

Arithmetic coding against unknown small-prime signs

Brice Pouly

With AI assistance

Abstract

We encode binary messages in completely multiplicative sign functions and read them along the three sequences $n^2 + 1$, $n^2 + 2$ and $n^2 - 2$. Two observations with the same exact cofactor remove the unknown smaller-prime signs when multiplied. Allocating shared large primes before pairing embeds ordinary binary error-correcting codes in each trace. Under an explicit large-factor mass hypothesis, a computable schedule gives one infinite encoding with positive prefix rate and a linear errors-and-erasures allowance. A first-occurrence filter makes its raw readings depend causally on a linear message prefix, while encoding and decoding remain polynomial in the numerical observation length. Its supporting primes have arbitrarily small harmonic cost. A finite construction, independent of the analytic hypothesis, stores 273 bits in 8736 readings of each trace and corrects every pattern of 499 reading errors.

Complementary constructions give persistent sign prescriptions, uniform laws for the active coordinates of words, and compact multiplicative subgroups whose nontrivial quotients have positive quadratic entropy. The word-law and entropy conclusions are not asserted for the polynomially decoded encoding. A separate matrix construction gives another causal positive-rate map. Fixed-scale balancing provides an unconditional quarter-sign bound against every small-prime assignment. All coding statements concern deliberately constructed families, not recovery of an arbitrary multiplicative source. The sparse constructions at superlinear cutoffs keep their large-factor mass input explicit.

Contents

1	Encoding a message without knowing the small signs	2
2	Arithmetic resources and the mass input	5
3	Removing the cofactor before decoding	8
4	A constructive binary code and its decoder	10
5	One infinite message and every sufficiently large scale	13
6	A finite code with 273 message bits	15
7	Persistent prescriptions before the small signs are known	18
8	Uniform signs on the active coordinates of words	21
9	A compact group with every nontrivial quotient separated	25
10	An alternative causal encoding by row selection	28
11	What is recovered and what remains unspecified	31
A	Fixed-scale balancing without the superlinear mass input	32
B	Finite data and reproduction	35

1 Encoding a message without knowing the small signs

A completely multiplicative function with values in $\{-1, 1\}$ is determined by its signs at the primes. This makes it natural to place information at large primes. It does not make that information directly readable: an observation of $h(aq)$ is the intended large-prime sign $h(q)$ multiplied by an unknown cofactor sign $h(a)$. The difficulty is not ordinary bit noise. A single change at a small prime can alter many readings at once, and the receiver is not given those changes.

Here is the cancellation that will be used throughout the main construction:

$$2067^2 + 1 = 10 \cdot 427249, \quad 2073^2 + 1 = 10 \cdot 429733. \quad (1.1)$$

The two displayed large factors are prime. For every completely multiplicative sign function,

$$h(2067^2 + 1)h(2073^2 + 1) = h(427249)h(429733).$$

Prescribing the latter two signs as $(-1)^b$ and 1 makes the product reveal the bit b , irrespective of $h(2)$ and $h(5)$. The same argument works for any exact common cofactor. It does not require the cofactors attached to different pairs to be independent, known, or distinct.

One pair transmits one bit, but provides no protection against a reading error. We therefore put an error-correcting code on the pair products. The arithmetic tasks are to find enough pairs, to allocate a prime shared by several traces only once, and to preserve every earlier assignment while extending the same infinite message. These are separate from the classical coding tasks.

1.1 The observation model and the principal theorem

Write

$$\mathcal{D} = \{-1, -2, 2\}, \quad Q_d(n) = n^2 - d,$$

and let $\mathcal{M}_\pm$ be the group of completely multiplicative functions $\mathbb{N}_{\geq 1} \rightarrow \{-1, 1\}$. We retain the label (n, d) whenever $Q_d(n) > 0$; repeated integer values at distinct labels remain distinct observations. Logs are natural unless their base is indicated. The distance between two prime assignments is the possibly infinite quantity

$$d_\sigma(f, g) = \sum_{q \text{ prime: } f(q) \neq g(q)} q^{-\sigma}, \quad 0 < \sigma \leq 1. \quad (1.2)$$

An error changes a received sign, whereas an erasure marks a missing reading. Their numbers refer to labels, not to the number of modified prime signs.

The analytic input is a uniform population statement. For a fixed rational $1 < \gamma < 6/5$, set

$$R_\gamma(X) = \max\{2X, \lfloor X^\gamma \rfloor\}, \quad (1.3)$$

$$M_d^\gamma(X) = \#\{X/2 < n \leq X : P^+(Q_d(n)) > R_\gamma(X)\}. \quad (1.4)$$

Here $P^+(1) = 1$. The hypothesis used below is

$$\text{there exist } a > 0, X_a \text{ such that } \min_{d \in \mathcal{D}} M_d^\gamma(X) \geq aX \quad \text{for every integer } X \geq X_a. \quad (\text{GF}_\gamma)$$

Section 2.1 gives a sufficient weighted input and its counting consequence; Section 2.2 discusses a possible analytic route to that input. One large prime at each scale cannot replace (GF_γ) .

Theorem 1.1 (Causal prefix dependence with polynomial decoding). *Fix rational numbers $1 < \vartheta < \gamma < 6/5$ and assume (GF_γ) . For every rational $\varepsilon > 0$ and integer $B \geq 1$, there exist a decidable set of primes $E \subset (B, \infty)$ and a continuous injective homomorphism*

$$\Omega : \mathbb{F}_2^{\mathbb{N}_{\geq 1}} \longrightarrow \mathcal{M}_\pm, \quad \Omega_{x+y} = \Omega_x \Omega_y,$$

computable relative to x , such that $\Omega_x(q) = 1$ off E and $\sum_{q \in E} q^{-1} < \varepsilon$.

There are constants $r, c > 0$ and N_0 such that, for all $N \geq N_0$, each one of the three traces up to N recovers the first $\lfloor rN \rfloor$ message bits under the following conditions: all prime signs up to N^ϑ may have been reassigned arbitrarily; later message bits are unknown; and the received trace has t errors and s erasures with

$$2t + s < cN. \tag{1.5}$$

Before prime reassignment, the entire labelled raw prefix

$$(\Omega_x(Q_d(n)))_{1 \leq n \leq N, d \in \mathcal{D}, Q_d(n) > 0}$$

depends only on $x_1, \dots, x_{\lfloor 3N/4 \rfloor}$, for every $N \geq 1$. The encoding tables, the requested encoded readings, and the decoder are computable in time polynomial in the numerical parameter N , for fixed construction parameters. The threshold N_0 is common to all messages and reassignments.

The proof supplies the stronger finite fact that only the two reserved signs of a used pair matter. Every other prime sign may be arbitrary. Neither a factorisation oracle nor an enumeration of all small-prime assignments is used by the decoder. The causal statement concerns message dependence, not bounded computational delay: the public arithmetic tables can require finite anticipation of later epochs. Nor does it restrict an adversary's later, possibly adaptive, prime reassignment. Polynomial in N is not polynomial in the bit length $\log N$ of a single index.

A binary code of length L uses $2L$ arithmetic readings on one trace. The words transmitted on the three traces encode the same prefix, not three unrelated messages. In the finite implementation of Section 6, $L = 4368$, the message has 273 bits, and the guaranteed binary distance after pairing is 1000. Consequently every 499 reading errors can be corrected from any one trace. This finite assertion does not use (GF_γ) and does not determine r, c or N_0 .

The small harmonic support has an elementary additional consequence. For every two messages and every $T \geq 1$,

$$\#\{m \leq T : \Omega_x(m) \neq \Omega_y(m)\} \leq T \sum_{q \in E} q^{-1} < \varepsilon T. \tag{1.6}$$

Thus a code can be almost constant on ordinary integers while carrying a positive amount of recoverable information on the quadratic observations. The latter are not uniform samples of the former.

1.2 Three complementary ways to use the large primes

Exact cancellation is the route to polynomial decoding. Other uses of the same arithmetic resource retain properties which are not part of [theorem 1.1](#).

Persistent prescriptions in Section 7 balance the large signs before an arbitrary small-prime assignment is known. Their finite discrepancy estimate is uniform in the observation

window. Under the stated mass input, a single computable prescription keeps both signs present on each trace at all large scales. On $n^2 + 1$ alone, exact alternation is possible and has a wider analytic input, associated with the weighted estimate at exponent $13/10$.

Sections 8–9 instead make the signs on an arithmetically determined set of active word coordinates nearly uniform. A finite information budget pays for every small-prime assignment and every currently visible message. One deterministic selection then gives a compact multiplicative subgroup in an arbitrarily small prime ball; each fixed nontrivial quotient has positive entropy on the quadratic traces but zero entropy on the ordinary integers. The identity is excluded from the positive-entropy assertion. These groups are not asserted to have the polynomial decoder of [theorem 1.1](#).

Section 10 gives an alternative positive-rate construction by selecting a matrix whose row lengths are fixed by first admissible occurrences. It has causal raw dependence and computable selection and decoding, but no polynomial bound for those algorithms is proved. This is not a necessary computational price of causality: the first-occurrence filter in [Section 5](#) already gives that property to the polynomially decoded paired construction. The alternative retains a different, concentration-based organisation of the information budget.

Finally, [Appendix A](#) works at the smaller cutoff $2N$, without $(\mathbf{GF}_\gamma)$. It gives a quarter of each sign, up to $O(N/\sqrt{\log N})$, against every small-prime assignment. It also describes the small-dimensional cofactor code near which an almost-positive source must lie. A small set of possible tails is not an empty set of possible tails.

1.3 Relation to earlier methods and a route through the proof

Large prime factors as private coordinates of polynomial observations occur in earlier work on multiplicative correlations, including Klurman [[13](#), Section 2, Lemma 2.1]. In that lemma, the small-prime-power data are fixed first, and the remaining prime values are then chosen to produce a large quadratic mean. Here the arithmetic families are completely multiplicative and sign-valued, and the message is read despite small signs chosen after the encoding. The order of choices and the information recovered differ; we do not identify the two statements. The required simultaneous allocation is important here: a prime is private to one location on a fixed trace, but can occur on all three traces. The same paper’s proof of Proposition 1.1 also uses widely spaced scales and a summable reciprocal-prime cost; [Appendix A](#) identifies these antecedents locally. The local threshold $2n$ and first appearance of a large factor are also classical in the literature on primitive divisors of quadratic values, as recorded by Everest–Harman [[6](#), Proposition 1.3]. The contribution of the paired construction lies in combining this arithmetic resource with disjoint two-reading cancellation, compatible allocation between traces, and one persistent causal encoding. A dense elimination of the small-sign variables would not by itself preserve the same reading-error budget; the disjoint pairs do.

The binary codes are classical concatenated codes: a short linear inner code, Reed–Solomon evaluation [[24](#)] outside, and Forney’s generalised minimum-distance decoding [[7](#)]. We use the exposition in Guruswami’s notes [[11](#), Sections 1 and 4], [[12](#), Sections 3.3 and 5], and give the errors-and-erasures argument in the precise form needed here. We do not claim a new rate–distance tradeoff for abstract binary codes. The balancing argument uses floating variables of Beck–Fiala type [[1](#), [18](#)]. For the ordered lists arising here, it also belongs to the classical discrepancy problem for three permutations [[3](#), [17](#)]; [Section 7](#) gives the reduction. The effective selections use the classical conditional-expectation principle; constructive statistical selection has a broader antecedent in Galatolo–Hoyrup–Rojas [[9](#)]. [Section 8.3](#) compares the exact finite-cost selections in the companion papers [[22](#), Section 5] and [[21](#), Section 7] with the integer bad-event potential used here. The proof below is autonomous.

The recovery problem in [20, Theorems 1.1–1.2] concerns arbitrary multiplicative sources and the complete set of compatible readings. Here the sender chooses a family with additional coded redundancy. A linear corruption allowance for this family does not give the same allowance for every source.

For the principal theorem, Sections 2–5 are a complete route: count candidate primes, control their first occurrences, allocate and pair them, construct a classical code, and prove causal message dependence, with bounded multiplicative gaps under $(\mathbf{GF}_\gamma)$. Section 6 turns the same finite mechanism into inspectable arithmetic data. The later sections can be read independently after the elementary geometry and support estimates.

2 Arithmetic resources and the mass input

The coding argument needs many independent prime assignments, not independence of the observed integers. Two elementary facts identify which prime assignments can be used and the cost of their cofactors.

Lemma 2.1 (Uniqueness on one trace). *Let $X \geq 5$ and $1 \leq n \leq X$. If a prime $q > 2X$ divides $Q_d(n) > 0$, it is the unique prime factor greater than $2X$, has valuation one, and*

$$Q_d(n) = qa, \quad 1 \leq a < X/2.$$

For a fixed d , the same q divides at most one of these values. Across the three traces it divides at most three labelled observations.

Proof. Two factors greater than $2X$, including a repeated one, have product at least $(2X + 1)^2 > X^2 + 2$. The cofactor satisfies

$$a \leq \frac{X^2 + 2}{2X + 1} < X/2.$$

If q divides both $n^2 - d$ and $m^2 - d$, then it divides $(n - m)(n + m)$. For distinct positive $m, n \leq X$, neither nonzero factor is divisible by $q > 2X$. $\square$

The local version of this geometry has a classical primitive-divisor interpretation. For a fixed integer b with $-b$ not a square and $n > |b|$, Everest–Harman [6, Proposition 1.3] record that $n^2 + b$ has a primitive divisor exactly when $P^+(n^2 + b) > 2n$; it is then a unique prime. Here a primitive divisor is coprime to every earlier nonzero term. Our common window threshold $2X$, and especially $R_\gamma(X)$, impose additional requirements. The cited local criterion is not the population hypothesis $(\mathbf{GF}_\gamma)$, and it does not separate primes shared by different polynomials.

For example, the prime 41 occurs in $9^2 + 1 = 2 \cdot 41$, $11^2 + 2 = 3 \cdot 41$ and $17^2 - 2 = 7 \cdot 41$. It must receive only one sign. Treating those three occurrences as independent prime bits would change the source model.

At cutoff $R_\gamma(X)$ the exact cofactors belong to

$$1 \leq a \leq A_\gamma(X) := \left\lfloor \frac{X^2 + 2}{R_\gamma(X) + 1} \right\rfloor = O_\gamma(X^{2-\gamma}) = o(X). \quad (2.1)$$

This is the gain used in pairing. A linear population of readings split among $o(X)$ cofactor classes loses only $o(X)$ readings when at most one is left unpaired in each class.

2.1 From a weighted deficit to a population

The exact role of (GF_γ) is worth isolating. Fix a nonnegative smooth function b , supported in $(1, 2)$, with $b \leq 1$ and $I_b = \int b > 0$. For each of the three fixed d , put

$$\mathcal{A}_{d,q}(x) = \sum_n b(n/x) \mathbf{1}_{q|n^2-d}, \quad W = xI_b.$$

A sufficient analytic input at an exponent $u > 1$ is

$$\sum_{q > x^u} \mathcal{A}_{d,q}(x) \log q \geq \eta_d W \log x \quad (x \geq x_d), \quad \eta_d > 0. \quad (2.2)$$

All sums over q in this section are over primes.

Lemma 2.2 (Weighted mass implies the required counting mass). *If (2.2) holds for every $d \in \mathcal{D}$, then (GF_γ) holds for every fixed $1 < \gamma < u$.*

Proof. For one integer n in the support of $b(n/x)$, the sum of the logarithms of its distinct prime divisors is at most $\log(n^2 - d) \leq 3 \log x$ for all large x . Hence (2.2) gives at least $(\eta_d I_b / 3)x$ contributing indices in $(x, 2x)$. Put $X = 2x$. For all sufficiently large X ,

$$(X/2)^u > \max\{2X, X^\gamma\}.$$

Taking the minimum of the three positive constants and slightly decreasing it for integer endpoints proves the claim. No constant is required to be uniform in a moving polynomial. $\square$

This elementary last step requires a weighted lower bound of order $x \log x$. The inequality $P^+(\prod_{x < n \leq 2x} Q_d(n)) > x^u$ alone does not provide it.

2.2 A possible analytic route to the mass hypothesis

This subsection describes a possible route to (GF_γ) , not an independent proof of it. The simultaneous sparse constructions retain that hypothesis. The weighted Chebyshev argument in de la Bretèche–Drappeau [4, Theorem 1.2 and Section 8.4], together with the sieve main-term calculation of Deshouillers–Iwaniec [5, Section 8], motivates the estimates below. The references identify the published editions; the locators for de la Bretèche–Drappeau refer to arXiv:1703.03197v3. We write D_s for the sieve level, to distinguish it from the nonsquare integer defining the polynomial.

Section 8.4 of [4] obtains, with $P = x^t$ and spectral parameter $s = 7/64$, an error estimate of the shape

$$R_H \ll x^{\epsilon + O(\eta)} (x^{-1} D_s P + (D_s P)^{1/2} + x^{1/2} P^s D_s^{1-2s}).$$

The admissible level relevant to the proposed transfer is

$$D_s = x^{\delta_s(t) - \eta}, \quad \delta_s(t) = \frac{1 - 2st}{2(1 - 2s)}, \quad 1 \leq t \leq 6/5, \quad (2.3)$$

with a sufficiently enlarged fixed loss in η when inserted into the error estimate. In this interval $\delta_s(t) \leq 1/2$ and $\delta_s(t) < 2 - t$, so the ancillary constraints $D_s \leq x^{1/2}$ and $D_s \ll x^2/P$ are compatible.

The smoothed upper-bound sieve estimate needed in addition is

$$\limsup_{x \rightarrow \infty} \frac{1}{W \log x} \sum_{x < q \leq x^v} \mathcal{A}_{d,q}(x) \log q \leq F_s(v), \quad F_s(v) = 4(1 - 2s) \int_1^v \frac{t \, dt}{1 - 2st}. \quad (2.4)$$

At $v = 6/5$ its proposed main-term coefficient has a strict margin:

$$F_{7/64}(6/5) = \frac{3200}{49} \log \frac{125}{118} - \frac{20}{7} < \frac{907}{1000} < 1. \quad (2.5)$$

The corresponding large-prime total required by this route is

$$\mathcal{T}_d(x) := \sum_{q>x} \mathcal{A}_{d,q}(x) \log q = W \log x + O_{d,b}(x). \quad (2.6)$$

This is not the logarithm of the full product: the full weighted logarithmic sum is

$$\sum_n b(n/x) \log(n^2 - d) = 2W \log x + O_{d,b}(x).$$

The Chebyshev decomposition removes the contribution of primes at most x and accounts for repeated prime powers before reaching (2.6). Combining (2.6) with the smoothed bound (2.4) and the strict margin (2.5) would yield (2.2) at $u = 6/5$. The numerical margin does not prove the required smoothed main-term estimate for each polynomial. That estimate, rather than the numerical comparison, is the additional analytic obligation in this proposed route.

For $n^2 + 1$ alone, the larger sufficient weighted input at $u = 13/10$ is located in Pascadi [19, Sections 6.1 and 6.3, equation (6.20)]. The final sieve calculation has a strict deficit even at 1.30008. Its use here is precisely (2.2), followed by lemma 2.2, not an inference from the infinitude formulation of the theorem. Section 7.1 records the construction under this single-trace input. No transfer of Pascadi's range to the other two polynomials is assumed.

Larger prime-factor exponents versus the mass used here. Two preprints from 2025 refine the greatest-prime-factor problem. Grimmelt–Merikoski [10, Theorem 1.1] obtain exponent 1.312 for $an^2 + h$ in their stated positive-parameter range, with a real-character hypothesis for the moving-parameter uniformity; fixed admissible positive a, h satisfy their hypothesis for sufficiently large scales. This includes the fixed cases $n^2 + 1$ and $n^2 + 2$. Li [14, Theorem 1.1] states exponent 1.317 in a further refinement, in the specifically identified second version of that preprint. These are context, not inputs whose full analytic proofs or numerical sieve calculations are reverified here. Their positive- h statements are not automatically assertions for $n^2 - 2$. Nor can their existence conclusions alone replace the simultaneous weighted population (GF_γ). We therefore leave the range and conditional status of the coding theorems unchanged.

2.3 A common sparse support

The later constructions use the cutoff

$$R_\vartheta(n) = \max\{2n, \lfloor n^\vartheta \rfloor\}, \quad 1 < \vartheta < 2.$$

An occurrence (n, d) is *admissible* if $q \mid Q_d(n)$ and $q > R_\vartheta(n)$. The same elementary root argument as in lemma 2.1 shows that a prime has at most one admissible occurrence on each trace, even over the entire infinite sequence: both indices of two putative occurrences would be less than $q/2$.

For $L \geq 2$ let E_L be the set of primes having an admissible occurrence with $n \geq L$. One can assign each prime a single such occurrence, so

$$\sum_{q \in E_L} q^{-\sigma} < 3 \sum_{n \geq L} n^{-\vartheta\sigma} \leq \frac{3(L-1)^{1-\vartheta\sigma}}{\vartheta\sigma - 1} \quad (\vartheta\sigma > 1), \quad (2.7)$$

and $\#(E_L \cap [1, y]) \leq 3y^{1/\vartheta}$. This is a bound on the entire possible support, before any message or sign is chosen.

Membership is decidable. For a given prime q , inspect the centred roots of $n^2 \equiv d \pmod{q}$ with $1 \leq n < q/2$, and check $n \geq L$ and $q > R_\vartheta(n)$. Rational exponents can be tested by integer powers. Thus neither decidability nor the choice of L in (2.7) uses an unknown threshold in (GF_γ) .

2.4 First occurrences control every raw reading

To limit the message bits that a raw reading can consult, the first admissible occurrence is more useful than the epoch at which a prime is reserved. For the cutoff R_ϑ above, define

$$\nu(q) = \min\{n \geq 2 : \text{for some } d \in \mathcal{D}, q \mid Q_d(n), q > R_\vartheta(n)\} \quad (2.8)$$

whenever this set is nonempty. This minimum is taken across all three traces. It is a common-cutoff refinement of the first-appearance geometry discussed after lemma 2.1, not a new primitive-divisor criterion.

Lemma 2.3 (First occurrence and all later uses). *If $\nu(q)$ is defined and q divides a positive value $Q_d(n)$ with $n \geq 2$, then*

$$\nu(q) \leq n. \quad (2.9)$$

There are at most $3(s-2)_+$ primes with $\nu(q) < s$, for any integer $s \geq 1$.

Proof. If $q > R_\vartheta(n)$, then n participates in the defining minimum. Otherwise $R_\vartheta(\nu(q)) < q \leq R_\vartheta(n)$, so monotonicity of R_ϑ forces $\nu(q) < n$. Each index $2 \leq n < s$ has three labels, and each label has at most one admissible large prime. Assigning every prime to a label of its first occurrence proves the count. $\square$

For a candidate prime at epoch X , $q > R_\gamma(X)$ and an occurrence is already known at an index at most X . Every smaller index $n \leq X$ satisfies $q > R_\vartheta(n)$ when $\vartheta < \gamma$. Thus $\nu(q)$ can be found by inspecting the three traces only up to X . In particular, the test $\nu(q) < s$ requires only indices $2 \leq n < s$ when $s \leq X$. This test needs no mass constant and no analytic threshold. It also explains why the code implementing the filter can compare ordinary first divisibility indices without taking ϑ as an additional parameter.

3 Removing the cofactor before decoding

The cofactor is a nuisance only as long as each reading is used separately. Pairing equal exact cofactors turns the multiplicative observation into an ordinary binary channel. We first state this finite reduction without any density assumption.

Proposition 3.1 (Exact binary channel). *Suppose that, on one trace, L disjoint pairs of labels satisfy*

$$Q_d(n_{i,1}) = a_i q_{i,1}, \quad Q_d(n_{i,2}) = a_i q_{i,2} \quad (1 \leq i \leq L),$$

where all $2L$ reserved primes $q_{i,j}$ are distinct. Let $C : \mathbb{F}_2^L \rightarrow \mathbb{F}_2^L$ be an injective linear code of minimum distance at least D . Set

$$h(q_{i,1}) = (-1)^{C(x)_i}, \quad h(q_{i,2}) = 1.$$

All unreserved prime signs may be arbitrary. Then the pair products recover $C(x)$ exactly. A decoder correcting binary errors and erasures whenever $2e + f < D$ recovers x from the

readings whenever $2t + s < D$, where t is the number of erroneous readings and s the number of erased readings.

Proof. Complete multiplicativity and the sign-valued hypothesis give

$$h(Q_d(n_{i,1}))h(Q_d(n_{i,2})) = h(a_i)^2 h(q_{i,1})h(q_{i,2}) = (-1)^{C(x)_i}. \quad (3.1)$$

Even if a cofactor contains a reserved prime from another pair, its sign is squared and disappears. No compatibility requirement on the cofactor signs is needed.

Declare a pair erased if either reading is erased. Otherwise multiply the two received signs and convert $+1, -1$ to $0, 1$. A wrong resulting bit requires at least one erroneous reading in that pair. The disjointness of the labels gives $e \leq t$ and $f \leq s$. Two sign errors in one pair cancel rather than cause two output errors. Hence $2e + f \leq 2t + s < D$, and the binary decoder applies. $\square$

The proposition asks that reserved signs remain fixed; it does not infer their values from arbitrary data. Its use is to implement an encoding which the sender has already chosen. It is not a reconstruction identity for an unknown prime sign outside this encoding.

3.1 Shared primes are allocated, not duplicated

At a scale X , start from candidate observations in $(X/2, X]$ with prime factor greater than $R_\gamma(X)$. Delete every prime reserved at a previous accepted scale and, for the causal schedule of Section 5, every prime rejected by its first-occurrence filter. Let $\mathcal{P}_d$ be the remaining prime set for trace d , and put $m_d = |\mathcal{P}_d|$. By lemma 2.1, there is exactly one candidate label per member of $\mathcal{P}_d$ on that trace.

Lemma 3.2 (Simultaneous allocation). *The primes in $\bigcup_d \mathcal{P}_d$ can be assigned to incident traces, one trace per prime, so that trace d receives at least $\lfloor m_d/3 \rfloor$ primes. Such an assignment is found by an integral flow computation.*

Proof. For every subset $T \subseteq \mathcal{D}$, a prime belongs to at most three of its incident prime sets, so

$$\left| \bigcup_{d \in T} \mathcal{P}_d \right| \geq \frac{1}{3} \sum_{d \in T} m_d \geq \sum_{d \in T} \lfloor m_d/3 \rfloor.$$

These are precisely the Hall inequalities for matching the prescribed number of copies of each trace to distinct primes. An integral maximum flow supplies the matching; unused primes can be assigned to an incident trace by any fixed rule. The network may group primes according to the seven nonempty subsets of traces in which they occur, and then distribute the actual primes in numerical order. $\square$

This allocation sacrifices at most a fixed fraction of the available resources. It makes the subsequent encodings on the three traces compatible with one global prime assignment.

3.2 The cofactor count pays for pairing

Suppose trace d receives u_d primes. Write its associated readings as $Q_d(n) = aq$. Sort them by (a, n) and pair consecutive readings within each exact cofactor class. Each class discards at most one reading. Consequently the number P_d of available pairs satisfies

$$P_d \geq \frac{u_d - A_\gamma(X)}{2}. \quad (3.2)$$

The paired readings and their reserved primes are all distinct, both within the trace and across the allocation to different traces. Let $P = \min_d P_d$. We can use the same length- L binary code, for any $L \leq P$, on each of the three traces. Each trace then gives its own complete copy of the message; one does not add three separate distance estimates to obtain a decoding radius.

The decisive arithmetic comparison is

$$\text{available readings} \asymp X, \quad \text{possible exact cofactors} = O(X^{2-\gamma}) = o(X).$$

It yields linearly many pairs without requiring infinitely many primes in any specified cofactor class. In particular, no prime-values conjecture for an individual polynomial $(n^2 - d)/a$ is used.

4 A constructive binary code and its decoder

After [proposition 3.1](#), the arithmetic has produced P available binary coordinates. An existence theorem for a good code would not yet provide a polynomial construction. We use a short inner code which can be built by exact conditional expectations, followed by a Reed–Solomon outer code. The underlying coding methods are classical [\[24, 7, 11, 12\]](#); the details below fix the complexity parameter and the errors-and-erasures guarantee.

4.1 Constructing the inner code

Lemma 4.1 (A small inner alphabet). *For each $m \geq 2$, a deterministic procedure constructs an m -dimensional binary linear code of length $8m$ and distance $d_{\text{in}} \geq m + 1$. Its generator matrix and exact distance can be obtained in a number of bit operations polynomial in 2^m and m .*

Proof. Choose the columns of an $m \times 8m$ generator matrix successively. If the columns chosen so far give the nonzero message $z \in \mathbb{F}_2^m$ a word of weight w_z , put

$$U = \sum_{z \neq 0} 2^{-w_z}.$$

For a uniformly distributed new column $v \in \mathbb{F}_2^m$, the bit $\langle z, v \rangle$ is fair for every fixed $z \neq 0$. Thus the mean new potential is $(3/4)U$. Choose the first numerical column minimising that potential. After $8m$ columns,

$$U \leq (2^m - 1)(3/4)^{8m} < 2^{-m}, \tag{4.1}$$

because $(3/4)^8 < 1/4$. Each individual term is smaller than 2^{-m} , so every nonzero word has weight greater than m . This also proves injectivity.

All comparisons are exact: multiplying by 2^{8m} makes every potential an integer. Direct enumeration of the 2^m possible columns and $2^m - 1$ nonzero messages at each step is already polynomial in 2^m . A Walsh transform evaluates all column costs in $O(m2^m)$ integer operations per step, hence $O(m^2 2^m)$ over the construction. The integers have $O(m)$ bits. Finally, enumerating the 2^m inner words gives their exact minimum nonzero weight. $\square$

The enumeration is over a short inner alphabet, not over all messages transmitted by the final code. This distinction is what makes the construction useful at large observation length.

4.2 The outer code and the rate

For $P \geq 48$, take the least $m \geq 2$ for which

$$n = \left\lfloor \frac{P}{8m} \right\rfloor \leq 2^m - 1, \quad k = \min \left\{ \lfloor n/2 \rfloor, \left\lfloor \frac{X}{64m} \right\rfloor \right\}. \quad (4.2)$$

In our application $P \leq X/4$ and $X \geq 4096$, so $k \geq 1$. Use the field $\mathbb{F}_{2^m}$ in a fixed polynomial basis, choosing the first monic irreducible polynomial in numerical order. Existence of such a polynomial is the usual finite-field construction; trial irreducibility tests and enumeration are polynomial in 2^m . Evaluate each polynomial of degree less than k at n distinct nonzero field elements. A nonzero polynomial of degree less than k has at most $k - 1$ roots. The resulting Reed–Solomon code has distance

$$d_{\text{out}} = n - k + 1.$$

Replace each of its symbols, represented by m bits, by the inner codeword of [lemma 4.1](#). The binary parameters are

$$L = 8mn \leq P, \quad b = mk, \quad D = d_{\text{in}}(n - k + 1). \quad (4.3)$$

Here D is a guaranteed lower bound for the binary minimum distance, not a claim that the minimum of the concatenated code has been determined exactly.

Lemma 4.2 (Linear parameters at a linear supply of pairs). *If $P \geq \beta X$ for a fixed $\beta > 0$ and $P \leq X/4$, then, for all sufficiently large X ,*

$$m = O(\log X), \quad 2^m = O(P), \quad L \geq P/2, \quad b \geq P/32, \quad D \geq P/32.$$

The implied constants and the threshold may depend on β .

Proof. Minimality in (4.2) gives $m = O(\log P)$ and $2^m = O(P)$ (the finitely many cases $m = 2$ can be absorbed). Also $P - 8m < L \leq P$. Since

$$b \geq \min\{L/16 - m, X/64 - m\}$$

and $X \geq 4P$, the asserted lower bound for b follows when $m = o(P)$. For the distance, $d_{\text{in}} \geq m + 1$ and $n - k + 1 \geq n/2$ give $D \geq mn/2 = L/16$. These estimates prove the displayed bounds. $\square$

All tables needed for this code are polynomial in X . In particular, the possibility of enumerating $2^m = O(P)$ inner words does not permit, or require, enumerating the 2^b full messages.

4.3 Reliability scores for inner decoding

An uncertain inner block should not automatically become a hard outer error: an incorrect symbol costs two in the outer correction budget, whereas an erasure costs one. Forney’s generalised minimum-distance method [7, 12] uses reliability scores to decide which blocks to erase. The twenty-block example in Section 6.2 shows why this distinction can recover a message that hard decoding misses.

In inner block i , let s_i be the number of erased binary coordinates. Choose an inner codeword $\hat{c}_i$ closest to the known received bits, breaking ties numerically, and let e_i be that distance. With $d = d_{\text{in}}$, define

$$\alpha_i = \max\{0, 1 - (2e_i + s_i)/d\}.$$

Let c_i be the true inner codeword, and let t_i count its errors on the known positions. Then

$$\begin{cases} 1 - \alpha_i, & \hat{c}_i = c_i, \\ 1 + \alpha_i, & \hat{c}_i \neq c_i \end{cases} \leq \frac{2t_i + s_i}{d}. \quad (4.4)$$

If the chosen word is correct, $e_i = t_i$ proves this immediately. If it is wrong, the two distinct inner words differ in at least d coordinates, so $t_i + e_i + s_i \geq d$, while minimality gives $e_i \leq t_i$. When $2e_i + s_i < d$, these inequalities yield

$$1 + \alpha_i = 2 - \frac{2e_i + s_i}{d} \leq \frac{2t_i + s_i}{d}.$$

When $2e_i + s_i \geq d$, we instead have $\alpha_i = 0$ and $2t_i + s_i \geq d$. Thus (4.4) covers errors and erasures simultaneously.

4.4 Why some threshold decodes

For $u \in [0, 1]$, erase every outer symbol whose reliability satisfies $\alpha_i \leq u$. A wrong remaining symbol costs two in the outer errors-and-erasures budget, and an erased symbol costs one. Integrating this cost over u gives $1 - \alpha_i$ for a correct inner choice and $1 + \alpha_i$ for a wrong one. Hence (4.4) implies

$$\int_0^1 (2e_{\text{out}}(u) + s_{\text{out}}(u)) du \leq \frac{2t + s}{d_{\text{in}}}.$$

If $2t + s < d_{\text{in}}d_{\text{out}}$, at least one threshold has outer cost less than d_{out} . Only finitely many sets of erasures occur. The nonzero reliabilities are integer multiples of $1/d_{\text{in}}$, and $d_{\text{in}} \leq 8m$, so $O(m)$ trials suffice. The comparisons are rational; no approximate logarithms are involved.

For completeness, the outer decoder can be written as a linear system. Delete the erased symbols. If n' evaluations remain and $n' \geq k$, set $T = \lfloor (n' - k)/2 \rfloor$ and solve

$$Q(a_i) = y_i E(a_i), \quad \deg Q < k + T, \quad E \text{ monic of degree } T. \quad (4.5)$$

If at most T symbols are wrong, their error locator, multiplied by a monic polynomial to reach degree T , supplies a solution with $Q = PE$, where P is the sent polynomial. Conversely, every solution has $Q - PE$ vanishing at at least $n' - T \geq k + T$ distinct points, more than its degree. Therefore $Q = PE$. Divide, check that the quotient has degree less than k , and re-evaluate it to verify its distance. If these tests fail, reject this trial. This is the Welch–Berlekamp formulation described in [12].

Proposition 4.3 (Polynomial decoding of the concatenation). *The code (4.3) has a deterministic decoder that corrects every binary error-and-erasure pattern satisfying $2t + s < D$, in time polynomial in P and X .*

Proof. Perform the inner searches and all reliability-threshold trials. At least one trial succeeds by the integral argument. For each candidate, reconstruct the concatenated binary word and check $2t' + s < D$ against the actual received bits. There is at most one such candidate: two distinct codewords would have distance at most $t'_1 + t'_2 + s < D$, contradicting their distance lower bound. Inner searches enumerate $2^m = O(P)$ words per block; the number of thresholds is $O(m)$; and the outer systems have polynomial size and field-operation cost. The construction and the decoder are thus polynomial in the numerical input scale. $\square$

Combining proposition 4.3 with proposition 3.1 proves a finite arithmetic error-correction statement. It remains to place these codes at compatible scales and to control the gaps between those scales.

5 One infinite message and every sufficiently large scale

A separate finite code at each X is not yet an infinite multiplicative encoding. The same prime might be requested twice, and a large gap between successful scales could destroy the prefix rate. A third issue is message dependence: a newly reserved prime can already occur on another trace at a much earlier index. A first-occurrence filter prevents such a prime from introducing too many message bits there. The finite admission test guarantees consistency, and (GF_γ) guarantees that the filtered test eventually succeeds frequently enough.

5.1 An admission test using only finite data

Fix $X_0 \geq 4096$ and try the integer scales $X_0, 2X_0, 4X_0, \dots$. A rejected scale reserves nothing. At a candidate scale X , let

$$M = \min_d M_d^\gamma(X), \quad Q = \#\{\text{primes already reserved}\}.$$

A first-occurrence threshold must be large enough to limit the message prefix carried by an early reading, but small enough to leave a fixed fraction of the candidate primes available. We use the initial population M to set $s_X = \lceil M/24 \rceil$; it is not recomputed after deletions. Delete old reserved primes and all candidate primes with $\nu(q) < s_X$, where ν is defined in (2.8). Allocate the remaining primes by lemma 3.2, and pair their readings by exact cofactors. Let P be the minimum pair count on the three traces. Accept precisely when

$$M \geq \max\{96, 2Q\}, \quad P \geq \max\{48, \lceil M/24 \rceil\}. \quad (5.1)$$

Use (4.2) to determine L, b, D . On L pairs of each trace encode the first b bits of the same infinite message x . Exactly $6L$ distinct primes are now reserved. All later scales exclude these primes, and every previously chosen sign remains fixed.

This definition uses neither the constant a nor the threshold X_a in (GF_γ) . Even without that hypothesis it is a well-defined algorithm, possibly with only finitely many accepted scales. The hypothesis will prove that this possibility does not occur. One can omit the first-occurrence filter if only cancellation and positive prefix rate are sought; the code and the admission inequalities are otherwise identical. We use the filtered construction for Ω throughout the theorem.

Lemma 5.1 (Eventually bounded gaps). *Under (GF_γ) , infinitely many scales are accepted. Writing their sizes as X_j , there is a constant C such that $X_{j+1} \leq CX_j$ for all sufficiently large j . One may take $C = 24/a$ after decreasing a to at most one.*

Proof. At scale X_i at most $3X_i$ primes are reserved, a deliberately generous bound. Accepted sizes at least double, so after X_j the number already reserved satisfies

$$Q \leq 3 \sum_{i \leq j} X_i \leq 6X_j.$$

At any sufficiently large candidate $X \geq 12X_j/a$, we have $M \geq aX \geq 2Q$. Deleting old primes costs at most Q candidates on each trace. The first-occurrence filter deletes at most $3s_X \leq M/8 + 3$ further primes, by lemma 2.3. Since $M \geq 2Q$, each trace retains at least

$$M - Q - 3s_X \geq 3M/8 - 3$$

candidates. Allocation gives at least $M/8 - 2$ primes per trace, and pairing therefore gives

$$P \geq M/16 - A_\gamma(X)/2 - 1. \quad (5.2)$$

Since $A_\gamma(X) = o(X)$ and $M \geq aX$, this eventually exceeds both 48 and $\lceil M/24 \rceil$. The next dyadic candidate after $12X_j/a$ is at most $24X_j/a$. Thus there must be an accepted scale within that bound. If a last accepted scale existed, the same argument with fixed Q would give a later one, a contradiction. The case of no accepted scale is treated by $Q = 0$. $\square$

The finite resource count P , rather than a presumed distribution of prime factors, controls acceptance. The analytic hypothesis is consumed only in the proof that rejected candidates cannot create unbounded relative gaps at large scales.

5.2 Rate and resistance to moving small signs

At every sufficiently large accepted scale, (5.1) and lemma 4.2 imply the conservative bounds

$$b_X \geq aX/1536, \quad D_X \geq aX/1536. \quad (5.3)$$

For a large observation length N , take the last accepted $X \leq N$. By lemma 5.1, $N < CX$ eventually. All reserved primes of this block satisfy $q > X^\gamma$, so

$$q > X^\gamma \geq (N/C)^\gamma > N^\vartheta \quad (5.4)$$

for all sufficiently large N , because $\gamma > \vartheta$ is fixed. The first b_X bits are consequently recoverable even after arbitrary changes of all prime signs at most N^ϑ . Cofactor signs need not be preserved; they cancel by (3.1).

Equations (5.3)–(5.4) and the finite decoder prove the recovery assertion of theorem 1.1, with, for instance, any fixed

$$0 < r, c < \frac{a}{1536C}.$$

The threshold includes the constants in (GF_γ) , the starting scale, the gap bound, and the strict comparison in (5.4). It is independent of the message, the prime reassignment and the corruption pattern. The actual block parameters used by the algorithm are computed from finite data, not from r or c .

5.3 The global map and its support

For every prime reserved in a pair, its sign is either a linear function of a finite prefix of x or the constant 1. Off the reserved set E , set the sign to 1 and extend completely multiplicatively. This defines Ω_x and gives

$$\Omega_{x+y} = \Omega_x \Omega_y.$$

The values b_X tend to infinity along accepted scales by (5.3). Therefore every nonzero message prefix is eventually encoded by an injective block code, proving injectivity of $x \mapsto \Omega_x$.

For any $\sigma > 0$ with $\gamma\sigma > 1$, the support satisfies

$$\sum_{q \in E} q^{-\sigma} < 3 \sum_{j \geq 0} (2^j X_0)^{1-\gamma\sigma} = \frac{3X_0^{1-\gamma\sigma}}{1-2^{1-\gamma\sigma}}. \quad (5.5)$$

Increasing X_0 makes this smaller than a prescribed positive budget and places every reserved prime above B . For $\sigma = 1$ this gives the support assertion of theorem 1.1; the same construction permits the stronger budget $1/\gamma < \sigma \leq 1$.

To prove continuity in d_σ , choose finitely many reserved primes whose complement has total weight less than η . Their signs depend on a common finite message prefix. Two messages agreeing on that prefix have d_σ -distance less than η . The image is thus compact and, by injectivity, homeomorphic to the Cantor group $\mathbb{F}_2^{\mathbb{N}_{\geq 1}}$. Formula (1.6) follows because every integer disagreement has a divisor in the set of disagreeing primes.

5.4 What the algorithm must compute

Membership of a prime q in E is decidable: only candidate scales with $X < q^{1/\gamma}$ can reserve it. Simulating those finitely many scales decides the question. At each scale, trial division or a sieve factors the integers of size at most $X^2 + 2$; the flow, grouping, matrix construction and decoding all have polynomial cost in X . The number of tested scales before N is logarithmic. This includes the cost of preparing the tables used by the decoder.

The filter can be computed within the same polynomial budget: inspect the factorizations at indices below s_X and remove every candidate prime that occurs there. For these candidates this is exactly the ν -test, by Section 2.4. Only the pair products are used by the decoder, and each depends exactly on the block's encoded prefix, not on its suffix.

An unselected raw reading can still contain a prime reserved at a later epoch. To compute all raw readings up to N , determine assignments at primes at most $N^2 + 2$ by simulating epochs only to $O(N^{2/\gamma})$. The work is polynomial in N and depends solely on public arithmetic data until the message bits are queried. The following proposition bounds those queries even when a relevant prime is reserved at a future epoch.

Proposition 5.2 (Causal raw dependence of the paired code). *For the filtered schedule, every raw encoded reading $\Omega_x(Q_d(n))$ depends only on $x_1, \dots, x_{\lfloor 3n/4 \rfloor}$. This property holds without (GF_γ) , even if the schedule has only finitely many accepted epochs.*

Proof. At epoch X , choose a trace attaining the original minimum M . It has only M candidates before any deletions or allocation, so the minimum pair count satisfies $P \leq M/2$. The code parameters (4.2)–(4.3) give

$$b_X = mk \leq \frac{mn}{2} = \frac{L}{16} \leq \frac{P}{16} \leq \frac{M}{32} \leq \frac{3}{4}s_X. \quad (5.6)$$

Here n in mn is the outer block length. Every prime q reserved at this epoch passes the filter, hence $\nu(q) \geq s_X$. Its sign depends on at most $b_X \leq 3\nu(q)/4$ message bits. If it divides any observation $Q_d(j)$ with $j \geq 2$, lemma 2.3 gives $\nu(q) \leq j$. Every reserved factor of that observation therefore consults only the first $\lfloor 3j/4 \rfloor$ bits. All unreserved prime signs are constant. At index $j = 1$, the positive values are 2 and 3, neither of which is reserved because $X_0 \geq 4096$. These readings are constant as well. $\square$

This completes the causal clause of theorem 1.1. Causality is a dependence statement for the encoder before reassignment; an adaptive adversary need not preserve it. The polynomial table construction can involve arithmetic anticipation, so no real-time or bounded-delay algorithm follows. Nor is the complexity polynomial in $\log N$ or asserted to be near-linear.

A known reference f_0 may be included by using $g_x = f_0\Omega_x$. On each pair the receiver divides out the known sign $f_0(q_{i,1})f_0(q_{i,2})$ before binary decoding. The ordinary proximity and the pairwise separation are unchanged. An arbitrary unknown reference cannot be removed in this way and is not part of the decoding promise.

6 A finite code with 273 message bits

The mass hypothesis is unnecessary for checking a specific set of pairs. This section gives the finite instance supplied with the article. Its arithmetic identities, prime assignments and decoding parameters are directly inspectable. The infinite extension is a separate rule with a separately paid support cost.

6.1 Two compatible epochs

Take $\gamma = 7/6$ and try the four candidate scales below. The table uses the admission test (5.1); Q is the number of primes reserved before the candidate.

X	$R_\gamma(X)$	Q	$\min_d M_d^\gamma(X)$	Decision
4096	16384	0	1177	accept
8192	36780	3120	2374	reject
16384	82570	3120	4706	reject
32768	185363	3120	9348	accept

The filter uses $s_{4096} = 50$ and $s_{32768} = 390$ at the two accepted epochs. Since

$$50^2 + 2 < 16384, \quad 390^2 + 2 < 185363,$$

no prime above the corresponding cutoff divides a value at an index below s_X . Thus the first-occurrence filter removes no candidate at either accepted epoch. The saved pair data and code parameters are also the data of the filtered schedule. The rejected candidates do not reserve any additional primes. At the second accepted epoch, the same first message bits are encoded again, together with a longer prefix. No earlier prime assignment changes.

Parameter, common to the three traces	$X = 4096$	$X = 32768$
Available pairs on the least supplied trace P	535	4382
Pairs used per trace L	520	4368
Readings used per trace $2L$	1040	8736
Field degree m	5	7
Outer length n	13	78
Outer dimension k	6	39
Exact inner distance d_{in}	19	25
Outer distance $n - k + 1$	8	40
Message bits $b = mk$	30	273
Guaranteed binary distance D	152	1000
Guaranteed reading errors, no erasures	75	499

There are $3(520 + 4368) = 14664$ pairs and 29328 distinct reserved primes across both epochs and all three traces. The prime factors are not declared private to every other unselected observation. What matters is that none is reused as a reserved prime in another selected pair.

Theorem 6.1 (Finite arithmetic transmission). *The supplied integer pairs realise a common 273-bit message using 8736 labelled readings of each trace, at indices at most 32768. Each trace alone determines the message despite arbitrary signs at every unreserved prime and every reading corruption satisfying*

$$2t + s < 1000.$$

In particular, every pattern of at most 499 reading errors is corrected. The same assignment also realises its first 30 message bits at the earlier epoch, with guarantee $2t + s < 152$.

Proof. The supplied pair identities have the form in [proposition 3.1](#), with distinct prime factors and no reused labels within a trace. The verified inner distances are 19 and 25,

obtained by enumerating respectively 32 and 128 inner words. The Reed–Solomon root bound gives outer distances 8 and 40. Thus (4.3) gives $D = 152$ and $D = 1000$. Apply [propositions 3.1](#) and [4.3](#). The first epoch uses the first 30 of the same 273 message bits, and its reserved primes are disjoint from those of the second epoch. $\square$

The proof of this theorem uses the finite identities and the designed code distances. It is not a calibration of the constants in $(\mathbf{GF}_\gamma)$, and 499 is not claimed to be the optimal correction radius of this binary code. In particular, no ambiguous pair of messages at 500 errors is asserted.

6.2 Decoding examples

The supplied verifier checks the pair identities, reserved-prime assignments and inner distances used in [theorem 6.1](#). [Appendix B](#) identifies the data and commands; the execution record gives the file comparisons. Here we examine what three decoding examples demonstrate.

The second epoch uses $\mathbb{F}_{128}$ defined by $T^7 + T + 1$, with 78 inner blocks of length 56 and 39 outer coefficients. The decoder compares $78 \cdot 128 = 9984$ short inner words, not 2^{273} full messages. The following tests are included in the finite data.

Test	Errors	Erasures	Recovered bits
Guaranteed error radius	499	0	273
Reliability threshold needed	260	0	273
Mixed strict budget	200	599	273

In the 260-error test, thirteen bits are flipped in each of twenty inner blocks. Hard decisions give twenty wrong outer symbols, exceeding the hard outer radius 19. A subsequent reliability threshold erases these twenty blocks and succeeds: twenty wrong symbols cost 40 in the outer budget, while twenty erasures cost only 20. Thus the example exercises the generalised minimum-distance step.

In the reading-level mixed example, 200 signs are flipped and 599 individual readings are erased, on distinct pairs, before products are formed. All 273 bits are recovered from each trace. The selected readings are values of one common completely multiplicative source, including any reserved factors of a cofactor. The checks also use a second message and three different small-prime assignments. Uniformity over all assignments comes from (3.1), not from these choices; the guarantee against every 499-error pattern comes from [proposition 4.3](#) and the distance lower bound, not their enumeration.

The standalone decoder receives only the pair table, the code and the labelled received signs, not the expected message. The arithmetic verifier follows a different checking path from the constructor. All these tests nevertheless reuse the same decoding module, rather than an independently implemented decoder.

6.3 Paying for an uncomputed extension

The least reserved prime in the pilot is 16547, so every encoded source equals 1 on $[1, 16000]$. An upward rational bound for the pilot's harmonic cost is

$$\sum_{q \in E_{\text{pilot}}} \frac{1}{q} \leq \frac{27697555482416339}{10^{18}}.$$

To append an infinite schedule without changing the pilot, resume the filtered admission test at

$$X_* = 1920^6 = 50096498540544000000,$$

and then at $2X_*, 4X_*, \dots$, retaining all old prime reservations. Since $(10/9)^6 < 2$, we have $1 - 2^{-1/6} > 1/10$. Formula (5.5) bounds the future cost by

$$\frac{3X_*^{-1/6}}{1 - 2^{-1/6}} < \frac{30}{1920} = \frac{1}{64}.$$

Consequently the complete possible support, including the uncomputed continuation, has

$$\sum_{q \in E} \frac{1}{q} < \frac{27697555482416339}{10^{18}} + \frac{1}{64} = 0.043322555482416339 < \frac{1}{20}. \quad (6.1)$$

The bound holds whether or not later candidates succeed. Infinitely many successes and positive rate at all sufficiently large N require (GF_γ) . This pilot uses $\gamma = 7/6$ and therefore supports the asymptotic resistance range $1 < \vartheta < 7/6$, under that input. The general theorem uses a different chosen γ when ϑ is larger. No future epoch at or beyond X_* was executed for the finite instance. The long gap also prevents interpolating the pilot's numerical rate across intermediate scales.

7 Persistent prescriptions before the small signs are known

Encoding a message is one use of a large-prime assignment. Another is to force both signs to remain present, irrespective of how the small signs are later chosen. A finite balancing rule must now be made consistent over every observation window and every scale. Exact cofactors again supply the order needed for that consistency.

7.1 One trace: exact alternation

Fix $1 < \vartheta < 2$ and retain the occurrences

$$n \geq L, \quad n^2 + 1 = aq, \quad q > R_\vartheta(n).$$

On this one trace, a prime has only one retained occurrence in the entire sequence. For each exact cofactor a , list the retained indices increasingly as $n_{a,1} < n_{a,2} < \dots$, and prescribe

$$\tau(q_{a,j}) = (-1)^j, \quad q_{a,j} = (n_{a,j}^2 + 1)/a. \quad (7.1)$$

The list may be finite. Distinct lists have disjoint primes, so this defines a single table.

For $0 \leq u \leq N$, take the retained indices in $(u, N]$ whose prime exceeds $R_\vartheta(N)$. In each cofactor list these are a consecutive interval: both $n_{a,j}$ and $q_{a,j}$ increase. If B_a is the sum of their prescribed signs, then $|B_a| \leq 1$. Writing M for their total number and $\Lambda = \sum_a |B_a|$, we obtain

$$\Lambda \leq \left\lfloor \frac{N^2 + 1}{R_\vartheta(N) + 1} \right\rfloor, \quad \min_{v=\pm 1} \#\{u < n \leq N : h(n^2 + 1) = v\} \geq \frac{M - \Lambda}{2} \quad (7.2)$$

for every source taking the prescribed prime signs at these occurrences. Indeed, their signed sum is $\sum_a B_a h(a)$, whose absolute value is at most Λ . The actual $M - \Lambda$ is nonnegative and even, by the parity of each class sum.

For $\vartheta\sigma > 1$ the support costs at most $(L - 1)^{1-\vartheta\sigma}/(\vartheta\sigma - 1)$. Membership and the parity of the rank j in (7.1) are decidable by finite arithmetic searches. This constructs the same prescription at all sizes, rather than a new optimised table for each N .

Corollary 7.1 (Single-trace range). *Assume the weighted input (2.2) for $d = -1$ at $u = 13/10$. For rational parameters $1 < \vartheta < 13/10$ and $1/\vartheta < \sigma \leq 1$, there is a computable prescription of arbitrarily small σ -cost, above any prescribed prime prefix, such that every sufficiently large interval $(N/2, N]$ contains at least $c_\vartheta N$ occurrences of each sign after any reassignment of the prime signs up to $R_\vartheta(N)$.*

Proof. The single-trace version of lemma 2.2, using an intermediate exponent between ϑ and $13/10$, gives $M \gg N$ on this interval. The defect in (7.2) is $O(N^{2-\vartheta}) = o(N)$. Increasing L pays the desired support cost and protects the prefix, without changing the positive asymptotic density constant. $\square$

The weighted provenance is Pascadi's calculation specified in Section 2.2. This wider range is a one-trace statement. It does not prescribe the other two traces, nor follow from the three-trace input at $\gamma < 6/5$.

7.2 Why three independent alternations fail

A prime shared between traces cannot be given separate signs. This obstruction already occurs in the following lists, valid for cutoff exponent $7/6$:

(d, a)	n	q
$(-2, 3)$	7, 11, 17	17, 41, 97
$(2, 2)$	6, 8, 12, 14	17, 31, 71, 97.

The first alternation would require $\tau(17) = \tau(97)$; the second would require $\tau(17) = -\tau(97)$. No initial phase choice fixes their five-edge cycle. Instead of perfect alternation, we permit a small discrepancy in every list interval.

Lemma 7.2 (Ordered-list rounding). *Let a finite family of ordered lists have m distinct elements, with no element repeated in a list and each element belonging to at most t lists. A deterministic rational procedure chooses signs such that every consecutive interval J of every list satisfies*

$$\left| \sum_{q \in J} \tau(q) \right| \leq 8t(1 + \lceil \log_2(m+1) \rceil).$$

Proof. Begin with values $x_q = 0$ in $[-1, 1]$. A coordinate reaching either endpoint remains fixed. At the beginning of a phase let u coordinates be free. Filter each list to these coordinates, and partition it into consecutive groups of $2t$, discarding the last incomplete group. Preserve each complete group's sum throughout the phase. The number a of equations satisfies $2ta \leq tu$, hence $a \leq \lfloor u/2 \rfloor$.

As long as more than $\lfloor u/2 \rfloor$ coordinates are free, rational elimination finds a nonzero direction preserving all the group sums. Move along it until a coordinate reaches the boundary. Eventually at least half of the phase's free coordinates have become fixed. For any original list interval, its intersection with the free list consists of whole groups and at most two partial boundary groups. The whole-group sums are preserved. There are fewer than $4t$ boundary coordinates, each changing by less than two, so its sum changes by less than $8t$ in this phase. After at most $1 + \lceil \log_2(m+1) \rceil$ phases all coordinates are fixed. Summing the phase defects proves the bound. $\square$

This is a floating-variable discrepancy argument of Beck–Fiala type [1, 18], not a new general discrepancy theorem. For the arithmetic application, the sharper comparison is with

discrepancy of permutations. Bohus [3] gives a logarithmic bound and an efficient colouring algorithm for a fixed number of orders. Newman–Nikolov [17] give three permutations with discrepancy of order at least $\log m$. That lower bound concerns general permutations; it proves no optimality for the particular arithmetic lists used here.

7.3 Fixing the blocks before the windows

Use the common support E_L of Section 2.3. For each (d, a) , order its admissible primes by magnitude. They are also ordered by their corresponding indices. Divide the prime axis once and for all into blocks

$$(A_j, A_{j+1}], \quad A_j = 2^{2^j}, \quad A_{j+1} = A_j^2.$$

On a finite prime block, the cofactor lists for one trace are disjoint. Concatenate them and append primes absent from that trace to obtain a permutation of the block's support. Repeating this for the three traces gives three permutations, and every original list interval is still an interval of its corresponding permutation. This is the precise reduction to the classical setting just cited. The arithmetic task is to fix these blocks irreversibly and retain their control when the observation window changes.

On every block, form all the restricted lists and apply lemma 7.2 with $t = 3$. Each block is finite; choosing deterministic rules in rational elimination defines a computable global table $\tau : E_L \rightarrow \{-1, 1\}$. No later calculation recolours an earlier block.

For $0 \leq u \leq N$, let $J_d(N, u)$ be the indices $u < n \leq N$, $n \geq L$, with prime factor $q_n > R_\vartheta(N)$, and write $Q_d(n) = q_n a_n$. Define

$$B_{d,a}(N, u) = \sum_{\substack{n \in J_d(N, u) \\ a_n = a}} \tau(q_n), \quad \Lambda_d(N, u) = \sum_a |B_{d,a}(N, u)|.$$

Put $V_N = N^2 + 2$ and $A(N) = \lfloor V_N / (R_\vartheta(N) + 1) \rfloor$. For $N \geq 5$, the interval of possible primes $(R_\vartheta(N), V_N]$ meets at most two prime blocks: $V_N < R_\vartheta(N)^2$. Each block it meets has upper endpoint at most V_N^2 . A restriction by $u < n \leq N$ and by the prime cutoff is still a list interval. Thus lemma 7.2 gives

$$\Lambda_d(N, u) \leq A(N) 96(1 + \lceil \log_2 V_N \rceil) = O_\vartheta(N^{2-\vartheta} \log N). \quad (7.3)$$

For every source with the prescribed signs at these primes,

$$\begin{aligned} \sum_{n \in J_d(N, u)} h(Q_d(n)) &= \sum_a B_{d,a}(N, u) h(a), \\ \min_{v=\pm 1} \#\{u < n \leq N : Q_d(n) > 0, h(Q_d(n)) = v\} &\geq \frac{|J_d(N, u)| - \Lambda_d(N, u)}{2}. \end{aligned} \quad (7.4)$$

The unselected observations can only increase the two sign counts. The bound uses the actual Λ_d ; replacing it by (7.3) may give a negative, hence uninformative, lower bound at small sizes.

Theorem 7.3 (Persistent simultaneous prescription). *Fix rational $1 < \vartheta < \gamma < 6/5$, $1/\vartheta < \sigma \leq 1$, $\varepsilon > 0$ and $B \geq 1$, and assume (GF_γ) . There is a decidable prime set $E \subset (B, \infty)$ with $\sum_{q \in E} q^{-\sigma} < \varepsilon$, and a computable sign table τ fixed independently of the input source, with the following property. For every sufficiently large N , every source h satisfying*

$$h(q) = \tau(q) \quad (q \in E, q > R_\vartheta(N))$$

has at least cN occurrences of each sign on $(N/2, N]$ on each of the three traces. The positive constant c is independent of the input and of the reassigned signs.

Proof. Choose L by (2.7). For large N , all indices in $(N/2, N]$ exceed L , and (GF_γ) supplies at least aN primes above $R_\gamma(N) > R_\vartheta(N)$. Hence $|J_d(N, N/2)| \geq aN$. Apply (7.4), using the $o(N)$ defect in (7.3), and take $c = a/4$ after increasing the threshold. $\square$

For an arbitrary input f_0 , overwrite its signs on E by τ and keep all others. This operation protects $[1, B]$, is continuous in the product topology of prime signs, is idempotent, and is computable relative to f_0 . It is a prescription operation, not multiplication by a fixed factor independent of f_0 . All but fewer than εT integers in each ordinary prefix $[1, T]$ retain their signs.

The same finite estimate also clarifies its limitation. For an unchanged source h , let U_d count the relevant primes at which $h(q) \neq \tau(q)$. Each such prime changes at most one selected reading on trace d , so

$$\min_{v=\pm 1} \#\{u < n \leq N : Q_d(n) > 0, h(Q_d(n)) = v\} + U_d \geq \frac{|J_d(N, u)| - \Lambda_d(N, u)}{2}.$$

This is a necessary test on an arbitrary source. It does not prohibit that source from disagreeing with a linear number of prescribed prime signs: the entire prescribed support itself has small harmonic cost.

8 Uniform signs on the active coordinates of words

Balancing each sign count does not force complicated words: an alternating sequence has balanced signs and zero entropy. To control words, we retain coordinates having a sufficiently large prime factor and mask the rest. The reference law keeps the actual arithmetic mask, rather than replacing its positions by independent random choices.

8.1 The mask and the collision bound

Fix $1 < \vartheta < 2$ and the support E_L of Section 2.3. A finite set of additional primes and finitely many prescribed signs can also be allowed; their support cost is paid separately. Write the resulting support as E . At the N origins $n = 2, \dots, N + 1$, consider words of length H on each trace:

$$W_h^H(n) = (h(Q_d(n + a)))_{0 \leq a < H, d \in \mathcal{D}} \in \{-1, 1\}^{3H}.$$

Let $\mu_{N,h}^H$ denote their empirical law and $\mu_{N,h}^{H,d}$ its projection to trace d . The shift of the starting index from 1 to 2 ensures positive polynomial values throughout.

Set

$$M = N + H, \quad V = M^2 + 2, \quad R = R_\vartheta(M), \quad A = \left\lfloor \frac{V}{R + 1} \right\rfloor, \quad m = 3H.$$

Coordinate $i = (a, d)$ is active at origin n when

$$Q_d(n + a) = q_i(n)a_i(n), \quad q_i(n) > R, \quad q_i(n) \in E.$$

Its cofactor is at most A . Let $S(n) \subseteq [m]$ be the set of active coordinates. It depends on the arithmetic, not on any sign choices.

Lemma 8.1 (Distinct active primes in a word). *If $H \leq N$ and $R \geq (H^2 + 4)^2$, the active primes within one word are all distinct. Across the N origins, one active prime affects at most $3H$ words.*

Proof. Two distinct coordinates correspond to distinct monic irreducible polynomials $T^2 - d$ and $(T + t)^2 - e$, with $|t| < H$. Their resultant is

$$t^4 - 2(d + e)t^2 + (d - e)^2 \neq 0. \quad (8.1)$$

Its absolute value is at most $(H^2 + 4)^2$. A prime dividing both specialised values would divide this nonzero resultant, impossible for an active prime exceeding R . A fixed prime has at most one active occurrence on each trace by [lemma 2.1](#); an occurrence can lie in at most H words. This proves the second assertion. $\square$

The resultant condition matters. The values $9^2 + 1 = 2 \cdot 41$, $11^2 + 2 = 3 \cdot 41$ and $17^2 - 2 = 7 \cdot 41$ can all depend on a single prime sign. Without a collision exclusion, three apparently separate signs would have only two possible joint states, not eight.

Let $\rho_{N,H}(S) = N^{-1} \#\{n : S(n) = S\}$. The target on the disjoint union of masked cubes is

$$U(\rho_{N,H})(S, w) = \rho_{N,H}(S) 2^{-|S|}, \quad w \in \{-1, 1\}^S. \quad (8.2)$$

Only the signs conditional on each mask are made uniform. We assume neither convergence of $\rho_{N,H}$ nor independence between its coordinates.

8.2 The cost of uniformity over all small assignments

Temporarily choose independent fair signs $\tau(q)$ on E . For an assignment κ at primes at most A , extend it multiplicatively to cofactors. For a mask S and $\emptyset \neq J \subseteq S$, define

$$F_{S,J,\kappa} = \frac{1}{N} \sum_{n=2}^{N+1} \mathbf{1}_{S(n)=S} \prod_{i \in J} \kappa(a_i(n)) \tau(q_i(n)).$$

Under [lemma 8.1](#), each summand has expectation zero. Cofactors contain no active prime, since $a_i \leq A < R$. Changing one prime sign changes at most $3H$ summands, each by at most two, so the normalised change is at most $6H/N$. There are at most $3M \leq 6N$ relevant independent prime signs.

The bounded-differences inequality, in the form

$$\Pr\{|F - \mathbb{E}F| > \eta\} \leq 2 \exp\left(-\frac{2\eta^2}{\sum_j c_j^2}\right),$$

for independent inputs and coordinate oscillations at most c_j , therefore gives

$$\Pr\{|F_{S,J,\kappa}| > \eta\} \leq 2 \exp\left(-\frac{N\eta^2}{108H^2}\right). \quad (8.3)$$

This is McDiarmid's bounded-differences inequality [\[15\]](#); the stated form is Theorem 2 of Raginsky's notes [\[23\]](#). Independence of overlapping words is not required.

There are at most $2^{\pi(A)}$ heads κ , and

$$\sum_{S \subseteq [m]} (2^{|S|} - 1) = 3^m - 2^m < 2^{6H}$$

mask-character tests. Although all small-prime assignments number $2^{\pi(A)}$, their cost in the exponent of the union bound is only $\pi(A)$. For $H \leq N$, the cofactor bound $A = O(N^{2-\vartheta})$ makes this cost sublinear; the mask-character tests add a cost proportional to H . This is

the room needed to choose one table before any of the small signs. To accommodate k_0 arbitrarily prescribed prime-sign bits, put $\ell_N = \lceil \log_2(N+1) \rceil$ and

$$\eta_{N,H,k_0}^2 = \frac{108H^2}{N}(\pi(A) + 7H + 4 + 4\ell_N + k_0). \quad (8.4)$$

Define the bad event $B_{N,H}$ to mean that one of the tests exceeds this threshold. If $H > N$, the collision condition fails, or $\eta \geq 1$, set $B_{N,H} = \emptyset$. The union bound and $e^{-u} \leq 2^{-u}$ imply

$$\Pr(B_{N,H}) \leq 2^{-k_0-H-3}(N+1)^{-4}. \quad (8.5)$$

The union is taken before selecting the table. It thus covers a head chosen after the entire table is known. Arbitrary unrelated weights at all N origins would have a much larger information cost and are not covered by this argument.

8.3 Selecting one computable table

An almost-sure existence statement alone does not supply a computable prescription. The method of conditional expectations has the broader constructive statistical context of [9]. More specifically, the selections in [22, Section 5, Proposition 5.1] and [21, Section 7] use exact rational costs, irreversible decisions, and positive tails controlled after conditioning. Their potentials are weighted sums of normalised nonnegative costs. Here we instead count bad events and keep the conditional expectation strictly below one. Once a finite collection of indicators is determined, its sum is an integer smaller than one and must vanish. We give this specialised argument in full, including its tail budget.

Lemma 8.2 (Effective elimination of all bad events). *Suppose the events $B_{N,H}$ depend on finitely many binary coordinates, can be decided by finite exact computation, and satisfy (8.5) under the fair product measure. After any prescribed values of k_0 coordinates, the remaining coordinates can be chosen computably so that no bad event occurs.*

Proof. Consider the nonnegative integer-valued potential

$$U = \sum_{N,H \geq 1} \mathbf{1}_{B_{N,H}}.$$

Its expectation is at most $2^{-k_0}/8$. Conditioning on any k_0 specified fair bits multiplies this upper bound by at most 2^{k_0} , leaving at most $1/8$.

For an integer $r \geq 1$, the sum of the probabilities outside the finite rectangle $H \leq r$, $N \leq 2^r$ is at most 2^{-k_0-r-2} . Indeed, the H -tail is geometric, and

$$\sum_{N > 2^r} (N+1)^{-4} \leq 2^{-3r}/3.$$

After s further bits are fixed, each branch for the next bit has unconditional probability $2^{-(k_0+s+1)}$. Positivity bounds its conditional omitted expectation by 2^{s-r-1} .

At step s , take $r = 2s + 6$. Compute exactly the two conditional expectations of the finite-rectangle sum and choose the smaller, with a fixed tie rule. The full conditional expectation then increases by at most 2^{-s-7} , since the omitted expectation is nonnegative and bounded by that quantity in both branches. These losses sum to $1/64$.

Any finite subcollection of the bad indicators becomes determined once all its finitely many bits have been fixed. Its value is bounded by the full conditional expectation, hence by $1/8 + 1/64 < 1$. It is an integer, so it is zero. No bad event survives.

Each computation is finite: enumerate the finitely many unchosen bits involved in the rectangle and average the exact indicators. This proves computability, not a practical complexity bound. $\square$

In the arithmetic application, event membership is exact: after finite factorisations and enumerations, compare the square of each rational $F_{S,J,\kappa}$ with the rational threshold in (8.4). The algorithm does not have to recognise the onset of an asymptotic estimate.

8.4 Recovering the masked law from its characters

For the selected table, let $p_S(w)$ be the empirical joint mass of mask S and signs w for a head κ . Fourier inversion on $\{-1, 1\}^S$ gives

$$p_S(w) - \rho_{N,H}(S)2^{-|S|} = 2^{-|S|} \sum_{\emptyset \neq J \subseteq S} F_{S,J,\kappa} \prod_{i \in J} w_i.$$

Summing absolute values over w and S , and dividing by two, yields

$$\sup_{\kappa} d_{\text{TV}}(\mathcal{L}_N(S, W_{\kappa}^H | S), U(\rho_{N,H})) \leq \Delta_{N,H,k_0}, \quad (8.6)$$

where, when the collision condition holds and $H \leq N$,

$$\Delta_{N,H,k_0} = \min \left\{ 1, \frac{3^{3H} - 2^{3H}}{2} \eta_{N,H,k_0} \right\}.$$

Outside this range we use the trivial bound $\Delta = 1$. Here d_{TV} is one half of the ℓ^1 distance. The same bound applies to projection onto one trace.

Since $A = O_{\vartheta}(N^{2-\vartheta})$ for $H \leq N$, the error tends to zero at each fixed H :

$$\Delta_{N,H,k_0} = O_{\vartheta,H,k_0}(N^{-(\vartheta-1)/2}). \quad (8.7)$$

It also tends to zero uniformly for

$$1 \leq H \leq c \log N, \quad c < \frac{\vartheta - 1}{6 \log 3}, \quad (8.8)$$

because the upper bound is $O_{\vartheta,k_0,c}(HN^{-(\vartheta-1)/2+3c \log 3})$. The collision condition holds eventually throughout this range. These are statements about a single table and the real masks at all the indicated sizes, not about independently chosen tables or model masks.

8.5 Entropy without subtracting the mask entropy

We measure entropy in bits, writing $h_2(t) = -t \log_2 t - (1-t) \log_2(1-t)$ with the usual endpoint conventions.

Lemma 8.3 (Entropy transfer). *Let $Y \in \{-1, 1\}^m$ and let $S \subseteq [m]$ be any possibly dependent mask. If the law of (S, Y_S) is at total variation distance D_0 from $\rho(S)2^{-|S|}$, with the same mask marginal ρ , then*

$$H(Y) \geq \mathbb{E}|S| - mD_0 - h_2(D_0). \quad (8.9)$$

Proof. For each mask s , couple Y_s conditional on $S = s$ to a uniform word of length $|s|$ with disagreement probability δ_s equal to their total variation distance. Revealing whether they disagree and, on disagreement, the uniform word, proves

$$H(Y_s | S = s) \geq |s| - |s|\delta_s - h_2(\delta_s).$$

The average of δ_s is D_0 . Average the inequality, apply concavity of h_2 , and use

$$H(Y) \geq H(Y | S) \geq H(Y_S | S).$$

This proves (8.9). In particular, no independence of the mask and the full word was used. $\square$

When only $D_0 \leq \Delta$ is known, we use $m\Delta + h_2(\Delta)$ once $\Delta \leq 1/2$; otherwise the trivial entropy lower bound zero is sufficient. Monotonicity of h_2 on all of $[0, 1]$ is not assumed.

Theorem 8.4 (An entropy-resistant prescription). *Fix rational $1 < \vartheta < \gamma < 6/5$, $1/\vartheta < \sigma \leq 1$, and assume $(\mathbf{GF}_\gamma)$. Above any fixed prefix, an arbitrarily small σ -cost support admits a computable prescription τ such that, for some $\delta > 0$ and every fixed H ,*

$$\liminf_{N \rightarrow \infty} \inf_{h \in \mathcal{H}_N} \frac{H(\mu_{N,h}^{H,d})}{H} \geq \delta \quad (d \in \mathcal{D}), \quad \liminf_{N \rightarrow \infty} \inf_{h \in \mathcal{H}_N} \frac{H(\mu_{N,h}^H)}{H} \geq 3\delta,$$

where $\mathcal{H}_N$ consists of all sources agreeing with τ at $q \in E$, $q > R_\vartheta(N)$. Every coherent subsequential limit of these empirical word laws, including those obtained from moving choices $h_N \in \mathcal{H}_N$, has metric entropy at least δ on each trace and 3δ jointly.

Proof. Use the support bound (2.7) and the selection above. Under $(\mathbf{GF}_\gamma)$, for each trace at least aN indices in $(N/2, N]$ have a prime greater than $R_\gamma(N)$. For $H = O(\log N)$, this exceeds $R_\vartheta(N + H)$ eventually. All these indices exceed L , so their primes lie in E . For each shift $0 \leq a' < H$, these indices are read at origins $n - a'$ within $\{2, \dots, N + 1\}$ for all sufficiently large N . Therefore

$$\mathbb{E}|S_d| \geq \delta H, \quad \mathbb{E}|S| \geq 3\delta H$$

with a fixed positive δ , decreased from a if needed for initial endpoints. A source in $\mathcal{H}_N$ has the prescribed active prime signs, since active primes exceed $R_\vartheta(N + H)$. Its cofactor signs are one of the heads already covered in (8.6). Apply lemma 8.3 and (8.7).

For each fixed horizon, finite-alphabet entropy is continuous. Every coherent limit law thus satisfies the same lower bound at all horizons. Shifting origins changes empirical laws only at $O(H)$ boundary positions, even for a moving source h_N , so every such limit is stationary. Its metric entropy is $\inf_{H \geq 1} H(\mu^H)/H$, which proves the assertion. $\square$

For the logarithmic horizons in (8.8), the same proof gives at least $2^{(3\delta - o(1))H}$ distinct joint words, because entropy is at most the logarithm of the support size. Neither this lower bound nor theorem 8.4 asserts full normality, uniqueness of the limiting law, or independence of active signs from the unmasked inactive values.

9 A compact group with every nontrivial quotient separated

A collection of individually good prescriptions need not be closed under multiplication: two of them might cancel. To make every nontrivial quotient retain quadratic complexity, we choose a matrix of prime prescriptions and pay for all the message projections visible at each finite scale. This gives a different encoding from the paired code.

9.1 Finite rows for infinite messages

Let $\text{bl}(u) = \lfloor \log_2 u \rfloor + 1$ for integers $u \geq 1$. On E_L assign to each prime a row

$$v(q) \in \mathbb{F}_2^{\text{bl}(q)}, \quad \Psi_x(q) = (-1)^{\sum_{j \leq \text{bl}(q)} v_j(q)x_j}, \quad \Psi_x(q) = 1 \quad (q \notin E_L).$$

Extend completely multiplicatively. For any matrix of these rows,

$$\Psi_{x+y} = \Psi_x \Psi_y. \tag{9.1}$$

At a word scale (N, H) use M, V, R, A as in Section 8, and put

$$K = \text{bl}(V), \quad b_N = \text{bl}(R_\vartheta(N) + 1).$$

Every active prime has row length between b_N and K . Call a message visible when its first b_N bits are not all zero. Its projection on every active row is then nonzero. At this scale only the first K message bits occur, so there are at most 2^K visible possibilities to control.

Temporarily choose the rows independently and uniformly. For a fixed visible message z and a fixed head κ , replace $\tau(q_i)$ in the character sum of Section 8 by $(-1)^{\langle v(q_i), z \rangle}$. Distinct active primes give independent fair signs within a word. Changing a row still affects at most $3H$ origins. Thus the same bounded-differences estimate holds, with one additional factor 2^K in the union bound.

If k_0 individual matrix coefficients are prescribed beforehand, use

$$\eta_{N,H,k_0}^2 = \frac{108H^2}{N} (\pi(A) + K + 7H + 4 + 4\ell_N + k_0). \tag{9.2}$$

The bad-event probabilities again satisfy (8.5). Enumerate individual coefficients by increasing prime and then increasing column, and apply lemma 8.2. It supplies one computable matrix for which (8.6) holds uniformly over all visible messages and heads, with (9.2) replacing (8.4). Since $K = O(\log N)$, both the fixed-horizon rate and the logarithmic-horizon range remain valid.

This finite union is the simultaneous argument. We have not intersected an uncountable collection of almost-sure events indexed by infinite messages. Every fixed nonzero infinite message eventually becomes visible and is then covered by the finite statements.

Theorem 9.1 (Small compact group, separated quadratic readings). *Assume (GF_γ) for rational $1 < \vartheta < \gamma < 6/5$. Fix $1/\vartheta < \sigma \leq 1$, a positive rational ε and $B \geq 2$. There is a computable row matrix on a decidable support $E \subset (B, \infty)$, $\sum_{q \in E} q^{-\sigma} < \varepsilon$, giving a continuous injective homomorphism*

$$\Psi : \mathbb{F}_2^{\mathbb{N}_{\geq 1}} \longrightarrow \mathcal{M}_\pm.$$

Its image is compact in d_σ and has diameter less than ε . All its members equal 1 on $[1, B]$. There are $c, \delta > 0$, independent of B, ε , with the following properties.

For every fixed $x \neq y$ and every trace d ,

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \# \{2 \leq n \leq N+1 : \Psi_x(Q_d(n)) \neq \Psi_y(Q_d(n))\} \geq c. \tag{9.3}$$

For each nonzero message, every limiting frequency law on a quadratic trace has metric entropy at least δ , and every joint limiting law has entropy at least 3δ . These entropy and sign-count guarantees also hold uniformly for visible messages after arbitrary reassignment of all prime signs other than the prescribed ones in E above $R_\vartheta(N)$.

Proof. Choose L to pay the support cost, and choose the matrix as above. The mass and entropy arguments in [theorem 8.4](#) apply uniformly to every visible message. At horizon one, the reference masked law assigns exactly half of its active mass to each sign. Thus, on each trace, the number of negative active readings of any such source is at least

$$\delta N/2 - N\Delta_{N,1,k_0} \geq \delta N/4$$

for every sufficiently large N . The same lower bound holds for positive readings. Take $c = \delta/4$.

If $x \neq y$, the message $x + y$ has a first nonzero coordinate, which is visible eventually. Applying the preceding sign bound to the quotient [\(9.1\)](#) proves [\(9.3\)](#) and injectivity. For continuity, choose finitely many primes whose omitted σ -weight is less than η , and then a message prefix long enough to determine all of their rows. Agreement on this prefix implies $d_\sigma < \eta$. The domain is compact and the target metric space is Hausdorff, so the image is compact and homeomorphic to the domain. The diameter bound follows from the common support. $\square$

The identity $\Psi_0 = 1$ necessarily has zero quadratic entropy and is excluded from the positive-entropy assertion. A computable message gives a computable source; an arbitrary message gives a source computable relative to that message. The uncountable image is not asserted to consist entirely of computable functions.

9.2 Ordinary frequencies have entropy zero

A common harmonic support gives more than ordinary closeness. The following fact applies to each member of all the constructed groups, including the paired code.

Proposition 9.2 (Ordinary entropy under a summable prime support). *Suppose $f \in \mathcal{M}_\pm$ and $\sum_{q:f(q)=-1} q^{-1} < \infty$. Then the ordinary sequence $f(n)$ has frequencies for every finite sign word. The resulting stationary measure has metric entropy zero.*

Proof. First truncate the negative-prime set to a finite subset. The upper density of disagreement is bounded by the reciprocal sum of the omitted primes, since every disagreement is a multiple of at least one of them. For the remaining primes, truncate valuations at a level J by assigning a fixed value to multiples of q^J . This adds at most $\sum q^{-J}$ to the disagreement density and gives a periodic sequence, of period dividing $\prod q^J$.

Thus f admits periodic approximations arbitrarily close in upper density. For a fixed word length H , the discrepancy between empirical word laws is at most H times the single-site discrepancy. The limiting laws of the periodic approximations are Cauchy, and the empirical law of f converges to their limit.

Take a periodic approximation of period T and error density at most $\eta < 1/2$. A word of f is determined by the phase of the periodic sequence and the vector of disagreements. Averaging over origins and passing to a stationary limit gives

$$H(\mu_f^H) \leq \log_2 T + Hh_2(\eta).$$

Divide by H , let $H \rightarrow \infty$, and then let $\eta \downarrow 0$. The entropy rate is zero. $\square$

This is an elementary specialization of the theory of rationally almost periodic sequences. For a finite alphabet, Bergelson–Kulaga–Przymus–Lemańczyk–Richter [[2](#), Theorem 1.7 in the cited author version] prove genericity for an ergodic measure with rational discrete spectrum, a conclusion stronger than the entropy-zero consequence used here. The proof above is

included to make its summable-prime application self-contained, not as a new entropy principle. A broader multiplicative-function context is provided by Frantzikinakis–Lemańczyk–de la Rue [8]; no result from that broader theory is needed for this proof. The proposition does not assert zero topological entropy.

For a prime set A of finite harmonic sum, the ordinary density of negative values of $\prod_{q \in A} (-1)^{v_q(n)}$ also exists and equals

$$\frac{1}{2} \left(1 - \prod_{q \in A} \frac{q-1}{q+1} \right).$$

For finite A , the local factor is the mean of $(-1)^{v_q(n)}$, computed by a geometric series; the Chinese remainder theorem multiplies the local means. The uniform tail bound $\sum_{q \in A, q > Y} q^{-1}$ passes to infinite A . This ordinary calculation does not imply convergence of a quadratic mean.

9.3 Two topologies and the order of quantifiers

Define the aligned observation pseudometric

$$D_d(f, g) = \limsup_{N \rightarrow \infty} \frac{1}{N} \# \{ 2 \leq n \leq N+1 : f(Q_d(n)) \neq g(Q_d(n)) \}.$$

It satisfies the triangle inequality. By (9.3), the compact prime-metric group contains an uncountable set separated by at least c in every D_d . After identifying pairs of zero pseudodistance, this set is nonseparable in the observation metric. No countable collection of prototypes can approximate all its members within a fixed radius smaller than $c/2$.

Translation by any reference f_0 preserves all pairwise observation distances and ordinary agreements. It does not give each translated member positive quadratic entropy: the translate contains f_0 itself, and f_0 is arbitrary. The positive-entropy assertion applies to the nontrivial quotients, or to the nonidentity members over base 1.

For the unit messages e_j , continuity gives $d_\sigma(\Psi_{e_j}, 1) \rightarrow 0$, while $D_d(\Psi_{e_j}, 1) \geq c$ for every j . These are aligned sequence distances, not distances between word laws. Two different sequences can have identical word frequencies.

There is no scale which separates all nonzero infinite messages simultaneously. At any fixed N , choose a nonzero message whose first $\text{bl}((N+1)^2 + 2)$ bits vanish. Every raw reading up to $N+1$ then equals 1. Thus the infimum over nonzero messages of the normalised finite discrepancy is zero at every N , even though every fixed nonzero message has positive limiting lower discrepancy. Visibility is indispensable.

At a fixed large N , the visible prefix has length $b_N = \vartheta \log_2 N + O(1)$. Distinct such prefixes remain at least cN apart even after independently changing their cofactor heads. Consequently they are uniquely recoverable under $2t + s < cN$. A finite search over the $K = 2 \log_2 N + O(1)$ consulted message bits and the small-prime heads performs the decoding. Its head enumeration need not be polynomial, and the guaranteed number of message bits is only logarithmic. Positive entropy of the group is therefore not by itself the positive prefix rate of the main theorem.

10 An alternative causal encoding by row selection

The filtered paired code already combines causal raw dependence with polynomial decoding. We retain a different construction in which row lengths are prescribed by first occurrences

and a concentration argument pays directly for linearly many visible message bits. Its explicit selection and decoder use unrestricted finite searches. This is a property of these algorithms, not a lower bound on the cost of causal coding. In particular, [proposition 5.2](#) rules out reading this section as a necessary efficiency–causality tradeoff.

10.1 A prime’s first admissible occurrence

Fix $1 < \vartheta < 2$ and use the common first admissible occurrence $\nu(q)$ of (2.8). By [lemma 2.3](#), $\nu(q) \leq n$ whenever this prime divides any positive value $Q_d(n)$ with $n \geq 2$, even where it is no longer the large factor. The primitive-divisor context of this first-appearance geometry was specified after [lemma 2.1](#); the minimum across three traces and the superlinear cutoff remain part of our definition.

Take $E = \{q : \nu(q) \geq L\}$. This is a first-occurrence support, not the support defined by the existence of any occurrence above L in [Section 2.3](#). Each index can be the first occurrence of at most three primes, so the same bound

$$\sum_{q \in E} q^{-\sigma} \leq \frac{3(L-1)^{1-\vartheta\sigma}}{\vartheta\sigma - 1}$$

holds. The centred-root tests decide $\nu(q)$ and membership in E .

Choose explicitly $k(t) = \lfloor \rho t \rfloor$, for a positive rational ρ . This function is computable and nondecreasing, with $k(t) = \rho t + O(1)$. At prime $q \in E$ choose a row of length $k(\nu(q))$ and set

$$\Phi_x(q) = (-1)^{\langle \nu(q), x \leq k(\nu(q)) \rangle};$$

set other prime signs to 1. The source is a homomorphism of the message by linearity. Equation (2.9) proves that every raw reading $\Phi_x(Q_d(n))$ with $n \leq N$ depends only on the first $k(N)$ bits.

10.2 Discarding a controlled number of old primes

For $0 < \alpha \leq 1/2$, let $s_N = \lceil \alpha N \rceil$ and select

$$J_{d,N} = \{N/2 < n \leq N : Q_d(n) = qa, \ q > R_\vartheta(N), \ q \in E, \ \nu(q) \geq s_N\}.$$

Write $M_{d,N} = |J_{d,N}|$, $M_N = \min_d M_{d,N}$ and

$$A_N = \left\lfloor \frac{N^2 + 2}{R_\vartheta(N) + 1} \right\rfloor = O(N^{2-\vartheta}).$$

There are at most $3(s-2)_+$ primes whose first occurrence is less than s , because each earlier index supplies at most three. Each such prime removes at most one candidate on each trace. Under [\(GF \$_\gamma\$ \)](#), choose a rational

$$0 < a < \min\{a_\gamma, 1/8\}, \quad \alpha = a/12, \quad \rho = a/64,$$

where a_γ is any mass constant in that hypothesis. For every sufficiently large N , with $s_N \geq L$, the candidate population from exponent $\gamma > \vartheta$ gives

$$M_N \geq aN - 3 \lceil \alpha N \rceil \geq aN/2. \tag{10.1}$$

Thus one can insist that every remaining row starts recently enough to see a common prefix of length proportional to N , without exhausting the supply.

10.3 Paying for linearly many message bits

Put $K_N = k(N)$ and $b_N = k(s_N)$. Call $z \in \mathbb{F}_2^{K_N}$ visible when $z_{\leq b_N} \neq 0$. Every row in $J_{d,N}$ has length between b_N and K_N , so it sees a nonzero projection of z .

Choose rows independently and uniformly for the existence argument. For a fixed visible z and a small-prime assignment κ up to A_N , the sum

$$S_{d,z,\kappa} = \sum_{n \in J_{d,N}} \kappa(a_n) (-1)^{\langle v(q_n), z \rangle}$$

is a sum of $M_{d,N}$ independent fair signs. The primes are distinct on this trace, although primes shared between traces keep the same row. The elementary bound $\cosh t \leq e^{t^2/2}$ and exponential Markov give

$$\Pr\{|S_{d,z,\kappa}| > M_{d,N}/2\} \leq 2e^{-M_{d,N}/8}.$$

Taking the union over three traces, messages and heads yields

$$\Pr(B_N) \leq 6 \cdot 2^{K_N + \pi(A_N)} e^{-M_N/8}. \quad (10.2)$$

Unlike Section 9, the number of visible message bits is now linear. A vanishing discrepancy error would not automatically pay this cost. The constant tolerance $|S| \leq M/2$ leaves a positive distance, which is all the present decoding statement needs.

10.4 A computable selection without an analytic stopping test

Allow k_0 predetermined coefficients. Define a bad event only at sizes with $b_N \geq 1$ and satisfying the finite test

$$M_N \geq 8(K_N + \pi(A_N) + k_0 + 8 + 4 \text{bl}(N + 1)). \quad (10.3)$$

Otherwise put $B_N = \emptyset$. From (10.2),

$$\Pr(B_N) \leq 2^{-k_0-5} (N + 1)^{-4}.$$

The potential $U = \sum_{N \geq 2} \mathbf{1}_{B_N}$ has conditional expectation below $1/16$ after any values of the k_0 prescribed bits. After s further coefficients have been fixed, the omitted expectation of the events $N > 2^R$ in either next-bit branch is at most

$$\frac{2^{s-4-3R}}{3}.$$

This follows by summing the displayed probability bound and dividing by the branch probability $2^{-(k_0+s+1)}$. Taking $R = s + 4$ makes this less than 2^{-s-6} . At each step minimise the exact finite conditional expectation over $N \leq 2^R$. The increases sum to at most $1/32$. As in lemma 8.2, every determined finite sum of bad indicators is an integer smaller than $1/16 + 1/32 < 1$, hence zero.

This defines one computable matrix. Its algorithm takes the chosen rationals α, ρ as parameters; it does not need a numerical threshold at which (10.1) starts. The existence of admissible positive parameters, rather than their numerical calibration, uses (GF $_\gamma$).

Theorem 10.1 (Causal prefix dependence and positive rate). *Under (GF $_\gamma$), the first-occurrence construction gives a continuous injective homomorphism $x \mapsto \Phi_x$ on an arbitrarily small harmonic support above any fixed prefix. All raw readings at indices at most N depend*

on at most $K_N = \rho N + O(1)$ message bits. For every sufficiently large N , each trace alone determines the first $b_N = \rho \alpha N + O(1)$ message bits after arbitrary reassignment of all prime signs other than those prescribed at $q \in E$, $q > R_\vartheta(N)$, and after any errors and erasures with

$$2t + s < cN$$

for some $c > 0$. Both the selection and the decoder are computable; no polynomial bound is asserted for them.

Proof. Since $K_N = \rho N + O(1)$ and $\pi(A_N) = o(N)$, (10.1) eventually implies (10.3); indeed, $8\rho = a/8 < a/2$. The selected matrix then gives $|S_{d,z,\kappa}| \leq M_{d,N}/2$ for every visible difference and every head. The corresponding quotient has at least $M_{d,N}/4 \geq aN/8$ negative readings on each selected trace. Thus two candidates with different first b_N bits are separated by that distance, even if their later bits and small signs differ.

Any two candidates fitting the received data with t errors and s erasures have at most $2t + s$ disagreements. Taking, for example, $c = a/16$ proves uniqueness of the desired prefix. A decoder enumerates the K_N consulted bits and the heads on primes up to A_N , computes the selected readings, and retains candidates of smallest compatible distance. The intermediate bits after b_N are unrestricted; they are not set to zero. No bits after K_N act on the raw prefix by (2.9).

Here $b_N = a^2 N / 768 + O(1)$, so a conservative positive bit rate is $a^2 / 1536$. Each fixed nonzero message eventually has a visible coordinate, proving injectivity. The support bound gives continuity and compactness exactly as in theorem 9.1. $\square$

The raw dependence is causal in the message index, but computing a coefficient can require a large finite anticipation. The theorem gives latency proportional to the observation scale, not a distance bound at every delay after two messages first diverge. These are different guarantees from the probabilistic anytime-reliability codes considered, for example, by Sukhvasi–Hassibi [25]. Neither a bounded-delay code nor a cryptographic security claim follows.

The constructions retain distinct roles. The filtered paired group Ω provides polynomial decoding, positive rate and causal raw dependence. The row-length-bl(q) group Ψ provides masked word laws and entropy of all nontrivial quotients, without a proved positive-rate polynomial decoder. The first-occurrence group Φ is an alternative causal positive-rate map with its own concentration-based proof. No argument above transfers the word-law and quotient-entropy conclusions of Ψ to Ω or Φ .

11 What is recovered and what remains unspecified

The main construction separates arithmetic availability from coding. Its finite input is a list of disjoint prime pairs with equal exact cofactors; its finite output is an ordinary binary error-correcting channel. The availability estimate (GF_γ) is used only to obtain positive rate and a linear correction budget at all sufficiently large observation lengths. A numerical improvement in an inner code does not improve that analytic input.

All messages in this paper are chosen coordinates of constructed families. They are not the arbitrary prime signs of an unknown source. For comparison, the free arithmetic minimum is

$$m_p(N) = \min_{\substack{h \in \mathcal{M}_\pm \\ h(p) = -1}} \sum_{\substack{1 \leq n \leq N, \\ Q_d(n) > 0}} \mathbf{1}_{h(Q_d(n)) = -1}.$$

A positive distance between different members of our code says nothing by itself about this minimum. Even a translate by $\beta_p(n) = (-1)^{v_p(n)}$ with $p \leq B$ contains β_p itself, while all its members have sign -1 at p . The code's prescribed large signs are additional constraints which the free minimum is not required to obey.

The same distinction applies to statistical complexity. Positive entropy of all limiting quadratic laws is weaker than complete Bernoulli normality, and does not assert a unique limiting law. Conversely, a code's aligned Hamming separation does not require different word laws. These observation types were kept separate in the constructions because they require different estimates.

Several natural refinements are left as separate questions by these proofs: a quantitative calibration of the weighted mass input and its threshold; better rate and distance constants for the same arithmetic resource; and the addition of the masked word-law and quotient-entropy guarantees to the already causal, polynomially decoded paired encoding. None is a premise of the finite theorem or of the conditional implications proved above.

Assistance and verification

GPT-6 Astra assisted the research notes and initial consolidation; GPT-6 Astra Pro assisted the subsequent proof exposition, reviews, code and finite checks, bibliography and manuscript revisions. The author is responsible for the article. No complete independent human verification, Lean formalisation or independent external replication is claimed. Appendix B identifies the reproducible finite checks and their limits; they do not establish (GF_γ) .

A Fixed-scale balancing without the superlinear mass input

The superlinear cutoff makes the number of exact cofactors small and their support summable. At the smaller cutoff $2N$, a different count of square classes gives a useful unconditional result with an explicit leading constant. It does not produce the same all-scale sparse encoding.

For $N \geq 5$ let

$$J_d(N) = \{n \leq N : Q_d(n) > 0, P^+(Q_d(n)) > 2N\}, \quad M_d(N) = |J_d(N)|,$$

and let $\mathcal{Q}_N$ be the set of all primes greater than $2N$ occurring in these observations. Write $Q_d(n) = q_n a_n = q_n b_n s_n^2$, where $b_n = \text{core}(a_n)$ is squarefree. By lemma 2.1, $a_n < N/2$ and the q_n on any one trace are distinct.

A.1 A half-population of large factors

The private-prime argument and the coefficient $1/2$ for $n^2 + 1$ already occur in the proof of Klurman's Lemma 2.1 [13, Section 2]. We include the elementary fixed-form calculation for the three traces, with its error term, so that the subsequent adversarial balancing statement has an explicit arithmetic input.

Lemma A.1. *For each of the three fixed traces,*

$$M_d(N) \geq N/2 - O(N/\log(2N)).$$

Proof. The logarithm of the product of the positive polynomial values is

$$\sum_{\substack{n \leq N \\ Q_d(n) > 0}} \log Q_d(n) = 2N \log N + O(N).$$

For an odd split prime q , the congruence $n^2 \equiv d \pmod{q^j}$ has exactly two roots for every $j \geq 1$. With $V = N^2 + 2$, summing their counts gives

$$\sum_{n \leq N} v_q(Q_d(n)) \leq \frac{2N}{q-1} + \frac{2 \log V}{\log q}.$$

Odd inert primes do not divide these polynomial values. The prime 2 contributes $O(N)$ to the logarithmic mass. The split primes for these fixed d are unions of residue classes modulo 4 or 8. The fixed-progression weighted Mertens formula [16, Section 27.11, equation (27.11.11)] gives

$$\sum_{\substack{q \leq x \\ (d/q)=1}} \frac{\log q}{q-1} = \frac{1}{2} \log x + O(1).$$

Replacing q by $q-1$ in the denominator adds a bounded sum. Summing the error terms using $\pi(2N) \ll N/\log(2N)$ shows that primes at most $2N$ contribute at most $N \log N + O(N)$. Thus primes greater than $2N$ contribute at least $N \log N - O(N)$. Each corresponding observation contributes at most $\log V = 2 \log N + O(N^{-2})$, which proves the lemma. $\square$

This elementary weighted argument at the cutoff $2N$ is not a proof of (GF_γ) at a fixed exponent greater than one. In particular, its prime-counting error cannot simply be used at a power-sized larger cutoff.

A.2 Few square classes and a common assignment

Let $s_d(N)$ be the number of distinct squarefree cofactors b_n on trace d . For every integer $U \geq 1$,

$$s_d(N) \leq U + \pi((N^2 + 2)/U). \quad (\text{A.1})$$

There are at most U classes with $b \leq U$. For each remaining class, choose one representative. Its cofactor $a \geq b > U$ gives $q < (N^2 + 2)/U$, and distinct classes on this trace give distinct primes. This proves (A.1). Taking U of order $N/\sqrt{\log(2N)}$ yields

$$s_d(N) \ll N/\sqrt{\log(2N)}.$$

The class sets

$$V_{d,b} = \{q_n : n \in J_d(N), \text{ core}(a_n) = b\}$$

have incidence at most three per prime. We use the classical Beck–Fiala bound $2t - 1$ for incidence at most t [1, 18]. Its short proof fixes the exact finite construction. Start at $x_q = 0$ and preserve the sum of every row having more than t free coordinates. If u coordinates are free and a rows are constrained, $(t+1)a \leq tu$, hence $a < u$. A rational null direction exists; move until a coordinate reaches ± 1 and freeze it. When a row is released, its sum is still zero and it has at most t free coordinates. Each can subsequently change by less than two. Its final integer sum therefore has absolute value at most $2t - 1$. Rows never constrained satisfy the same argument from their initial zero sum.

With $t = 3$, this produces one table $\tau_N : \mathcal{Q}_N \rightarrow \{-1, 1\}$, shared by all traces, for which every class sum has absolute value at most five. For an arbitrary table define

$$B_{d,b} = \sum_{q \in V_{d,b}} \tau_N(q), \quad \Lambda_d = \sum_b |B_{d,b}|.$$

Every source extending the table satisfies

$$\sum_{n \in J_d(N)} h(Q_d(n)) = \sum_b B_{d,b} h(b),$$

$$\min_{v=\pm 1} \#\{n \leq N : Q_d(n) > 0, h(Q_d(n)) = v\} \geq (M_d(N) - \Lambda_d)/2.$$

For the constructed table, $\Lambda_d \leq 5s_d(N)$. We have proved the following unconditional assertion.

Theorem A.2 (A quarter of each sign). *For every $N \geq 5$, a deterministic finite procedure chooses one assignment on the primes in $\mathcal{Q}_N$ such that every completely multiplicative extension has, on each trace, at least*

$$\frac{N}{4} - O\left(\frac{N}{\sqrt{\log(2N)}}\right)$$

occurrences of each sign. The prime signs at most $2N$ may be chosen after the table and without restriction.

The table can change with N . Changing u reserved prime signs of this table decreases either sign-count guarantee by at most u on each trace, by uniqueness of their trace occurrences.

A.3 A necessary cofactor code for any almost-positive source

We now drop the balancing prescription. For every $q \in \mathcal{Q}_N$, choose one canonical observation $Q_{d_q}(n_q) = qa_q$. Let $K_N = |\mathcal{Q}_N|$, and form the binary matrix

$$T_N(q, \ell) = v_\ell(a_q) \pmod{2}, \quad \ell < N/2 \text{ prime}.$$

Its image $\mathcal{C}_N$ has dimension $r_N \leq \pi(\lfloor N/2 \rfloor) = O(N/\log(2N))$. If $h(q) = (-1)^{y_q}$ on $\mathcal{Q}_N$ and x is the small-prime vector, the negative-reading vector on the representatives is exactly $y + T_N x$. Hence

$$\text{dist}_H(y, \mathcal{C}_N) \leq E_N(h) := \sum_{\substack{n \leq N, d \in \mathcal{D} \\ Q_d(n) > 0}} \mathbf{1}_{h(Q_d(n)) = -1}. \quad (\text{A.2})$$

Consequently the number of large-prime tables admitting any extension with $E_N(h) \leq t$ is at most

$$\min \left\{ 2^{K_N}, 2^{r_N} \sum_{j=0}^{\min(t, K_N)} \binom{K_N}{j} \right\}. \quad (\text{A.3})$$

No disjointness of the balls is needed. Moreover,

$$N/2 - O(N/\log N) \leq \frac{1}{3} \sum_d M_d(N) \leq K_N \leq 3N.$$

If $t = o(N)$, the logarithm of the ball volume is $o(N)$, by the usual estimate $\sum_{j \leq t} \binom{K}{j} \leq (t+1)(eK/t)^t$ when $1 \leq t \leq K/2$. Thus at most $2^{o(N)}$ large-prime tables admit such an extension, a fraction at most $2^{-(1/2-o(1))N}$ of all tables.

This is a condition on the real prime signs of any almost-positive source. It is not sufficient for extension to all remaining observations, and the small set it identifies need not be empty. In particular, a low-dimensional cofactor code cannot itself exclude the constant table or supply the missing negative-prime constraint in $m_p(N)$.

A.4 One random source, or a computable sparse subsequence

There are two distinct ways to make the quarter-sign conclusion persistent. Neither should be confused with the superlinear-cutoff constructions in the main text.

First, choose independent fair signs $F(q)$ once at all primes. For a fixed trace and a fixed head on primes below $N/2$, the sum on $J_d(N)$ is a sum of $M_d(N)$ independent fair large-prime signs. If $R'_N = \pi(\lfloor N/2 \rfloor)$, exponential Markov and a union over the three traces and all heads give

$$\Pr \left\{ \max_{d,\kappa} |S_{d,\kappa}| > u \right\} \leq 6 2^{R'_N} e^{-u^2/(2N)}.$$

Taking

$$u_N = \sqrt{2N(R'_N \log 2 + 4 \log N + \log 6)} = O(N/\sqrt{\log N})$$

makes this probability at most N^{-4} . Borel–Cantelli and [lemma A.1](#) show that, almost surely, this one F has the quarter-sign guarantee at every sufficiently large N , even after an arbitrary reassignment at primes at most $2N$. A finite prime prefix may be prescribed before the experiment. No small harmonic support or computability of the resulting F is asserted.

Second, the deterministic tables of [theorem A.2](#) can be embedded into one computable source close to a reference, but only on a selected sequence of sizes. Uniformly over sets $\mathcal{A}_N$ of at most $3N$ primes greater than $2N$,

$$\sum_{q \in \mathcal{A}_N} \frac{1}{q} \ll \frac{\log \log(3N)}{\log(2N)} \rightarrow 0. \quad (\text{A.4})$$

This sparse-scale summability mechanism is already used in the proof of Klurman’s Proposition 1.1 [13, Section 2], including the reciprocal-prime estimate of order $\log \log N / \log N$. For the uniform form (A.4), split at $N(\log N)^2$. The upper portion costs at most $3/(\log N)^2$, and partial summation of $\pi(x) \ll x/\log x$ bounds the lower portion as displayed.

Given a reference f_0 , a protected prefix B and a rational $\varepsilon > 0$, find successively integers N_j satisfying

$$2N_j > \max\{B, N_{j-1}^2 + 2\}, \quad \sum_{q \in \mathcal{Q}_{N_j}} q^{-1} < \varepsilon 2^{-j-1}.$$

Finite search terminates by (A.4). Assign the corresponding table τ_{N_j} on each of the disjoint prime intervals $(2N_j, N_j^2 + 2]$, and keep f_0 elsewhere. Later stages cannot change earlier tables. The total harmonic cost is less than ε , and the quarter-sign conclusion holds at every N_j against all head reassignments up to $2N_j$. To evaluate a fixed integer, search only until $2N_j$ exceeds it. The source is computable relative to f_0 , but the guarantee is not asserted at intervening sizes.

The quantifiers explain why these statements do not solve the free minimum: selecting a large-prime table which defeats every head is a $\max_\tau \min_\kappa$ construction. The free minimum also minimises over τ , under its prescribed negative prime. Interchanging those operations would discard the additional information on which the construction depends.

B Finite data and reproduction

The publication package `N5_arithmetic_coding_publication_v1_3.zip` contains the integer data for [theorem 6.1](#), the LaTeX source, the frozen bibliography and the reproduction programs. Its edition identifier is `N5-AC-1.3-2026-10-12`. Its root files `VERSION`, `MANIFEST.json` and `SHA256SUMS.txt` identify the edition, manuscript hash, execution record

and supplied files. The arithmetic implementation uses Python 3.10 or later, the standard library, exact integers, rational comparisons and binary finite-field arithmetic.

B.1 Files that define the finite instance

The file `data/pairs.csv` records the epoch, trace, pair index, the two integer indices, their common cofactor and reserved primes. The field basis, inner generator, outer parameters and message-prefix lengths are in `data/epochs.json`. The decoder reconstructs the code from these tables; no message or prime-head assignment is supplied to it.

The module `code/arithmetic_code.py` implements the inner code, field operations, outer linear-system decoder and reliability thresholds. The scripts `build.py`, `verify.py` and `check_theory.py` construct the instance, check its saved arithmetic data and exercise the finite lemmas. The first-occurrence filter in `causal_filter.py` is enabled by default; `-unfiltered` selects the cancellation-only variant and does not promise causal dependence. Filter decisions are recorded in `data/causal_filter.json`. Validation uses explicit exceptions, including under `python -O`.

From the code directory, decode the 499-error example on $n^2 + 1$ by

```
python decode_readings.py --data ../data --epoch 32768 \
  --trace -1 --received ../data/received.csv
```

Use `received_gmd.csv` for the reliability-threshold example. For the reading-level mixed example, run

```
python decode_readings.py --data ../data --epoch 32768 \
  --trace -1 \
  --received ../verification/revision/data/received_mixed.csv
```

Expected-message files are separate test fixtures and are never read by this decoder.

B.2 Rebuilding and checking the finite construction

From the archive root, `bash reproduce.sh` rebuilds into a new directory, refuses to overwrite an existing destination, and compares the outputs with the supplied data. The individual commands, from `code`, include

```
python build.py --out ../recomputed
python verify.py --data ../recomputed
python check_theory.py --out ../recomputed/theory_checks.json
python check_revision.py --data ../recomputed \
  --out ../revision_recomputed
```

The checks cover 14664 pair identities, 29328 reserved primes, the inner distances, the first-occurrence filter and 98303 positive raw labels through 32768. They include 24 reading-level decodings using two messages and three multiplicative head assignments at the two accepted epochs. All use the decoding module of Section 4; this is not a second independent decoder. Counts for the small exhaustive tests and byte comparisons between normal Python and `python -O` are recorded in `verification/publication_v1_3/execution.json` and its logs, rather than interpreted as independent proofs.

The diagnostic windows 65536 and 131072 remove respectively zero and ten candidate primes, with first-occurrence witnesses in `verification/revision/data/revision_checks`.

`json`. They do not add accepted epochs to the pilot’s continuation schedule. The future epoch at 1920⁶, all infinite selectors and the auxiliary balancing and matrix-selection pilots are outside this replay. The harmonic bound in Section 6 pays for an uncomputed continuation; it is not an execution of it.

Checksums establish file identity, not mathematical correctness. The finite tests do not establish (GF_γ) or enumerate all admissible corruption patterns. The `research` directory contains the revision responses and comparison history; earlier logs retain their own dates and scope. The current execution record distinguishes this replay from those earlier runs.

References

- [1] József Beck and Tibor Fiala. “Integer-making” theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981. doi:[10.1016/0166-218X\(81\)90022-6](https://doi.org/10.1016/0166-218X(81)90022-6). The classical floating-variable proof is also given in Nikolov’s Lecture 3.
- [2] Vitaly Bergelson, Joanna Kułaga-Przymus, Mariusz Lemańczyk, and Florian K. Richter. Rationally almost periodic sequences, polynomial multiple recurrence and symbolic dynamics. *Ergodic Theory and Dynamical Systems*, 39(9):2332–2383, 2019. doi:[10.1017/etds.2017.130](https://doi.org/10.1017/etds.2017.130). Author version: [arXiv:1611.08392v2](https://arxiv.org/abs/1611.08392v2), Theorem 1.7.
- [3] Géza Bohus. On the discrepancy of 3 permutations. *Random Structures & Algorithms*, 1(2):215–220, 1990. doi:[10.1002/rsa.3240010208](https://doi.org/10.1002/rsa.3240010208).
- [4] Régis de la Bretèche and Sary Drapeau. Niveau de répartition des polynômes quadratiques et crible majorant pour les entiers friables. *Journal of the European Mathematical Society*, 22(5):1577–1624, 2020. doi:[10.4171/JEMS/951](https://doi.org/10.4171/JEMS/951). Locators: Theorem 1.2 and Section 8.4 of [arXiv:1703.03197v3](https://arxiv.org/abs/1703.03197v3), 7 May 2019.
- [5] Jean-Marc Deshouillers and Henryk Iwaniec. On the greatest prime factor of $n^2 + 1$. *Annales de l’Institut Fourier*, 32(4):1–11, 1982. Section 8. https://www.numdam.org/item/AIF_1982__32_4_1_0/.
- [6] Graham Everest and Glyn Harman. On primitive divisors of $n^2 + b$. In James McKee and Chris Smyth, editors, *Number Theory and Polynomials*, volume 352 of *London Mathematical Society Lecture Note Series*, pages 142–154. Cambridge University Press, 2008. doi:[10.1017/CBO9780511721274.011](https://doi.org/10.1017/CBO9780511721274.011). Proposition 1.3; author preprint [arXiv:math/0701234](https://arxiv.org/abs/math/0701234) (2007).
- [7] G. David Forney, Jr. Generalized minimum distance decoding. *IEEE Transactions on Information Theory*, 12(2):125–131, 1966. doi:[10.1109/TIT.1966.1053873](https://doi.org/10.1109/TIT.1966.1053873).
- [8] Nikos Frantzikinakis, Mariusz Lemańczyk, and Thierry de la Rue. Furstenberg systems of pretentious and MRT multiplicative functions. *Ergodic Theory and Dynamical Systems*, 45(9):2765–2844, 2025. doi:[10.1017/etds.2024.140](https://doi.org/10.1017/etds.2024.140). Broader context; not a proof input here.
- [9] Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. A constructive Borel–Cantelli lemma. Constructing orbits with required statistical properties. *Theoretical Computer Science*, 410(21–23):2207–2222, 2009. doi:[10.1016/j.tcs.2009.02.010](https://doi.org/10.1016/j.tcs.2009.02.010). Author version: [arXiv:0711.1478v2](https://arxiv.org/abs/0711.1478v2). Context for effective statistical selection, not a black-box input.

- [10] Lasse Grimmelt and Jori Merikoski. On the greatest prime factor and uniform equidistribution of quadratic polynomials. [arXiv:2505.00493v2](#), 30 May 2025, 2025. Theorem 1.1. Preprint cited for context, not as an input to the coding theorem.
- [11] Venkatesan Guruswami. Introduction to coding theory. Notes 6: Reed–Solomon, BCH, Reed–Muller, and concatenated codes. Carnegie Mellon University, February 2010, 2010. Scribes: Eric Blais and Venkat Guruswami. Sections 1 and 4. <https://www.cs.cmu.edu/~venkatg/teaching/codingtheory/notes/notes6.pdf>.
- [12] Venkatesan Guruswami. Introduction to coding theory. Notes 7: Justesen codes, Reed–Solomon and concatenated codes decoding. Carnegie Mellon University, March 2010, 2010. Scribes: Venkat Guruswami and Balakrishnan Narayanaswamy. Sections 3.3 and 5. <https://www.cs.cmu.edu/~venkatg/teaching/codingtheory/notes/notes7.pdf>.
- [13] Oleksiy Klurman. Correlations of multiplicative functions and applications. *Compositio Mathematica*, 153(8):1622–1657, 2017. doi:10.1112/S0010437X17007163. Locator for the large-prime argument: [arXiv:1603.08453v1](#), Section 2, proofs of Lemma 2.1 and Proposition 1.1.
- [14] Runbo Li. On the largest prime factor of quadratic polynomials. [arXiv:2406.07575v2](#), 31 May 2025, 2025. Theorem 1.1. This title and statement are those of version 2, not the different version 1 of 2024; context only.
- [15] Colin McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics, 1989*, pages 148–188. Cambridge University Press, 1989. doi:10.1017/CBO9781107359949.008.
- [16] National Institute of Standards and Technology. Digital Library of Mathematical Functions, section 27.11: Asymptotic formulas: Partial sums. <https://dlmf.nist.gov/27.11>. Equation (27.11.11) and the fixed-progression prime number theorem. Accessed 27 September 2026.
- [17] Alantha Newman and Aleksandar Nikolov. A counterexample to Beck’s conjecture on the discrepancy of three permutations. [arXiv:1104.2922v1](#), 14 April 2011, 2011.
- [18] Aleksandar Nikolov. Prague Summer School: Discrepancy Theory. Lecture 3, 2018. Theorem 2, pages 1–2. <https://www.cs.toronto.edu/~anikolov/Prague18/Lectures/lect3.pdf>.
- [19] Alexandru Pascadi. Large sieve inequalities for exceptional Maass forms and the greatest prime factor of $n^2 + 1$. *Forum of Mathematics, Pi*, 14:e8, 2026. doi:10.1017/fmp.2026.10025. Locators refer to [arXiv:2404.04239v3](#), 15 January 2026, Sections 6.1 and 6.3, equation (6.20).
- [20] Brice Pouly. Arithmetic reconstruction and error recovery from three quadratic traces. Zenodo, 2026. doi:10.5281/zenodo.23317860. Theorems 1.1–1.2.
- [21] Brice Pouly. Correlation laws and identifiability of multiplicative functions on quadratic traces. Zenodo, 2026. doi:10.5281/zenodo.23318918. Section 7.
- [22] Brice Pouly. Selective normality and algebraic rigidity of multiplicative traces. Zenodo, 2026. doi:10.5281/zenodo.23318326. Section 5, especially Proposition 5.1.
- [23] Maxim Raginsky. Concentration inequalities. Lecture notes, 29 August 2013, 2013. Theorem 2, page 6. <https://maxim.ece.illinois.edu/teaching/fall13/notes/concentration.pdf>.

- [24] Irving S. Reed and Gustave Solomon. Polynomial codes over certain finite fields. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):300–304, 1960. doi:10.1137/0108018.
- [25] Ravi Teja Sukhavasi and Babak Hassibi. Linear error correcting codes with anytime reliability. In *2011 IEEE International Symposium on Information Theory Proceedings*, pages 1748–1752, 2011. doi:10.1109/ISIT.2011.6033847. Author record: <https://authors.library.caltech.edu/records/4jyv2-64h44>.