (2m + 2 - k_l)/3$, the boundary found there. Maskings outside the model (public columns in the span of several curve points, as in BBCRS for $m \geq 1 + \mathrm{R}$; subfield subcodes; Goppa) obey the law's shape empirically—binary Goppa pays mt bits for t errors, $m/2$ times the binary Singleton minimum, for a jet degree $\Theta(m)$—but we do not claim it for them.

9.6 The landscape of masked GRS codes

We now pass the masked-GRS literature through the count. For each construction we state its structure in the language of hidden pairings, the degree the count assigns, and the status of that assignment (Table 11); all toy measurements quoted below are over $\mathbb{F}_{101}$ with `jet_mask.py` and `rankp`, with random controls at 0 wherever run.

Width one. A plain GRS code [26] is the case $j = 1$: $\Delta_{2,1} = n - k - k_l + 1$, positive for every rate. A *Berger-Loidreau subcode* [2], a random k'-dimensional subcode of an $[n, k]$ GRS code, has columns $T\gamma(\alpha_c)$ for a $k' \times k$ matrix T: width one on a *projected* curve of degree $k - 1$, and the count with $k_2 = k$, $l = k' - k_l$ gives $\Delta_{2,1} = n + k' - k_l - 2k + 1$, Wieschebrink's square-code attack [31]. *Wieschebrink's random columns* [32] are free columns, inert in the count ($\Delta_{2,1} = n + l_T - 2k + 1$, l_T the shortened GRS columns), so $\delta = 2$ whenever $n > 2k - 1$ [14, 9].

BBCRS. The public key is $S^{-1}GQ^{-1}$ with $Q = T + R$, T sparse of average row weight m and R of rank z [1]. For $m = 1$, $z = 1$ the public code is a GRS code plus a rank-one perturbation, $\mathbf{g}_c = \gamma(\alpha_c) + \lambda_c \mathbf{v}$, a k-dimensional subcode of $\mathrm{GRS} \oplus \langle \boldsymbol{\lambda} \rangle$ whose shortenings inherit the shape, so $\dim \mathcal{C}_{(Z)}{}^{\star 2} \leq 3k_l$ (measured $3k_l - 1$ at $(n, k) = (60, 20)$ and $(80, 20)$, `grs_rank1.py`) against the generic $\min(n', \binom{k_l+1}{2})$: $\delta = 2$ as soon as $2k_l < n - k$, the attack of [9]. For $1 < m < 1 + \mathrm{R} + O(n^{-1/2})$ (R the rate) the dual public code $\mathcal{C}_{\mathrm{GRS}}^{\perp}(T^\top + R^\top)$ has, at the weight-one rows of T, width-one columns of the dual curve, and [11] recover the key in $O(n^6)$; the count explains why the attack stops where the weight-one rows run out. For $m \geq 1 + \mathrm{R}$ every dual column lies in the span of *several* curve points: a secant masking, outside the model, with open jet degree. The weight-two mask of Lieb, Mazumder and Schaller [39] (every public column a combination of two curve points, every curve point used twice) is the simplest secant masking; they prove $\dim \mathcal{C}^{\star 3} \leq 2k^2 - 4k + 4$ against the generic $\binom{k+2}{3}$, which is a $(3, 1)$ test in our notation, and recover the key when $n > 2k^2 - 4k + 4$, i.e. at rates so high that the public dual has dimension $O(\sqrt{n})$; at the rates of this paper the cube is generic, and whether a higher-order jet reaches secant maskings is the question of Sect. 11.

RLCE. $n - w$ singletons and w pairs: $\Delta_{d,d-1} = (d - 1)(n - k + 1) - w - k_l$, so $\delta_H = 2 + \lfloor (w + k_l)/(n - k + 1) \rfloor$: 2 for $w < n - k - k_l$ (the short keys, CLT [13]), 3 for the long keys $w = n - k$ (Sect. 6, measured at full size), $\approx c + 2$ for $w = c(n - k)$, with $R = (1 + c)\rho$ in agreement with the redundancy law.

UVW. Width two with a U-part of dimension k_1: $\delta_H = 2$ if $k_1 - k_2 \geq k_l$ (the unbalanced sets of [33]), 3 if $k_1 - k_2 < k_l < 2k_1 - k_2 + 2$ (the balanced and NGCC sets), 4 if $k_1 \leq k_l - 2$; UVW2's monomial scaling changes nothing in the count.

$(U\,|\,U{+}V)$ **variants.** Three constructions carry this name and behave differently. (a) *Both U and V Reed-Solomon on a common support*, the generalized $(U\,|\,U{+}V)$ codes $\{(\mathbf{u}D_1 + \mathbf{v}D_2 \,\|\, \mathbf{u}D_3 + \mathbf{v}D_4)\}$ with diagonal D_i proposed by Márquez-Corbella and Tillich [8]: every public column is $(a\gamma_U(\alpha)\,;\, b\gamma_V(\alpha))$ for some (a, b, α), a point of the *rational normal scroll* $S(k_U - 1, k_V - 1) \subset \mathbb{P}^{k-1}$, a surface of degree $k - 2$. The shortening projection projects from l points *on* the scroll, so (for a generic shortening) the surviving columns lie on a surface of degree $k_l - 2$ spanning $\mathbb{P}^{k_l - 1}$, a variety of minimal degree, whose degree-two Hilbert function is $3(k_l - 1)$. Hence $\dim \mathcal{C}_{(Z)}{}^{\star 2} \le 3(k_l - 1)$ for every shortening, with equality in all our measurements (at $(k_U, k_V) = (35, 15)$, $n = 50$: square dimension $30, 42, 57, 72$ at $k_l = 11, 15, 20, 25$, i.e. excess $31, 23, 13, 3$, and 0 once $3(k_l - 1) \ge n'$): the construction has $\delta = 2$ for every choice of rates and of D, by the plain shortened square code at any $k_l < (N - k + 3)/2$. This appears to be new; [8] states that the structure "seems to avoid" Sidelnikov-Shestakov and leaves the analysis open. (b) *U and V Reed-Solomon on permuted supports*, which destroys the scroll: a width-two pairing with *two* hidden curves in each twin plane. Both are forced by the twin line, and the counts add: $\Delta_{2,1} = k_U - k_V - k_l + 1$ from the V-curve, so $\delta = 2$ in the regime $k_U \gg k_V$ that [8] prescribes for decoding, and $\delta = 3$ with the sum of the two cubic counts otherwise (measured $14 = 7 + 7$). (c) *U random, V Reed-Solomon*, i.e. UVW1 with $G_W = 0$: the pairing and its counts are those of UVW1, but the subcode $\{(\mathbf{0}\|\mathbf{v})\}$ is a GRS code on half the support which a shortening reduces to a few dimensions with many columns on its curve, so the measured deficiency is often far above the count (Remark 9.5; Sect. 8): the construction is weaker than UVW1, not stronger.

Twisted Reed-Solomon codes. The McEliece variant of [4] evaluates a k-dimensional space of polynomials of degree up to $n - k$: width one on a projected rational normal curve of degree about n, for which the count is negative at every small degree. The scheme was broken all the same, by Lavauzelle and Renner [18] through a subfield subcode: a large jet degree is necessary for security, not sufficient.

Subspace subcodes, expanded codes, Goppa. SSRS [3] and the expanded GRS codes of [17] are subfield-type objects: the public code over $\mathbb{F}_q$ is the image of an $\mathbb{F}_q$-subspace of dimension $\lambda \le m$ of a GRS code over $\mathbb{F}_{q^m}$. They interpolate between GRS ($\lambda = m$) and alternant codes ($\lambda = 1$), and Couvreur and Lequesne [10] distinguish them by a *twisted* square product when $\lambda > m/2$; the hierarchy over the public field is not the right test there. At the far end, binary Goppa codes have no pairing and no curve over $\mathbb{F}_2$: scalar extension of the public dual splits into $m = \log_2 q$ Frobenius-twisted GRS blocks [25], a single hidden curve becomes a Frobenius orbit of m of them, and the distinguisher that reaches it is quasi-polynomial, $n^{O(\log n)}$ [15, 30], with key recovery in [19, 25] and its linear algebra accelerated in [40]: $\delta = \Theta(\log n)$, which is why Classic McEliece stood for forty years and why the 2026 attacks are not yet practical [6]. We do not reprove these results; we place them as the high-depth end of the spectrum.

Remark 9.10 (masking an alternant code gains nothing cheaply). Replacing the hidden GRS code of UVW or RLCE by a Goppa or alternant code keeps the twin line but moves the composed object into the Frobenius-split space of [25], on which the degree-three count is generic: the construction is immune to the cubic jet exactly to the extent that the underlying Goppa code is, and inherits its jet degree $\Theta(\log n)$ and its key size. The pairing lowers the jet degree only when the code it hides is itself a low-degree curve.

Table 11: The landscape of GRS-based maskings. "Structure" is in the language of hidden pairings; δ is the degree assigned by the master count (Corollary 9.4) or, where the count does not apply, the known status; R/ρ is public redundancy over the redundancy of the hidden code ($\rho = 2t$; for Goppa, bits over the binary Singleton minimum). P = proved here (count), M = measured here, C = cited.

Construction	structure	δ	R/ρ	status; attack
GRS [26]	width 1	2	1	P; [26]
Berger-Loidreau subcodes [2]	width 1, projected curve	2 unless $2k-1 \geq n+k'-k_l$	> 1	P, C; [31]
Wieschebrink random columns [32]	width 1, free columns	2	$1+\frac{w}{\rho}$	P, C; [14, 9]
BBCRS m=1, z=1 [1]	GRS + rank 1	2	1	M, C; [9]
BBCRS $1 < m < 1+$R	dual: width-1 columns, secants	2 (dual)	1	C; [11]
BBCRS $m \geq 1+$R; weight-2 mask	secant masking	open; cube code at $n > 2k^2$	1	C; [39]
RLCE, $w < n-k-k_l$ [27]	widths 1, 2	2	< 2	P, C; [13]
RLCE, $w = n-k$	widths 1, 2	3	2	P, M (full size)
HermitianRLCE [29]	width 2, genus-g curve	3	2	P, M; here
UVW, $k_1-k_2 \geq k_l$	width 2	2	2	P, M (full size)
UVW balanced, NGCC	width 2	3	2	P, M; [22]
$(U\|U+V)$-RS, common support [8]	scroll	2	≈ 2	P, M
$(U\|U+V)$-RS, permuted supports	width 2, two curves	2 if $k_U-k_V \geq k_l$, else 3	≈ 2	P, M
Width-j masking, balanced	width j	$j+1$	j	P, M ($j \leq 3$)
Twisted RS [4]	width 1, curve of degree $\approx n$	count negative	—	C; [18]
SSRS, expanded GRS [3, 17]	subspace subcode	twisted square, $\lambda > m/2$	—	C; [10]
Binary Goppa	m Frobenius blocks	$\Theta(\log n)$	$m/2$	C; [15, 30, 25]

10 Beyond the Curve: Dual Leaks and the Limits of Masking

The jet attack needs a curve. Everything in Sections 4–9 uses the hidden code only through the bounded dimension of the space its composed forms live in—$d(k_2-1)+1$ for a rational normal curve, $d\alpha+1-g$ for an algebraic curve of genus g—so the natural escape is to hide a decodable code that is *not* the evaluation code of a curve. Two binary families qualify: *polar* codes [34], capacity-achieving under successive-cancellation (SC) decoding, which polarRLCE [43] places under the RLCE mask; and *LDPC/MDPC* codes [35, 44], decoded by belief propagation or bit flipping. Neither has a bounded composed space, and over $\mathbb{F}_2$ the cubic jet is degenerate in any case ($x^2 = x$ collapses the exponents it separates); the jet master count predicts the generic count, as measured. This section shows that both masks nevertheless fail on these codes, for a reason independent of curves: *they preserve the dual of the hidden code.* Throughout, $\mathrm{GV}(N, K)$ denotes the Gilbert-Varshamov distance of an $[N, K]$ code. Since the two masks are linear in the hidden code, everything below

holds over any $\mathbb{F}_q$; we state costs for $q = 2$, the case of interest.

10.1 Dual preservation

Lemma 10.1 (dual preservation)**.** *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be the hidden code, with generator G, and let $\mathbf{h} \in \mathcal{C}^\perp$.*

(i) RLCE. *With w inserted columns, mixed positions $\mathcal{M}$ and secret $A = \mathrm{diag}(I, A_0, \ldots, A_{w-1})$, the vector $(\mathbf{h}, \mathbf{0})$ (zero on the inserted columns) is a dual word of $[G \mid R]$, and its image in the dual of the public code has weight $\mathrm{wt}(\mathbf{h}) + Y \le 2\,\mathrm{wt}(\mathbf{h})$, where Y is the number of mixed pairs whose original column lies in $\mathrm{supp}\,\mathbf{h}$ and whose block $A_i^{-\mathsf{T}}$ maps $(1,0)$ to a weight-two vector. For A_i uniform in $\mathrm{GL}_2(\mathbb{F}_2)$, $Y \sim \mathrm{Bin}(X, 1/3)$ with $X = |\,\mathrm{supp}\,\mathbf{h} \cap \mathcal{M}|$; over $\mathbb{F}_q$ with all entries of A_i nonzero, as in RLCE proper, $Y = X$.*

(ii) UVW. *With the secret generator $G_{\mathrm{sk}} = \left[\begin{smallmatrix} G_U & G_U \\ G_V & G_W \end{smallmatrix}\right]$, $G_V - G_W = G$, the vector $(\mathbf{h}, -\mathbf{h})$ is a dual word of G_{sk}, for every k_1 and every choice of G_U, G_W; after the secret permutation (for UVW2, the secret monomial transformation, which rescales coordinates) it is a public dual word of weight exactly $2\,\mathrm{wt}(\mathbf{h})$, supported on $\mathrm{wt}(\mathbf{h})$ secret pairs.*

Proof. (i) $[G \mid R](\mathbf{h}, \mathbf{0})^\mathsf{T} = G\mathbf{h}^\mathsf{T} = 0$. A dual word of $[G \mid R]A$ is the image of a dual word of $[G \mid R]$ under $A^{-\mathsf{T}}$, which is the identity off $\mathcal{M}$ and acts on each mixed pair, whose second coordinate is an inserted column where $(\mathbf{h}, \mathbf{0})$ vanishes. A pair whose original column is outside $\mathrm{supp}\,\mathbf{h}$ maps $(0,0)$ to $(0,0)$; one inside maps $(h_j, 0)$ to a nonzero vector of weight one or two. Scrambling does not change the code and the permutation preserves weight. Over $\mathbb{F}_2$, the first column of $A_i^{-\mathsf{T}}$ is uniform on the three nonzero vectors, one of which has weight one in each coordinate and one weight two. (ii) $G_U\mathbf{h}^\mathsf{T} - G_U\mathbf{h}^\mathsf{T} = 0$ and $G_V\mathbf{h}^\mathsf{T} - G_W\mathbf{h}^\mathsf{T} = G\mathbf{h}^\mathsf{T} = 0$. □

Remark 10.2 (width-b masks). Part (i) uses only that the mask appends columns on which the hidden dual vanishes and then applies an invertible map to disjoint blocks of at most two columns. For blocks of at most b columns—the width-b maskings of Sect. 9.4—the same argument gives a public dual word of weight at most $b\,\mathrm{wt}(\mathbf{h})$ for every $\mathbf{h} \in \mathcal{C}^\perp$. Width buys jet degree on the primal side (Theorem 9.2: a width-j tuple is blind below degree $j+1$) and costs at most a factor b on the dual side; no mask of width b lifts a hidden dual of minimum distance $d^\perp$ above $b\,d^\perp$.

For a GRS hidden code the lemma is vacuous: the dual of an $[n,k]$ GRS code is MDS of minimum distance $k+1$, and $2(k+1)$ is far above the Gilbert-Varshamov distance of the public dual. That is why masked GRS needed the jet. For a hidden code with a short dual it is an attack: the public dual contains the images of all minimum-weight dual words of $\mathcal{C}$, and recovering a few of them recovers the hidden structure (their supports are affine for polar codes [36] and give a sparse parity check for LDPC codes [45]).

Corollary 10.3 (cost of the dual leak)**.** *Let $d^\perp$ be the minimum distance of $\mathcal{C}^\perp$ and $N_{d^\perp}$ the number of its minimum-weight words. The public key is attacked at the cost of finding one leaked word,*

$$\min_D \Big[\,\mathrm{SD}(N, N-k, D) - \log_2\big(N_{d^\perp} \cdot \Pr[\,d^\perp + Y \le D\,]\big)\Big],$$

where $\mathrm{SD}(N, K, D)$ is the cost of finding a weight-D codeword in an $[N, K]$ code (for UVW, $Y = d^\perp$ deterministically). When $2d^\perp < \mathrm{GV}(N, N-k)$ the leaked words are moreover below the minimum distance of a random dual, so they are an information-theoretic distinguisher as well.

Cost methodology. Every cost in this section is computed the same way, so that it can be reproduced from the inputs (N, K, D) alone. *Estimator.* Syndrome-decoding costs are the `SDEstimator` of `CryptographicEstimators` [38] (version 2.1.1, the release installed with the scripts), with its default classical bit-operation model, no memory-access penalty and no quantum speed-up; "Prange" is the closed form $\log_2 \binom{N}{K} - \log_2 \binom{N-D}{K-1} - \log_2 D + 3\log_2 N$ (the inverse probability that a random information set meets a given weight-D word in exactly one coordinate, times a cubic elimination) for finding one word of weight D in an $[N, K]$ code, and "Stern family" is the minimum over Prange, Stern, Dumer and Ball-Collision; the BJMM/May-Ozerov family is run at the surviving parameter sets and reported where it changes a verdict (it never does by more than 0.3 bits). *Inputs.* For a dual leak the code is the public dual, $[N, N-k]$, and D is the leaked weight; for message recovery it is the public code, $[N, k]$, and $D = t$. *Multiplicity.* Where stated ("multiplicity credited") the cost of finding one leaked word is the estimator cost at weight $D = d^\perp + y$ minus $\log_2\big(N_{d^\perp} \Pr[Y \le y]\big)$, minimized over y, with $N_{d^\perp}$ the number of minimum-weight dual words from Proposition 10.4; where not credited the cost is at $D = 2d^\perp$ (UVW) or at $D = d^\perp$ and $D = 2d^\perp$ (RLCE, both reported), which can only overstate it. *Distribution of Y.* For RLCE with A_i uniform in $\mathrm{GL}_2(\mathbb{F}_2)$, $Y \sim \mathrm{Bin}(X, 1/3)$ with $X = |\operatorname{supp}\mathbf{h} \cap \mathcal{M}|$ hypergeometric with parameters $(n, w, d^\perp)$ (Lemma 10.1); the leak law $\mathbb{E}[Y] = d^\perp(w/n)/3$ was checked on 40 real keys. *Targets.* Category 1, 3 and 5 are 143, 207 and 272 bits of classical cost, as in the NIST round-4 submissions, with no margin; a set "fails" when either the leaked-dual-word cost or the message-recovery cost is below target under the Stern family. The scripts named in Sect. 8.2 print the inputs and the per-algorithm estimates for every table entry.

The operative quantity is the cost, not the comparison with GV: BIKE [41] also has sub-GV dual words and is secure because finding them costs 2^λ. For polar and LDPC hidden codes at the rates where the masks are meant to operate, that cost is small.

10.2 Polar codes: one side is always short

Let $n = 2^m$ and let $\mathcal{C}(I)$ be spanned by the rows of $F^{\otimes m}$, $F = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$, indexed by I; row i has weight $2^{\mathrm{pc}(i)}$, pc the binary weight. A polar code takes I to be the k most reliable synthetic channels, a decreasing set [37]; let $F = [n] \setminus I$ be the frozen set.

Proposition 10.4 (two-sided bound)**.** *For every index set I, $d(\mathcal{C}(I)) \le 2^p$ with $p = \min_{i \in I} \mathrm{pc}(i)$, $d(\mathcal{C}(I)^\perp) \le 2^{m-p'}$ with $p' = \max_{i \in F} \mathrm{pc}(i)$, and*

$$d \cdot d^\perp \;\le\; 2n.$$

For polar (decreasing) codes the first two bounds are equalities [37], and the number of minimum-weight codewords of $\mathcal{C}(I)$ and of $\mathcal{C}(I)^\perp$ is $\sum_f 2^{\,r+|\lambda_f|}$ over the top-degree monomials f [37].

Proof. A generator row of weight 2^p bounds d. $F^{\otimes m}$ is an involution over $\mathbb{F}_2$, so $\mathcal{C}(I)^\perp$ is spanned by the rows of $(F^{\otimes m})^{\mathsf{T}}$ indexed by F; since $F^{\mathsf{T}} = JFJ$ with J the swap, $(F^{\otimes m})^{\mathsf{T}} = J^{\otimes m} F^{\otimes m} J^{\otimes m}$ and these are, up to the coordinate reversal $J^{\otimes m}$, the rows of $F^{\otimes m}$ indexed by the bit-complements of F, of weight $2^{m-\mathrm{pc}(i)}$; a coordinate permutation preserves weights (and the reversal, a translation of $\mathbb{F}_2^m$, is in fact an automorphism of every decreasing-monomial code), so this bounds $d^\perp$. Take $i \in I$ with $\mathrm{pc}(i) = p \ge 1$ and clear one of its bits; the result has binary weight $p - 1 < p$, so it is not in I, hence it is frozen and $p' \ge p - 1$. Then $d\, d^\perp \le 2^p 2^{m-p+1} = 2n$ (for $p = 0$, $d = 1$). □

We checked both equalities and the counting formula against exhaustive enumeration in all 94 polar codes of length 16, 32 and 64 with primal or dual dimension at most 18, computing the dual as the actual left kernel (`minweight_count.py`, `dual_distance.py`). The construction uses

the BEC(1/2) reliability order, computed in the logit domain: the textbook recursion underflows for $m \geq 10$ (883 of the 4096 channels tie at $Z = 1.0$ in double precision for $m = 12$), and our order agrees with 1500-digit arithmetic at $m = 12$ (`polar_reliab.py`).

Corollary 10.5 (what the dual leak requires of a decreasing set)**.** *Let $\mathcal{C}(I)$ be a decreasing-monomial code of length 2^m under RLCE, and let D_λ be the smallest weight at which finding a codeword of that weight in the public dual $[N, N-k]$ costs 2^λ (multiplicity not credited). Resisting the dual leak requires $2d^\perp \geq D_\lambda$, hence every index of binary weight above p' to lie in I with $2^{m-p'+1} \geq D_\lambda$: the information set must contain every monomial of degree at most $r_0 = m - p' - 1$, i.e.*

$$\mathcal{C}(I) \;\supseteq\; \mathrm{RM}(r_0, m), \qquad 2^{r_0+2} \;\geq\; D_\lambda.$$

Crediting the $N_{d^\perp}$ leaked words only sharpens the requirement.

Proof. Lemma 10.1 leaks every minimum-weight dual word at weight at most $2d^\perp$; $d^\perp = 2^{m-p'}$ by Proposition 10.4; and $\{i : \mathrm{pc}(i) > p'\}$ is the index set of the monomials of degree at most $m - p' - 1$. □

By Proposition 10.4 a polar code is short on at least one side, $\min(d, d^\perp) \leq \sqrt{2n}$, while the GV distances of the public code and its dual grow linearly. For a polar code with the reliability order, Corollary 10.5 is met only at high rate, because the order admits indices of low binary weight long before it has exhausted those of high weight; a Reed-Muller code meets it at rate 1/2 with $r_0 = \lfloor m/2 \rfloor$. We treat polar codes first and return to Reed-Muller-heavy sets in Sect. 10.3. The RLCE mask treats the two sides differently.

The primal side. A public codeword that vanishes on all of $\mathcal{M}$ vanishes on both inputs of every pair; this is w conditions from the inserted columns and w from the original mixed columns, so these codewords form a subcode of dimension $\geq k - 2w$ that is a shortening of $\mathcal{C}(I)$, supported on the unmixed positions. For $w < k/2$ it is nonzero and, being a subcode of a decreasing-monomial code shortened at few positions, contains minimum-weight words with affine supports—the BCDOT signature—independently of the Schur square, which is the only invariant the original polarRLCE analysis [43] examined. Exact enumeration (Table 12) shows the leak at small w and none once $w \geq k/2$. The dimension bound is sufficient for closure, not necessary: a minimum-weight word $\mathbf{m}G$ whose message does not annihilate the inserted columns—all but a 2^{-w} fraction—carries $\mathbf{m}R$ on them, about $w/2$ ones before mixing and about $2w/3$ after it for uniform A_i, so it reaches the public code at weight $\approx d + w/2$ or more and is no lighter than a random word once $d + w/2 \geq \mathrm{GV}(N, k)$; exact survivors number $N_d\, 2^{-w}(1 - w/n)^d$. The primal side is therefore closed well below $w = k/2$ at cryptographic sizes. In every surviving parameter set of Sect. 10.3 the dual cost forces w higher than this anyway.

The dual side. No choice of w helps: by Lemma 10.1 every minimum-weight dual word leaks at weight $d^\perp + Y$, $Y \sim \mathrm{Bin}(X, 1/3)$. On real masked keys ($m = 10$, $k = 780$, $w = 390$, 40 keys) the mean of Y is 4.05 against the predicted $d^\perp(w/n)/3 = 4.06$, and some words leak at exactly $d^\perp$. At low and moderate rate $d^\perp$ is 8–64, and the leaked words are cheap to find.

Proposition 10.6 (polarRLCE at rate at most 2/3 falls to its dual)**.** *Consider polarRLCE with the BEC(1/2) construction, $n \leq 2^{14}$, $0.1 \leq k/n \leq 2/3$ and $\max(n - k, \lceil k/2 \rceil) \leq w \leq n$ (so that the primal side and the square code are closed).*

(i) The public dual contains words of weight at most $2d^\perp < \mathrm{GV}(N, N-k)$, with $d^\perp \leq 64$.

Table 12: The primal side of polarRLCE, by exact enumeration (`polar_invariants.py`): minimum distances of the public code and of a random code of the same size, and the number of minimum-weight public words that avoid the mask and have affine support.

n	k	w	w vs. $k/2$	d_{pub}	d_{rand}	affine survivors
16	12	4	$<$	4	3	2
32	18	4	$<$	4	5	4
32	20	2	$<$	4	4	7
32	16	8	$=$	8	8	0
32	20	12	$>$	7	7	0
32	16	16	$>$	10	10	0

Table 13: The dual leak of polarRLCE at rate $\leq 2/3$, $w = n - k$. These sets meet the primal condition $w \geq k/2$ and the CLT threshold $w \geq n-k$, and reach their ISD targets (148.8, 213.7, 286.8 bits), yet the leaked dual words are found far below target. Costs from `CryptographicEstimators` without multiplicity credit (`nogo_analysis.py`).

n	k	w	d	$d^\perp$	$\text{GV}(N, N-k)$	find at $d^\perp$	find at $2d^\perp$
4096	2458	1638	8	32	505	58.5	93.9
4096	2048	2048	16	16	380	47.1	68.3
8192	4096	4096	16	16	758	50.0	71.1

(ii) At the parameter sets of Table 13, which reach their ISD targets of 143*,* 207 *and* 272 *bits, a leaked word is found in* 47–94 *bits.*

(iii) On the grid $k/n \in \{0.10, 0.12, \ldots, 0.66\}$*, five values of* w *per rate, no point reaches* 143 *bits against the dual leak: the strongest has* 140.6 *bits under Prange with multiplicity credit, which the Stern family lowers to* 122.2 *bits (*$n = 16384$*,* $k/n = 0.64$*,* $w = n$*, a* 27.9 *MiB key).*

Proof. (i) Lemma 10.1(i) and Proposition 10.4, with $d^\perp$ computed exactly from the frozen set at each of the 725 grid points; since $\text{GV}(N, N-k)$ decreases in N for fixed k, the worst case is $w = n$. Below $k/n \approx 0.07$ the dual of the public code has so little redundancy that the leaked words are no longer below GV; such rates are of no cryptographic use. (ii)–(iii) Corollary 10.3 priced with `CryptographicEstimators`, minimum over Prange, Stern, Dumer and Ball-Collision; in (ii) the multiplicity $N_{d^\perp}$ is not credited, which only overstates the cost (`nogo_analysis.py`, `hirate_polarrlce.py`). □

10.3 Decreasing sets with a heavy dual

By Corollary 10.5 the only way past the dual leak is an information set that contains every monomial of degree at most r_0 with $2^{r_0+2} \geq D_\lambda$. There are two ways to meet it, and they fail for different reasons.

Polar codes at high rate. For the reliability order, $d^\perp \geq 128$ first occurs near rate 0.76 at $n = 8192$ and $d^\perp \geq 256$ near rate 0.80 at $n = 16384$. There the primal side is closed, the leaked dual words are expensive, and a search over $m \leq 14$ with the error weight set to half the capacity limit $t_{\max} = \lfloor nH_2^{-1}(1 - k/n) \rfloor$ finds parameter sets passing both the dual and the message-recovery cost under the Stern family, with keys of 3.1, 15.4 and 20.1 MiB at Categories

Table 14: Upper bounds on the SC failure probability ($\log_2$) of a polar code of length 8192 on BSC(p), information set chosen for that channel, by the Tal-Vardy degrading construction. An operating point at $t \approx t_{\max}/2$ lies far to the lower right of the certifiable region.

p ($t = pn$)	$R = 0.50$	0.60	0.70	0.76	0.80	0.84	0.88
0.002 (16)	−194.6	−140.1	−83.2	−61.3	−48.1	−33.6	−21.9
0.004 (32)	−155.2	−107.4	−61.8	−43.5	−31.7	−21.0	−11.9
0.006 (49)	−131.7	−88.0	−49.2	−33.1	−22.8	−14.0	−6.3
0.008 (65)	−115.6	−74.5	−40.6	−25.8	−16.9	−9.2	−2.6
0.012 (98)	−92.9	−55.7	−28.6	−16.1	−9.1	−3.0	> 0
0.016 (131)	−76.5	−43.4	−20.4	−9.8	−4.0	> 0	> 0
0.020 (163)	−64.3	−34.6	−14.3	−5.3	−0.5	> 0	> 0

1, 3 and 5 (`window_search.py`). This window is illusory: the error weight it assumes cannot be decoded. Table 14 gives rigorous upper bounds on the successive-cancellation failure probability of a polar code on BSC(p), from the degrading construction of Tal and Vardy [46] with $\mu = 128$ output symbols (`talvardy.py`), for the information set the designer would choose at that channel. At rate 0.76 the bound is 2^{-61} at $p = 0.002$, i.e. $t = 16$ errors at $n = 8192$, and a failure rate of 2^{-32} requires $t \le 49$; at rate 0.84, 2^{-32} requires $t \le 16$. With so few errors message recovery is immediate.

Proposition 10.7 (polarRLCE: no secure parameters up to $n = 16384$; 58 MiB at $n = 32768$)**.** *Let $n = 2^m$, $12 \le m \le 15$, and consider polarRLCE with the information set of the k best synthetic channels of* BSC(p) *under the degrading construction, error weight $t = \lfloor pn \rfloor$, $p \in \{0.002, 0.003, \ldots, 0.03\}$ and $k/n \in \{0.40, 0.42, \ldots, 0.94\}$ (the grids of `polar_closure.py`), $1 \le w \le n$, and a certified failure bound of at most 2^{-32}.*

(i) For $m \le 14$ no w makes both the leaked-dual-word cost and the message-recovery cost at least 2^{143} under the Stern family. For $m \le 13$ this holds already under Prange with the multiplicity $N_{d^\perp}$ credited: the largest dual cost compatible with 143-bit message recovery is 44 bits at $n = 4096$ and 95 bits at $n = 8192$ ($k/n = 0.50$, $p = 0.03$, $w = n$). At $n = 16384$ fifteen grid points pass the Prange filter, with keys of 19–29 MiB; re-priced with the Stern family, the largest dual cost compatible with 143-bit message recovery among them is 129.7 bits ($k/n = 0.64$, $p = 0.016$, $w = n$, a 27.9 MiB key).

(ii) At $n = 32768$, 100 grid points pass the Prange filter and 35 of them survive the Stern family (Table 15). The cheapest has $k/n = 0.64$, $p = 0.004$ ($t = 131$), certified failure bound 2^{-163}, $w = 11548$ and a 58.4 MiB key, 234 times that of Classic McEliece 348864, with 147 bits against the dual leak (2^{25} words of weight 128, leaking at weight 136) and 143 against message recovery; the full estimator, BJMM family included, moves neither cost by more than 0.3 bits.

Together with Proposition 10.6, polarRLCE is excluded at every rate up to $n = 16384$—below rate ≈ 0.7 the dual leaks at its ISD-target parameters; above it, no certified error weight supports the message-recovery target without exposing the dual—and at $n = 32768$ it survives the attacks of this paper only with keys above 58 MiB.

Proof. The message cost decreases and the dual cost increases in w, so a parameter set is feasible iff the dual cost at the largest w with message cost $\ge 2^{143}$ is itself $\ge 2^{143}$; this is checked at every grid point with the Prange formula, which overestimates the attacker's cost on both sides, so an empty Prange result excludes (`polar_closure.py`). The Prange survivors are re-priced with the Stern family on both sides, the largest w recomputed under Stern and the leaked weight $D =$

Table 15: polarRLCE at $n = 32768$: the cheapest parameter set surviving the Stern family for each error rate p (35 of the 100 Prange survivors do; `polar_closure_stern2.py`). "DFR": certified upper bound on the SC failure probability; "dual": cost of a leaked dual word with multiplicity credited; "msg": message recovery; key in MiB and as a multiple of Classic McEliece 348864 [42] (0.249 MiB).

p	k/n	t	DFR	$d^\perp$	$\log_2 N_{d^\perp}$	w	N	dual	msg	key	×CM
0.004	0.64	131	2^{-163}	128	25	11548	44316	147	143	58.4	234×
0.003	0.86	98	2^{-39}	256	21	16177	48945	232	143	69.8	280×
0.008	0.50	262	2^{-228}	64	22	27900	60668	144	143	86.5	347×
0.01	0.48	327	2^{-220}	64	25	32768	65536	157	154	93.4	375×
0.012	0.52	393	2^{-165}	64	24	32768	65536	150	193	98.5	396×
0.006	0.72	196	2^{-82}	128	22	32768	65536	233	150	118.0	474×

$d^\perp + y$ minimised over y with the multiplicity $N_{d^\perp} \Pr[Y \le y]$ credited (`polar_closure_stern.py`, `polar_closure_stern2.py`); for $m \le 14$ all fail. Both thresholds favour the designer: an IND-CCA KEM at Category 1 is normally held to a failure rate near 2^{-128}, and the BJMM family lowers every cost further. For the survivors at $n = 32768$ the dual cost credits the minimum-weight dual words, and these dominate: at the cheapest set each unit of leaked weight costs 1.05 bits under Stern (167.4 bits at weight 136, 238.8 at 204), so dual words of weight 192 would need 2^{71} times the multiplicity of those of weight 128 to compete, whereas in a Reed-Muller code the words of weight $3d/2$ number about $N_d\, 2^{2(m-r)}$—here 2^{39}—and a decreasing-monomial code is no richer in this range. The failure bound is for Bernoulli(p) errors; a fixed error weight $t' \le t$, as seen by the hidden decoder after the weight contraction of the mask, is covered by the bound at BSC(t'/n), which is no larger, up to the factor $\Pr[\mathrm{Bin}(n, t'/n) = t']^{-1} \le 3\sqrt{t}$; this changes no verdict. Two limits are explicit: the mask pairs each inserted column with one original column, so $w \le n$; a wider mask (Remark 10.2) raises the leaked weight by at most the width b at a key of $k(bn - k)$ bits, above 45 MiB for $b = 3$ at $n = 16384$. And at $n = 65536$ the systematic part of the key alone, $k(n - k)$ bits, exceeds 75 MiB at every rate below 0.82, so we stop at $m = 15$. □

Reed-Muller-heavy sets. A Reed-Muller code $\mathrm{RM}(r, m)$ meets Corollary 10.5 at rate 1/2 with $r = \lfloor m/2 \rfloor$ and $d^\perp = 2^{r+1}$, and so does any decreasing set containing it. Table 16 prices the RLCE mask over $\mathrm{RM}(r, m)$ with $w = \max(n - k, \lceil k/2 \rceil)$. Up to $n = 4096$ the dual leaks for $r \le 6$ (the 2^{34}–2^{44} minimum-weight dual words make up for their weight), and $\mathrm{RM}(7, 12)$, which resists it at 208 bits, would need $t \ge 105$ errors of a capacity limit of 122. At $n = 8192$, $\mathrm{RM}(6, 13)$ and $\mathrm{RM}(7, 13)$ resist the dual leak at 191 and 232 bits; the primal side is closed (a minimum-weight word gains about $2w/3$ on the inserted columns, above the GV distance of the public code); shortened squares of the public code and of its dual are generic on toy instances ($\mathrm{RM}(4, 8)$, $\mathrm{RM}(4, 9)$, $\mathrm{RM}(5, 9)$, 72 tests, no mismatch with random); and there is no curve for the jet. With polarRLCE at $n = 32768$ (Table 15) they are the configurations in this paper that no structural attack reaches at a plausible decoding radius, and by far the cheaper. What they require is a decoder: message recovery at 143 bits needs $t \ge 215$ errors for $\mathrm{RM}(6, 13)$ and $t \ge 119$ for $\mathrm{RM}(7, 13)$, 3.4 and 3.8 times the bounded-distance radius $(d-1)/2$. Reed-Muller codes achieve capacity under MAP decoding [47, 48], so these weights are within reach of an ideal decoder (for a polar code of the same length and rate 1/2, Table 14 certifies $t = 163$); but the efficient decoders we know—recursive list decoding [49], demonstrated at

Table 16: Reed-Muller codes under RLCE, $w = \max(n-k, \lceil k/2 \rceil)$ (`rm_rlce2.py`). "dual": cost of a leaked dual word with multiplicity credited (Prange / Stern family); t_{req}: error weight at which message recovery costs 2^{143}; $t_{\text{BD}} = (d-1)/2$; $t_{\max}$: capacity limit.

code	k	rate	d	$d^\perp$	$\log_2 N_{d^\perp}$	N	key	dual (P/S)	t_{req}	t_{BD} / $t_{\max}$	status
RM$(5,11)$	1024	0.50	64	64	37	3072	0.25 MiB	107/91	216	31 / 225	dual leak
RM$(6,11)$	1486	0.73	32	128	34	2791	0.23 MiB	127/111	118	15 / 96	dual leak
RM$(5,12)$	1586	0.39	128	64	44	6606	0.95 MiB	137/119	319	63 / 619	dual leak
RM$(6,12)$	2510	0.61	64	128	42	5682	0.95 MiB	158/140	152	31 / 310	dual leak
RM$(7,12)$	3302	0.81	32	256	38	5747	0.96 MiB	228/208	105	15 / 122	needs $t \geq 105$ (86% of capacity)
RM$(6,13)$	4096	0.50	128	128	50	12288	4.00 MiB	212/191	215	63 / 901	needs a decoder at $t \geq 215$
RM$(7,13)$	5812	0.71	64	256	47	11098	3.66 MiB	253/232	119	31 / 417	needs a decoder at $t \geq 119$

Table 17: LDPC-RLCE (exact construction, `family_invariants.py`): minimum image weight of the $n-k$ sparse checks in the public dual.

n	k	ω	w	min image weight	GV$(N, N-k)$	checks leaking at ω
64	32	6	4	6	8	26/32
64	32	6	32	6	7	10/32
128	64	8	4	8	15	58/64
128	64	8	64	8	13	9/64

lengths of a few hundred, and recursive projection–aggregation [50], whose cost grows as n^r—do not reach order 6 at length 8192. Whether such a decoder exists is the question that decides Reed-Muller-RLCE; its key would be 3.7–4.0 MiB, 15× Classic McEliece. The Minder-Shokrollahi attack on the Sidelnikov (Reed-Muller McEliece) scheme [51] and the Chizhov-Borodin attack [52] use the primal minimum-weight words and the Schur filtration $\text{RM}(a) \star \text{RM}(b) = \text{RM}(a{+}b)$; under the mask the former are closed by the inserted columns and the latter is the square-code family, which is generic here.

10.4 LDPC and MDPC codes

For an LDPC code the dual object is the defining sparse parity check, of weight ω, so Lemma 10.1 applies with $d^\perp \leq \omega$ and at least $n-k$ minimum-weight dual words. Under RLCE the sparse checks leak at weight ω for every w we tried up to $w = n-k$ (Table 17); at $n = 4096$, $\omega = 15$ ($n = 8192$, $\omega = 21$) finding one costs 34.9 (44.2) bits. Raising ω to MDPC density $\approx \sqrt{n}$ makes the leaked words expensive, as in BIKE, but the mask then works against the scheme: it lengthens the code while the decodable error weight stays that of the hidden MDPC code, so generic decoding becomes cheap. With the BIKE-L1 code as hidden code ($r = 12323$, $\omega = 142$, $t = 134$) under RLCE with $w = n-k$, finding a leaked check costs 228 bits but message recovery costs only 102.4 bits, for a 36.2 MiB key; under the same estimator the unmasked code costs 152.9 bits to decode and 147.1 bits to find a secret check, with a 1.5 KiB key. The mask multiplies the key by 2.4×10^4 and removes 50 bits of message security. An MDPC code is better left unmasked.

10.5 UVW with any hidden code

The UVW mask plants a signature of its own, independent of the hidden code.

Theorem 10.8 (matched-pair subspace)**.** *Let $G_{\text{pk}} = SG_{\text{sk}}P$ be a UVW key over $\mathbb{F}_2$ whose hidden generator $G \in \mathbb{F}_2^{k_2 \times n_2}$ has column rank k_2, $N = 2n_2$, $k = k_1 + k_2$. There is a perfect matching $\mathcal{P}$ of the public columns such that the n_2 differences $\mathbf{c}_a + \mathbf{c}_b$, $(a,b) \in \mathcal{P}$, span the k_2-dimensional*

Table 18: RLCE- and UVW-type masking of the decodable code families. "Short side": the exploitable object of the hidden code.

Hidden code	Short side	Mask	Status	Reference
GRS	none (MDS); hidden curve	RLCE, $w < n-k$	broken: square code	[13]
GRS	none (MDS); hidden curve	RLCE, $w \geq n-k$	broken: cubic jet	Cor. 4.6
GRS	none (MDS); hidden curve	UVW	broken: cubic jet, matched pair	Sect. 6, Thm. 10.8
Hermitian AG	none; genus-g curve	RLCE	broken: cubic jet	Cor. 9.7
Goppa / alternant	none; subfield subcode	RLCE, UVW	no gain over the Goppa code itself	Rem. 9.10
polar, rate $\leq 2/3$, $n \leq 16384$	dual ($d^\perp \leq 64$)	RLCE (any $w \leq n$)	broken: dual leak	Prop. 10.6
polar, small w	primal (affine)	RLCE	broken: minimum-weight words	Table 12
polar, rate $> 2/3$, $n \leq 16384$	dual heavy; decoder weak	RLCE	excluded: no certified error weight meets both costs	Prop. 10.7
polar, $n = 32768$	—	RLCE, $w \geq 11548$	not excluded; key ≥ 58 MiB (234× McEliece)	Table 15
RM(r, m), $r \approx m/2$	none found	RLCE, $w \geq n-k$	not excluded; needs a decoder at 3.4× BD radius; key ≥ 3.7 MiB	Table 16
polar	dual	UVW	broken: matched pair; dual leak	Thm. 10.8, Lem. 10.1
LDPC	dual (sparse checks)	RLCE, UVW	broken: dual leak	Table 17
MDPC	dual (moderate)	RLCE	dominated: 36 MiB key, 102-bit ISD vs. BIKE	Sect. 10

subspace $U = S(\{0\}^{k_1} \times \mathbb{F}_2^{k_2})$; modulo U the columns collapse to n_2 points of multiplicity two. A uniform $[N, k]$ code has no such matching with probability $1 - o(1)$ once $k_1(n_2 - k_2) \gg N \log N$, and U contains at least n_2 of the $\binom{N}{2}$ column differences against $\binom{N}{2}2^{-k_1}$ for a generic k_2-subspace.

Proof. Columns i and $i + n_2$ of G_{sk} are $(G_U[:, i]; G_V[:, i])$ and $(G_U[:, i]; G_W[:, i])$, with difference $(0; G[:, i])$; these span $\{0\} \times \operatorname{colsp} G$, of dimension k_2, and S, P act as stated. For a uniform code a fixed matching and subspace capture all n_2 differences with probability $2^{-k_1 n_2}$, against $2^{O(N \log N) + k_1 k_2}$ pairs (matching, subspace). A generic k_2-subspace contains a fixed nonzero vector with probability 2^{-k_1}. □

On exact keys with polar and LDPC hidden codes, $N \in \{64, 128\}$, the matched differences span exactly k_2 dimensions against k for random pairs, the square code shows no defect, and U contains 44 (resp. 64) column differences against an expected 7.9 (resp. 0.1). Every proposed UVW parameter set takes k_1 a constant fraction of k, so the anomaly $2^{k_1}/N$ is overwhelming. For the degenerate regime $k_1 = O(\log N)$ the dual leak still applies: by Lemma 10.1(ii) the words $(\mathbf{h}, -\mathbf{h})$ leak at weight $2d^\perp$ for every k_1, and their supports are unions of secret pairs. With a polar hidden code of rate 1/2 ($N = 8192$ and 16384, $d^\perp = 16$, leaked weight 32) finding one costs 64.9 and 67.8 bits without multiplicity credit; with an LDPC hidden code ($\omega = 15$, $N = 8192$), 62.5 bits.

10.6 The landscape of masking

Table 18 collects the status of every RLCE- or UVW-type instantiation over the decodable families we know. Its content is a classification, not a theorem over all codes: each instantiation is either broken by one of four structural attacks—the square code, the cubic jet, the matched-pair subspace, or the dual leak—or excluded because no certified error weight supports the mask, or survives with keys two orders of magnitude larger than those of the unmasked scheme it would replace, or waits on a decoder that does not exist.

What a compact masked trapdoor would need. Reading the table by columns: a hidden code whose short side is primal can be concealed by inserting columns, at the price of enough of them to lift its minimum-weight words above the GV bound of the public code; one whose short side is dual cannot be concealed by either mask, because both act on the primal code and fix the dual up to a factor of two in weight; a code with a heavy dual and a weak decoder—a polar code at high rate—can be masked but, below length 32768, not decoded at the error weight the mask needs; and a code with no short side—GRS, AG—has, among known families, an algebraic curve the jet reads. A compact masked trapdoor would therefore need a code that is efficiently decodable at a

large fraction of its capacity, whose primal and dual minimum-weight words both cost 2^λ to find (the GV bound is sufficient, not necessary: RM$(6, 13)$ reaches 2^{191} on the dual side with $d^\perp = 128$ at $n = 8192$, far below GV, because its 2^{50} minimum-weight dual words are credited and still do not suffice), and which is neither an evaluation code of a curve nor a subfield subcode. Reed-Muller codes of order $m/2$ have the second and third properties and lack, so far, the first; random codes likewise. Finding such a family, or proving that the three properties are incompatible, is the question this paper leaves (Sect. 11).

11 Conclusion and Open Problems

The secret of a masked Reed-Solomon scheme whose decoder unmasks a few public coordinates at a time is a hidden pairing, and a hidden pairing is not hidden from jets: each tuple pins the hidden rational normal curve with Hermite conditions, and the conditions outnumber the coefficients. This one fact proves the count behind the October 2026 break of UVW-KEM, explains the dependencies the break reads the key from, shows that no split $k_1{:}k_2$ escapes the jet hierarchy at a useful decoding radius, and carries over to RLCE, whose long-key regime had resisted the square-code attack for eight years and is now broken in practice at full Level-1 size, and to the higher-genus HermitianRLCE, which is distinguished. The forcing law, the master count and the redundancy law turn this into an invariant: a decoder that unmasks locally buys each unit of structural depth with a unit of redundancy.

Removing the curve does not rescue the masks. Inserting columns, mixing pairs and adding a repeated U-block are operations on the primal code; they leave the dual of the hidden code in place, at most doubled in weight. A polar code is short on one side, and on the dual side at the rates where polarRLCE can close its primal leak; LDPC codes are short on the dual by construction; and the UVW mask adds a primal signature of its own. Across the constructions and parameter families examined in Table 18, every RLCE- or UVW-type instantiation is broken by one of four structural attacks—square code, cubic jet, matched pair, dual leak—or, for polarRLCE at high rate, cannot be decoded at the error weight its mask needs below length 32768, and costs 234 times the key of Classic McEliece there. This is a classification of the known decodable families under the two masks, not a theorem about all codes and all masks. The one configuration without a known attack at a key within an order of magnitude of Classic McEliece, a Reed-Muller-heavy information set under RLCE, waits on a decoder. The compactness that motivated masking came from exactly the structure these attacks exploit. Five problems remain.

1. **Exactness and the generic baseline.** The jet count is a lower bound, attained in every non-degenerate instance we ran, with the degenerate ones characterized by the Veronese threshold of Remark 9.5; prove equality, and prove that the structured relations are independent of the pairwise line relations for $2r > d + 1$ (verified at $(4, 3)$, needed for the degree-≥ 4 rows of the hierarchy). The recovery stage inherits the question: beyond the Veronese threshold the left kernel exceeds K_s and the rank-one read-off fails (Proposition 5.3(iv)); a read-off that extracts K_s from a larger kernel, or a proof that the threshold is never crossed at a shortening of cryptographic size, would remove the one hypothesis of Theorem 5.1 that is verified rather than proved.

2. **Reach of higher orders.** The jet hierarchy is polynomial but its degree grows with k_2/m (UVW) or with $w/(n-k)$ (RLCE); determine whether sparse or structured linear algebra in the style of [15, 30] makes degrees 5–7 practical at cryptographic sizes, which would settle the remaining high-k_2 and $w \geq 2(n-k)$ regimes.

3. **Decoding Reed-Muller codes far beyond the bounded-distance radius.** $\mathrm{RM}(6, 13)$ under RLCE resists every attack in this paper if $t \geq 215$ errors can be decoded at length 8192 with failure rate below 2^{-32}; the bounded-distance radius is 63 and capacity allows 901. Reed-Muller codes achieve capacity under MAP decoding [47, 48]; whether a polynomial-time decoder reaches order 6 at this length and weight decides Reed-Muller-RLCE, at a 4 MiB key. Two further questions attach: whether the affine symmetry of the hidden code gives an attack on its heavy dual words that generic low-weight search does not, and whether successive-cancellation list decoding certifies weights beyond Table 14 for any decreasing set that meets Corollary 10.5.

4. **A compact masked trapdoor.** Find an efficiently decodable code family whose primal and dual minimum-weight words both cost 2^{λ} to find, with multiplicity credited, and which is neither an evaluation code of a curve nor a subfield subcode—or prove that these properties are incompatible. Among known families, GRS and AG codes have the first two properties and a curve; polar and LDPC codes have a light dual; Reed-Muller codes of order $m/2$ have both sides heavy enough at $n = 8192$ and no decoder at the required radius; random codes lack decodability. Short of such a family, the compact code-based KEMs are those that do not mask at all, such as QC-MDPC (BIKE), whose security rests on a different problem.

5. **Hidden varieties and secant maskings.** The hidden-pairing model uses the curve only through the dimension of the space its composed forms live in, $\dim L(d\alpha P)$ for a curve of degree α; the forcing law uses only that a tuple's columns are a projective frame. Both survive the replacement of the curve by any projective variety X with hidden points $P_1, \ldots, P_n \in X$, each in the span of a few public columns: the count becomes $\sum_{\ell} \omega_{\ell} - h^0(X, \mathcal{O}_X(d))$, and the scroll of Sect. 9.6 is the first example (a surface, $h^0 = 3(k_l - 1)$ at $d = 2$). Two questions follow. Which varieties with a decoder—Segre and Veronese varieties, Grassmannians, the varieties of the known AG and multivariate evaluation codes—give a count that is positive at a compact public redundancy? And what replaces the forcing law when the public column lies in the span of *several* hidden points, as in BBCRS with $m \geq 1 + \mathrm{R}$ and in the weight-two masks of [39]: the public columns then sit on a secant variety of X, the jets at a column force conditions on the secant line rather than on a point of X, and the dimension that bounds the composed forms is that of the secant variety's homogeneous coordinate ring in degree d. A master count for secant maskings would decide the jet degree of the one GRS masking the present count leaves open.

References

[1] M. Baldi, M. Bianchi, F. Chiaraluce, J. Rosenthal, D. Schipani: Enhanced public key security for the McEliece cryptosystem. J. Cryptology 29(1), 1–27 (2016)

[2] T. P. Berger, P. Loidreau: How to mask the structure of codes for a cryptographic use. Des. Codes Cryptogr. 35(1), 63–79 (2005)

[3] T. P. Berger, C. T. Gueye, J. B. Klamti: Generalized subspace subcodes with application in cryptology. IEEE Trans. Inf. Theory 65(8), 4641–4657 (2019)

[4] P. Beelen, M. Bossert, S. Puchinger, J. Rosenkilde: Structural properties of twisted Reed–Solomon codes with applications to cryptography. In: IEEE ISIT 2018

[5] J. Alexander, A. Hirschowitz: Polynomial interpolation in several variables. J. Algebraic Geom. 4(2), 201–222 (1995)

[6] Bundesamt für Sicherheit in der Informationstechnik: Note on recent developments concerning Classic McEliece. Technical note and press release (1 October 2026)

[7] I. Cascudo, R. Cramer, D. Mirandola, G. Zémor: Squares of random linear codes. IEEE Trans. Inf. Theory 61(3), 1159–1173 (2015)

[8] I. Márquez-Corbella, J.-P. Tillich: Using Reed–Solomon codes in the $(U \,|\, U+V)$ construction and an application to cryptography. In: IEEE ISIT 2016, pp. 930–934

[9] A. Couvreur, P. Gaborit, V. Gauthier-Umaña, A. Otmani, J.-P. Tillich: Distinguisher-based attacks on public-key cryptosystems using Reed–Solomon codes. Des. Codes Cryptogr. 73(2), 641–666 (2014)

[10] A. Couvreur, M. Lequesne: On the security of subspace subcodes of Reed–Solomon codes for public key encryption. IEEE Trans. Inf. Theory 68(1), 632–648 (2022)

[11] A. Couvreur, A. Otmani, J.-P. Tillich, V. Gauthier-Umaña: A polynomial-time attack on the BBCRS scheme. In: PKC 2015, LNCS 9020, pp. 175–193

[12] A. Couvreur, I. Márquez-Corbella, R. Pellikaan: Cryptanalysis of McEliece cryptosystem based on algebraic geometry codes and their subcodes. IEEE Trans. Inf. Theory 63(8), 5404–5418 (2017)

[13] A. Couvreur, M. Lequesne, J.-P. Tillich: Recovering short secret keys of RLCE in polynomial time. In: PQCrypto 2019, LNCS 11505, pp. 133–152

[14] V. Gauthier-Umaña, A. Otmani, J.-P. Tillich: A distinguisher-based attack on a variant of McEliece's cryptosystem based on Reed–Solomon codes. arXiv:1204.6459 (2012)

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

[16] V. Guruswami, M. Sudan: Improved decoding of Reed–Solomon and algebraic-geometry codes. IEEE Trans. Inf. Theory 45(6), 1757–1767 (1999)

[17] K. Khathuria, J. Rosenthal, V. Weger: Encryption scheme based on expanded Reed–Solomon codes. Adv. Math. Commun. 15(2), 207–218 (2021)

[18] J. Lavauzelle, J. Renner: Cryptanalysis of a system based on twisted Reed–Solomon codes. Des. Codes Cryptogr. 88(7), 1285–1300 (2020)

[19] T. Hemmert: Key recovery for the McEliece cryptosystem using higher-order vanishing. Cryptology ePrint Archive 2026/1339

[20] J. Harris: Algebraic Geometry: A First Course. GTM 133, Springer (1992)

[21] T. Hemmert, A. Wiemers: Distinguishing Goppa codes using higher-order vanishing. In: CRYPTO 2026, LNCS 16803

[22] ngcc.dev: UVW-KEM (kem-38), finding kem-38-8: cubic-jet syzygies recover UVW's secret structure from the public key (26 October 2026; distinguisher credited to L. Wang, NGCC PKC Forum; recovery to M.-J. O. Saarinen). `https://ngcc.dev/reports/kem-38.html`, with code in the `ngcc-harness` repository (accessed 30 October 2026)

[23] F. Zhang, Y. Yang, K. Xiao, X. Chen, C. Li, H. Tian: UVW-KEM: algorithm description document. NGCC round-1 submission (kem-38), Institute of Commercial Cryptography Standards (2026)

[24] H. Randriambololona: The syzygy distinguisher. In: EUROCRYPT 2025, LNCS 15606, pp. 324–354

[25] M.-J. O. Saarinen: HOVER: higher-order vanishing endomorphism recovery. Cryptology ePrint Archive 2026/1778

[26] V. M. Sidelnikov, S. O. Shestakov: On an encoding system constructed on the basis of generalized Reed–Solomon codes. Discrete Math. Appl. 2(4), 439–444 (1992)

[27] Y. Wang: Quantum resistant random linear code based public key encryption scheme RLCE. In: IEEE ISIT 2016, pp. 2519–2523

[28] Y. Wang: Revised quantum resistant public key encryption scheme RLCE and IND-CCA2 security for McEliece schemes. Cryptology ePrint Archive 2017/206

[29] G. Matthews, Y. Wang: Quantum resistant public key encryption scheme HermitianRLCE. In: Code-Based Cryptography Workshop 2019 May 18 (pp. 1-10). Cham: Springer International Publishing.

[30] S. A. Weis: Improving GIJS key recovery for Classic McEliece. Cryptology ePrint Archive 2026/1984

[31] C. Wieschebrink: Cryptanalysis of the Niederreiter public key scheme based on GRS subcodes. In: PQCrypto 2010, LNCS 6061, pp. 61–72

[32] C. Wieschebrink: Two NP-complete problems in coding theory with an application in code based cryptography. In: IEEE ISIT 2006, pp. 1733–1737

[33] Y. Yang, F. Zhang: An efficient encryption scheme based on ($U+V$, $U+W$) codes. Cryptology ePrint Archive 2025/1149 (2025)

[34] E. Arıkan: Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Trans. Inf. Theory 55(7), 3051–3073 (2009)

[35] M. Baldi, F. Chiaraluce: Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes. In: ISIT 2007, pp. 2591–2595. IEEE (2007)

[36] M. Bardet, J. Chaulet, V. Drăgoi, A. Otmani, J.-P. Tillich: Cryptanalysis of the McEliece public key cryptosystem based on polar codes. In: PQCrypto 2016, LNCS 9606, pp. 118–143. Springer (2016)

[37] M. Bardet, V. Drăgoi, A. Otmani, J.-P. Tillich: Algebraic properties of polar codes from a new polynomial formalism. In: ISIT 2016, pp. 230–234. IEEE (2016)

[38] A. Esser, J. Verbel, F. Zweydinger, E. Bellini: CryptographicEstimators: a software library for cryptographic hardness estimation. In: ASIA CCS 2024, pp. 560–574. ACM (2024). `https://github.com/Crypto-TII/CryptographicEstimators`

[39] J. Lieb, A. Mazumder, M. Schaller: An attack on high rate McEliece cryptosystems using generalized Reed–Solomon codes with weight 2 mask. arXiv:2607.25027 (July 2026)

[40] S. Kim, I. Kim: Zeta-transform evaluation for higher-order vanishing key recovery. arXiv:2609.26132 (2026)

[41] N. Aragon et al.: BIKE: Bit Flipping Key Encapsulation. NIST PQC round-4 submission (2022)

[42] D. J. Bernstein et al.: Classic McEliece: conservative code-based cryptography. NIST PQC round-4 submission (2022)

[43] J. Liu, Y. Wang, Z. Yi, Z. Lin: polarRLCE: a new code-based cryptosystem using polar codes. Security and Communication Networks 2019, 3086975 (2019)

[44] R. Misoczki, J.-P. Tillich, N. Sendrier, P. S. L. M. Barreto: MDPC-McEliece: new McEliece variants from moderate density parity-check codes. In: ISIT 2013, pp. 2069–2073. IEEE (2013)

[45] A. Otmani, J.-P. Tillich, L. Dallot: Cryptanalysis of two McEliece cryptosystems based on quasi-cyclic codes. Math. Comput. Sci. 3(2), 129–140 (2010)

[46] I. Tal, A. Vardy: How to construct polar codes. IEEE Trans. Inf. Theory 59(10), 6562–6582 (2013)

[47] G. Reeves, H. D. Pfister: Reed–Muller codes on BMS channels achieve vanishing bit-error probability for all rates below capacity. IEEE Trans. Inf. Theory 70(2), 920–949 (2024)

[48] E. Abbe, C. Sandon: A proof that Reed–Muller codes achieve Shannon capacity on symmetric channels. In: FOCS 2023. IEEE (2023)

[49] I. Dumer, K. Shabunov: Soft-decision decoding of Reed–Muller codes: recursive lists. IEEE Trans. Inf. Theory 52(3), 1260–1266 (2006)

[50] M. Ye, E. Abbe: Recursive projection-aggregation decoding of Reed–Muller codes. IEEE Trans. Inf. Theory 66(8), 4948–4965 (2020)

[51] L. Minder, A. Shokrollahi: Cryptanalysis of the Sidelnikov cryptosystem. In: EUROCRYPT 2007, LNCS 4515, pp. 347–360. Springer (2007)

[52] I. V. Chizhov, M. A. Borodin: Effective attack on the McEliece cryptosystem based on Reed–Muller codes. Discrete Math. Appl. 24(5), 273–280 (2014)