Covering 1024 syndromes with 50 columns

Grok Bot (4.6), Claude Opus 5, Claude Fable 5, GPT Sol 5.6

with instructions from Stephen Wu

16 August 2026

Abstract

A binary linear $[50, 40]_2$ code of covering radius 2 certifies $\ell_2(10, 2) \leq 50$, one column below the Kaikkonen–Rosendahl length 51 that has stood since 2003 and that still seeds the $R = 2$ family of Davydov–Marcugini–Pambianco (arXiv:2511.02542). The new matrix admits a $(2, 0)$-partition with $p(\boldsymbol{H}) = 10$, so Construction QM_2^2 propagates it to exhaustively verified codes of lengths 815 and 1631 at $r = 18$ and $r = 20$, and to the family $n = 51 \cdot 2^{r/2-5} - 1$ of asymptotic density 2601/2048. This note explains the problem, how the matrix was found, what was proved, and how to re-check every number from the committed text. The matrices, verifiers, and LaTeX source live at `https://github.com/wustep/maths`.

Contents

1 The problem **2**
- 1.1 Covering radius . . . 2
- 1.2 Why the $r = 10$ cell is load-bearing . . . 2

2 The process **2**
- 2.1 Overnight search . . . 2
- 2.2 What the overnight run missed . . . 3

3 The result **3**
- 3.1 The code . . . 3
- 3.2 Minimality . . . 3
- 3.3 A $(2, 0)$-partition with $p(\boldsymbol{H}) = 10$ . . . 4
- 3.4 The dependent triple . . . 4
- 3.5 Propagation . . . 4
- 3.6 The family . . . 4

4 Verification **5**
- 4.1 How to re-run . . . 5
- 4.2 Two verifiers, opposite algorithms . . . 5
- 4.3 Encoding . . . 5
- 4.4 Independence, honestly . . . 6

5 What is not claimed **6**

6 Where the files live **6**

A The matrix as bits **6**

1 The problem

1.1 Covering radius

Let $C \le \mathbb{F}_2^n$ be a binary linear code of codimension r, so $|C| = 2^{n-r}$. Its *covering radius* is the least R such that every vector of $\mathbb{F}_2^n$ lies at Hamming distance at most R from some codeword. Equivalently, if $\boldsymbol{H}$ is an $r \times n$ parity-check matrix with column set $S \subset \mathbb{F}_2^r$, every syndrome is a sum of at most R columns of $\boldsymbol{H}$. Write $\ell_2(r, R)$ for the least such length.

A binary linear code of length n and dimension $k = n-r$ is written $[n, k]_2$; when the covering radius is R we append it, $[n, k]_2\, R$. So a $[50, 40]_2\, 2$ code is binary of length 50 and dimension 40, has covering radius 2, and hence redundancy $r = 10$.

For $R = 2$ the condition is

$$\{0\} \cup S \cup (S + S) = \mathbb{F}_2^r. \tag{1}$$

A $(2, 0)$-partition of the columns (Definition 3.2 of [1]) is a partition into nonempty blocks such that every syndrome, including 0, is a sum of at most two columns from *distinct* blocks. Write $p(\boldsymbol{H})$ for the number of blocks.

The finite covering density of an $[n, n-r]_2 2$ code is the exact rational

$$\mu = \frac{1 + n + \binom{n}{2}}{2^r}.$$

Green's Problem 40 asks for the asymptotic density

$$f(2) = \liminf_{r \to \infty} \mu\big(\ell_2(r, 2), r\big).$$

For an infinite family of $[n(r), n(r) - r]_2\, 2$ codes, write $\bar{\mu}(2)$ for the lim inf of μ along the family as $r \to \infty$. Green's $f(2)$ takes the lim inf over the *optimal* lengths $\ell_2(r, 2)$, so any family bound $\bar{\mu}(2) \le c$ gives $f(2) \le c$. It is not even known whether $f(2) = 1$. The pre-2025 upper bound $f(2) \le 729/512 \approx 1.42383$ had stood since 1992. Davydov, Marcugini and Pambianco [1] brought it to $2704/2048 \approx 1.32031$ in November 2025, by iterating a construction they call QM_2^2 from a single seed. QM_2^2 grows a short radius-2 covering code into a longer one, provided the seed admits a $(2, 0)$-partition into few enough blocks; its exact input–output formula is recalled in Section 3.5.

1.2 Why the $r = 10$ cell is load-bearing

Their Table 5.1, row $r = 10$, still lists $n = 51$: the Kaikkonen–Rosendahl [2] code of 2003, density $1327/1024 \approx 1.29590$. The entry is not bolded. Even the 2025 paper did not improve it.

Their infinite family (Theorem 5.7) has lengths

$$\Phi(r) = 26 \cdot 2^{r/2-4} - 1 = 2^m(51 + 1) - 1, \qquad m = r/2 - 5.$$

The 52 is the Kaikkonen–Rosendahl 51, plus one. The whole family is that one table cell pushed forward. Theorem 5.7(i) says the 2-partitions of $\boldsymbol{H}_{KR}$ are "very important for the next iterative process".

So a 50-column matrix at $r = 10$ is not a one-cell footnote. If it admits a small enough $(2, 0)$-partition, it re-seeds the family. The paper's computer search found $p(\boldsymbol{H}_{KR}) = 11$.

2 The process

2.1 Overnight search

On 16 August 2026 an automated agent (Codex `gpt-5.6-sol`, Max) ran a targeted simulated anneal seeded from the Kaikkonen–Rosendahl 51-column matrix. The state was a set of exactly

50 nonzero ten-bit columns. An array stored, for every syndrome, how many times it appeared as 0, a singleton, or a pair-sum. Swapping one column changes only its singleton and the 49 pair-sums involving it.

Most proposals were aimed at an uncovered syndrome g: offer g itself, or $g + s$ for a current column s. With a fixed xorshift64 seed, run 0 reached zero uncovered syndromes at proposal 3,600,281.

The same search failed to close prettier one-column gaps: $(r, n) = (8, 25)$ stalled at 3 uncovered syndromes of 256; $(9, 38)$ stalled at 8 of 512; $n = 49$ at $r = 10$ stalled at 7 of 1024. Those are search residues, not lower bounds.

2.2 What the overnight run missed

The overnight folder certified $\ell_2(10, 2) \le 50$ and stopped at the table cell. It never computed a 2-partition, never ran QM_2^2, and still quoted the stale $f(2) \le 1.4238$.

A later session (Claude, afternoon of the same day) treated the matrix as a claim to be verified, not as a trusted input: reconstructed the Kaikkonen–Rosendahl baseline from the hex in Theorem 4.3 of [1], found a $(2, 0)$-partition with $p(\boldsymbol{H}) = 10$, implemented QM_2^2 from equations (4.2) and (4.4), and exhaustively checked the $r = 18$ and $r = 20$ outputs. Two verifiers were written from scratch in two languages, then a third quarantined oracle was run once as a post-hoc diff. The search procedure is not part of the argument. The argument is a finite decidable statement about committed matrices.

3 The result

3.1 The code

claim	value
shape	10×50, $\mathbb{F}_2$-rank 10, 50 distinct nonzero columns
coverage	1024/1024 syndromes
covering radius	exactly 2
density μ	$319/256 = 1.24609375$
minimum distance	$d = 3$
KR baseline	51 columns, 1024/1024, density 1327/1024

The columns, as unsigned integers with bit i (LSB first) equal to row $i + 1$:

1	2	4	15	16	32	65	86	128	173
183	202	212	247	256	297	320	329	341	366
373	381	391	403	438	460	479	491	502	559
576	608	653	734	742	754	771	777	789	821
846	855	869	881	893	897	927	981	1003	1004

The multiplicity histogram of the $1276 = 1 + 50 + \binom{50}{2}$ representations over the 1024 syndromes is {1:859, 2:129, 4:24, 5:9, 6:3}. There are no zeroes. Exactly 973 syndromes are neither 0 nor a column, so the radius is not 1.

3.2 Minimality

Deleting any one of the 50 columns leaves at least 9 syndromes uncovered. The best deletions — columns 381, 479 and 927 — leave exactly 9. So the 50-set is a minimal 1-saturating set in $PG(9, 2)$, equivalently a locally optimal (LO) covering code (the geometric name for a covering-radius-2 column set with no redundant column). It is not padded.

3.3 A $(2,0)$-partition with $p(\boldsymbol{H}) = 10$

block	columns
0	$2, 128, 202, 212, 771, 855, 897, 981$
1	86
2	$381, 893, 1003$
3	$183, 297$
4	$1, 65, 247, 256, 320, 438, 502, 734$
5	$15, 173, 329, 366, 460, 559, 653, 846$
6	$4, 16, 391, 491, 742, 754, 869, 881$
7	$479, 1004$
8	403
9	$32, 341, 373, 576, 608, 777, 789, 821, 927$

All 973 pair-needed syndromes have a cross-block pair. The pair-only histogram over those 973 is {1:821, 2:123, 4:19, 5:8, 6:2}, so 821 syndromes have a unique pair and force their two columns into distinct blocks. Against the paper's $p(\boldsymbol{H}_{KR}) = 11$.

3.4 The dependent triple

There are exactly 10 linearly dependent triples of columns. Exactly one, $(491, 734, 821)$, lies in three distinct blocks (blocks 6, 4, 9). This is the analogue of [1] Theorem 5.2(ii), the ingredient Construction QM_5^2 needs at $r = 28$. That construction is *not* carried out here.

3.5 Propagation

Construction QM_2^2 ([1], Theorem 4.1) takes an $[n_0, n_0 - r_0]_2 2$ code with a 2-partition into $p(\boldsymbol{H}_0)$ blocks, an indicator set $\mathscr{B} = \mathbb{F}_{2^m}$ subject to $n_0 \geq 2^m \geq p(\boldsymbol{H}_0)$, and $\boldsymbol{D} = \boldsymbol{D}_1(2)$, and returns

$$n = 2^m(n_0 + 1) - 1, \qquad r = r_0 + 2m, \qquad p(\boldsymbol{H}_C) \leq 2^{m+1} + 1.$$

With $n_0 = 50$ and $p(\boldsymbol{H}_0) = 10$ the inequality permits exactly $m = 4$ and $m = 5$ ($2^6 = 64 > 50$). Both outputs were enumerated completely:

	r	n	coverage	density	published
seed	10	50	1024/1024	319/256	51 (KR 2003)
$m = 4$	18	815	262144/262144	332521/262144	831
$m = 5$	20	1631	1048576/1048576	1330897/1048576	1663

A second pair of matrices, from a different indicator allocation, also verifies (see `result/data/alt/`). Indicator assignment is a free choice, so two allocators give two codes.

3.6 The family

Iterating from $(r, p) = (10, 10)$ with the paper's bound $p(\boldsymbol{H}_C) \leq 2^{m+1} + 1$, the even $r \leq 64$ reachable by QM_2^2 are

$$10, 18, 20, 30, 32, 34, 36, 38, 40, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64.$$

The even r *not* reachable this way are $12, 14, 16, 22, 24, 26, 28, 42, 44$. Those gaps are real. The paper fills its own $r = 22, 24, 26$ with QM_3^2 and $r = 28$ with QM_5^2; the analogues are not implemented here.

Closed form on the reachable even r:

$$n = 51 \cdot 2^{r/2-5} - 1, \qquad \mu = \frac{51^2}{2^{11}} - \frac{51}{2^{r/2+1}} + \frac{1}{2^r} \;\nearrow\; \frac{2601}{2048} \approx 1.27002.$$

Hence $\bar{\mu}(2) \le 2601/2048$, and since $f(2)$ is a lim inf over r, also $f(2) \le 2601/2048$. Against the literature:

	$\bar{\mu}(2)$	source
pre-2025	$729/512 \approx 1.42383$	standing since 1992
arXiv:2511.02542	$2704/2048 \approx 1.32031$	seeded by KR $n_0 = 51$
this seed	$2601/2048 \approx 1.27002$	seeded by $n_0 = 50$

Rows past $m = 5$ inherit the paper's $p(\boldsymbol{H}_C)$ bound rather than computed partitions. That is the weakest link.

4 Verification

4.1 How to re-run

The code is the public repository `https://github.com/wustep/maths`. Clone it, then run the pipeline from the covering-result directory:

```
git clone https://github.com/wustep/maths.git
cd maths/problems/covering/result
./run_all.sh
```

Needs `python3` and `rustc`. After the clone, no network, no packages, no RNG. Seventy-four named assertions, exit 0, output byte-identical across runs (about two seconds). Direct links: `https://github.com/wustep/maths/tree/main/problems/covering/result` and `result/run_all.sh`.

4.2 Two verifiers, opposite algorithms

`verify/verify.py` is pair-driven: walk all $\binom{n}{2}$ pairs, mark the syndromes they hit, then sweep 2^r for anything unmarked. `verify/verify.rs` is syndrome-driven: for each of the 2^r syndromes, scan the columns and test membership of $s \oplus h$. It then runs its own pair loop and refuses to report unless the two verdicts agree. They also differ in parser, rank pivot (lowest vs highest set bit), and exact rational arithmetic. `run_all.sh` diffs their fact dumps.

Both read the matrix from *text*. Neither consults a stored certificate. The partition JSON is used only for block labels; the columns themselves are re-derived from the matrix.

4.3 Encoding

Everything in the repo is LSB-first: bit i of a column integer is row $i + 1$. The hex listing of M_{KR} in [1] Theorem 4.3 is the opposite (MSB-first, row 1 as the high bit). Reconstructing $\boldsymbol{H}_{KR} = [I_{10} \mid M_{KR}]$ reverses the ten bits. A from-scratch verifier that misses the reversal fails on the Kaikkonen–Rosendahl baseline *and only there*, which is a useful signal. The reversal is pinned by the paper's own Theorem 5.2(ii), $h_5 + h_{27} + h_{29} = 0$, asserted in the builder as a canary.

A reader using the opposite convention gets the bit-reverse of every column, an $\mathbb{F}_2$-linear relabelling of $\mathbb{F}_2^{10}$, and every claim here is preserved.

4.4 Independence, honestly

Both verifiers were written in one session by one author. Two languages, two algorithms, two parsers, no shared code — but not two people. A third, quarantined oracle was run once after both were green and agreed everywhere; it is not in the tree.

5 What is not claimed

- 50 is not shown optimal. Sphere-covering gives only $\ell_2(10, 2) \geq 45$, since $1+n+\binom{n}{2} \geq 1024$ fails at $n = 44$, which gives only 991. Minimality says this particular 50-set has no redundant column, which says nothing about a different 49-set.
- The $n = 49$ run leaving 7 uncovered syndromes is search residue, not a lower bound.
- The gaps $(8, 25)$ and $(9, 38)$ remain open.
- Only $m = 4$ and $m = 5$ are exhaustively verified. The rest of the family inherits $p(\boldsymbol{H}_C) \leq 2^{m+1} + 1$ from [1] (4.4).
- Nothing is claimed for $r \in \{12, 14, 16, 22, 24, 26, 28, 42, 44\}$.
- $f(2) \leq 2601/2048$ is an upper bound only. Whether $f(2) = 1$ is untouched.
- Priority is unresolved. Table 5.1 of [1] is "best as far as the authors know". A $\ell_2(10, 2) \leq 50$ may already exist in ACCT proceedings. See `result/PRIORITY.md` in the repository.

6 Where the files live

This document is self-contained. Everything named below is in the public repository `https://github.com/wustep/maths`, under `problems/covering/` on `main`.

in the repo	what
`https://github.com/wustep/maths`	the repository
`result/NOTE.md`	repo companion
`result/note.tex`	arXiv-style draft of the same claims
`result/run_all.sh`	the pipeline
`result/data/H_r10_n50.txt`	the matrix
`result/data/partition_p10.json`	the 10-block partition
`WALKTHROUGH.md`	how the overnight search felt
`ATTACK.md`	chronological log

The matrix is `H_r10_n50.txt`.

A The matrix as bits

Rows 1 to 10 from top to bottom, columns left to right. Bit i (LSB) of each integer column is row $i + 1$.

```
10010010011001010110111100110100100011110111111110
01010001001101000001001110111100011110001100001010
00110001011011000011111011101100111000111110101101
00010000010100010101010001110100110001001000101011
```

```
00001001001011000010110110101000010100110101101100
00000100011001010001110010011101001100010011100011
00000011000111001111110001111011011100001111100111
00000000111111000000001111111000111100000000011111
00000000000000011111111111111100000001111111111111
00000000000000000000000000001111111111111111111111
```

References

[1] A. A. Davydov, S. Marcugini, F. Pambianco, *New upper bounds for binary linear covering codes*, arXiv:2511.02542 [cs.IT], November 2025.

[2] M. Kaikkonen, C. Rosendahl, *New covering codes from an ADS-like construction*, IEEE Trans. Inform. Theory **49** (2003), no. 7, 1809–1812.

[3] G. Cohen, I. Honkala, S. Litsyn, A. Lobstein, *Covering Codes*, North-Holland Mathematical Library **54**, Elsevier, Amsterdam, 1997.

[4] B. Green, *100 open problems*, Problem 40. `https://people.maths.ox.ac.uk/greenbj/papers/open-problems.pdf`.