

Fast Linear Codes over Large Fields: Addition-Only and Small-Coefficient Encoders

Stanislav Peceny and Peter Rindal

Abstract. We construct fast linear codes over large fields using addition-only encoders and variants with small integer coefficients. A local precode is followed by rounds of a random permutation and a scan, a linear recurrence with a short feedback window. At dimension 2^{20} , rate $1/2$, and $p = 2^{128} - 159$, two plain accumulators (unweighted prefix sums) give relative distance above 11.7%, except with probability below $2^{-49.64}$ over the sampling of the encoder. A binary subset-sum precode instead gives distance above 12.56%, except with probability below $2^{-22.05}$. These addition-only encoders use fewer than $6.8125n$ and $3.3125n$ field additions, respectively, for output length n . Both exceed the binary rate-1/2 Gilbert–Varshamov benchmark. The proof counts the supports and ranks of the zero equations a codeword must satisfy, retaining the dependence between cancellations.

Coefficient size and scan width extend the tradeoff. On one core, with a common implementation and setup failure below 2^{-40} , encoding time ranges from 12.171 ms at proved distance above 11.7% to 32.342 ms above 42.987%. A coefficient-independent upper bound limits every encoder with two width-one scans after the Newton precode to 39.6855%.

The distance and setup-failure bounds hold over $\mathbb{F}_p$ and every finite extension, covering all nonzero messages after setup. Timings are for $\mathbb{F}_p$. Extension preserves distance exactly because the encoder coefficients lie in the prime subfield. Changing characteristic requires a separate argument; we give field-transfer results in both directions, including counterexamples.

Keywords: linear codes · pseudorandom correlation generators · learning parity with noise · minimum distance · finite fields

1 Introduction

Pseudorandom correlation generators (PCGs) expand short seeds into long correlated strings, often through a linear encoder. Encoding over a large field makes the arithmetic cost of that encoder important. Addition-only circuits avoid field multiplications; small integer coefficients offer another way to trade arithmetic cost for distance. We ask:

What distance can addition-only encoders achieve over large fields, and what do small coefficients and wider scans buy?

A natural starting point is a binary encoding circuit evaluated over the large field, but its distance can change with the characteristic. Moreover, lifting a binary generator matrix entrywise can differ from lifting its factors and multiplying them over the new field. We instead give direct distance guarantees for a family of sparse encoding circuits with tunable scan width and coefficient size. These guarantees cover messages chosen after the encoder is sampled; a cancellation bound for each fixed message does not by itself establish minimum distance.

The construction and its tradeoffs. A local precode expands each block of eight input symbols to sixteen. The encoder then applies a number of rounds, two in all our results. Each round permutes the coordinates and applies a scan, in which each output is the current input plus w weighted preceding outputs. Each coefficient is sampled independently from $\{1, \dots, q\}$. We call the scan with $w = q = 1$, an ordinary prefix sum, a *plain accumulator*. For input length k , the precode output and all later vectors have length $n = 2k$; the final vector is the codeword. This extends the block-and-accumulate (BAA) architecture [4].

Our finite guarantees use $k = 2^{20}$ and $p = 2^{128} - 159$. Setup samples the permutations and coefficients; a setup fails if the resulting code has smaller distance than claimed. With the [16, 8, 9] Newton precode and two independently permuted plain accumulators, the relative distance exceeds 11.7% except with probability below $2^{-49.64}$. We call this the *all-one* encoder, after its scan coefficients. Encoding uses $6.8125n - 2$ field additions, including the precode, which is evaluated with finite differences, and takes 12.171 ms on one core of an AMD Ryzen 9 7950X.

The same precode with width one and coefficients up to 256 gives distance above 30.524% in 17.897 ms. Width four with coefficients up to 65536 gives above 42.987% in 32.342 ms. Both use two permutations and have failure below 2^{-40} . Tables 1 and 2 report all eleven configurations, timed with a common implementation. Products by small integers can be computed with additions, but our implementations for $q > 1$ use native field multiplication.

A binary subset-sum precode with two plain accumulators gives distance above 12.56%, failure below $2^{-22.05}$, and $3.3125n - 2$ additions. Both addition-only guarantees exceed the binary rate-1/2 Gilbert–Varshamov benchmark of approximately 11.003%, the usual reference point for the binary encoders used in PCGs. The Gilbert–Varshamov bound over $\mathbb{F}_p$ itself is close to 1/2 at this rate and remains far above these guarantees. All guarantees fix the characteristic before setup and cover every nonzero message after a successful setup.

The same circuits work over every finite extension $\mathbb{F}_{p^m}$ with exactly the same symbol distance and setup-failure probability. All scalar coefficients lie in $\mathbb{F}_p$, so extension-field encoding splits into component encodings over $\mathbb{F}_p$. This applies to both additions and small integer coefficients. The numerical guarantees hold for characteristic $p = 2^{128} - 159$ and every extension degree m , but do not transfer to another characteristic. The reported timings are for $\mathbb{F}_p$.

Counting dependence and bounding distance from above. The proof tracks which scan outputs vanish and how much rank the resulting zero equations add, avoiding

a union bound over field-valued messages. For plain accumulators, a zero output that ends a run of nonzero outputs may impose no new equation; counting whole runs pairs these free returns with the returns that do increase rank. Bounds that hold uniformly over the linear forms not yet constrained let the scans compose despite these dependencies.

A separate upper bound limits every two-scan width-one Newton encoder to 39.6855%, regardless of coefficient size. Increasing the alphabet alone therefore cannot reach the proved width-four point. A gap remains between the width-one lower and upper bounds.

What survives a field change. Scalar extension preserves distance. Changing characteristic can reduce it, even for an addition-only BAA encoder over a prime exceeding the code length. A binary-matrix family has positive binary relative distance and vanishing ternary relative distance, though its rate vanishes. Conversely, dense random binary generators give distance near $1/2$ at sufficiently small constant rate over prescribed finite families of fields. Their failure bound depends on the number of fields, not their sizes, but direct encoding of these dense codes has quadratic cost.

Related work. Unique-neighbor expansion proves field-independent distance for some sparse matrices [2]. Binary expand-accumulate and expand-convolute analyses track binary weight processes [1, 7]. Security analyses can accommodate a smaller distance through noise parameters, but still need a justified distance bound [2, 5]. Walsh–Hadamard PCGs analyze codes directly over a large prime field [6]. The appendices study repetition-based BAA and a fixed-alphabet Walsh–Hadamard variant, distinguishing their unresolved predictions from proved guarantees.

Organization. Sections 2–3 give transfer baselines; Section 4 gives the construction and tradeoffs. Section 5 develops the counting argument, and Section 6 discusses open questions. The appendices provide supporting studies, including a four-input precode (Appendix C), proof refinements, and finite verification.

2 Field changes and their security role

Let F be a field and let $G \in F^{k \times n}$. The code generated by G is

$$\mathcal{C}_F(G) := \{xG : x \in F^k\}.$$

For a code subspace $C \subseteq F^n$, define

$$d(C) := \min\{\text{wt}(c) : c \in C \setminus \{0\}\}, \quad d(\{0\}) := +\infty.$$

We write $d(\mathcal{C}_F(G))$ for a row generator and $d_F(E) := d(\text{im}_F E)$ for a column encoder $x \mapsto Ex$. When the field is fixed, $d(E)$ abbreviates $d_F(E)$. The matrix need not have full rank. The rate of $\mathcal{C}_F(G)$ is $\text{rank}_F(G)/n$.

2.1 Scalar extension is exact

Suppose K/F is a field extension and G has entries in F . The same matrix defines codes over F and K .

Theorem 2.1 (Scalar-extension invariance). *For every field extension K/F and every matrix $G \in F^{k \times n}$,*

$$d(\mathcal{C}_K(G)) = d(\mathcal{C}_F(G)).$$

Proof. Fix an F -basis $\beta_1, \dots, \beta_s$ of the finite-dimensional subspace of K spanned by the coordinates of a message $x \in K^k$. Write $x = \sum_i \beta_i x_i$ with $x_i \in F^k$. Then

$$xG = \sum_i \beta_i (x_i G).$$

A coordinate of xG is zero exactly when the same coordinate is zero in every $x_i G$. Hence the support of xG is the union of the supports of its nonzero F -components. Every nonzero K -codeword therefore has weight at least $d(\mathcal{C}_F(G))$. The reverse inequality follows from $\mathcal{C}_F(G) \subseteq \mathcal{C}_K(G)$. If $G = 0$, both distances are $+\infty$.

Theorem 2.1 covers, for example, a binary matrix interpreted over $\mathbb{F}_{2^s}$. It does not cover $\mathbb{F}_p$ for odd p , because no field embedding $\mathbb{F}_2 \hookrightarrow \mathbb{F}_p$ exists.

2.2 Changing characteristic

For a specified binary factorization $G = G_1 \cdots G_t$, define

$$G'_p := \text{lift}_p(G_1) \cdots \text{lift}_p(G_t),$$

where lift_p sends each binary entry to 0 or 1 in $\mathbb{F}_p$. The product is evaluated over $\mathbb{F}_p$. In general, it differs from the entrywise lift of the product first computed over $\mathbb{F}_2$. Equivalently, the stagewise construction reduces the integer product of the lifted factors modulo each characteristic. The dense constructions below have one factor; the BAA construction keeps its repetition, permutation, and accumulator stages.

Let $G_{\mathbb{Z}} \in \mathbb{Z}^{k \times n}$ be an integer matrix. Reduction modulo a prime p gives a matrix G_p . Select r columns, where $1 \leq r \leq \min(k, n)$. These columns are independent over $\mathbb{F}_2$ and dependent over $\mathbb{F}_p$ exactly when some r -by- r minor of the selected columns is odd and every such minor is divisible by p . Thus the obstruction is torsion in an integer cokernel, not a failure of scalar extension.

The same distinction appears at the codeword level. A binary sum vanishes when an even number of active terms meet. Over $\mathbb{F}_p$, the values of those terms matter. New cancellations can appear, while old parity cancellations can disappear.

2.3 Distance, noise, and PCG security

Minimum distance is not itself an LPN assumption. It prevents linear tests from finding an unusually short dual relation and supports parameter choices against known attacks. Silver makes both points explicit. Its unique-neighbor argument applies to binary sparse matrices over arbitrary fields, while its security discussion observes that additional noise can compensate for an overestimate of distance against linear attacks [2].

Expand-Convolute demonstrates why empirical extrapolation is insufficient. It gives efficient attacks against the proposed Silver codes and develops a new binary code with a direct distance proof [7]. Its Markov-chain analysis is binary; it is not a change-of-characteristic theorem.

Stationary Syndrome Decoding gives a complementary security result. Against linear tests, its bias depends explicitly on code distance and noise weight, and a larger noise weight can offset a smaller distance in that model [5]. This tradeoff does not establish the distance of the matrix over a new field. A code theorem and a security compensation argument answer different questions.

3 Positive and negative baselines

The field-transfer problem is meaningful only if some binary matrices work over many characteristics. This section first gives such a construction. It then shows why binary distance alone does not certify an arbitrary matrix.

3.1 A dense field-universal code

Let $G_\square$ be the k -by- 2^k binary matrix whose columns are all vectors in $\{0, 1\}^k$.

Proposition 3.1 (Boolean evaluation code). *For every field F , the code $\mathcal{C}_F(G_\square)$ has dimension k and minimum distance 2^{k-1} .*

Proof. Fix $x \in F^k \setminus \{0\}$ and an index i with $x_i \neq 0$. Pair each $a \in \{0, 1\}^k$ with the vector obtained by flipping coordinate i . Their inner products with x differ by x_i , so at most one inner product in each pair is zero. Every nonzero word has weight at least 2^{k-1} . A message supported on one coordinate attains this weight.

The rate of this code vanishes. Sampling only linearly many columns recovers constant rate.

For a field F and nonzero $x \in F^k$, define

$$Z_x := \{a \in \{0, 1\}^k : \langle x, a \rangle = 0\}.$$

The proof of Proposition 3.1 gives $|Z_x|/2^k \leq 1/2$.

Lemma 3.2 (Kernel sections). *The set system $\mathcal{Z}_F := \{Z_x : x \in F^k \setminus \{0\}\}$ has VC dimension at most $k - 1$.*

Proof. Suppose $a_1, \dots, a_m$ are shattered. One labeling places every a_j in the kernel of a nonzero functional, so their span is proper. For each j , another labeling contains every a_i with $i \neq j$ but excludes a_j . Hence no a_j lies in the span of the other points. The points are linearly independent and span a proper subspace, which implies $m \leq k - 1$.

Sample independent columns $A_1, \dots, A_n \leftarrow \{0, 1\}^k$ and set $G = [A_1 | \dots | A_n]$. A standard VC inequality and Lemma 3.2 give the following theorem [8].

Theorem 3.3 (Dense simultaneous baseline). *Fix fields $F_1, \dots, F_t$, a real $0 < \varepsilon < 1/2$, and integers $k \geq 2$ and n with $2n \geq k - 1$. Except with probability at most*

$$8t \left(\frac{2en}{k-1} \right)^{k-1} \exp \left(-\frac{n\varepsilon^2}{32} \right),$$

the sampled matrix G satisfies, simultaneously for every $s \in [t]$,

$$\text{rank}_{F_s}(G) = k \quad \text{and} \quad d(\mathcal{C}_{F_s}(G)) \geq (1/2 - \varepsilon)n.$$

Proof. Apply the VC inequality to the uniform distribution on $\{0, 1\}^k$ and the class $\mathcal{Z}_{F_s}$. Uniformly for every nonzero $x \in F_s^k$, the empirical fraction of columns in Z_x is at most $1/2 + \varepsilon$ outside the stated failure event. The word xG therefore has weight at least $(1/2 - \varepsilon)n$. Its positive weight also proves full row rank. A union bound over the t fields completes the proof.

For fixed ε , a sufficiently small constant rate makes the failure probability exponentially small in k . Theorem 3.3 also handles growing primes. Taking all primes below 2^λ adds at most λ bits to the union bound. If $k = \Omega(\lambda)$, one binary matrix can retain constant rate and distance over $\mathbb{F}_2$ and every such prime field.

Direct encoding computes n dense subset sums and costs $\Theta(kn)$ field additions. At constant rate this is quadratic. Theorem 3.3 settles existence but not efficient encoding.

3.2 Why binary distance is insufficient

A small explicit family shows the danger. Let $M := J_4 - I_4$. For $L \geq 1$, form a 4-by- $(4L + 1)$ binary matrix G_L by repeating each column of M L times and appending one all-one column. Since $\det(M) = -3$, the matrix M is invertible over $\mathbb{F}_2$ but not over $\mathbb{F}_3$; over $\mathbb{F}_3$, the all-one message annihilates every column of M and leaves only the final column. Hence

$$d(\mathcal{C}_{\mathbb{F}_2}(G_L)) = L + 1, \quad d(\mathcal{C}_{\mathbb{F}_3}(G_L)) = 1.$$

Both reductions have row rank four. Consequently,

$$\frac{d(\mathcal{C}_{\mathbb{F}_2}(G_L))}{4L + 1} \longrightarrow \frac{1}{4}, \quad \frac{d(\mathcal{C}_{\mathbb{F}_3}(G_L))}{4L + 1} \longrightarrow 0.$$

The family has vanishing rate, so it does not rule out a theorem with an explicit positive-rate hypothesis. It does rule out an inheritance theorem for unrestricted binary matrices based only on their binary relative distance.

The BAA case gives positive-rate finite counterexamples. For code length six, one rate-1/3 repetition-three encoder has distance 4 over $\mathbb{F}_2$ and 3 over $\mathbb{F}_7$. Thus even a prime exceeding the code length does not guarantee that every sampled encoder keeps its binary distance. These examples motivate ensemble statements rather than equality for every realized encoder.

4 Addition-only and small-coefficient encoders

We construct the encoders over $\mathbb{F}_p$, then extend them to $\mathbb{F}_{p^m}$ without changing their distance guarantees. The Newton precode expands the message, then each round permutes the coordinates and applies a scan. This extends the block-and-accumulate architecture of BAA [4]. The two main parameters are the scan width, the number of preceding outputs fed back at each step, and the coefficient alphabet size. Increasing either parameter preserves the vector lengths and the rate. Their costs differ. A wider scan performs more arithmetic and reads more coefficients. Moving from all-one coefficients to the alphabet $\{1, 2\}$ adds a small-integer product to every scan step; beyond that, a larger alphabet can keep the same machine instruction sequence until a coefficient or metadata representation changes.

4.1 A common family

Fix $p = 2^{128} - 159$, $k = 2^{20}$, and $n = 2k$. Unless stated otherwise, the precode in this section is the Newton precode

$$f_x(t) = \sum_{j=0}^7 \binom{t}{j} x_j, \quad B_8(x) = (f_x(0), \dots, f_x(14), x_7). \quad (1)$$

Apply B_8 independently to each block of eight message symbols. This gives a length- n vector. The code B_8 is $[16, 8, 9]$ MDS: if $x_7 \neq 0$, at most seven finite evaluations vanish; if $x_7 = 0$, the polynomial has degree at most six. In either case a nonzero block has weight at least nine. Its first eight evaluations also establish injectivity. Finite differences evaluate the map with 77 additions per block, including only updates that can reach a remaining output. Thus this precode has an addition-only circuit even though its matrix contains binomial coefficients larger than one.

For $1 \leq w < n$, a width- w scan $A_{w,a}$ maps u to c by

$$c_i = u_i + \sum_{j=1}^{\min(w,i)} a_{i,j} c_{i-j}, \quad 0 \leq i < n. \quad (2)$$

Here $1 \leq q < p$, and every coefficient $a_{i,j}$ is sampled independently and uniformly from $\{1, \dots, q\}$. A width-one scan with $q = 1$ has every coefficient equal to

one and computes prefix sums; we call it a *plain accumulator*. For L rounds, sample independent uniform permutation matrices $P_1, \dots, P_L$, independently of all coefficients. With column-vector notation, the encoder is

$$E = A_{w,a_L} P_L \cdots A_{w,a_1} P_1 B_8. \quad (3)$$

Every scan is invertible. The encoder therefore has dimension k for every setup, and every vector after the precode has length n . The output is the complete vector produced by the L th scan; no further unpermuted scan follows it. Every timed configuration in this section has this form.

The comparison grid fixes $L = 2$ and uses widths $w \in \{1, 2, 4, 8\}$. The *all-one encoder* is the case $L = 2$, $w = 1$, $q = 1$: the Newton precode followed by two permuted plain accumulators. For a target distance d_0 , each distance guarantee below bounds the setup-failure probability

$$\Pr_{P,a} [\exists x \in \mathbb{F}_p^k \setminus \{0\} : \text{wt}(Ex) < d_0]$$

over the permutations and coefficients. The prime is fixed first, the setup is then sampled, and the message x is chosen last, with knowledge of the setup.

Corollary 4.1 (Extension-field guarantees). *Let M be the encoding matrix of any fixed setup of (3). For every integer $m \geq 1$, interpreting the same circuit over $\mathbb{F}_{p^m}$ preserves its rate and minimum symbol distance:*

$$d_{\mathbb{F}_{p^m}}(M) = d_{\mathbb{F}_p}(M).$$

Consequently, each setup-failure bound over $\mathbb{F}_p$ holds simultaneously over all finite extensions of $\mathbb{F}_p$ for the same sampled circuit.

Proof. The precode and all scan coefficients have entries in $\mathbb{F}_p$. Theorem 2.1, applied to the transpose of the encoding matrix, gives equality of distances for each setup. Matrix rank is also unchanged by extension. The distance-failure event is therefore the same for every m .

The corollary also applies when the Newton precode is replaced by another matrix over $\mathbb{F}_p$, including the subset-sum precode below. The operation counts then refer to extension-field additions and products by the same integer coefficients. They do not imply unchanged machine latency: in an $\mathbb{F}_p$ -basis, one extension-field addition uses m base-field additions. Our measurements use $m = 1$.

4.2 Two width-one scans with small coefficients

For $L = 2$ and $w = 1$, the coefficient alphabet shared by the two scans determines whether encoding uses only additions or also small-integer products. At $q = 1$, both scans are plain accumulators and setup samples only the permutations. Table 1 gives the achieved finite guarantees.

Table 1. Proved width-one guarantees for the Newton precode B_8 and two permutations. Failure exponents are truncated toward zero, so the stated bounds are conservative. Times are measured as described in Section 4.3. These are achieved points, not optimal frontiers.

q	Distance $\geq$	Relative bound	Failure below	Time (ms)
1	245367	$> 11.7\%$	$2^{-49.64}$	12.171
2	299893	$> 14.3\%$	$2^{-45.82}$	17.911
4	390595	$> 18.625\%$	$2^{-41.50}$	17.876
8	471860	$> 22.5\%$	$2^{-51.89}$	17.924
16	554173	$> 26.425\%$	$2^{-54.31}$	17.942
32	610796	$> 29.125\%$	$2^{-56.31}$	17.897
64	637220	$> 30.385\%$	$2^{-56.92}$	17.982
128	640051	$> 30.52\%$	$2^{-58.68}$	17.909
256	640135	$> 30.524\%$	$2^{-59.13}$	17.897

Theorem 4.2. *For each row of Table 1, the encoder in (3) with $L = 2$, $w = 1$, and the listed q has minimum distance at least the listed integer, except with probability less than the listed setup-failure bound.*

Section 5 gives the common argument, and its last subsection identifies the appendix that completes the proof for each row.

The all-one row exceeds the binary rate-1/2 Gilbert–Varshamov (GV) benchmark $H_2^{-1}(1/2) \approx 0.110028$. This compares a proved prime-field distance with a binary benchmark; it does not compare the realized distances of the same circuit over two fields. Section 5 explains why a count of supports and equation ranks covers messages chosen after setup, and describes the refinement for plain accumulators used in the all-one row.

A different precode gives a complementary addition-only point. The systematic subset-sum precode B_{sub} of Appendix L has a binary matrix and uses 21 additions per eight inputs. With the same two plain accumulators it gives distance above 12.56%, failure below $2^{-22.05}$, and $3.3125n - 2$ additions (Theorem L.1). Its smaller failure margin and different precode prevent a direct comparison within the table. We did not time this variant.

Each width-one scan has $n - 1$ feedback terms $a_{i,1}c_{i-1}$, so every row of Table 1 has $2n - 2$ feedback terms. Together with the precode, adding them in takes $6.8125n - 2$ field additions. For $q = 1$ there are no coefficient products, so these additions are the complete cost: exactly 14286846 field additions. For $q \geq 2$, a native implementation may use $2n - 2$ multiplications of field elements by small integer coefficients. Each product acts on a full 128-bit field element; it is not one machine multiplication. The measured implementation multiplies both 64-bit limbs and reduces the result. Measured alphabets have $q \leq 2^{16}$; the numerical grid extends to $q = 2^{20}$. With coefficients in $\{1, 2\}$, each product requires at most one doubling. The entire encoder then uses at most $8.8125n - 4$ additions, with expectation $7.8125n - 3$ over setup. For general q , computing each product by

Table 2. Proved guarantees and measured times for wider scans with the Newton precode B_8 and two permutations. Failure exponents are truncated toward zero, so the stated bounds are conservative.

w	q	Distance $\geq$	Relative bound	Failure below	Time (ms)
2	2048	829634	$> 39.56\%$	$2^{-57.87}$	22.994
4	65536	901503	$> 42.987\%$	$2^{-61.76}$	32.342

double-and-add gives the expected addition count

$$4.8125n + 2(n-1) \left(1 + \frac{1}{q} \sum_{a=1}^q [\log_2 a] + \text{popcount}(a) - 1 \right).$$

Two length- n field buffers occupy 64 MiB. A 32-bit word stores a 21-bit permutation index and each listed coefficient-minus-one, giving 16 MiB for the two metadata arrays. This 80 MiB model excludes separately retained input. The measured implementation also retains the 16 MiB input, giving 96.008 MiB including padded metadata.

4.3 Proved distance and measured encoding time

The same precode and two-permutation family also gives the width-two and width-four points in Table 2.

Theorem 4.3. *For each row of Table 2, the encoder in (3) with $L = 2$ and the listed w and q has minimum distance at least the listed integer, except with probability less than the listed setup-failure bound.*

Appendix M proves both rows. It extends the width-one rank bound of Section 5 to wider scans and applies Lemma 5.1.

All eleven configurations were timed on one core of an AMD Ryzen 9 7950X. GCC 15.2.0 used `-O3 -march=native -DNDEBUG` and C++20. Each configuration used six setup seeds, two warmups, and 31 timed encodings per seed. The tables report the median over seeds of the per-seed median time. These are warm encoding latencies for k field elements to $2k$ field elements. Setup uses reproducible pseudorandom streams; the distance theorems concern the specified ideal sampling distribution. Further measurement details are in Appendix N.

The Newton precode is evaluated by batched finite differences. Each scan fuses permuted reads, arithmetic, and sequential writes, with sixteen gathered inputs per group and prefetching ahead. At $q = 1$, both scans use a branchless field-addition kernel. For $q \geq 2$, each scan step also multiplies by a small integer coefficient and reduces. This raises the width-one time from 12.171 ms at $q = 1$ to 17.911 ms at $q = 2$. For $2 \leq q \leq 256$ the arithmetic sequence changes little, and the times stay between 17.876 ms and 17.982 ms.

In the all-one encoder, the median phase times are 2.289 ms for the precode and 4.916 ms and 4.968 ms for the two scans. Widths two and four take 22.994 ms and 32.342 ms and retain 112.016 MiB and 144.031 MiB of encoding payload.

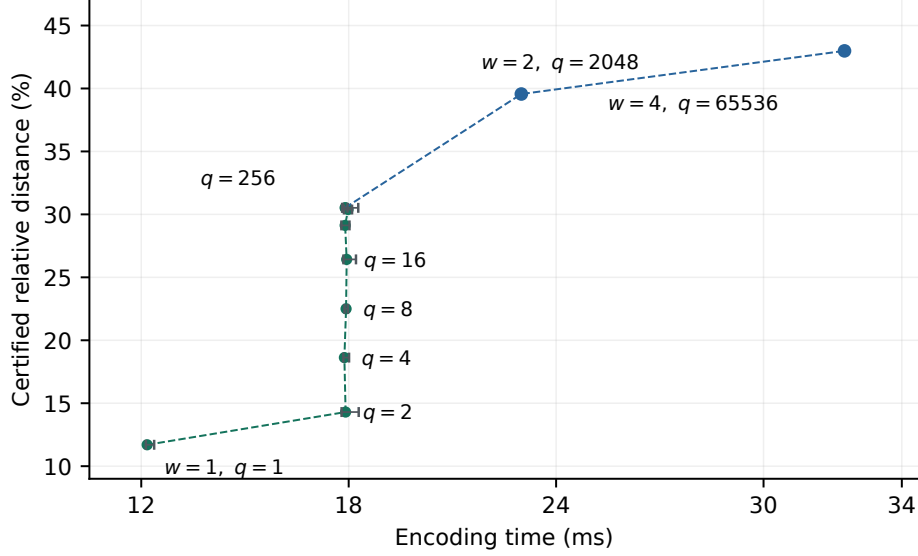


Fig. 1. Proved finite distance guarantees versus measured encoding time. All eleven points use the Newton precode and the same machine, compiler and seeds, with at least forty bits of setup-failure margin. Margins need not be equal. Whiskers show the range of the six per-seed medians. Only some coefficients are labeled; Table 1 lists every width-one point. Connecting lines assert neither interpolation guarantees nor optimality.

4.4 The width and coefficient grid

The grid fixes the Newton precode, two permutations, and widths $w \in \{1, 2, 4, 8\}$. Its common alphabet sizes are $q \in \{16, 64, 256, 2048, 2^{16}, 2^{20}\}$; the width-one study also includes the smaller alphabets in Table 1. Every grid cell has the same field, dimension, and rate.

Table 3 records numerical dense-support thresholds from the direct rank-counting bound. Each entry is a numerical zero of the exponent after local maximization over source and intermediate densities, with both densities restricted to at least 0.02. These estimates omit sparse supports and finite prefactors. They are neither complete distance certificates nor upper bounds on the actual code distance. The grid also omits the projective pairs with small coefficients of Appendix J, which the certified width-one rows with $q \geq 2$ may use on each part of their cover. This is why the certified value 26.425% at $q = 16$ can exceed the grid estimate 26.382%.

At width eight, the estimates reach about 43.9%; we have neither a finite guarantee nor a timing for width eight. A finite guarantee must cover all sparse weights and dense support profiles, including their finite counting factors.

Table 3. Numerical dense-support thresholds, as percentages of output length. These estimates have no certified finite setup-failure margin. Proved finite guarantees and measured times appear separately in Tables 1 and 2.

q	$w = 1$	$w = 2$	$w = 4$	$w = 8$
16	26.382	27.040	26.111	25.665
64	30.391	34.126	33.966	33.881
256	30.528	37.937	37.907	37.892
2048	30.528	39.569	41.106	41.104
65536	30.528	39.570	42.993	43.819
1048576	30.528	39.570	42.993	43.911

4.5 A coefficient-independent obstruction

Coefficient size cannot overcome every limitation of a narrow scan. The following count gives an upper bound for each fixed encoder.

Lemma 4.4. *Let $n = 2\rho g$, and let H be a parity-check matrix with g groups of ρ rows each. Suppose each column touches at most Δ groups, where $2 \leq \Delta \leq g$. If $\Delta \leq b \leq g$ and*

$$2 \binom{b-1}{\Delta-1} > \binom{g-1}{\Delta-1}, \quad (4)$$

then $d(\ker H) \leq \rho b + 1$.

Proof. Fix H and sample a uniform set of b row groups. A column touching $j \leq \Delta$ groups is supported entirely in the selected rows with probability $\binom{b}{j} / \binom{g}{j} \geq \binom{b}{\Delta} / \binom{g}{\Delta}$. The expected number of such columns is at least $n \binom{b}{\Delta} / \binom{g}{\Delta}$. Condition (4) makes this number exceed ρb . Some selection therefore contains $\rho b + 1$ columns supported on at most ρb rows. These columns are dependent, and their dependence gives a nonzero word of weight at most $\rho b + 1$ in $\ker H$.

The Newton precode has a block diagonal parity check H_0 , with eight rows per block. Write $G_j = A_{w,a_j}^{-1}$. With the convention in (3), an output parity check is

$$H = H_0 P_1^{-1} G_1 \cdots P_L^{-1} G_L.$$

Each G_j has bandwidth w and at most $w + 1$ nonzeros per column. Thus every column of H touches at most $(w + 1)^L$ row groups of H_0 . Consecutive inverse scans without a permutation have additive bandwidth: a sequence of widths $u_1, \dots, u_j$ has at most $1 + \sum_i u_i$ nonzeros per column. This yields tighter counts than multiplying their individual column weights.

For width one, the same count applies to variants of (3) with more scans. Let $(t_1, \dots, t_L)$ give the number of consecutive width-one scans after each permutation; the family (3) with $w = 1$ is the profile $(1, \dots, 1)$, and the other profiles below add a permutation or an unpermuted scan. We may take $\Delta = \prod_j (t_j + 1)$. Applying the exact integer condition with $\rho = 8$ and $g = 131072$ gives

Consecutive scan counts	Δ	Distance at most	Relative ceiling
(1, 1)	4	832265	39.6855%
(1, 2)	6	912849	43.5281%
(1, 1, 1)	8	949729	45.2867%
(2, 2)	9	961553	45.8505%

The percentages are rounded upward. These bounds hold for every permutation and every coefficient choice, including full-field coefficients. In particular, two width-one scans cannot reach 44%, even after one additional width-one scan without another permutation. For fixed Δ and growing g , the relative ceiling tends to $\frac{1}{2}2^{-1/(\Delta-1)}$. The bound does not determine typical distance: the best achieved width-one lower bound and its ceiling still leave a gap between 30.524% and 39.6855%.

5 How the distance bounds are proved

The minimum-distance event quantifies over every message after setup. A bound for a fixed message cannot establish this event without an additional argument. We instead enumerate supports and track the rank of their zero equations. This section gives the common argument for Theorems 4.2 and 4.3; its last subsection states which appendix completes each row of Tables 1 and 2.

5.1 Branches and equation rank

For a message x , call B_8x its source vector; the support and weight of B_8x are its *source support* and *source weight*. Fix a source support $S \subseteq \{0, \dots, n-1\}$ and define its shortened source space

$$V_S = \{x \in \mathbb{F}_p^k : (B_8x)_i = 0 \text{ for } i \notin S\}, \quad D_S = \dim V_S.$$

Require every source coordinate in S to be nonzero on the candidate message. Discard S if $D_S = 0$. A *branch* specifies which outputs are zero in each scan. Fix a forbidden final weight e ; a branch is *bad* if its final output has weight at most e .

Each permutation assigns to every position of the following scan an *input label*, the index of the coordinate that position reads. A label is *active* if its coordinate is required to be nonzero. We reveal (*expose*) input labels and tap coefficients in scan order. Every exposed coordinate is a linear form on V_S . Let $W \subseteq V_S^*$ span the zero equations imposed so far, starting with $W = 0$. The same span is retained across both scans.

A required nonzero coordinate must remain a nonzero form in V_S^*/W . Discard a branch whenever this necessary condition fails. Keeping other infeasible branches only enlarges the count. A zero output whose form is outside W increases $\dim W$ by one; a zero output whose form lies in W adds no rank. A *run* is a maximal block of consecutive nonzero outputs of a scan, and the zero output that ends it is a *return*. A return whose form already lies in W is a *free return*.

For $\tau \geq 1$, weight a branch of rank r by τ^{-r} . Any branch realized by a nonzero message has $r \leq D_S - 1$, because that message belongs to $\ker W$. Thus, for each fixed setup,

$$\mathbf{1}_{\{\text{a nonzero message realizes a bad branch from } S\}} \leq \tau^{D_S-1} \sum_{\gamma \text{ retained and bad}} \tau^{-r(\gamma)}. \quad (5)$$

Several messages may realize the same branch. The right side counts that branch once, so it introduces no factor depending on p^{D_S} .

5.2 A uniform bound for one scan

First consider width one. At any point of the exposure, the *pool* is the multiset of forms at the input labels not yet exposed. A *continuation bound* is an upper bound on the weighted count of branch completions from a given feedback state. It must hold for every quotient V_S^*/W , every pool, and every earlier nonzero requirement. Fresh coefficients remain independent of this exposure history. Nonzero requirements restrict the branches being counted; they do not condition the setup distribution.

There are two feedback states: zero (state 0) and nonzero (state 1). Let $x, y > 0$ be formal markers for nonzero inputs and outputs; from here on, x and y denote these markers, not a message or a scan output. Let $v_0, v_1 > 0$ be continuation bounds for the two states. The following positive homogeneous operator bounds the two states:

$$\begin{aligned} T_0(v) &= v_0 + xyv_1, \\ T_1(v) &= yv_1 + x \max \left\{ yv_1 + \frac{v_0}{\tau}, \left(1 - \frac{1}{q}\right) yv_1 + \frac{v_0}{q} \right\}. \end{aligned} \quad (6)$$

To justify the maximum, consider a nonzero input form and a nonzero feedback form in the quotient by W , before exposing the fresh tap. If these forms are independent, requiring a zero output adds one independent equation. This gives the first alternative. If the forms are dependent, at most one coefficient produces the zero form. The zero and nonzero outcomes are then mutually exclusive, giving the second alternative when this coefficient belongs to the alphabet. If it does not, the first alternative bounds the nonzero outcome. For $q = 1$ the only coefficient is one, so a dependent zero output is deterministic, the second alternative is v_0 , and only the two permutations are random.

Normalize $v_0 = 1$, set $\lambda = \max_i T_i(v)/v_i$, and $C = \max_i 1/v_i$. Monotonicity and homogeneity imply $[T^n(\mathbf{1})]_0 \leq C\lambda^n$. For sparse supports we instead use

$$[T^n(\mathbf{1})]_0 \leq \lambda^{n-16} \max_i \frac{[T^{16}(\mathbf{1})]_i}{v_i}. \quad (7)$$

The terminal iterations are evaluated exactly over the rationals.

5.3 Composing the scans

For the Newton precode, put

$$Q_8(z, \tau) = 1 + \sum_{j=9}^{16} \binom{16}{j} z^j \tau^{j-8}, \quad F_s(\tau) = [z^s] Q_8(z, \tau)^{n/16}. \quad (8)$$

The coefficient $F_s(\tau)$ sums τ^{Ds} over source supports of size s . Indeed, a support of size $j \geq 9$ within one block has shortening dimension $j - 8$ by the MDS property, and there are $\binom{16}{j}$ such supports. The constant term represents a zero block.

Fix the incoming forms of a scan, and suppose κ of its input labels are active. Give a branch with output weight h' and incremental rank r' the weight $x^\kappa y^{h'} \tau^{-r'}$. Average this weight over the order of the distinct active labels and over all fresh coefficients, then sum over placements of the active labels and over output-support branches. A *uniform scan bound* is a function $K(\kappa)$ that bounds this sum for every incoming quotient and every set of earlier nonzero requirements. The two-state bound above gives $K(\kappa) = C\lambda^n$. The bounds for plain accumulators also average over the active labels and may include a κ -dependent factor.

Lemma 5.1 (Composition of uniform scan bounds). *Fix source weight s , intermediate weight h , and forbidden final weight e . For scan j , let K_j be a uniform scan bound, with markers $x_j, y_j > 0$, $y_2 \leq 1$, and the same $\tau \geq 1$ in both scans. The probability of a nonzero message with these source and intermediate weights and final weight at most e is at most*

$$\frac{F_s(\tau)}{\tau} \frac{K_1(s) K_2(h)}{\binom{n}{s} \binom{n}{h} x_1^s (y_1 x_2)^h y_2^e}. \quad (9)$$

Proof. Fix S of size s . A uniform permutation sends its active labels to a uniform s -subset and gives them a uniform order on that subset. These choices are independent. The first normalization is therefore $1/\binom{n}{s}$; the label-order average is already included in $K_1(s)$.

Fix a complete first-scan history with intermediate weight h . The second permutation and its coefficients remain independent of that history. Apply $K_2(h)$ in the quotient left by the first scan. The second support-placement normalization is $1/\binom{n}{h}$. Rank increments add, so the product of the two incremental weights is the single weight τ^{-r} in (5). This conditional argument does not assume independent zero equations.

All terms are nonnegative. Extracting source and intermediate weights divides by $x_1^s (y_1 x_2)^h$. For a final weight $j \leq e$, $y_2^j \geq y_2^e$, which gives the remaining divisor. Apply inequality (5), then sum over S ; this supplies $F_s(\tau)/\tau$.

For $K_j = C_j \lambda_j^n$, the lemma gives the width-one profile bound

$$\frac{F_s(\tau)}{\tau} \frac{C_1 \lambda_1^n C_2 \lambda_2^n}{\binom{n}{s} \binom{n}{h} x_1^s (y_1 x_2)^h y_2^e}. \quad (10)$$

Uniformity is essential here: a bound proved only for a fresh, unconstrained input space would not justify the second application.

5.4 Finite evaluation and the plain-accumulator refinements

The inverse width-one scan gives $h \leq \min(n, 2e)$ and $s \leq \min(n, 2h)$. The *direct bound* evaluates (10) with the operator (6) on every such integer profile. For source weights below 32, it uses the exact coefficients $F_s(\tau)$ and either the terminal bound (7) or the hypergeometric support bound of Appendix I. The remaining weights use $F_s \leq Q_s(z, \tau)^{n/16}/z^s$ and convex bounds on rectangles in the (s, h) plane. These rectangles come from a recursive subdivision of the profile domain; each final rectangle is a *leaf*. Each leaf is charged its bound times its number of integer profiles. The parameters for each leaf are chosen numerically, and every resulting inequality is checked in rational arithmetic with directed logarithm bounds. The check confirms that the leaves cover the domain and sums their upward-rounded contributions. The rows of Table 1 also use the stronger bounds described next.

The main difficulty at $q = 1$ is the dependent case of (6). For some pools, a single return is free with probability one. Maximizing that probability separately at every step ignores the cost of forming such pools. The refinements for plain accumulators instead charge several exposures jointly.

For example, draw two nonzero forms independently from one pool. Let $P_{\parallel}$ be the probability that they are proportional and $P_{\perp}$ the probability that they are exact opposites. Then $P_{\perp} \leq P_{\parallel}/2$: in odd characteristic the nonzero forms on each line split into disjoint pairs $\{f, -f\}$, so a line of total mass σ_{ℓ} yields an opposite pair with probability at most $\sigma_{\ell}^2/2$, while $P_{\parallel} = \sum_{\ell} \sigma_{\ell}^2$. A proportional nonopposite pair cannot sum to zero on a surviving branch: imposing that equation would kill both required inputs. The *projective-pair bound* (Appendix G) uses this restriction for the first two inputs of each run. The *joint-renewal bound* (Appendix H) retains it through neighboring returns. The *complete-run bound* (Appendix K) counts a whole run, with its free returns and rank-gaining returns, in a fixed quotient, then lets the quotient change when a new zero equation is imposed. Each bound is uniform over every pool, so Lemma 5.1 applies after each scan.

Independent draws are an auxiliary counting device. The actual permutation draws labels without replacement. The complete-run bound absorbs the resulting correction in an explicit factor κ^{16} for a scan with κ active inputs. Thus its two-scan bound uses $K_1(s)K_2(h) = (sh)^{16}C_1\lambda_1^n C_2\lambda_2^n$. It also includes the correction terms that arise because p is finite and the runs still open at the end of a scan.

5.5 Which appendix proves each row

The $q = 1$ row of Table 1 is proved in Appendix K. Its cover assigns each leaf the projective-pair bound of Appendix G, the joint-renewal bound of Appendix H, or the complete-run bound, and it bounds source weights below 32 by the hypergeometric support bound of Appendix I. The rows with $q \geq 2$ are proved in Appendix J, which extends projective pairs to small coefficients and chooses, on each leaf, either that bound or the direct bound. For source weights below 32 it uses the hypergeometric support bound at $q = 2$ and the direct bound with (7)

for $q \geq 4$. The two rows of Table 2 are proved in Appendix M, which replaces (6) by an operator on support masks and ranks of the preceding w outputs and treats source weights below 32 with exact source coefficients and sixteen terminal iterations. Each certificate covers every sparse weight and dense support profile.

6 Discussion

The direct construction gives a practical range of proved distances over $\mathbb{F}_p$ and every extension $\mathbb{F}_{p^m}$, for $p = 2^{128} - 159$: from above 11.7% at 12.171 ms with plain accumulators to above 42.987% at 32.342 ms with width four, all with setup failure below 2^{-40} and timed over $\mathbb{F}_p$. These are proved lower bounds, not measurements of the codes' actual minimum distances.

The design variables have different effects. Moving from plain accumulators to native products at $q = 2$ raises the time from 12.171 ms to about 17.9 ms. Beyond that switch, a larger coefficient alphabet improves the lower bound at little additional measured cost, but the gains shrink across the reported width-one range. Larger width costs more arithmetic and metadata, but permits stronger distance guarantees. The deterministic ceiling at 39.6855% shows that coefficient size alone cannot take two width-one scans to the width-four point. It does not explain the entire gap between that ceiling and the best width-one lower bound.

The local precode supplies another tradeoff. The binary subset-sum block followed by two plain accumulators proves distance above 12.56% with fewer than half as many additions as the Newton all-one encoder. Its failure bound is below $2^{-22.05}$; we did not time it. Despite its smaller local minimum distance, its exact catalog of shortened subspaces V_S (Section 5) gives a stronger bound for dense supports. Comparisons between precodes must therefore account for their support catalogs and failure probabilities, in addition to arithmetic cost.

These results give codes in one fixed characteristic, with identical distance guarantees over its extension fields. They do not prove that binary distance transfers to every prime, or that one sampled circuit works simultaneously across characteristics. A PCG using such an encoder needs a distance argument for its deployed field and a security analysis with the corresponding noise parameters. Scalar extension, a field-independent combinatorial certificate, or a direct prime-field theorem can supply that argument; a binary theorem alone cannot justify a change of characteristic.

Several mathematical gaps remain. A sharper upper bound for the all-one ensemble would locate its typical distance between the proved lower bound and the general width-one ceiling. Width eight has only numerical dense-support estimates (Table 3); a finite certificate would also have to control sparse profiles and finite prefactors, and we have not timed it. A simultaneous guarantee across a prescribed family of prime fields would require a further argument.

The supporting case studies expose different open questions. For repetition-based BAA, the quinary prediction near 0.2606 still needs global first-moment control and an argument relating expected counts to typical distance. For growing primes, determinant bounds restrict small rank-loss witnesses but leave large

residual cores uncontrolled. For Walsh–Hadamard codes, the fixed-alphabet obstruction leaves open whether a growing multiplier alphabet can guarantee distance at a useful addition cost. These questions remain separate from the finite certificates for the direct construction.

References

1. Boyle, E., Couteau, G., Gilboa, N., Ishai, Y., Kohl, L., Resch, N., Scholl, P.: Correlated pseudorandomness from expand-accumulate codes. In: *Advances in Cryptology — CRYPTO 2022, Part II*. LNCS, vol. 13508, pp. 603–633 (2022)
2. Couteau, G., Rindal, P., Raghuraman, S.: Silver: Silent VOLE and oblivious transfer from hardness of decoding structured LDPC codes. In: *Advances in Cryptology — CRYPTO 2021, Part III*. LNCS, vol. 12827, pp. 502–534 (2021)
3. Kliever, J., Zigangirov, K.S., Koller, C., Costello, Jr., D.J.: Coding theorems for repeat multiple accumulate codes. arXiv:0810.3422 (2008)
4. Kolesnikov, V., Peceny, S., Rachuri, R., Raghuraman, S., Rindal, P., Shah, H.: Block-accumulate codes: Accelerated linear codes for PCGs and ZK. In: *Advances in Cryptology — CRYPTO 2026, Part VIII*. LNCS, vol. 16807, pp. 397–429 (2026), also IACR Cryptology ePrint Archive, Report 2025/1828
5. Kolesnikov, V., Peceny, S., Raghuraman, S., Rindal, P.: Stationary syndrome decoding for improved PCGs. In: *Advances in Cryptology — CRYPTO 2025*. pp. 284–317 (2025), also IACR Cryptology ePrint Archive, Report 2025/295
6. Li, Z., Liu, H., Xing, C., Yao, Y., Yuan, C.: Faster pseudorandom correlation generators via Walsh–Hadamard transform. IACR Cryptology ePrint Archive, Report 2026/196 (2026)
7. Raghuraman, S., Rindal, P., Tanguy, T.: Expand-convolute codes for pseudorandom correlation generators from LPN. In: *Advances in Cryptology — CRYPTO 2023, Part IV*. LNCS, vol. 14084, pp. 602–632 (2023)
8. Vapnik, V.N., Chervonenkis, A.Y.: On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications* **16**(2), 264–280 (1971)

A BAA as a case study

Block-and-accumulate (BAA) codes were introduced as fast linear encoders for PCGs and zero-knowledge systems [4]. This appendix studies their addition-only specialization with repetition three.

Fix $N = 3K$. Let $B_3 : F^K \rightarrow F^N$ repeat each message coordinate three times, let A be the prefix-sum accumulator, and sample independent uniform coordinate permutations P_1, P_2 . The encoder is

$$E_{P_1, P_2} := AP_2AP_1B_3.$$

It has rate $1/3$ over every field and uses only additions and data movement. Its matrices have integer entries, so one pair (P_1, P_2) defines an encoder over every field. We call the resulting random family the repetition-three BAA ensemble.

The question is whether this ensemble, taken over an odd prime field, retains the good distance of the binary construction, rather than merely a positive asymptotic distance. Positive distance holds over every fixed prime; Appendix D records these baseline theorems. The ternary lower bound 0.196 of Theorem D.1 already exceeds the binary value 0.1323 reported in [3], but the quinary evidence is substantially ahead of the proof. This appendix explains that evidence and its limitations.

A.1 The first-moment prediction

The first-moment analysis tracks the compositions of symbol values in $\mathbb{F}_p$, not only supports, through both accumulators. A separate spacing argument treats sparse messages. For a fixed prime p , the stationary type model consists of message, middle, and output distributions μ, β, ζ on $\mathbb{F}_p$, and two transition laws Q_1, Q_2 on increments and states (D, S) . Under Q_1 , the increment has law μ and both $S, S + D$ have law β . Under Q_2 , the increment has law β and both states have law ζ . Write $\tau = (\mu, \beta, \zeta, Q_1, Q_2)$ for such a type. Its first-moment objective is

$$\Phi_p(\tau) = H(Q_1) + H(Q_2) - \frac{2}{3}H(\mu) - 2H(\beta) - H(\zeta),$$

where H is entropy with natural logarithms. Its normalized message and output weights are $a(\tau) = 1 - \mu(0)$ and $c(\tau) = 1 - \zeta(0)$. Define the first-moment threshold of this stationary model by

$$\delta_{\text{FM}}(p) := \inf \mathcal{D}_p, \quad \mathcal{D}_p := \left\{ \delta > 0 : \begin{array}{l} \text{some admissible } \tau \text{ satisfies } a(\tau) > 0, \\ c(\tau) \leq \delta, \text{ and } \Phi_p(\tau) > 0 \end{array} \right\}.$$

The restriction $a(\tau) > 0$ isolates linear-weight messages. A proof of a distance lower bound must also control the sparse range uniformly.

Numerical optimization gives the following sign brackets. The negative side of the $\mathbb{F}_3$ row is certified by Theorem D.1, which excludes codewords of relative weight up to 0.196. The other signs refer to returned local optima, not certified global maxima; the quinary evaluation restricts message support to at least $0.002K$.

field	negative exponent	positive exponent
$\mathbb{F}_2$	$\delta = 0.132$	$\delta = 0.133$
$\mathbb{F}_3$	$\delta = 0.195$	$\delta = 0.197$
$\mathbb{F}_5$	$\delta = 0.260600$	$\delta = 0.260612$

The binary bracket reproduces the binary value 0.1323 reported in [3]. The odd-prime brackets suggest an improvement in the ensemble threshold, not equality of distances for each encoder across fields. For $\mathbb{F}_5$, the candidate is near 0.2606, without a confidence interval.

For a finite abelian group G and distributions α, β on G , let R range over laws on $(D, S) \in G^2$ and define

$$\Gamma_G(\alpha, \beta) := \{R : D \sim \alpha, S \sim \beta, D + S \sim \beta\},$$

$$I_{\text{stat}}^G(\alpha, \beta) := \inf_{R \in \Gamma_G(\alpha, \beta)} I_R(D; S).$$

The infimum of the empty set is $+\infty$. Otherwise compactness and continuity give a minimum. We omit the superscript when $G = \mathbb{F}_5$, and call $I_{\text{stat}}^G(\alpha, \beta)$ the stationary information cost of (α, β) . Since $H(Q_1) = H(\mu) + H(\beta) - I_{Q_1}(D; S)$, and similarly for Q_2 , the objective is $\Phi_p(\tau) = \frac{1}{3}H(\mu) - I_{Q_1}(D; S) - I_{Q_2}(D; S)$. Each accumulator therefore lowers the exponent by at least the stationary information cost of its increment and state compositions.

For $p = 5$, the minimization defining I_{stat} can be reduced without relaxing its objective. Let α be an increment composition, let β be a stationary state composition, and let $R \in \Gamma_{\mathbb{F}_5}(\alpha, \beta)$. Set $U := \mathbf{1}[D \neq 0]$ and $V := \mathbf{1}[S \neq 0]$. The admissible crossings are

$$X(\alpha, \beta) := \{R : \Pr[U = 1, V = 0] : R \in \Gamma_{\mathbb{F}_5}(\alpha, \beta)\}.$$

Fix $x \in X(\alpha, \beta)$, and let J_x be the law of (U, V) determined by x and the support marginals. The conditional symbol laws are

$$\alpha_u(d) := \Pr_{\alpha}[D = d \mid U = u], \quad \beta_v(s) := \Pr_{\beta}[S = s \mid V = v].$$

On a conditioning event of zero mass, fix any distribution supported on the corresponding zero or nonzero symbol class; its contribution below is zero. The support-preserving product law is

$$Q_x(d, s) := J_x(u, v)\alpha_u(d)\beta_v(s), \quad u = \mathbf{1}[d \neq 0], \quad v = \mathbf{1}[s \neq 0].$$

For every feasible R with crossing x , the chain rule for relative entropy gives

$$D_{\text{KL}}(R \parallel \alpha \otimes \beta) = I_{J_x}(U; V) + D_{\text{KL}}(R \parallel Q_x).$$

Since R has marginals α and β , the left-hand side equals $I_R(D; S)$. Minimizing over feasible R therefore gives

$$I_{\text{stat}}(\alpha, \beta) = \inf_{x \in X(\alpha, \beta)} \left(I_{J_x}(U; V) + \min_R D_{\text{KL}}(R \parallel Q_x) \right),$$

where the inner minimum ranges over $R \in \Gamma_{\mathbb{F}_5}(\alpha, \beta)$ whose support indicators (U, V) have law J_x . In the first-moment optimization, applying the reduction to both accumulators replaces the two 5-by-5 transition laws with the three compositions μ, β, ζ and two crossings.

As a numerical consistency check, the reduced objective and the full transition objective agree within $1.9 \cdot 10^{-7}$ at output fractions 0.22, 0.24, 0.25, 0.26, 0.260600, 0.260612, and 0.261. The discrepancy is below $1.2 \cdot 10^{-13}$ at the last four points. These calculations produce the refined quinary sign bracket in the table. The reduced optimization starts from the full optimizer's output and shares its numerical routines, so their agreement is not an independent global optimization.

A.2 What is established near the candidate threshold

Sparse messages are controlled beyond the candidate threshold. For $1 \leq s \leq K$, let $Z_s(d)$ count messages $x \in \mathbb{F}_5^K$ with $\text{wt}(x) = s$ and $\text{wt}(E_{P_1, P_2}x) \leq d$.

Proposition A.1 (Sparse quinary messages). *Over $\mathbb{F}_5$,*

$$\sum_{s=1}^{\lfloor 10^{-5}K \rfloor} \mathbb{E}_{P_1, P_2} Z_s(\lfloor 0.262N \rfloor) \longrightarrow 0 \quad \text{as } K \rightarrow \infty.$$

Since E_{P_1, P_2} is injective, the sum is the expected number of nonzero codewords of weight at most $\lfloor 0.262N \rfloor$ whose messages have weight at most $10^{-5}K$. The proof, at the end of Appendix D, applies Lemma D.2 with $\eta = 10^{-5}$ and $\delta = 131/500 = 0.262$, using

$$4\delta(1 - \delta) \leq \frac{48339}{62500} < \left(\frac{22}{25}\right)^2 \quad (0 < \delta \leq 0.262)$$

and the geometric summation in that lemma. Proposition A.1 controls a small linear message-weight range, not the distance of the whole code.

The message weights $10^{-5} < \text{wt}(x)/K < 0.002$ lie between Proposition A.1 and the support restriction of the quinary evaluation. The converged local optimizations performed in this interval return negative exponents at output cutoff 0.26 or 0.261. Neither this interval nor the range above $0.002K$ has a global upper bound.

The stationary information cost also admits a fixed-dual lower bound directly in the two compositions. A feasible solution of the entropy dual of the minimization defining I_{stat} determines a constant λ_0 , a vector λ , and a positive matrix M such that

$$I_{\text{stat}}(\alpha, \beta) \geq \lambda_0 + \langle \lambda, (\alpha, \beta) \rangle - \log(\alpha^\top M \beta).$$

The dual data do not depend on the compositions. A finite family of such inequalities, fixed in advance, therefore bounds both accumulator terms at one common middle composition. Bounding the two terms by separate minimizations over the middle composition would allow a different middle composition in each term and weaken the bound.

At output cutoff 0.24, fixed-dual evaluation over a partition into 101 boxes returns only negative local values, the largest being $-5.93 \cdot 10^{-4}$. At 0.22, directed arithmetic certifies particular boxes; in one, a rational LP dual gives the upper bound $-8.72 \cdot 10^{-6}$. Neither these certificates nor the floating-point evaluations cover the full domain, and a certificate for a restricted family at 0.22 does not resolve the message-weight gap near 0.26.

A global certificate by this method faces a structural limitation. Replacing products of composition entries by separate moment variables permits mixtures that need not arise from one composition. A positive relaxed optimum can therefore coexist with negative values at the tested compositions. Consistency of the numerical model and validity of local certificates leave a separate global problem.

A.3 From an expected count to typical distance

Conjecture A.2 (First-moment threshold). For each fixed prime p , the minimum distance of the repetition-three BAA ensemble satisfies

$$\frac{d_{\mathbb{F}_p}(E_{P_1, P_2})}{N} \xrightarrow{\text{Pr}} \delta_{\text{FM}}(p) \quad \text{as } K \rightarrow \infty.$$

The first moment addresses only the lower side of Conjecture A.2. A negative bulk exponent, combined with the spacing argument, excludes smaller codewords with probability tending to one. A positive exponent proves only that their expected number is exponential. Rare permutation pairs could dominate that expectation. The reverse inequality therefore requires concentration, a second-moment argument, or another typical-existence theorem. In particular, the numerical quinary transition near 0.2606 is a conjectural threshold, not a claimed distance theorem or a partially certified global bound at that value.

There is, however, initial evidence that a second-moment argument may be possible. Count one-dimensional message subspaces rather than nonzero messages. Each quinary message line has four nonzero representatives. Injectivity identifies message lines with projective codewords; distinct lines remain distinct even if their codeword supports agree. A pair of distinct lines is described coordinatewise over the additive group $\mathbb{F}_5^2$. We call the two coordinates of this group the replicas; the replica marginals of a pair composition are its two coordinate projections. The two accumulator terms of a pair are stationary information costs over $\mathbb{F}_5^2$. The following identity identifies the product-pair benchmark.

Proposition A.3 (Product tensorization). *For distributions α and β on G ,*

$$I_{\text{stat}}^{G^2}(\alpha \otimes \alpha, \beta \otimes \beta) = 2I_{\text{stat}}^G(\alpha, \beta).$$

Proof. If $\Gamma_G(\alpha, \beta)$ is empty, a feasible law on G^2 would project to a feasible law on G , a contradiction. Both sides are then $+\infty$. Otherwise a minimizing law exists by compactness. The product of two minimizing one-coordinate laws proves the upper bound. For the reverse bound, let Q be any feasible law for the left-hand side and group its variables as $X_i = (D_i, S_i)$. If $R_i = \alpha \otimes \beta$, the relative-entropy chain rule gives

$$D_{\text{KL}}(Q_{X_1 X_2} \parallel R_1 \otimes R_2) = D_{\text{KL}}(Q_{X_1} \parallel R_1) + D_{\text{KL}}(Q_{X_2} \parallel R_2) + I_Q(X_1; X_2).$$

Each coordinate projection of Q is feasible for the one-coordinate problem. The first two terms are therefore at least $I_{\text{stat}}^G(\alpha, \beta)$, and the last term is nonnegative.

Proposition A.3 shows that product compositions cannot reduce an accumulator's stationary information cost below twice the single-message value. A larger stationary pair exponent must come from nonproduct compositions; finite-length second moments require additional control even when no such improvement occurs.

At output fraction 0.261, a local test fixes both replica marginals in each of the message, middle, and output compositions. This leaves 16 correlation directions

per composition. The resulting 48-dimensional numerical Hessian is negative definite, with eigenvalues from approximately $-1.46 \cdot 10^{-2}$ to $-2.51 \cdot 10^{-5}$ in the chosen coordinates. An inverse-covariance calculation and directional second differences agree on the negative curvature; both use one numerical transition solver and neither is a directed-arithmetic certificate.

A further numerical optimization frees both replica marginals and optimizes the two pair transition laws jointly. Its starting points include products, mixtures toward proportional messages, and invertible linear changes of pair coordinates; some optimizations also require the messages to differ from every scalar multiple on at least $0.001K$ coordinates. No returned candidate beats the product exponent beyond rounding. Some unrestricted optimizations approach the proportional boundary. Continued to the numerical small-step criterion, the highest-valued of these end within $4.6 \cdot 10^{-11}$ of the product exponent, and their two transition laws differ from the products of their replica marginals by less than $2.5 \cdot 10^{-13}$ in L1 distance. Their correlations disappear before their exponents finish improving; the remaining slow motion lies in the individual marginal laws. This supports the product prediction from those starting points, without establishing global optimality.

The evidence therefore has a specific scope. Sparse messages are controlled analytically; the symbol-composition computations locate a candidate transition; and the pair tests find no competing correlated type in the optimizations performed. A proof still needs global first-moment control, a treatment of nearly proportional pairs and other possible pair maxima, and finite-length estimates strong enough to establish concentration. None of these requirements is replaced by a positive expected count.

A.4 What large characteristic proves

The cryptographic regime instead has message length $K \approx 2^{20}$ and $\log_2 p \approx 128$. Here p is much larger than N , but its bit length is much smaller than N . Two deterministic statements explain why large characteristic limits small witnesses of rank loss.

Write $L := AP_2AP_1$ over the integers, so $E = LB_3$ in the column convention. Let $\mathcal{H}_0 \in \mathbb{Z}^{2K \times N}$ have rows $(1, -1, 0)$ and $(0, 1, -1)$ on each repetition block, and put $\mathcal{H} := \mathcal{H}_0 L^{-1}$. The matrices A and P_i are unimodular, so $\ker_F \mathcal{H} = \text{im}_F E$ over every field F . Write $d_p := d_{\mathbb{F}_p}(E)$ for this fixed setup.

Proposition A.4 (Absolute distance floor). *For every realized repetition-three BAA encoder,*

$$d_p \geq \min \{d_2, \lceil \log_{\sqrt{20}} p \rceil \}.$$

Proof. Right multiplication of a row by A^{-1} takes adjacent differences. After the first difference, a row of $\mathcal{H}_0$ has either four unit entries, or entries of magnitudes $(1, 2, 1)$, or a truncated version at the boundary. Permutations preserve these magnitudes. A second difference has squared norm at most 20: its squared norm is at most twice the original squared norm minus twice the sum of adjacent

products. Four unit entries give at most $8 + 6 = 14$; magnitudes $(1, 2, 1)$ give at most $12 + 8 = 20$. Thus each row of $\mathcal{H}$ has norm at most $\sqrt{20}$.

Suppose $d_p = d < d_2$ and let W be the support of a minimum word over $\mathbb{F}_p$. The columns $\mathcal{H}_W$ are independent over $\mathbb{F}_2$ but dependent over $\mathbb{F}_p$. Some d -by- d minor is therefore an odd nonzero integer divisible by p . Hadamard's inequality gives $p \leq 20^{d/2}$, proving the claim.

A prime near 2^{128} gives a distance floor of about 60 whenever the binary distance is at least 60. This floor is not linear in N .

For the second bound, retain the circuit's internal variables rather than forming $\mathcal{H}$. Choose permutation indices so that $(P_1 z)_i = z_{\pi_1(i)}$ and $(P_2 v)_i = v_{\pi_2(i)}$. The homogeneous circuit equations are

$$z_i - x_{\lceil i/3 \rceil} = 0, \quad v_i - v_{i-1} - z_{\pi_1(i)} = 0, \quad y_i - y_{i-1} - v_{\pi_2(i)} = 0, \quad 1 \leq i \leq N,$$

where $v_0 = y_0 = 0$. To test an output support W , set $y_i = 0$ for $i \notin W$. Repeatedly delete a variable forced to zero by a one-variable equation, and discard zero equations. These operations preserve the solution space on the remaining variables over every field. They do not take linear combinations of equations. Consequently, the residual integer matrix still has entries in $\{0, 1, -1\}$ and at most three nonzeros in each row.

Its support graph has a vertex for each remaining equation and variable, and an edge for each nonzero entry. For a connected component, define its cycle rank $\kappa := \#\text{edges} - \#\text{vertices} + 1$.

Proposition A.5 (Cycle-rank barrier). *Suppose the matrix R of a connected residual component has $\text{rank}_{\mathbb{F}_p} R < \text{rank}_{\mathbb{Q}} R$. Its cycle rank is at least $\lceil \log_2 p \rceil$.*

Proof. Put $r = \text{rank}_{\mathbb{Q}} R$. Some r -by- r minor has nonzero integer determinant, and every such determinant vanishes modulo p . Expand one nonzero determinant over perfect matchings of its support graph. Fix a spanning forest of that graph. Two matchings with the same nonforest edges would have a symmetric difference consisting of cycles inside a forest, so they must coincide. The graph has at most 2^κ matchings: deleting vertices or edges cannot increase its cycle rank. Each determinant term is $+1$ or -1 , because all nonzero entries are units. Hence $p \leq |\det| \leq 2^\kappa$.

This argument concerns the unweighted circuit above. Arbitrary integer coefficients or elimination that creates larger entries need not satisfy it.

Propositions A.4 and A.5 exclude rank-loss witnesses below their respective size bounds. They do not control a support that creates a large, nearly square residual core. The missing theorem is a uniform random-saturation statement over every support of size below the binary threshold. Fixed-support experiments show substantial rank surplus, but a minimum-distance witness selects its support after seeing both permutations.

B The Walsh–Hadamard alternative

Walsh–Hadamard PCGs take a different route. They define a quasi-abelian code directly over a large prime field and use a transform containing only additions and subtractions [6]. This avoids an unsupported binary-distance transfer. Random spectral multipliers still supply a large family of projective types.

Drawing the multipliers from a fixed finite alphabet, such as signs, would make the encoder still cheaper. Such fixed-alphabet variants, however, have a structural obstruction. Fix a field F of odd characteristic, put $V = \mathbb{F}_2^r$ and $k = 2^r$, and fix $c \geq 2$. The Walsh matrix is $W_{u,\xi} = (-1)^{\langle u,\xi \rangle}$ for $u, \xi \in V$, with $W^{-1} = k^{-1}W$. For each frequency choose $a(\xi) = (a_0(\xi), \dots, a_{c-1}(\xi)) \in F^c \setminus \{0\}$. Define the encoder and its image by

$$E_a(x) := (W^{-1}(a_j \odot Wx))_{j=0}^{c-1}, \quad C_a := \text{im}_F E_a,$$

where $\odot$ denotes coordinatewise multiplication. Each frequency has a nonzero multiplier in some block, so E_a is injective. Thus C_a has length ck , dimension k , and rate $1/c$. Write $\sigma(\xi) := [a(\xi)] \in \mathbb{P}^{c-1}(F)$ for the projective type of ξ , and $\Delta := d(C_a)/(ck)$.

Proposition B.1 (Affine-flat obstruction). *If σ is constant on an affine flat of dimension s , then the relative minimum distance of the length- ck code is at most 2^{-s} .*

Proof. On the flat, write the multiplier vector as a fixed projective vector times a nonzero scalar. Choose the message spectrum to cancel that scalar and equal a character on the flat. Its inverse Walsh transform is supported on one coset of the orthogonal subspace. Each nonzero output block has weight $k/2^s$.

Suppose the types $\sigma(\xi)$ are independent and identically distributed, with support of size at most M . A partition into parallel s -flats gives

$$\Pr[\Delta > 2^{-s}] \leq \exp\left(-\frac{k}{2^s} M^{1-2^s}\right).$$

Choosing $s \asymp \log_2(\log k / \log M)$ yields

$$\Delta = O\left(\frac{\log M}{\log k}\right)$$

with probability tending to one for fixed $M \geq 2$. If $M = 1$, the type is constant on all of V , and Proposition B.1 with $s = r$ applies directly. Thus this random fixed-alphabet model fails asymptotically at fixed block count. The argument does not assert failure for every deterministic assignment of finitely many spectral types.

The obstruction is asymptotic. At the practical dimension 2^{20} , the following computation shows that an alphabet of 1024 small integers already makes the three-flat witnesses of this kind unlikely. Let $r = 20$, $c = 4$, and $\text{char } F > 1024$.

Set $a_0(\xi) = 1$ and sample every $a_j(\xi)$, for $j = 1, 2, 3$ and $\xi \in V$, independently and uniformly from $\mathcal{A} := \{\pm 1, \dots, \pm 512\} \subset F^\times$. The characteristic condition makes these $|\mathcal{A}| = 1024$ elements distinct. Multiplication by these small integers can use addition chains.

Call an affine three-flat $L = \xi_0 + U$ dangerous if, for each $j = 1, 2, 3$, there are $b_j \in \mathcal{A}$ and a character $\chi_j : U \rightarrow \{1, -1\}$ with $a_j(\xi_0 + u) = b_j \chi_j(u)$ for every $u \in U$. The message spectrum 1_L then gives a word of relative weight $1/8$: each output block is supported on one coset of $U^\perp$. For a fixed flat and block, exactly $8|\mathcal{A}|$ of the $|\mathcal{A}|^8$ assignments have this form. Independence and a union bound therefore give

$$\Pr[\text{some dangerous three-flat}] \leq 2^{17} \begin{bmatrix} 20 \\ 3 \end{bmatrix}_2 (8|\mathcal{A}|^{-7})^3 < 2^{-131.39} < 2^{-128}. \quad (11)$$

Here the number of linear three-dimensional subspaces is

$$\begin{bmatrix} 20 \\ 3 \end{bmatrix}_2 := \prod_{i=0}^2 \frac{2^{20} - 2^i}{2^3 - 2^i}.$$

The strict 2^{-128} comparison is an integer inequality after multiplying by $|\mathcal{A}|^{21}$; its arithmetic check is listed in Appendix N. Equation (11) excludes only this witness family; it does not establish a global distance lower bound.

C The four-input precode and its implementation

This appendix analyzes the two-scan construction with the four-input precode, an $[8, 4, 5]$ MDS code, in place of the Newton precode. With coefficients from $\{1, \dots, 2048\}$ it proves relative distance above 14.5% at rate $1/2$, with setup failure below $2^{-56.90}$. The Newton precode B_8 of Section 4 supersedes this family: with $q = 4$ it already proves 18.625% (Table 1), and at each common alphabet size $q \in \{32, 64, 128, 256\}$ its proved distance is more than twice as large. For $q \in \{64, 128, 256\}$, separate timing runs on the same machine measured the Newton encoder about one millisecond slower (Tables 1 and 4).

The encoder retains the two-scan architecture of BAA [4]. Its random coefficients multiply the running scan states. The analysis counts support patterns and the ranks of their zero equations, without taking a union over field-valued messages.

C.1 Construction and finite guarantee

Fix a prime $p > q \geq 32$, let $k = 2^{20}$ and $n = 2k$, and work over $\mathbb{F}_p$. Partition the input into blocks (a, b, c, d) of four symbols. For each block, define

$$f(t) = a + bt + c \binom{t}{2} + d \binom{t}{3}, \quad B(a, b, c, d) = (f(0), \dots, f(6), d).$$

Finite differences evaluate this map with 15 additions. Any four output coordinates determine the input: four evaluations determine a cubic; three evaluations together with d determine its remaining coefficients. Thus the four-input precode B is an $[8, 4, 5]$ MDS code. Apply B independently to all blocks, obtaining a vector of length n .

At setup, sample two independent uniform permutations π_1, π_2 of $\{0, \dots, n-1\}$. Independently sample each $a_{j,i}$ uniformly from $\{1, \dots, q\}$, for $j \in \{1, 2\}$ and $1 \leq i < n$. Round j maps its input v to y by

$$u_i = v_{\pi_j(i)}, \quad y_0 = u_0, \quad y_i = u_i + a_{j,i}y_{i-1} \quad (1 \leq i < n). \quad (12)$$

Output the complete second-round vector. Each round is invertible, so the code has dimension k for every setup. All vectors after B have length n ; there is no systematic copy or final compression.

Theorem C.1 (Finite small-coefficient guarantee). *For each fixed prime $p > q$, the construction above satisfies the following bounds over its setup randomness:*

q	$d_{\min}$ at least	$\Pr[\text{failure}]$ less than
32	209716	$2^{-24.85}$
40	230897	$2^{-44.11}$
64	251659	$2^{-45.65}$
128	277873	$2^{-47.87}$
256	288359	$2^{-50.24}$
512	298845	$2^{-52.46}$
1024	301990	$2^{-54.72}$
2048	304088	$2^{-56.90}$

Failure means that the minimum distance is below the listed threshold. On a successful setup the distance bound holds for every nonzero message, including messages selected after the setup is revealed. Failure exponents are truncated toward zero, so the stated bounds are conservative.

The $q = 32$ row uses the sparse bound with $\tau = 1$, which ignores coefficient randomness for source weights at most 128 (Appendix E); this alphabet was not implemented or timed. Every other row retains coefficient randomness in the sparse-support calculation.

For $q = 40$ the relative distance exceeds 11.01%, above the binary rate-1/2 Gilbert–Varshamov (GV) benchmark $H_2^{-1}(1/2) \approx 0.110028$. This comparison is not a binary-code guarantee, and it is far below the prime-field GV distance. The theorem fixes the prime before setup; it does not claim that one sampled circuit works simultaneously over primes. The larger alphabets retain the same precode, two scans, and vector lengths. Sizes 128, 512, and 2048 give relative distances above 13.25%, 14.25%, and 14.5%, respectively. All rows from $q = 40$ through $q = 2048$ use the same measured scan kernel and memory footprint. A separate certificate for $q = 2048$ gives minimum distance at least 230897 with failure below $2^{-58.45}$. These are achieved finite bounds; the table does not claim optimal distance or failure probability for an alphabet.

Multiplication by a positive integer m has a binary double-and-add circuit of length $\lceil \log_2 m \rceil + \text{popcount}(m) - 1$. For $1 \leq m \leq 40$, these lengths sum to 205 and have maximum eight. Including B and the input additions in (12), the $q = 40$ encoder therefore uses at most $19.875n - 18$ additions, with expectation $14.125n - 12.25$. The expectation is over setup, not over messages. The coefficient alphabet may instead be any q distinct nonzero field elements for the distance proof; the addition-cost claim uses the positive integer alphabet. For the measured $q = 2048$ configuration, the expected addition count is $(15815/512)n - 14855/512$, about $30.889n$, and the worst-case count is $(351/8)n - 42$. These circuit counts grow with the alphabet. The native small-integer multiply/add implementation below uses the same arithmetic instruction sequence for all measured alphabets.

C.2 Equation ranks instead of a message union

Fix a source support S and support patterns for both rounds, and condition on permutations that map these supports as prescribed. The shortened source space V_S of inputs satisfying the source zero coordinates has dimension D_S . Expose coefficients in round order. Let $W \subseteq V_S^*$ be the span of all zero equations exposed so far. A branch remains alive only if $\ker W$ contains a message satisfying every required nonzero condition whose form is already known.

A *return* is a zero output that ends a run of nonzero outputs. Its input must be nonzero. The new zero equation has the form $\ell + ar = 0$, where ℓ is the current input form and r the preceding state form. Both are known before exposing a . On an alive branch, $\ell \notin W$. At most one coefficient can make $\ell + ar \in W$: two distinct such coefficients would imply $r \in W$ and then $\ell \in W$.

For $\tau \geq 1$, put $g = 1/q + (1 - 1/q)/\tau$. Give an alive branch the *rank potential* $\tau - \text{rank } W$ and a dead branch potential zero. Each return multiplies the conditional expected potential by at most g ; other steps cannot increase it. A realizable nonzero path has $\text{rank } W \leq D_S - 1$. Consequently, for a path with c returns,

$$\Pr[\text{prescribed support path is realizable}] \leq \min\{1, \tau^{D_S-1} g^c\}. \quad (13)$$

This argument is adaptive across both rounds. It requires no independence among the zero equations and incurs no factor p^{D_S} .

For a local output support of size $s \in \{5, 6, 7, 8\}$, the source dimension is $s - 4$. The dimension-weighted support polynomial is therefore

$$Q(z, \tau) = 1 + 56z^5\tau + 28z^6\tau^2 + 8z^7\tau^3 + z^8\tau^4. \quad (14)$$

Let $A_g(s, h)$ count admissible input/output support pairs of one scan, with weights s, h , assigning a factor g per return. It is the coefficient of $x^s y^h$ in

$$[1, 0] \begin{bmatrix} 1 & xy \\ gx & (1+x)y \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (15)$$

The initial state is zero and either final state is allowed. This count retains support patterns that no field values realize.

Table 4. Complete encoding at $k = 2^{20}$, $n = 2^{21}$ and $p = 2^{128} - 159$. Times are in milliseconds on one core of an AMD Ryzen 9 7950X.

q	Precode	B	Scan 1	Scan 2	Total
40	1.919	7.307	7.730	16.947	
64	1.924	7.317	7.741	16.989	
128	1.918	7.277	7.661	16.864	
256	1.919	7.315	7.737	16.964	
512	1.922	7.337	7.739	17.009	
1024	1.920	7.345	7.743	17.007	
2048	1.921	7.299	7.704	16.909	

For a forbidden output weight e , write $H_g(h, e) = \sum_{j=1}^e A_g(h, j)$ and $L = n/8$. A uniform permutation maps a fixed support of size s to a uniform s -subset, so each prescribed placement of the source support has probability $1/\binom{n}{s}$, and each placement of the intermediate support has probability $1/\binom{n}{h}$. Combining these normalizations with (13) gives

$$\Pr[d \leq e] \leq \frac{1}{\tau} \sum_{s=1}^{4e} \frac{[z^s]Q(z, \tau)^L}{\binom{n}{s}} \sum_{h=1}^{2e} \frac{A_g(s, h)H_g(h, e)}{\binom{n}{h}}. \quad (16)$$

The inverse scan is bidiagonal, giving $s \leq 2h \leq 4e$. A separate τ may be used for each source weight. Appendix E gives the finite evaluation of this bound, including the sparse range.

C.3 Implementation cost

We implemented alphabets $q \in \{40, 64, 128, 256, 512, 1024, 2048\}$ over $p = 2^{128} - 159$. The optimized precode kernel evaluates eight blocks together with AVX-512, postpones reduction in 43-bit limbs, and uses contiguous loads and stores. Each scan packs its 21-bit permutation index and eleven-bit coefficient-minus-one into one 32-bit word. Thus $q = 2048$ is the largest alphabet in this layout. The scan loop gathers 16 entries at a time and prefetches 256 entries ahead. Two length- n field buffers suffice; one holds the final output.

Table 4 reports, for each alphabet, the median of six process medians, each over 31 timed encodes after two warm-up encodes, with GCC 15.2 at `-O3 -march=native`; setup is not timed. Stage medians need not sum to the median total. All seven totals lie within 0.5% of the $q = 40$ total, and process medians range from 16.855 to 17.075 ms for $q = 40$ and from 16.866 to 17.056 ms for $q = 2048$. The larger alphabets therefore show no measurable runtime penalty.

The optimized implementation uses native small-integer multiplication; it evaluates exactly the same field map as the addition circuit. The measured code is therefore distinct from a strict addition-only implementation, although both have the same distance theorem. Retained input, work buffers, and metadata occupy 96.008 MiB, excluding validation temporaries and stack scratch.

D Fixed-prime BAA baseline bounds

The following results establish positive distance for the ensemble defined in Appendix A. They provide baselines, not a certificate for the quinary first-moment prediction $\delta \approx 0.2606$ of Appendix A.

Theorem D.1 (Fixed-prime BAA bounds). *For the repetition-three BAA ensemble,*

$$\begin{aligned} \Pr[d_{\mathbb{F}_3}(E_{P_1, P_2}) \leq \lfloor 0.196N \rfloor] &\longrightarrow 0, \\ \Pr[d_{\mathbb{F}_5}(E_{P_1, P_2}) \leq \lfloor N/20 \rfloor] &\longrightarrow 0, \\ \Pr[d_{\mathbb{F}_7}(E_{P_1, P_2}) \leq \lfloor 9N/500 \rfloor] &\longrightarrow 0. \end{aligned}$$

The probability is over (P_1, P_2) as $K \rightarrow \infty$.

The estimates combine sparse-message bounds with certified negative bounds for the bulk exponent. For the quinary bound $1/20 = 0.05$, the spacing argument covers message supports up to $K/60$. A sign-orbit relaxation covers larger supports, with certified exponent upper bound $-1.1386 \cdot 10^{-6}$. The verifier encloses logarithms by directed intervals and checks rational LP duals. This is a global positive-distance result, unlike the region-restricted certificates at larger cutoffs in Appendix A.

The ternary bound 0.196 exceeds the binary value 0.1323 reported for repetition three and two accumulators [3]. This compares ensemble bounds, not the distances of one encoder over $\mathbb{F}_2$ and $\mathbb{F}_3$.

D.1 From spacing and type counts to distance

This subsection supplies the analytic bounds used by Theorem D.1. All limits keep p fixed and take $N = 3K$. For $1 \leq s \leq K$, let $Z_s(d)$ count messages $x \in \mathbb{F}_p^K$ with $\text{wt}(x) = s$ and $\text{wt}(E_{P_1, P_2}x) \leq d$.

Lemma D.2 (Sparse messages). *Fix $0 < \delta < 1/2$ and $0 < \eta < 1$. Put*

$$c_\delta := -\frac{1}{2} \log(4\delta(1-\delta)), \quad L_\delta := e^{c_\delta} - 1, \quad R_0 := 3/\eta,$$

$$g(R) := \frac{(p-1)eR}{3} \left(\frac{3}{(R-3)L_\delta} \right)^{3/2}.$$

If $R_0 > 6$, $3/((R_0-3)L_\delta) < 1$, and $g(R_0) < 1$, then

$$\sum_{s=1}^{\lfloor \eta K \rfloor} \mathbb{E} Z_s(\lfloor \delta N \rfloor) \longrightarrow 0.$$

Proof. First fix the order of t nonzero increments entering one accumulator. Their positions under a uniform permutation form a uniform t -subset of $[N]$. Let a be

the number of nonzero partial sums after these increments. The first partial sum is nonzero, and two consecutive partial sums cannot both vanish. Thus $a \geq \lceil t/2 \rceil$.

The gap before the first increment is nonnegative. Each of the t gaps following an increment is positive, including the final gap through N . These gaps sum to N and are uniform over the $\binom{N}{t}$ possibilities. If Y is the output weight, counting the gaps at nonzero and zero states gives

$$\Pr[Y = w \mid \text{increment order}] = \frac{\binom{w-1}{a-1} \binom{N-w}{t-a}}{\binom{N}{t}}.$$

Summing to d is the probability that a hypergeometric variable with population N , d marked elements, and t draws is at least a . For $\theta > 0$, its moment generating function is at most $(1 - d/N + (d/N)e^\theta)^t$. This follows by applying the elementary symmetric mean inequality to the N weights, each equal to 1 or e^θ . The Chernoff bound at threshold $t/2$ therefore gives

$$\Pr[Y \leq \lfloor \delta N \rfloor] \leq e^{-c_\delta t}. \quad (17)$$

For any $c > 0$, the same gap count and the negative-binomial series give

$$\mathbb{E}e^{-cY} \leq \frac{(t)_a}{(N)_a} (e^c - 1)^{-a}, \quad (v)_a := v(v-1) \cdots (v-a+1). \quad (18)$$

Indeed, replace $\binom{N-w}{t-a}$ by $\binom{N-a}{t-a}$ and use $\sum_{w \geq a} \binom{w-1}{a-1} e^{-cw} = (e^c - 1)^{-a}$. Both bounds hold uniformly over the nonzero increment values and their order.

A message of support s has $t = 3s$ nonzero coordinates after repetition. Condition on its first accumulator output, of weight Y . Independence of P_2 and (17) bound its final low-weight probability by $e^{-c_\delta Y}$. For $s \leq \eta K$, applying (18) with $c = c_\delta$ then bounds the probability by

$$\left(\frac{t}{(N-t)L_\delta} \right)^{t/2}.$$

Here $(t)_a/(N)_a \leq (t/(N-t))^a$, and the assumed base is less than one. There are $\binom{K}{s} (p-1)^s \leq ((p-1)eK/s)^s$ such messages, so $\mathbb{E}Z_s(\lfloor \delta N \rfloor) \leq g(N/s)^s$. The function $g(R)$ decreases for $R > 3$ and tends to zero as $R \rightarrow \infty$. Each fixed- s term tends to zero, while all terms are dominated by the summable sequence $g(R_0)^s$. This proves the limit.

For the remaining messages, let $\mathcal{T}_p(\eta, \delta)$ be the stationary types from Appendix A with $1 - \mu(0) \geq \eta$ and $1 - \zeta(0) \leq \delta$. Set

$$\Lambda_p(\eta, \delta) := \max_{\tau \in \mathcal{T}_p(\eta, \delta)} \Phi_p(\tau).$$

The maximum of an empty set is $-\infty$.

Lemma D.3 (Bulk messages). *If $\Lambda_p(\eta, \delta) < 0$, then $\sum_{s=\lceil \eta K \rceil}^K \mathbb{E}Z_s(\lfloor \delta N \rfloor) \rightarrow 0$.*

Proof. Write $\text{Mult}_\ell(\alpha)$ for the number of length- ℓ words of type α , and $T_N(\alpha, \beta)$ for the number of increment words of type α whose accumulator output has type β . A permutation is uniform on the words of a fixed type. Thus the expected number for message, middle, and output types (μ, β, ζ) is

$$\text{Mult}_K(\mu) \frac{T_N(\mu, \beta) T_N(\beta, \zeta)}{\text{Mult}_N(\mu) \text{Mult}_N(\beta)}.$$

The first increment type is μ because repetition preserves its normalized composition. The zero message is excluded by $\eta > 0$.

For each accumulator, record the p^2 transition counts $(S, S + D)$. Given the initial state zero and the ordered successor lists at each state, the path is determined. The number of paths with a transition type Q is therefore at most $\exp(N(H(Q) - H(Q_S)))$, where Q_S is the law of the state S under Q . Use

$$(\ell + 1)^{-p} e^{\ell H(\alpha)} \leq \text{Mult}_\ell(\alpha) \leq e^{\ell H(\alpha)}.$$

There are at most $(N + 1)^{2p^2 + 3p}$ choices of the three compositions and two transition types. Together with the two denominators, the polynomial prefactor is at most $(N + 1)^{2p^2 + 5p}$.

The source and target state marginals of a path differ only at its endpoints, by at most $1/N$ in each coordinate. Entropy continuity on the fixed-dimensional probability simplex and compactness imply

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \sum_{s=\lceil \eta K \rceil}^K \mathbb{E} Z_s(\lfloor \delta N \rfloor) \leq A_p(\eta, \delta).$$

To see uniformity, any sequence of maximizing finite path types has a convergent subsequence; its limit satisfies the stationary constraints, and its objective converges to Φ_p at that limit. A strictly negative upper bound absorbs the polynomial prefactor.

D.2 The certified bulk objectives

For $p = 5, 7$, a sign-orbit relaxation bounds the full stationary objective. Let $[a] := \{a, -a\}$, let $\mathcal{O}_p$ be the set of these orbits, and put

$$\mathcal{C}_p := \{([d], [s], [t]) : d, s, t \in \mathbb{F}_p, t = s + d\}.$$

Optimize over laws γ_1, γ_2 supported on $\mathcal{C}_p$. Their increment, source, and target marginals must respectively be $(\bar{\mu}, \bar{\beta}, \bar{\beta})$ and $(\bar{\beta}, \bar{\zeta}, \bar{\zeta})$. Let π_i be the increment-source marginal of γ_i . The objective is

$$\hat{\Phi}_p := H(\pi_1) + H(\pi_2) - \frac{2}{3}H(\bar{\mu}) - 2H(\bar{\beta}) - H(\bar{\zeta}) + \frac{1}{3}(1 - \bar{\mu}([0])) \log 2. \quad (19)$$

Every full type projects to a feasible orbit type with the same support weights. Data processing gives $I(D; S) \geq I([D]; [S])$, and $H(\mu) \leq H(\bar{\mu}) + (1 - \bar{\mu}([0])) \log 2$. Since $\Phi_p = H(\mu)/3 - I_{Q_1}(D; S) - I_{Q_2}(D; S)$, these inequalities prove $\Phi_p \leq \hat{\Phi}_p$.

Over $\mathbb{F}_3$, a sharper symmetry reduction is exact. For a distribution ρ , let σ negate every symbol and define $\text{Asym}(\rho) := H((\rho + \sigma\rho)/2) - H(\rho)$.

Lemma D.4 (Ternary sign reduction). *A stationary transition law T on $\mathbb{F}_3^2$, with state marginal β and increment marginal α , satisfies*

$$\text{Asym}(T) \geq \text{Asym}(\alpha) + \text{Asym}(\beta).$$

Consequently, the maximum defining $\Lambda_3(\eta, \delta)$ is attained by sign-symmetric compositions and transition laws whenever its domain is nonempty.

Proof. For nonnegative u, v , put $j(u, v) := u \log(2u/(u+v)) + v \log(2v/(u+v))$, with continuous boundary values. This function is convex and positively homogeneous. The off-diagonal entries of a stationary three-state law form a circulation. After exchanging symbols 1 and 2 if necessary, write

$$T_{01} = c + q, \quad T_{10} = c, \quad T_{12} = h + q, \quad T_{21} = h, \quad T_{20} = b + q, \quad T_{02} = b,$$

with $b, c, h, q \geq 0$, and put $T_{11} = y, T_{22} = z$. The three asymmetries are

$$\begin{aligned} \text{Asym}(T) &= j(c+q, b) + j(c, b+q) + j(y, z) + j(h+q, h), \\ \text{Asym}(\alpha) &= j(b+c+h+3q, b+c+h), \\ \text{Asym}(\beta) &= j(c+y+h+q, b+z+h+q). \end{aligned}$$

Subadditivity of j bounds the last expression by $j(y, z) + j(c+h+q, b+h+q)$. Set $m = b+c$, $x = c-b$, and $J(s, d) := j((s+d)/2, (s-d)/2)$. It suffices to prove that

$$\begin{aligned} G(x) &:= J(m+q, x+q) + J(m+q, x-q) + J(2h+q, q) \\ &\quad - J(2m+2h+3q, 3q) - J(m+2h+2q, x) \geq 0. \end{aligned}$$

This is an even convex function of x . Indeed, $\partial_d^2 J(s, d) = s/(s^2 - d^2)$ is convex in d and decreases in s ; hence the first two terms' second derivatives dominate the last one's. The feasible interval $[-m, m]$ is symmetric, so the minimum is at zero. There, subadditivity of J gives

$$G(0) = 2J(m+q, q) + J(2h+q, q) - J(2m+2h+3q, 3q) \geq 0.$$

Continuity covers boundary points. Finally average both stationary laws and their compositions with their sign images. The change in Φ_3 is

$$\text{Asym}(T_1) + \text{Asym}(T_2) - \frac{2}{3} \text{Asym}(\mu) - 2 \text{Asym}(\beta) - \text{Asym}(\zeta) \geq \frac{1}{3} \text{Asym}(\mu) \geq 0.$$

The support constraints are preserved.

A symmetric transition with nonzero increment fraction ν and nonzero state fraction ν' has matrix

$$T(\nu, \nu', b) = \begin{pmatrix} 1 - \nu' - 2b & b & b \\ b & (\nu' - \nu + 2b)/2 & \nu/2 - 2b \\ b & \nu/2 - 2b & (\nu' - \nu + 2b)/2 \end{pmatrix},$$

where

$$\max(0, (\nu - \nu')/2) \leq b \leq \min((1 - \nu')/2, \nu/4).$$

Let ν_1, ν_2, ν_3 be the nonzero fractions of μ, β, ζ . Thus the exact ternary bulk objective maximizes Φ_3 using $T(\nu_1, \nu_2, b_1)$ and $T(\nu_2, \nu_3, b_2)$ over $\eta \leq \nu_1 \leq 1$, $\nu_1 \leq 2\nu_2$, $\nu_2 \leq 2\nu_3$, and $0 \leq \nu_3 \leq \delta$. For each (ν, ν') the entropy is concave in b ; an interior maximizer solves

$$(1 - \nu' - 2b)(\nu/2 - 2b)^2 = b^2(\nu' - \nu + 2b)/2.$$

This specifies the objective checked by the ternary verifier, including its feasible boundary.

The directed-arithmetic verifiers give the following global upper bounds. The displayed values are rounded upward.

p	η	δ	upper bound and objective
3	1/10000	49/250	$-3.6177 \cdot 10^{-7}$, exact symmetric
5	1/60	1/20	$-1.1386 \cdot 10^{-6}$, sign orbit
7	1/16	9/500	$-1.1596 \cdot 10^{-4}$, sign orbit

The orbit verifier covers the feasible composition domain by rational boxes. It upper-bounds positive joint entropies by cross-entropy at positive trial laws, and negative marginal entropies by secants of $x \log x$. The resulting linear program includes the orbit incidence and marginal constraints. A rational dual solution bounds its objective on the whole box; directed logarithm intervals make each bound outward. Boxes discarded as infeasible have exact dual witnesses. Multiplicative automorphisms supply additional valid cuts for $p = 7$. The ternary verifier instead bounds the concave transition entropies by supporting planes and the negative marginal entropies by their endpoint secants. Rational boxes and interval arithmetic enclose the maximizing values of b_1, b_2 and the resulting objective.

The certificates and their verifiers are listed in Appendix N.

Proof (Proof of Theorem D.1). Apply Lemma D.2 with the three (p, η, δ) rows above. For $p = 3$, $L_\delta \geq 1/4$ and $R_0 = 30000$, giving $g(R_0) < 1/2$ from $16 \cdot 30000^2 \cdot 12^3 < 29997^3$ and $e < 3$. For $p = 5$, $L_\delta > 9/7$ and $R_0 = 180$ give $g(R_0) < 660(7/531)^{3/2} < 1$, using $e < 11/4$ and $660^2 \cdot 7^3 < 531^3$. For $p = 7$, $L_\delta > 367/133$ and $R_0 = 48$ give $g(R_0) < 264(133/5505)^{3/2} < 1$, using $e < 11/4$ and $264^2 \cdot 133^3 < 5505^3$. The other sparse hypotheses hold in all three rows. The certified bulk bounds and Lemma D.3 handle every larger support. Markov's inequality on the total nonzero-message count proves the three distance claims.

Proof (Proof of Proposition A.1). Take $p = 5$, $\delta = 131/500 = 0.262$, $\eta = 10^{-5}$, and $R_0 = 300000$ in Lemma D.2. The comparison displayed after Proposition A.1 gives $L_\delta > 3/22$. Then $g(R_0) < 4R_0(22/(R_0 - 3))^{3/2} < 4/5$, using $e < 3$ and $25 \cdot 300000^2 \cdot 22^3 < 299997^3$. The other hypotheses of the lemma hold, and Lemma D.2 proves the stated vanishing expectation.

D.3 A bound for every fixed prime

The spacing argument also gives a bound for every fixed prime.

Theorem D.5 (Prime ladder). *Fix a prime p and an integer $m \geq 4$. If*

$$288(p-1)^2 m^4 < (m-2)^6 (m+2)^3,$$

then

$$\Pr[d_{\mathbb{F}_p}(E_{P_1, P_2}) \leq \lfloor N/m^2 \rfloor] \longrightarrow 0.$$

Proof. The inverse accumulator is a difference map, so $\text{wt}(u) \leq 2 \text{wt}(Au)$. Repetition and two accumulators therefore give $\text{wt}(Ex) \geq 3 \text{wt}(x)/4$. A word of weight at most N/m^2 can only come from a message of support at most $4K/m^2$. Apply Lemma D.2 with $\delta = 1/m^2$ and $\eta = 4/m^2$. Then $R_0 = 3m^2/4 > 6$ and

$$L_\delta = \frac{m}{2\sqrt{1-1/m^2}} - 1 > \frac{m-2}{2}, \quad \frac{3}{(R_0-3)L_\delta} < \frac{8}{(m-2)^2(m+2)} < 1.$$

Using $e < 3$ gives

$$g(R_0) < \frac{3(p-1)m^2}{4} \left(\frac{8}{(m-2)^2(m+2)} \right)^{3/2} < 1;$$

the last inequality is exactly the theorem's hypothesis after squaring. The sparse lemma handles all possible low-weight messages. Markov's inequality completes the proof.

For every fixed p , a sufficiently large m satisfies the inequality. Hence the addition-only ensemble is asymptotically good (positive rate and relative distance) over every fixed prime field. The resulting lower bound is of order $p^{-4/5}$ and does not give a useful relative distance when p grows exponentially in the security parameter.

E Finite certificate for the four-input precode

We evaluate (16) using rational inequalities. All notation refers to Appendix C, and binomial coefficients outside their usual range are zero. Throughout, $\rho = e/(n-e)$.

Last-round tail. Decompose an output support into runs of nonzero outputs. A run is *closed* if a return ends it and *open* if it reaches the last position. For $h, j > 0$, the number of support pairs with input weight h , output weight j , and exactly c returns is

$$\binom{n-j}{c} \left[\binom{j-1}{c-1} \binom{j-c}{h-2c} + \binom{j-1}{c} \binom{j-c-1}{h-2c-1} \right],$$

and $A_g(h, j)$ weights this number by g^c . The two terms correspond to a closed and an open final run; in the second term $c = 0$ is allowed. Summing over $j \leq e$ and extending the number of available zero positions gives

$$\frac{H_g(h, e)}{\binom{n}{h}} \leq 2 \sum_{c=0}^{\lfloor h/2 \rfloor} \binom{h}{c} \binom{h-c}{c} g^c \rho^{h-c} \leq 2(\rho + 2\sqrt{g\rho})^h, \quad h \leq 2e. \quad (20)$$

For the first inequality, the closed-run binomial comparison is

$$\frac{\binom{n}{c} \binom{e}{h-c}}{\binom{n}{h}} = \binom{h}{c} \frac{\binom{e}{h-c}}{\binom{n-c}{h-c}} \leq \binom{h}{c} \rho^{h-c},$$

and the open case has the same bound with c closed runs. The remaining coefficient in each case is at most $\binom{h-c}{c}$. For the second inequality, $h!/(c!(h-2c)!) = \binom{h}{c} \binom{h-c}{c}$, so the sum is the constant coefficient in u of the Laurent polynomial $(\rho + \sqrt{g\rho}(u + u^{-1}))^h$. All its coefficients are nonnegative, so this constant coefficient is at most its value at $u = 1$, namely $(\rho + 2\sqrt{g\rho})^h$.

Fix a rational $t < 1$ with $t > \rho$ and $(t - \rho)^2 \geq 4g\rho$, that is, $\rho + 2\sqrt{g\rho} \leq t$. Then (20) gives $H_g(h, e)/\binom{n}{h} \leq 2t^h$ for $h \leq 2e$. Since g depends on τ , so does the admissible range of t .

Bulk source weights. Choose positive rational x, v, z , rational $\tau \geq 1$, t as above, and $0 < y \leq t$. Set

$$\lambda = \max\{1 + xyv, gx/v + (1 + x)y\}, \quad C = \max\{1, 1/v\}.$$

The transfer matrix in (15) maps $[1, v]^T$ to a vector bounded coordinatewise by $\lambda[1, v]^T$. Its generating function is consequently at most $C\lambda^n$. Combining coefficient bounds with (20), the contribution of a source weight s to (16) is at most

$$\frac{2(n+1)C}{\tau} Q(z, \tau)^L \lambda^n (t/y)^{2e} (xz)^{-s} (s/n)^s (1 - s/n)^{n-s}. \quad (21)$$

The factor $(n+1)(s/n)^s(1 - s/n)^{n-s}$ bounds $1/\binom{n}{s}$, by the estimate $\binom{n}{s} \geq 2^{nH_2(s/n)}/(n+1)$. With witnesses fixed, the logarithm of (21) is convex in s . Its values at an interval's two endpoints therefore bound its value at every interior integer. Multiplying by the number of integers in the interval bounds the interval's total.

Sparse source weights. For $1 \leq s \leq 128$, retain the return weight in the sparse calculation. With $\tau \geq 1$, $g = 1/q + (1 - 1/q)/\tau$, and t as above, put $b_g = t/(1 - t)$ and define

$$\begin{aligned} M_s(t; g) &= \sum_{j=1}^{\lfloor s/2 \rfloor} \binom{n}{j} \binom{s-j-1}{j-1} g^j b_g^{s-j} \\ &\quad + \sum_{j=1}^{\lceil s/2 \rceil} \binom{n}{j-1} \binom{s-j}{j-1} g^{j-1} b_g^{s-j+1}. \end{aligned} \quad (22)$$

An open final run ends with no return, which explains the factor g^{j-1} in the second sum. To obtain the bound, extend each nonzero run to any positive length. For j closed runs the generating function is $g^j x^{2j} [t/(1-t(1+x))]^j$. An open final run has one fewer input boundary and one fewer return. Thus $M_s(t; g)$ bounds $\sum_h A_g(s, h)t^h$. The entire sparse contribution is at most

$$P_{\text{sparse}} := \frac{2}{\tau} \sum_{s=1}^{128} \frac{[z^s]Q(z, \tau)^L}{\binom{n}{s}} M_s(t; g). \quad (23)$$

At $\tau = 1$, $g = 1$ and this is the unweighted sparse bound, which ignores coefficient randomness. The value of τ is a proof parameter and does not change the encoder coefficients.

Certified values. For the row $(q, e) = (32, 209715)$, the unweighted sparse bound and 114 bulk intervals give failure below $2^{-24.85}$. For the implemented alphabets, the weighted sparse bound with $\tau = \tau_{\text{sparse}}$ gives the following values. Bulk intervals use their own witnesses in (21).

q	e	τ_{sparse}	intervals	$\log_2 \Pr[d \leq e]$	bound
40	230896	160	115		-44.11
64	251658	256	116		-45.65
128	277872	512	119		-47.87
256	288358	1024	118		-50.24
512	298844	2048	126		-52.46
1024	301989	4096	119		-54.72
2048	304087	8192	119		-56.90

Each bulk interval contributes less than 2^{-80} , and the intervals together cover every integer from 129 through $4e$. These certificates prove Theorem C.1. The configuration with the largest certified distance in this four-input family uses $e = 304087$, so its bulk cover ends at 1216348. Exact sparse evaluation and outward logarithm bounds give

$$\log_2(P_{\text{sparse}} + 119 \cdot 2^{-80}) \leq -56.907879669916.$$

At $q = 2048$ and $e = 230896$, a separate cover of 115 intervals gives the upper logarithm -58.451178036656 . Rational entropy bounds also verify

$$H_2(230897/2097152) > 1/2,$$

so the $q = 40$ distance bound exceeds the binary rate-1/2 Gilbert–Varshamov benchmark, as stated after Theorem C.1. The certificates and their verifiers are listed in Appendix N.

F Finite evaluation for two plain accumulators

This appendix evaluates the basic two-state bound of Section 5 for $q = 1$ and proves relative distance above 6.6%. Appendices G, H, and K refine it to 11.7%;

the last supplies the $q = 1$ row of Table 1. Fix $p = 2^{128} - 159$, $n = 2^{21}$, the Newton precode B_8 from (1), and forbidden weight $e = 138412$. Only the two independent uniform permutations are random. Every scan coefficient equals one.

Small source supports. Put $\rho = e/(n - e)$ and choose a rational t satisfying $\rho < t < 1$ and $(t - \rho)^2 \geq 4\rho$. Let $A(s, h)$ denote the plain accumulator's admissible support count, as in (15) with $g = 1$. Its normalized tail obeys

$$\frac{\sum_{j=1}^e A(h, j)}{\binom{n}{h}} \leq 2(\rho + 2\sqrt{\rho})^h \leq 2t^h.$$

The run decomposition used in Appendix E proves this inequality without coefficient randomness. Extending run lengths to geometric series also bounds $\sum_h A(s, h)t^h$ by

$$\begin{aligned} M_s(t) &= \sum_{j=1}^{\lfloor s/2 \rfloor} \binom{n}{j} \binom{s-j-1}{j-1} \left(\frac{t}{1-t} \right)^{s-j} \\ &\quad + \sum_{j=1}^{\lceil s/2 \rceil} \binom{n}{j-1} \binom{s-j}{j-1} \left(\frac{t}{1-t} \right)^{s-j+1}, \end{aligned}$$

which is $M_s(t; 1)$ in (22). The two sums allow a closed or open final run. Binomial coefficients outside their usual range are zero.

Let $U(z) = Q_8(z, 1)$ and $u_s = [z^s]U(z)^{n/16}$. These coefficients upper-bound the number of possible source supports. Permutation averaging therefore bounds the whole range $1 \leq s \leq 31$ by the *sparse bound*

$$\epsilon_{\text{small}} := 2 \sum_{s=1}^{31} \frac{u_s}{\binom{n}{s}} M_s(t). \quad (24)$$

This covers every intermediate weight h . The coefficients u_s are computed exactly by differentiating $U(z)^{n/16}$; absent source weights contribute zero. Rational evaluation of (24) gives $\log_2 \epsilon_{\text{small}} < -53.961375350693$.

The remaining supports. The inverse scans imply $s \leq 2h \leq 4e$. Thus it remains to cover

$$32 \leq s \leq 553648, \quad 16 \leq h \leq 276824, \quad s \leq 2h.$$

Apply Lemma 5.1 with $K_j = C_j \lambda_j^n$ from the operator (6) at $q = 1$, as in (10), and replace $F_s(\tau)$ by $Q_8(z, \tau)^{n/16}/z^s$.

For an integer weight h , the binomial distribution with parameters $(n, h/n)$ has mode h and variance $h(n-h)/n$. If an integer $\sigma \geq 1$ upper-bounds its standard deviation, Chebyshev's inequality places at least $3/4$ of its mass in at most $4\sigma + 1$ positions. Hence

$$\binom{n}{h} \geq \frac{3}{16\sigma + 4} 2^{nH_2(h/n)}.$$

On an interval $a \leq h \leq b$, the choice $\sigma = \lceil \sqrt{\min(b, n-a)} \rceil$ suffices. Apply this inequality to both support-binomial denominators.

For fixed rational witnesses, the logarithm of the resulting upper bound is a convex function of (s, h) . Over a rectangle, its maximum on the polygon cut out by the half-plane $s \leq 2h$ is attained at a vertex of that polygon. Multiplying by the rectangle's number of integer profiles upper-bounds its total contribution, including any excess count outside the polygon. The certificate partitions the domain recursively into rectangles, each with its own rational witnesses; the cells of the final partition are called *leaves*. Here there are 137 nonempty leaves. The verifier checks the root domain, each split, each rational witness, and every vertex inequality using directed logarithm bounds.

If ℓ_R is the verified upper logarithm for a leaf R , its contribution is rounded upward to $2^{\max(-160, \lceil \ell_R \rceil)}$. Adding these rationals to (24) gives

$$\log_2 \left(\epsilon_{\text{small}} + \sum_R 2^{\max(-160, \lceil \ell_R \rceil)} \right) \leq -\frac{13490343832277}{250000000000} < -53.96.$$

Consequently the complete code has distance at least 138413, that is, relative distance above 6.6%, except with probability less than $2^{-53.96}$. The certificate, with its witnesses and exact rational sums, and its verifier are listed in Appendix N.

G The projective-pair bound for plain accumulators

Plain accumulators have no random coefficients, but through the random permutations the bound below still charges for some dependent cancellations. We use the Newton precode B_8 and the prime $p = 2^{128} - 159$ of Appendix F. For the same encoder, the following refinement proves minimum distance at least 209716 ($> 10\%$), except with probability less than $2^{-47.91}$ over the two independent uniform permutations. The bound holds simultaneously for all nonzero messages, including a message chosen after the permutations are revealed. Appendix K later refines it to 11.7%.

As in Section 5, fix a source support S with shortened source space V_S , and let $W \subseteq V_S^*$ be the span of zero equations imposed so far. The *active labels* of a scan are the coordinates of its incoming vector that are required to be nonzero; the permutation places them in uniformly random order on the nonzero input positions. Each active label carries a linear form on V_S , its *required active form*, which must remain nonzero in V_S^*/W . At the start of a run, expose its first two active input labels jointly. Before either label is revealed, they form a uniform ordered pair from the $m \geq 2$ remaining active labels. If any required active form vanishes in V_S^*/W , the branch has no surviving message and can be discarded. Otherwise every remaining active form is nonzero. For the moment, draw two labels independently with replacement. Let $P_{\parallel}$ be the probability that their forms are proportional, and let $P_{\perp}$ be the probability that they are exact opposites. For each one-dimensional subspace ℓ , let μ_{ℓ} be its total probability mass, and

write μ_v for the mass of each form v . Partition the nonzero forms of ℓ into pairs $\{v, -v\}$. Odd characteristic makes these pairs disjoint, and their masses satisfy

$$2 \sum_{\{v, -v\} \subset \ell} \mu_v \mu_{-v} \leq \mu_\ell^2/2.$$

Indeed, orient each opposite pair and bound the sum of products by the product of the two total oriented masses, whose sum is μ_ℓ . Summing over lines gives $P_- \leq P_\parallel/2$ and $P_- \leq 1/2$.

Let $Y, Z > 0$ bound the continuations with nonzero and zero sum, respectively, uniformly over forms, previous zero equations, and nonzero requirements. For $\tau \geq 1$, an independent pair contributes at most $Y + Z/\tau$ to the rank potential $\tau - \text{rank } W$: its zero branch gains one rank. A dependent opposite pair contributes only Z ; a dependent nonopposite pair contributes only Y . In the latter case, imposing zero sum would force both required nonzero inputs to vanish.

The with-replacement average is bounded by

$$\begin{aligned} & (1 - P_\parallel)(Y + Z/\tau) + P_-Z + (P_\parallel - P_-)Y \\ &= Y + Z/\tau + P_-(Z - Y) - P_\parallel Z/\tau \\ &\leq Y + Z/\tau + P_-((1 - 2/\tau)Z - Y) \\ &\leq \max\{Y + Z/\tau, (Y + Z)/2\}. \end{aligned}$$

The last step maximizes an affine function on $0 \leq P_- \leq 1/2$. The labels are distinct with probability $(m - 1)/m$. Since all branch contributions are nonnegative, conditioning on this event gives

$$\begin{aligned} \mathbb{E}_{\text{distinct}}[\text{branch contribution}] &\leq \frac{m}{m-1} \Psi(Y, Z), \\ \Psi(Y, Z) &= \max \left\{ Y + \frac{Z}{\tau}, \frac{Y + Z}{2} \right\}. \end{aligned} \tag{25}$$

The inequality also holds for $Y = 0$ or $Z = 0$ by continuity. In particular, a forced zero sum costs at most $\frac{m}{m-1} \max\{1/\tau, 1/2\}$. The improvement comes from counting proportional nonopposite pairs as dead zero branches, rather than allowing them a rank increment. Further nonzero requirements can only remove branches.

For positive input and output markers x, y , define

$$\begin{aligned} T_0(v) &= v_0 + xyv_2, \\ T_1(v) &= yv_1 + x \max\{yv_1 + v_0/\tau, v_0\}, \\ T_2(v) &= yv_2 + x\Psi(yv_1, v_0). \end{aligned}$$

State 0 has zero feedback. State 1 has nonzero feedback after at least two active inputs in the current run; T_1 is the nonzero-feedback operator (6) at $q = 1$. State 2 is a *macro state*: it stands for nonzero feedback after exactly one active input in the current run, but T_2 bounds that input together with the next active input

through (25). It does not describe a bound conditional on a revealed first label. This distinction is essential: after revealing one label, every remaining form might be its opposite.

To justify the macro state, write $A_t = (T^t \mathbf{1})_0$ and $B_t = (T^t \mathbf{1})_1$. Expanding the state-2 recurrence gives

$$A_t = A_{t-1} + xy^t + \sum_{\gamma=0}^{t-2} x^2 y^{\gamma+1} \Psi(y B_{t-\gamma-2}, A_{t-\gamma-2}).$$

The first term has zero input. The second is a final run with one active input. In each summand, choose the gap γ of zero inputs before exposing either of the first two active labels. Apply (25), then the uniform continuation bound. State 1 uses the two-state bound of Section 5 until the feedback returns to zero. This proves the macro interpretation without conditioning on the first label alone.

Each completed pair consumes two labels. The values m at pair starts are distinct and decrease along a trajectory, even if single exposures intervene. Their sampling corrections therefore multiply to at most $\prod_{m=2}^s m/(m-1) = s$ for a scan with input weight s . More formally, the induction budget $f(m) = \max(1, m)$ satisfies both $f(m-1) \leq f(m)$ and $mf(m-2)/(m-1) \leq f(m)$. Final single-input runs require no pair correction.

Normalize $v_0 = 1$, choose $\lambda \geq \max_i T_i(v)/v_i$, and put $C = \max_i 1/v_i$. Monotonicity and homogeneity imply $(T^n \mathbf{1})_0 \leq C\lambda^n$. With the pair corrections, a scan with m active inputs therefore has the uniform bound $K(m) = mC\lambda^n$ required by Lemma 5.1. For source weight s and intermediate weight h , the lemma gives

$$\frac{F_s(\tau)}{\tau} \frac{sh C_1 \lambda_1^n C_2 \lambda_2^n}{\binom{n}{s} \binom{n}{h} x_1^s (y_1 x_2)^h y_2^e}, \quad 0 < y_2 \leq 1.$$

The lemma keeps the same zero-equation span and rank potential across both scans. Conditioning on the first trajectory leaves the second permutation independent and uniform, and the uniformity of the continuation bounds justifies applying the macro state to its intermediate forms.

For $e = 209715$, the sparse bound (24) of Appendix F bounds source weights 1 through 31 by a probability whose base-two logarithm is at most -47.919090471278 . The dense range $32 \leq s \leq 4e$, $\lceil s/2 \rceil \leq h \leq 2e$ is partitioned into 121 leaves, as in Appendix F. Each leaf bounds sh by the product of its rectangle's upper endpoints and uses the rational binomial and convex vertex bounds of that appendix. Exact rational evaluation, including all finite prefactors and upward-rounded leaf sums, gives total logarithmic upper bound -47.919090461115 . The certificate and its verifier are listed in Appendix N. This strengthens the proof without changing the encoder's $6.8125n - 2$ field additions.

H A joint-renewal bound across run boundaries

The projective-pair bound of Appendix G can be strengthened by retaining the pool of remaining active input labels across a return and the next run start. We

call the result the joint-renewal bound. For the same B_8 encoder at $n = 2^{21}$ and $p = 2^{128} - 159$, it proves minimum distance at least 224396 ($> 10.7\%$), except with probability less than $2^{-46.39}$ over the two independent uniform permutations. The arithmetic remains $6.8125n - 2$ field additions. This is an intermediate result: the complete-run bound of Appendix K refines it to distance at least 245367 ($> 11.7\%$).

A common pool for neighboring returns. Fix the quotient by the zero equations imposed so far and a required nonzero feedback form f . Let μ be the empirical distribution of the remaining active forms. Every such form is nonzero on a surviving branch. Write

$$u = \mu(-f), \quad v = \mu(\text{span}(f) \setminus \{-f\}), \quad w = 1 - u - v.$$

Let $t = 1/\tau$, where $\tau \geq 1$. For three nonnegative continuation bounds Y, U, Z , consider the following exposure. Reveal the next active input. If the feedback stays nonzero, use Y and stop this exposure. If it returns to zero, reveal the next two active inputs jointly and use U or Z according as their sum is nonzero or zero. The bounds are uniform over the quotient and previous nonzero requirements.

For iid labels, the rank-weighted contribution is at most

$$\begin{aligned} \mathcal{H}(Y, U, Z; u, v, w) &= (1 - u)Y + wt\Psi(U, Z) + u\mathcal{P}(U, Z; u, v, w), \\ \mathcal{P}(U, Z; u, v, w) &= U + t(1 - (u + v)^2)Z \\ &\quad + (Z - U)_+ \max\{2uv, v^2/2\} \\ &\quad + ((Z - U)/2 - tZ)_+ w^2, \end{aligned}$$

where $a_+ = \max(a, 0)$ and Ψ is from (25). To prove this, split the first input into its mass u at $-f$, its other mass v on the same line, and its independent mass w . A return is free in the first case, impossible on a surviving branch in the second, and gains one rank in the third. The independent case admits the uniform bound $t\Psi(U, Z)$ after the quotient update. The nonzero continuation has contribution at most $(1 - u)Y$.

After a free return the quotient is unchanged. If $P_{\parallel}, P_{-}$ are the following pair's proportional and opposite probabilities, its contribution is at most $U + tZ + P_{-}(Z - U) - P_{\parallel}tZ$. On $\text{span}(f)$, the proportional mass is $(u + v)^2$ and the opposite mass is at most $\max(2uv, v^2/2)$. Indeed, if mass $z \leq v$ is at f , the opposite mass is at most $2uz + (v - z)^2/2$, a convex quadratic maximized at an endpoint. Off the line, $P_{-, \text{out}} \leq P_{\parallel, \text{out}}/2$ and $P_{\parallel, \text{out}} \leq w^2$. Dropping negative terms yields $\mathcal{P}$.

Distinct labels. The nonzero continuation exposes only one label, with no correction. For a free return followed by the pair, condition the iid triple on distinct labels, costing $\kappa^2/((\kappa - 1)(\kappa - 2))$. For an independent return, expose its label first; the pair on the remaining $\kappa - 1$ labels costs only $(\kappa - 1)/(\kappa - 2)$. The pair exposure from zero feedback costs $\kappa/(\kappa - 1)$. The budget $\beta(\kappa) = \max(1, \kappa^2)$

absorbs all three exposure lengths:

$$\begin{aligned}\beta(\kappa - 1) &\leq \beta(\kappa), & \frac{\kappa}{\kappa - 1}\beta(\kappa - 2) &\leq \beta(\kappa), \\ \frac{\kappa^2}{(\kappa - 1)(\kappa - 2)}\beta(\kappa - 3) &\leq \beta(\kappa).\end{aligned}$$

Use each inequality only when enough labels remain. Thus a scan with s inputs has total correction at most s^2 , even with adaptive quotient updates. The one-label branch does not consume the two labels used by its alternative. Prefixes that terminate earlier are covered below.

Gap sums and terminal terms. Let $x, y > 0$ mark active inputs and outputs. Choose $\lambda > \max(1, y)$ and give each position a factor $1/\lambda$. Summing zero-input gaps gives

$$A = \frac{xy}{\lambda - 1}, \quad B = \frac{x}{\lambda - y}.$$

The factor A includes the nonzero output on leaving zero feedback; B reaches an active input from nonzero feedback before marking its output. As in Section 5, state 0 is zero feedback and state 1 is nonzero feedback. Normalize the zero-feedback potential to 1, and use $V > 0$ for nonzero feedback. The homogeneous bounds are

$$T_0 = AB\Psi(yV, 1), \quad T_1(u, v, w) = \mathcal{H}(ByV, AB^2yV, AB^2; u, v, w).$$

For a rational $0 < \epsilon < 1$, require

$$T_0 \leq 1 - \epsilon, \quad T_1(u, v, w) \leq (1 - \epsilon)V \quad (u, v, w \geq 0, u + v + w = 1). \quad (26)$$

From zero feedback, a terminal prefix has zero or one active input before the pair exposure completes. Its generating function is

$$R_0 = \frac{\lambda}{\lambda - 1} \left(1 + \frac{xy}{\lambda - y} \right).$$

From nonzero feedback, termination precedes the next active input or follows a return before the next pair completes. Its contribution is at most $R_1 = \lambda/(\lambda - y) + BR_0$. Put

$$C = \frac{\max(R_0, R_1/V)}{\epsilon}.$$

Monotonicity, positive homogeneity, and (26) make $C(1, V)$ a super-solution for every finite truncation of the trajectory tree: one further exposure step, together with its terminal terms, does not exceed it. Monotone extension of the gap sums bounds all terminal lengths. Extracting length n contributes λ^n ; fixed input weight s adds the correction s^2 . Both possible final feedback states are included.

Keep the same equation span and rank potential across both scans. For source weight s , intermediate weight h , and forbidden final weight e (so the distance is at least $e + 1$), permutation averaging gives

$$\frac{F_s(\tau)}{\tau} \frac{(sh)^2 C_1 \lambda_1^n C_2 \lambda_2^n}{\binom{n}{s} \binom{n}{h} x_1^s (y_1 x_2)^h y_2^e}, \quad 0 < y_2 \leq 1.$$

The source moment F_s is unchanged. The uniform continuation bounds permit the second scan to start in the quotient left by the first.

Rational verification. Witnesses are found numerically on a finite grid of pools; the grid plays no role in the proof of (26). For rational witnesses, the verifier splits $\max(2uv, v^2/2)$ into two cubic polynomials and homogenizes using $u + v + w = 1$. On any triangle of the simplex, the largest degree-three Bernstein coefficient upper-bounds the polynomial. The verifier subdivides a triangle into its four midpoint children until both polynomials satisfy the required bound throughout every child. It checks only rational inequalities and complete subdivisions.

At forbidden weight $e = 224395$, the sparse contribution has base-two logarithm at most -46.396269051235 . The dense domain of weight profiles is partitioned recursively into rectangles; each cell of the partition is a leaf. There are 154 leaves: 47 use the projective-pair bound of Appendix G and 107 use the joint-renewal bound above. Each leaf covers all trajectories on its weight rectangle, so selecting either bound per leaf is valid. The rational binomial, convex vertex, and upward-rounded summation checks of Appendix F apply unchanged. The total failure probability has base-two logarithm at most -46.396269051207 . The certificate and its verifier are listed in Appendix N. This finite all-message guarantee does not exceed the binary rate-1/2 Gilbert–Varshamov benchmark $H_2^{-1}(1/2) \approx 0.110028$; the complete-run bound of Appendix K does.

I A hypergeometric support bound

We first sharpen the sparse-source bound of Appendix F without using coefficient randomness. Consider a width-one scan with nonzero coefficients over any field, zero initial state, and unrestricted final state. At every position of a realized pair of input and output supports, the triple of previous-output, input, and current-output indicators lies in $\{000, 011, 101, 110, 111\}$. Let $A_n(h, u)$ count pairs of input and output supports of weights h, u that satisfy this rule at every position. Some counted pairs may not be realizable; this only weakens the upper bound.

Lemma I.1. *Let $0 \leq e < n$, $\delta = e/n$, and $1 \leq h \leq n$. For every $L > 1$ such that $L(L - 1) \geq 1$, put $t = \delta L + (1 - \delta)/(L - 1)$. Then*

$$B_n(h, e) := \frac{1}{\binom{n}{h}} \sum_{u=0}^e A_n(h, u) \leq t^h.$$

In particular, for $0 < \delta < 1/5$ one may take $t = \delta + 2\sqrt{\delta(1 - \delta)} < 1$.

Proof. Decomposing the output support into r nonzero runs gives

$$A_n(h, u) = \sum_r \binom{u-1}{r-1} \left[\binom{n-u}{r} \binom{u-r}{h-2r} + \binom{n-u}{r-1} \binom{u-r}{h-2r+1} \right].$$

The terms distinguish final zero and nonzero states. Invalid binomial coefficients are zero. Let J count elements at most e in a uniform h -subset of $\{1, \dots, n\}$. Counting its $(a+1)$ st smallest element yields

$$\sum_{u=1}^e \binom{u-1}{a} \binom{n-u}{h-a-1} = \binom{n}{h} \Pr[J \geq a+1].$$

Using this identity in the run count gives the exact formula

$$\begin{aligned} B_n(h, e) &= \sum_{r=1}^{\lfloor h/2 \rfloor} \binom{h-r-1}{r-1} \Pr[J \geq h-r] \\ &\quad + \sum_{r=1}^{\lceil h/2 \rceil} \binom{h-r}{r-1} \Pr[J \geq h-r+1]. \end{aligned}$$

For $\zeta \geq 1$, the hypergeometric moment satisfies $\mathbb{E}[\zeta^J] \leq (1 - \delta + \delta\zeta)^h$. Indeed, it is the average product of h distinct entries of a list containing e copies of ζ and $n - e$ copies of one. Pairwise averaging entries with their sum fixed cannot decrease their elementary symmetric polynomial: its dependence on the pair's product has a nonnegative coefficient. Iterating this averaging bounds the moment by the product for the constant list. Markov's inequality now gives $\Pr[J \geq j] \leq (1 - \delta + \delta\zeta)^h \zeta^{-j}$.

Set $\zeta = L(L-1)$. The remaining polynomial after substituting these tails is

$$\kappa_h(\zeta) = \sum_{r=1}^{\lfloor h/2 \rfloor} \binom{h-r-1}{r-1} \zeta^r + \sum_{r=1}^{\lceil h/2 \rceil} \binom{h-r}{r-1} \zeta^{r-1}.$$

Pascal's identity gives $\kappa_h = \kappa_{h-1} + \zeta \kappa_{h-2}$, with $\kappa_0 = \kappa_1 = 1$ as recurrence initial values. Since $L^2 = L + \zeta$, induction gives $\kappa_h \leq L^h$. Consequently

$$B_n(h, e) \leq \left(\frac{(1 - \delta + \delta\zeta)L}{\zeta} \right)^h = t^h.$$

Minimizing over L gives $L = 1 + \sqrt{(1 - \delta)/\delta}$ and the stated special case.

For a fixed source support of size s , the first permutation gives each incoming support probability $1/\binom{n}{s}$. Union over its compatible intermediate supports, then apply the lemma to the independent second permutation. The result is

$$\Pr[\text{a compatible final support has weight at most } e] \leq \frac{\sum_h A_n(s, h) t^h}{\binom{n}{s}}.$$

This union counts support patterns and already covers every field-valued message realizing them. It assumes no independence of those messages. With $u_s = [z^s]Q_8(z, 1)^{n/16}$ and the moment $M_s(t)$ from Appendix F, the sparse-source contribution is at most

$$\sum_{s=1}^{31} \frac{u_s M_s(t)}{\binom{n}{s}}. \quad (27)$$

The sparse bound (24) carries a leading factor two; (27) does not. The verifier uses rational L , checks $L(L-1) \geq 1$, computes t exactly, and checks $t < 1$. The local support rule holds for arbitrary fixed nonzero scan coefficients, so this bound also applies to the sampled coefficients.

J Projective pairs with small coefficients

Fix the quotient by the zero equations already imposed. Required active forms must remain nonzero in this quotient. At zero feedback, expose the next two active input labels jointly. If g zero inputs separate them, their forms u, v cancel identically exactly when $v = -Cau$. Here C is the product of the intervening coefficients, and a is the fresh coefficient at the second active input. Both labels are exposed after fixing the gap and its coefficients.

Call a gap g short if $q^{3(g+1)} + 1 < p$. For a short gap, the integer representing Ca lies between one and $D = q^{g+1}$. In each one-dimensional subspace, form a graph on the nonzero forms, joining u and v , not necessarily distinct, when $v = -cu$ or $u = -cv$ for some integer $1 \leq c \leq D$. The graph has no loops: $u = -cu$ would require $c + 1 \equiv 0$ modulo p , but $2 \leq c + 1 \leq D + 1 < p$. It has no triangles: multiplying the oriented ratios around a triangle requires a sum of two positive integer products to vanish modulo p , while its value is at most $D^3 + 1 < p$.

Give the vertices nonnegative masses of total σ . The edge-product sum of a triangle-free graph without loops is at most $\sigma^2/4$. To prove this, maximize the sum and take a maximizer with minimal support. Nonadjacent positive-mass vertices can be merged without decreasing the objective, since it is linear in their mass split. Minimal support is therefore a clique of size at most two, where the claim follows from the arithmetic-geometric mean inequality.

Each ordered pair has at most one cancelling choice of a . If $P_{||}$ is the proportional-pair probability under independent draws from the remaining label pool, the identical-cancellation probability is therefore at most $P_{||}/(2q)$ for short gaps. For arbitrary gaps, each coefficient value has probability $1/q$, which instead gives $P_{||}/q$. For continuation bounds $Y, Z \geq 0$ and rank marker $\tau \geq 1$, put

$$\Psi_c(Y, Z) = \max\{Y + Z/\tau, (1-c)Y + cZ\}.$$

The pair contribution is bounded by $\Psi_{1/(2q)}$ for short gaps and $\Psi_{1/q}$ otherwise. Independent forms can add one rank on a zero branch; proportional forms contribute only one of the two continuations. This establishes the bound before either label is revealed.

For $q > 1$, let G be the largest short gap, that is, the largest integer $g \geq 0$ satisfying $q^{3(g+1)} + 1 < p$, or $G = -1$ if none does. Choose positive x, y, V , $\lambda > \max(1, y)$, and slack $0 < \epsilon < 1$. Define

$$A = \frac{xy}{\lambda - 1}, \quad B = \frac{x}{\lambda - y}, \quad T = (y/\lambda)^{G+1},$$

$$P_s = \Psi_{1/(2q)}(yV, 1), \quad P_l = \Psi_{1/q}(yV, 1).$$

Summing the geometric zero-input gaps yields sufficient inequalities

$$AB((1 - T)P_s + TP_l) \leq 1 - \epsilon, \quad BP_l \leq (1 - \epsilon)V.$$

The first inequality bounds a pair from zero feedback; the second bounds the next active input from nonzero feedback. For $q = 1$, set $T = 0$. Long gaps use only the bound $P_{||}/q$, so no estimate assumes that a product of coefficients over a long gap is less than p as an integer.

Conditioning independent labels to be distinct costs $m/(m - 1)$ for a pool of m labels. The induction budget $\max(1, m)$ absorbs each pair exposure, so a scan with s active inputs costs at most s . Its terminal terms are

$$R_0 = \frac{\lambda}{\lambda - 1} \left(1 + \frac{xy}{\lambda - y} \right), \quad R_1 = \frac{\lambda}{\lambda - y}, \quad C = \frac{\max(R_0, R_1/V)}{\epsilon}.$$

Induction on the remaining labels bounds the scan by $sC\lambda^n$. The two-scan profile bound is therefore (10) with an additional factor sh . All inequalities are rational.

For $q = 2$, use (27) for $s < 32$. For $q \geq 4$, use the coefficient-sensitive direct bound with exact $F_s(\tau)$ and the sixteen-step terminal factor in (7). For $s \geq 32$, the dense domain is partitioned recursively into rectangles, and each leaf uses either the direct or the projective-pair bound, uniformly over its profiles. The verifier checks complete coverage of $s \leq 2h$, $h \leq 2e$, and sums all contributions. Using different valid bounds on different leaves requires no independence assumption. These calculations give the $q \geq 2$ rows of Table 1.

K Counting complete runs

The complete-run bound counts a whole run at once, coupling free returns with returns that add a zero equation. It is uniform over the active forms and can be used with the Newton or subset-sum precode.

A bound for all free returns. Let μ be a distribution on nonzero vectors of a finite-dimensional space over $\mathbb{F}_p$, with p odd. Draw independent increments with law μ , set $S_0 = 0$, write S_j for their partial sum, and let T be the first positive return time to zero. For $0 < \omega < 1$, define

$$G_\mu(\omega) = \sum_{j \geq 0} \Pr[S_j = 0] \omega^j, \quad F_\mu(\omega) = \mathbb{E}[\omega^T; T < \infty].$$

The renewal identity gives $F_\mu = 1 - 1/G_\mu$. Uniformly over μ , including nonsymmetric distributions,

$$F_\mu(\omega) \leq F_p(\omega) := 1 - \frac{1}{G_p(\omega)}, \quad G_p(\omega) = \frac{1}{p} \sum_{j=0}^{p-1} \frac{1}{1 - \omega \cos(2\pi j/p)}. \quad (28)$$

To prove this, average over additive characters χ of the vector space and put $\phi_\chi = \sum_a \mu(a)\chi(a)$. Fourier inversion gives

$$\begin{aligned} G_\mu(\omega) &= \mathbb{E}_\chi \operatorname{Re} \frac{1}{1 - \omega \phi_\chi} \leq \mathbb{E}_\chi \frac{1}{1 - \omega \operatorname{Re} \phi_\chi} \\ &\leq \sum_a \mu(a) \mathbb{E}_\chi \frac{1}{1 - \omega \operatorname{Re} \chi(a)} = G_p(\omega). \end{aligned}$$

The second inequality is Jensen's inequality for the convex function $u \mapsto (1 - \omega u)^{-1}$. Every nonzero a has order p , so its character values are uniform over the p th roots of unity. Balanced increments $a, -a$ attain the bound. The Poisson-kernel series also gives

$$G_p(\omega) = \frac{1}{\sqrt{1 - \omega^2}} \frac{1 + r_\omega^p}{1 - r_\omega^p}, \quad r_\omega = \frac{\omega}{1 + \sqrt{1 - \omega^2}}. \quad (29)$$

Thus $0 \leq F_p(\omega) - (1 - \sqrt{1 - \omega^2}) \leq 2r_\omega^p$. The term $1 - \sqrt{1 - \omega^2}$ is the p -independent main term of $F_p(\omega)$.

Returns and rank gains in the same run. Fix a run start, its exposure history, and its quotient by earlier zero equations. Let $a_1, \dots, a_m$ be the forms of the remaining active input labels in this quotient. On a retained branch every a_i is nonzero. Define

$$\mu = \frac{1}{m} \sum_{i=1}^m \delta_{a_i}, \quad \mathcal{U} = \bigcup_{i=1}^m (\operatorname{span}(a_i) \setminus \{0\}).$$

Here δ_a denotes point mass at a . Until the next zero output, the quotient remains fixed. All these input forms remain required nonzero, including labels consumed during the run. Imposing a zero equation on a nonzero sum in $\mathcal{U}$ would make some required form a_i zero, so no such return occurs on a retained branch. Outside $\mathcal{U}$, allow a rank-gaining return with weight $t = 1/\tau$, where $\tau \geq 1$. Ignoring further nonzero requirements can only enlarge the count. For independent draws from the fixed distribution μ , the complete-run series is

$$\mathcal{R}_\mu(\omega, t) = F_\mu(\omega) + t \sum_{j \geq 1} \Pr[T > j, S_j \notin \mathcal{U}] \omega^j.$$

The event $T > j$ excludes the zero vector in the second term and every earlier free return. This is a bound on a weighted branch count, not the probability of one

event: a walk may allow rank-gaining endpoints at several lengths. Distinct-label sampling is restored below.

Assume $0 < \omega < 1$ and $\omega \leq c \leq 1$, where $c = t/(1 - \omega)$. For a one-dimensional subspace ℓ with $\mu(\ell) > 0$, write $a_\ell = \mu(\ell)$, $\theta_\ell = \omega a_\ell$, and let ν_ℓ be the conditional increment distribution on that line. Starting from $f \in \ell \setminus \{0\}$, let $H_\ell(f)$ be the θ_ℓ -discounted probability of hitting zero under ν_ℓ . A hit needs at least one step, so $0 \leq H_\ell(f) \leq \theta_\ell$. Put

$$b_\ell = \frac{c\omega(1 - a_\ell)}{1 - \omega a_\ell}, \quad V(f) = \begin{cases} b_\ell + (1 - b_\ell)H_\ell(f), & f \in \ell \setminus \{0\} \subset \mathcal{U}, \\ c, & f \neq 0, f \notin \mathcal{U}. \end{cases}$$

On an active line, $V(f) \leq \omega[c + (1 - c)a_\ell] \leq \omega \leq c$. Outside $\mathcal{U}$, no active increment cancels f ; its immediate rank-gaining return and subsequent continuations are bounded by $t + \omega c = c$. On a line, off-line increments have continuation at most c , and on-line transitions satisfy the hitting-probability recurrence. Substitution gives exactly $V(f)$ as their upper bound. Induction over finite truncations, followed by monotone convergence, proves

$$\mathcal{R}_\mu(\omega, t) \leq \sum_\ell [\theta_\ell b_\ell + (1 - b_\ell)F_p(\theta_\ell)].$$

Here $F_{\nu_\ell}(\theta_\ell) = \theta_\ell \mathbb{E}_{a \leftarrow \nu_\ell} H_\ell(a)$, and (28) bounds each line term.

Since $\sum_\ell \theta_\ell = \omega$, bound each $F_p(\theta_\ell)$ by its main term $1 - \sqrt{1 - \theta_\ell^2}$ plus $2r_{\theta_\ell}^p$. Substituting $\theta_\ell = 2r/(1 + r^2)$ into the main terms gives

$$\mathcal{R}_\mu(\omega, t) \leq \omega \max_{0 \leq r \leq r_\omega} \varphi(r) + 2r_\omega^p, \quad \varphi(r) = (1 - c\omega)r + c(2 - \omega) - \frac{2t}{1 - r}. \quad (30)$$

The finite-field error follows from $r_{\theta_\ell}/r_\omega \leq \theta_\ell/\omega$ and $\sum_\ell (\theta_\ell/\omega)^p \leq 1$. This bound retains the dependence between free returns and rank gains; it does not maximize their probabilities separately.

Distinct labels and terminal words. Choose $x > 0$, $0 < y \leq 1$, $\lambda > 1$, and set

$$A = \frac{xy}{\lambda - 1}, \quad B = \frac{x}{\lambda - y}, \quad \omega = By < 1.$$

After summing zero-input gaps, a complete run with j active inputs has factor $AB^{j-1}y^{j-2} = (A/(By^2))\omega^j$. The run-series multiplier $A/(By^2)$ therefore converts the complete-run series, whose j -input terms carry ω^j , into these run factors. Use cutoff $J = 16$ and budget $\beta(\kappa) = \max(1, \kappa^J)$. Write $(\kappa)_j = \kappa(\kappa - 1) \cdots (\kappa - j + 1)$ for the falling factorial. For $\kappa > J$ and $j \leq J$, conditioning independent labels to be distinct is absorbed by

$$\frac{\kappa^j}{(\kappa)_j} \beta(\kappa - j) \leq \beta(\kappa).$$

Indeed, after dividing by κ^J the left side is $((\kappa - j)^j/(\kappa)_j)((\kappa - j)/\kappa)^{J-j} \leq 1$. For $j > J$, work directly without replacement. Free and rank-gaining endpoints are

disjoint and have total rank-weighted probability at most one. Their geometric tail is at most $\omega^{J+1}/(1-\omega)$.

The finite certificate requires $\omega \leq 1/2$ and $p \geq 256$, so the finite-field term in (30) is at most 2^{-255} . With slack $\epsilon = 10^{-6}$, it suffices to check

$$\frac{A}{By^2} \left[\omega \max_{0 \leq r \leq r_\omega} \varphi(r) + 2^{-255} + \frac{\omega^{17}}{1-\omega} \right] \leq 1 - \epsilon. \quad (31)$$

The extra tail deliberately overcounts long runs already present in the independent-draw envelope. Terminal words and pools of at most sixteen active labels are bounded by

$$R_{\text{terminal}} = \frac{\lambda}{\lambda-1} + \frac{A\lambda}{(\lambda-y)(1-\omega)},$$

$$R_{\text{small}} = \frac{\lambda}{\lambda-1} \sum_{j=0}^{16} \left(\frac{x(1+y)}{\lambda-1} \right)^j, \quad C = \frac{\max(R_{\text{terminal}}, R_{\text{small}})}{\epsilon}.$$

To make the induction explicit, discount a word of length N by λ^{-N} . Its marker weight is $x^\kappa y^h$, where κ, h are the input and output weights, and each new zero equation contributes τ^{-1} . Average over the uniform order of the active labels and sum over support branches and placements, as in Lemma 5.1.

For $\kappa \leq 16$, $R_{\text{small}} \leq \epsilon C$ bounds the entire residual word, after summing its possible lengths. For $\kappa > 16$, split at the first completed run. The distinct-label inequalities and (31) bound the completed-run contribution by $(1-\epsilon)C\beta(\kappa)$, assuming the bound for smaller pools. The terminal contribution is at most $R_{\text{terminal}} \leq \epsilon C$. Since $\beta(\kappa) \geq 1$, their sum is at most $C\beta(\kappa)$. The bound is uniform over the new quotient after the return, which justifies its use in the induction.

Extracting length n gives $K(\kappa) = C\beta(\kappa)\lambda^n$, including either final feedback state. For positive s, h , Lemma 5.1 therefore gives the factor $(sh)^{16}$ multiplying (10). The source factor remains $F_s(\tau)/\tau$; it is charged only once.

Rational checks and the Newton instantiation. The certificate specifies, for each use of (31), rational τ, x, y, λ and $0 < \rho < 1$ satisfying $\omega(1+\rho^2) \leq 2\rho$. This implies $\rho \geq r_\omega$. After subtracting the explicit error and tail terms, inequality (31) reduces to $\varphi(r) \leq H$ on $[0, \rho]$, for a computed rational H . Multiplying by $1-r$ gives

$$(c\omega - H) + (1 - 2c + H)r + (c\omega - 1)r^2 \leq 0.$$

Its leading coefficient is negative. Exact evaluation at both endpoints and at its vertex, when inside the interval, proves the required uniform bound. All checks are exact rational comparisons.

For the Newton precode, retain $Q_8(z, \tau)$ from (8). At forbidden weight $e = 245366$ (so the distance is at least $e+1$), the hypergeometric bound (27) covers all $s < 32$. This sparse contribution has base-two logarithm at most -49.647382836382 . The dense domain is partitioned recursively into rectangles;

each cell is a leaf. There are 180 nonempty leaves: 44 use the projective-pair bound, three use the joint-renewal bound, and 133 use the complete-run bound. Every integer profile with $32 \leq s \leq \min(n, 4e)$ and $\lceil s/2 \rceil \leq h \leq \min(n, 2e)$ is covered. The total failure probability has base-two logarithm at most -49.647365460353 . Thus the unchanged addition-only Newton encoder has distance at least 245367 ($> 11.7\%$ of n) with failure probability below $2^{-49.64}$. The certificate and its verifier are listed in Appendix N.

The argument is uniform in the source forms. Substituting the catalog of Appendix L gives the stronger distance of Theorem L.1, with a smaller failure margin and fewer additions.

L A subset-sum precode above the binary Gilbert–Varshamov benchmark

The support catalog of the precode is another design parameter. The Newton precode B_8 is MDS, so it maximizes local distance, but it permits many distinct supports. The following smaller circuit trades local distance for fewer supports and gives a stronger complete distance bound with a smaller failure margin.

Index eight inputs by $j \in \mathbb{F}_2^3$ and define

$$y_i = \sum_{j: \langle i, j \rangle = 0} x_j, \quad B_{\text{sub}}(x) = (x_0, \dots, x_7, y_0, \dots, y_7).$$

The binary dot product selects the subsets; the sums are evaluated over $\mathbb{F}_p$. Thus y_0 sums all eight inputs and each other y_i sums four. Every entry of the matrix of B_{sub} is zero or one. Let E be the instance of (3) with $w = 1$, $q = 1$, and $L = 2$, and with B_8 replaced by the subset-sum precode B_{sub} applied blockwise. Thus E applies B_{sub} , then two independently permuted plain accumulators.

Theorem L.1. *For $p = 2^{128} - 159$, $k = 2^{20}$, and $n = 2k$, the encoder E has rate $1/2$ for every setup. Except with probability less than $2^{-22.05}$ over the two independent uniform permutations, its minimum distance is at least 263403, which exceeds 12.56% of n . Encoding uses exactly $3.3125n - 2 = 6946814$ field additions, with no subtractions or field multiplications.*

This relative distance exceeds the binary rate-1/2 Gilbert–Varshamov (GV) benchmark $H_2^{-1}(1/2) \approx 0.1100278644$. The theorem concerns the specified finite prime-field code. It does not compare its actual distance with that of the same circuit over $\mathbb{F}_2$, or give an asymptotic family theorem.

Circuit. The systematic copy gives injectivity. To compute the remaining outputs, first form $x_0 + x_j$ for $j \in \{1, 2, 3\}$, using the usual three-bit representation of the indices. Together with 0, these three indices form a two-dimensional subspace of $\mathbb{F}_2^3$, so every nonzero i is orthogonal to at least one of them. For such a j , the shared sum $x_0 + x_j$ and two further inputs compute y_i in two additions. Finally, $y_0 = y_1 + ((x_1 + x_3) + (x_5 + x_7))$ costs four additions. The block therefore uses $3 + 14 + 4 = 21$ additions, and the two scans use $2(n - 1)$.

Table 5. Exact shortening catalog of the eight-input subset-sum block.

$s [z^s]S(z, \tau)$	$s [z^s]S(z, \tau)$
0 1	10 $3045\tau^2 + 154\tau^3$
5 7τ	11 $3514\tau^3 + 7\tau^4$
6 77τ	12 $1785\tau^4$
7 28τ	13 $560\tau^5$
8 $154\tau + 84\tau^2$	14 $120\tau^6$
9 $772\tau + 511\tau^2$	15 $16\tau^7$
	16 τ^8

Support catalog. Let Z be a set of local output coordinates constrained to zero, and let V_Z be the kernel of those rows of B_{sub} . An exact zero set must be closed under row span: every local row in the span of Z also vanishes. Conversely, if Z is closed and has rank $j < 8$, its kernel has dimension $8 - j$ and each remaining coordinate is a nonzero linear functional on that kernel. Since $p > 16$, their proper hyperplanes cannot cover the kernel. Hence every proper closed set is realized as an exact zero set.

Enumerating the 2^{16} subsets by exact elimination over $\mathbb{F}_p$ gives the polynomial $S(z, \tau)$ in Table 5. Its coefficient of $z^s \tau^d$ counts supports of size s whose shortened source space has dimension d . The constant term represents the zero block.

Finite distance proof. For $n/16$ independent blocks, the dimension-weighted source moment at support s is $[z^s]S(z, \tau)^{n/16}$. The projective-pair and joint-renewal bounds of Appendices G and H are uniform in the source forms and their successive quotients. They therefore apply with this source polynomial; their scan inequalities and finite sampling corrections are unchanged.

These two bounds alone prove that E has distance at least 239076 ($> 11.4\%$ of n), except with probability less than $2^{-24.26}$. Take forbidden weight $e = 239075$. The sparse bound of Appendix F, with source counts from $S(z, 1)^{n/16}$, covers $1 \leq s \leq 31$. The remaining domain is $32 \leq s \leq 4e$ and $\lceil s/2 \rceil \leq h \leq 2e$. It is partitioned recursively into rectangles, and each cell is a leaf: 62 leaves use the projective-pair bound and 61 use the joint-renewal bound. The verifier recomputes the local catalog, checks every partition and vertex bound, and checks each joint-renewal inequality over the entire pool simplex. Adding all sparse and dense terms with upward rounding gives a failure probability with base-two logarithm at most -24.265906406170 .

For the stronger forbidden weight $e = 263402$ of Theorem L.1, use the complete-run bound of Appendix K with the same catalog S . The sparse support count again covers $1 \leq s \leq 31$. The dense cover has 65 projective-pair leaves and 131 complete-run leaves, covering every remaining feasible profile. The total failure probability has base-two logarithm at most -22.051928329751 , which proves the theorem. The verifier reconstructs the exact local catalog. Both certificates and their verifiers are listed in Appendix N.

The subset-sum encoder thus has a larger proved distance and fewer additions than the Newton encoder of Appendix K, whose distance exceeds 11.7% of n with failure probability below $2^{-49.64}$, but a smaller failure margin. We did not time this variant.

M Finite bounds for widths two and four

The measured wider encoders use the same Newton precode B_8 and two permutations, without a final unpermuted scan. We extend the local rank bound by recording the support mask and rank of the preceding w outputs. Fix the shortened source space V_S and quotient by the zero equations imposed so far. Let $M \subseteq \{1, \dots, w\}$ record the positions of the required nonzero feedback forms, and let R be their span dimension in this quotient. The states are $(\emptyset, 0)$ and (M, R) with $1 \leq R \leq |M|$. There are $1 + w2^{w-1}$ possible states. Choose positive continuation values ordered by rank: $v_{M,R+1} \geq v_{M,R}$. This is a condition on the witness, not a claim that the exact continuation count increases with rank.

Drop the oldest state and shift the remaining mask to obtain D , with its new-coordinate bit zero. Let $m = |D|$. With zero input, the output form lies in the span of the state forms. A *contraction* is a zero output whose form is a nonzero element of this span: the new zero equation lowers the rank of the state span by one. The following successor envelopes bound arbitrary linear dependencies:

Event	Successor mask	Rank bound
Zero output form (Z)	D	$\min(R, m)$
Nonzero output in state span (I)	$D \cup \{\text{new}\}$	$\min(R, m + 1)$
Output outside state span (O)	$D \cup \{\text{new}\}$	$\min(R, m) + 1$
Contract a state-span form (C)	D	$\min(R - 1, m)$

The last event is allowed only for $R \geq 2$ and a surviving contraction, one after which every required form remains nonzero. If the new zero equation makes a required form zero, that branch is discarded. A zero output whose form lies outside the state span preserves the old state rank and is bounded by the Z envelope. Rank ordering lets each envelope use its displayed upper rank. Write v_Z, v_I, v_O, v_C for the continuation values at the successor states of the four events.

At $R > 0$, fresh independent coefficients in $\{1, \dots, q\}$ make the output form zero with probability at most $\pi = q^{-R}$. Condition on all but R pivot coefficients: at most one pivot assignment cancels the form. With zero input and $R = |M|$, cancellation is impossible because the state forms are independent and every coefficient is nonzero. Define

$$\Theta_0 = yv_I + \mathbf{1}_{R \geq 2} \tau^{-1} v_C, \quad \Theta_{\text{ind}} = yv_O + \tau^{-1} v_Z, \quad \Theta_{\text{dep}} = (1 - \pi)\Theta_0 + \pi v_Z.$$

At rank one, the state span is spanned by a single required form, so a contraction would make that form zero; this is why Θ_0 has no zero branch when $R = 1$. The O and I successors share a mask and O has at least the rank of I , so rank

ordering gives $v_O \geq v_I$; likewise $v_Z \geq v_C$. Hence $\Theta_{\text{ind}} \geq \Theta_0$. Maximizing over the unknown cancellation probability in $[0, \pi]$ gives, in analogy with (6),

$$T_{M,R}(v) = \begin{cases} \Theta_0 + x \max(\Theta_{\text{ind}}, \Theta_{\text{dep}}), & R = |M|, \\ \max(\Theta_0, \Theta_{\text{dep}}) + x \max(\Theta_{\text{ind}}, \Theta_{\text{dep}}), & R < |M|, \end{cases}$$

and $T_0(v) = v_0 + xyv_{\{\text{new}\},1}$. After each step, take the maximum over smaller ranks at a fixed mask. This closure preserves the ordering. If $T(v) \leq \lambda v$, its closure satisfies the same inequality. With $v_0 = 1$, put $\Gamma(a) = \max_i [T^a(\mathbf{1})]_i / v_i$. Positive homogeneity gives $[T^n(\mathbf{1})]_0 \leq \Gamma(a) \lambda^{n-a}$, as in (7) for $a = 16$.

All successor bounds hold for arbitrary incoming forms and earlier nonzero requirements. Lemma 5.1 therefore applies with these scan bounds and the same $F_s(\tau)/\tau$ source factor. The inverse scan gives $h \leq (w+1)e$ and $s \leq (w+1)h$, capped by n . For $s < 32$, compute the source coefficients exactly, use $a = 16$, and cover all feasible intermediate weights by consecutive integer intervals. For $s \geq 32$, use $F_s(\tau) \leq Q_8(z, \tau)^{n/16}/z^s$ and $a = 0$. The rational binomial prefactors and convex vertex bounds of Appendix F then apply with the wider support constraint. The dense domain is partitioned recursively into rectangles; each cell is a leaf, and each leaf is charged its full number of integer profiles. The leaves cover the entire feasible domain; summing the interval and leaf bounds with upward rounding gives the complete failure bound.

For $w = 2$ and $q = 2048$, at forbidden weight $e = 829633$ the sparse range uses 185 intervals and the dense cover has 216 leaves; the total failure probability has base-two logarithm at most -57.875553147005 . For $w = 4$ and $q = 65536$, at $e = 901502$ the sparse range uses 180 intervals and the dense cover has 229 leaves; the total has base-two logarithm at most -61.763321686333 . These certificates yield the two rows of Table 2. Failure exponents there are truncated toward zero, so the stated bounds are conservative.

N Reproducing the finite results

The artifact provides one command that checks the direct construction’s headline results. From the repository root, run

```
python field-universality/paper/reproduce_results.py
```

The companion file `paper/REPRODUCING.md`, under `field-universality/`, gives dependencies, certificate paths, build commands, and the procedure for repeating the timing runs. Python assertions must remain enabled: the verifiers rely on `assert` statements, so the runner refuses to start under `python -O`. The runner reads a fixed list of certificates, the publication selection, recorded in `results/distance-time/certificate-audit.json` under `field-universality/expander-lift/`; it never searches for newer certificates. The inventory `paper/results-inventory.json` under `field-universality/`, generated by `paper/build_result_inventory.py`, records for every reported point its distance, failure bound, timing, and the hashes of its evidence files. Tables 1, 2, and 3 are generated from the same records.

What the runner checks. The runner performs four checks. First, it verifies twelve finite certificates: the eleven measured Newton points in Tables 1 and 2, and the subset-sum point of Theorem L.1. For each, it recomputes the verified summary from the certificate and compares it with the saved record. Second, it reconstructs the deterministic ceilings of Section 4.5 by exact integer counting. Third, it checks the archived benchmark source hashes and correctness logs, recomputes the timing summaries, and checks each timing’s association with its certificate and with the plotting data. Fourth, it regenerates the inventory and all three generated tables, including the numerical grid, and compares them with the manuscript files. The grid entries are compared with their stored records; they are not reoptimized. These checks verify numerical bounds under the paper’s mathematical arguments; they are not a formal verification of those arguments. Checking archived logs does not rerun the encoder or reproduce the timing experiment on another host.

Certificates and verifiers. Unless stated otherwise, paths are relative to `field-universality/expander-lift/`; certificate files without a directory are in `results/`, and verifier scripts without a directory are in the top level. Results marked $\dagger$ are checked by the runner; the others have separate verifiers.

- Width-one Newton points, $q = 1$ (Appendix K) $\dagger$: `all-one-newton-line-adaptive-e245366.json`, checked by `all_one_line_run_certificate.py`.
- Width-one Newton points, $q = 2, \dots, 256$ (Appendix J) $\dagger$: the rows of `results/distance-time/certificate-audit.json`, checked by `small_q_fresh_pair_certificate.py`.
- Widths two and four (Appendix M) $\dagger$: the rows of the same manifest under `results/convolution-wide-finite/`, checked by `convolution_width_finite.py`.
- Subset-sum precode, Theorem L.1 $\dagger$: `all-one-line-runs-e263402.json`, checked by `all_one_line_run_certificate.py`, including reconstruction of the exact local catalog.
- Subset-sum precode with the joint-renewal bound, $e = 239075$ (Appendix L): `all-one-binary-hadamard-e239075.json`, checked by `all_one_hadamard_certificate.py`.
- Deterministic ceilings (Section 4.5) $\dagger$: `results/convolution-width-one-ceilings.json`, checked by `convolution_width_one_ceiling.py`.
- Two-state bound, $e = 138412$ (Appendix F): `results/convolution-width-one-direct-r8-q1-e138412.json`, checked by `convolution_width_one_finite.py`.
- Projective-pair bound, $e = 209715$ (Appendix G): `results/all-one-projective-pairs-r8-e209715.json`, checked by `all_one_projective_pairs.py`.
- Joint-renewal bound, $e = 224395$ (Appendix H): `results/all-one-joint-renewal-r8-e224395.json`, checked by `all_one_renewal_certificate.py`.
- Four-input precode, Theorem C.1 (Appendix E): the summary `results/scalar-frontier/summary.json` and the certificates `results/scalar-frontie`

- r/distance-q2048-e304087.json and results/scalar-frontier/margin-q2048-e230896.json, generated by scalar_frontier.py and checked by distance_ladder.py, which verifies the rational side conditions, interval endpoints, coverage, sparse sum, total failure bound, and a 2^{-40} budget. The tests in test_scalar_frontier.py check every complete certificate in the summary. The $q = 32$ row is in results/signed-scan-certificate.json, with verifier signed_scan_certificate.py.
- Four-input timings (Table 4): the archive results/scalar-alphabet-benchmark/, with runner bench/run_scalar_alphabet.sh and archive checker bench/summarize_scalar_alphabet.py. The archive includes the measured sources, raw timing samples, host and compiler details, hashes, and validation outputs.
 - Fixed-prime BAA bounds, Theorem D.1 (Appendix D); paths relative to field-universality/: the summaries results/ternary-bulk-0.196-eta-0.0001-independent.json, results/quinary-sign-orbit-bulk-0.05-eta-1-over-60.json, and results/septic-sign-orbit-bulk-0.018-eta-0.0625.json, reconstructed and checked by scripts/verify_symmetric_ternary.py and scripts/verify_sign_orbit_bulk.py. Each summary’s command field gives the full cover reconstruction command; reading a summary alone does not check the cover.
 - Walsh–Hadamard union bound (11) and the sparse BAA parameter rows: paper/audit-supporting-checks.json, relative to field-universality/, written by paper/check_audit_arithmetic.py using rational arithmetic.

Measured curve. The timing summaries come from 72 processes and 2232 timed encodes: twelve kernels, each run with six setup seeds and 31 timed encodes per seed. For $q = 2$ there are two implementations, one with additions and one with small-integer multiplications; they produce identical output checksums, and the figure and tables use the one with the smaller median time. The reported latency for a kernel is the median of its six process medians, that is, the mean of the third and fourth; each process median is the median of 31 timed encodes. Figure 1 shows the range of those process medians.

Measurement environment. All timings used one core of an AMD Ryzen 9 7950X, pinned to CPU 2. For the Newton encoders, boost was disabled with a 4.5 GHz frequency setting, and GCC 15.2.0 compiled with `-O3 -march=native -DNDEBUG` and `C++20`. Each configuration used two warmups before its timed encodes. Runs were serial, with case order reversed on alternate seeds. Setup, allocation, reference checks, and output checksums were outside the timer. Setup uses reproducible pseudorandom streams; the distance theorems concern the specified ideal sampling distribution. Complete outputs were checked against independent references and integer arithmetic oracles, and sanitizer runs passed. The four-input timings of Table 4 used the same processor and core, GCC 15.2 with `-O3 -march=native`, 42 processes, and 1302 timed encodes. Its pages were resident and caches were not flushed. Its outputs matched a dense-local generic-field reference; boundary and tail tests, multiprecision arithmetic checks

through coefficient 2048, and address and undefined-behavior sanitizers passed. SplitMix64 generated its reproducible benchmark fixtures; these fixtures are not a cryptographic setup implementation.

Scope of the comparison. All eleven measured points use the Newton [16, 8, 9] precode B_8 , two permutations, no final unpermuted scan, and at least forty bits of setup-failure margin. The subset-sum point uses a different precode and a smaller margin, and was not timed. Width eight has numerical dense-support estimates only; it is not part of the finite-certificate or timing checks. The runner does not check the BAA and Walsh–Hadamard case studies (Sections A and B), the separate verifiers listed above, or any search for optimal parameters.