Counting Full-Rank Selections over $\mathbb{F}_2$

A Complexity Boundary between Counting Modulo 2 and Modulo 4

Zakaria Gharzouli
Residual Labs, Moissy-Cramayel, France
ORCID: 0009-0005-7604-4709 • contact@residual-labs.fr
Preprint v1.1 - November 26, 2026 - CC BY 4.0
DOI 10.5281/zenodo.24557901

Abstract

We count the ways to select one vector from each of several explicit binary sets so that the selected vectors span the whole space. The least significant bit can be computed in polynomial time by a single Gaussian elimination. Deciding whether the second bit is one is $\oplus$P-complete, even for the edge planes of a simple graph with no isolated vertices and at least two edges. The proof relates the count modulo four to the number of proper 3-colorings up to permutation of colors. For arbitrary domains, every fixed bit at index one or higher is $\oplus$P-complete. A polynomial-time algorithm for any of these bits would be equivalent to P = $\oplus$P. We give reproducible computations and a joining formula.

Keywords — Modular counting, matrix rank, graph 3-coloring, Möbius inversion, $\oplus$P.

1. The counting problem

Let $U = \mathbb{F}_2^r$. An instance supplies m explicit sets $A_1,\dots,A_m \subseteq U$. Each set specifies the admissible row vectors at one position in an m-by-r matrix. We count ordered tuples.

$$N = \#\{(v_1,\dots,v_m) \in A_1 \times \cdots \times A_m : \langle v_1,\dots,v_m\rangle = U\}. \quad (1)$$

The bit at index j is $\lfloor N / 2^j \rfloor \bmod 2$. Indices start at zero. An empty domain gives N = 0. If m = 0, N is 1 when r = 0 and 0 otherwise. If r = 0 and all domains are nonempty, every domain is {0} and N = 1. Vectors are given by dense binary coordinates.

Our main result concerns the transition from parity to counting modulo four. An odd N certifies that a full-rank selection exists. An even N can also be positive, as the examples in Section 5 illustrate.

2. Parity by Gaussian elimination

Theorem 1. Put $a_i = \bigoplus_{v \in A_i} v$ and $b_i = 1 + (|A_i| \bmod 2)$ in $\mathbb{F}_2$. The count N is odd exactly when the system $a_i \cdot u = b_i$, for $1 \le i \le m$, has a unique solution $u \in U$.

Proof. A tuple of rank s has exactly 2^{r-s} vectors u orthogonal to all its rows. This number is odd exactly when s = r. Interchanging the sums over tuples and over u gives

$$N \equiv \Sigma_{u \in U} \prod_{i=1}^{m} |\{v \in A_i : u \cdot v = 0\}| \pmod 2. \quad (2)$$

Over $\mathbb{F}_2$, the indicator of $u \cdot v = 0$ equals $1 + u \cdot v$. The ith factor is therefore $(|A_i| \bmod 2) + u \cdot a_i$. The product is one precisely at the solutions of the stated system. A consistent affine system of rank ρ has $2^{r-\rho}$ solutions. Its contribution modulo two is one exactly when ρ = r. □

Algorithm. Scan each set to compute its XOR and cardinality modulo two. Row-reduce the augmented matrix [a | b] over $\mathbb{F}_2$. Output 1 if the system is consistent and the coefficient matrix has rank r, and 0 otherwise. For dense vectors, the straightforward cost is $O(r \Sigma_i |A_i| + mr \min(m,r))$ bit operations. The square case also follows from multilinearity of the determinant.

Worked example. For r = 2, take $A_1 = \{e_1\}$, $A_2 = \{e_2\}$, $A_3 = \{e_1, e_2, e_1+e_2\}$. Their XORs are e_1, e_2, 0. All three cardinalities are odd. The equations are $u_1 = 0$, $u_2 = 0$, $0 = 0$. The unique solution gives N mod 2 = 1. The first two rows already span U, so N = 3.

Exact counting formula. For $W \le U$, the number of tuples whose rows all lie in W is the product of the $|A_i \cap W|$. Möbius inversion on the lattice of subspaces isolates the tuples spanning U [1].

$$N = \Sigma_{W \le U} (-1)^c \, 2^{c(c-1)/2} \prod_{i=1}^{m} |A_i \cap W|, \; c = r - \dim W. \quad (3)$$

The Möbius coefficient of [W,U] is the one for the subspace lattice of a c-dimensional quotient. In the product of the c factors $(1 + z\,2^j)$ for $j = 0,\dots,c-1$, the coefficient of z^k is the number of k-dimensional subspaces multiplied by $2^{k(k-1)/2}$. This is the Gaussian binomial identity, obtained by induction on c. At $z = -1$ the product vanishes for $c > 0$, as required by the Möbius recurrence. For c = 0 the coefficient is one.

3. The second bit and graph colorings

Theorem 2. Let $G = (V,E)$ be a simple graph with no isolated vertices and at least two edges. Associate to each edge uv the set $A_{uv} = \langle e_u, e_v \rangle = \{0, e_u, e_v, e_u+e_v\}$ in $U = \mathbb{F}_2^{|V|}$. If $\chi_G(3)$ counts proper 3-colorings with labeled colors, then

$$N(G) \equiv 2 \cdot ((\chi_G(3) / 6) \bmod 2) \pmod 4. \quad (4)$$

Division by six is performed over the integers before reduction modulo two. The labeled color count is always even on this family.

Proof. In (3), the terms of codimension at least three vanish modulo four by their coefficient. When $W = U$ each factor equals four. In codimension one each plane A_{uv} meets W in dimension at least one. With at least two edges, the product is divisible by four. The surviving candidate terms have codimension two and coefficient +2. Their product is odd exactly when $W \cap A_{uv} = \{0\}$ for every edge.

For such W, an isomorphism $U/W \to \mathbb{F}_2^2$ associates to each vertex the image of its basis vector. On an edge, the images are independent, hence distinct and nonzero. Every vertex belongs to an edge. The three nonzero vectors serve as the three colors. Conversely, a proper coloring defines a linear map through its values on the basis. It is surjective because an edge gives two independent images. Two surjective maps with the same kernel differ by an element of $GL(2,2)$, which realizes every permutation of the three nonzero vectors. Each admissible kernel therefore accounts for exactly six labeled colorings. □

With one edge and two vertices, $N = 0$ and $\chi_G(3)/6 = 1$. An isolated vertex forces a zero column and prevents full rank.

Worked example. For a triangle, the domains are the three coordinate planes in $\mathbb{F}_2^3$. A codimension-two kernel is a line. To avoid all three planes, its unique nonzero vector must be $(1,1,1)$. There is one admissible kernel, so (4) gives 2 modulo 4. For the exact count, codimensions 0, 1, 2 and 3 contribute 64, $-(3\cdot16+4\cdot8)$, $2\cdot(3\cdot4+3\cdot2+1)$, -8 respectively. Their sum is 14, also obtained by enumerating the 64 selections.

Computation modulo four. For arbitrary explicit sets, (3) at this precision uses only subspaces of codimensions 0, 1 and 2. For codimension two, one can enumerate independent pairs in the dual and keep one reduced basis for each dual subspace. The number of codimension-two candidates is $(2^r-1)(2^r-2)/6$. This method is exponential in r. More generally, modulo 2^s, a term of (3) can be dropped when $c(c-1)/2 \geq s$. For edge domains, (4) provides a separate check by enumerating colorings.

4. Complexity classification

Theorem 3. Deciding whether the bit of N at index 1 is one is $\oplus$P-complete under deterministic polynomial-time many-one reductions, including for the edge planes of Theorem 2.

Hardness. The class $\oplus$P decides the parity of the number of witnesses accepted by a polynomial-time verifier. The parity version of 3-SAT is complete for this class. Encode an AND, OR and NOT verification circuit with one equivalence per gate, and require the output to be true. Each equivalence is expressible in CNF of width at most three, allowing unit and binary clauses. Each witness uniquely determines all gate variables, preserving the count.

Start with a 3-CNF formula Φ. Construct a distinguished triangle T, F, B and fix the colors of the three vertices. For each variable x, create vertices x and ¬x, adjacent to each other and both adjacent to B. They have the colors of T and F in an order that encodes a truth assignment. For each clause, add an auxiliary vertex adjacent to F and every literal vertex in the clause. If the clause is true, the auxiliary vertex sees two colors and has one choice. If the clause is false, its neighbors all have the color of F and it has two choices. The auxiliary vertices form an independent set. An assignment that violates t clauses gives exactly 2^t extensions. The parity of the normalized color count is therefore the parity of #3SAT(Φ).

An empty clause gives two choices and a tautology gives one. Repeated literals do not change the colors seen by the auxiliary vertex. The triangle normalizes the orbits. Removing duplicate edges gives a simple graph with no isolated vertices and at least three edges. The dense encoding of the domains has polynomial size. By (4), the second bit gives the required parity.

Membership. The second bit of N equals C(N,2) mod 2. The verifier reads two selections in strict lexicographic order. It reads a fixed-width binary index per domain, rejects out-of-range codes, and checks each rank. Each unordered pair has exactly one accepted witness. □

Higher bits. In $\mathbb{F}_2[z]$, the binary expansion of N factors $(1+z)^N$. The factor for each set bit j is $1+z^{2^j}$. The coefficient of z^{2^j} gives bit j. For fixed j, the verifier reads 2^j distinct selections in canonical order. For $j \geq 2$, add $j-1$ disjoint dimensions with two domains {e} and {0,e} in each. The first forces the new dimension, and the second doubles the count. The total factor 2^{j-1} transfers the hardness of bit 1. This covers arbitrary domains and fixed indices.

Exact count. The rank verifier gives membership in #P. In a square matrix, restricting each domain to the basis vectors indicated by a 0–1 matrix realizes its permanent, since the selected indices must be distinct. Valiant proves #P-completeness under polynomial-time Turing reductions and gives a polynomial-time algorithm for the permanent modulo 2^k with fixed k [2].

Existence. If the domains are nonempty, intersect the linear matroid on pairs (i,v) with the partition matroid allowing at most one pair per row i. A maximum independent family is found in polynomial time [3,14]. If its size is r, fill the remaining positions with arbitrary admissible vectors.

5. Computations and reproduction

Graph	Vertices	Edges	$\chi(3)$	Orbits	Exact N	N mod 4
Triangle	3	3	6	1	14	2
Cycle C_4	4	4	18	3	30	2
Complete graph K_4	4	6	0	0	1,960	0

The table compares three computations. The first enumerates selections and checks rank. The second applies (3) over subspaces. The third enumerates colorings and applies (4). For the unit clause (x), the gadget has one satisfying assignment and gives three colorings with the triangle fixed, namely 1 + 2. Its full-rank count is 2 modulo four.

Rank 256. Fix $\{e_1\}$ and $\{e_2\}$, then set $A_i = \{e_i, e_i+e_1, e_i+e_2\}$ for $3 \le i \le 256$. Every selection is triangular with unit diagonal. The exact count is 3^{254}. The XOR of domain $i \ge 3$ is $e_i+e_1+e_2$. The 256 homogeneous equations have rank 256. Gaussian elimination recovers parity without traversing the 3^{254} selections.

Protocol. The Python scripts use the standard library, fixed random seeds and exact integers. Run python -B -S code/reproduire.py calculs, followed by python -B -S code/verifier_raccordement.py, from the preprint directory. The log covers 210 row models, 45 labeled graphs on three or four vertices, 160 gadget formulas and 5 complete reductions. It also covers the joining checks in the appendix. JSON and CSV files provide the computed values. Tests compare methods on finite cases. The proofs above establish the theorems.

6. Applications of the model

Coding and communications. For a binary coding matrix with selectable rows, N counts configurations that retain r independent directions. Under independent uniform selection in each nonempty A_i, $N / \prod_i |A_i|$ is the exact probability of full rank. Correlated losses and nonuniform selection require an appropriate probability model.

Sensing and control. If measurements are binary linear forms, full rank guarantees that exact measurements identify the binary state. The small example in Section 2 already models two fixed measurements and one selectable measurement. Physical sensors require an analysis of geometry, noise and numerical stability.

Distributed computing and combinatorial validation. The joining formula in the appendix counts compatible completions across two blocks of observations. The identities modulo two and four give consistency checks for counting solvers and structured matrix generators. These applications concern the binary model. The reported experiments use constructed instances.

Appendix A. Joining two blocks

Let $T \le U$ be the span of fixed shared rows. The left and right blocks have messages $m_L(R)$ and $m_R(S)$ on image spaces R,S containing T. Each weight counts local configurations with the indicated image. Assume the local choices are independent once shared observations are fixed. There are h additional unconstrained hidden rows in U. Put $d = \dim(R+S)$ and $q = 2^h$.

$$M(R,S) = q^d \prod_{i=0}^{r-d-1} (q - 2^i). \quad (5)$$

The h rows projected to $U/(R+S)$ form an h-by-$(r-d)$ matrix. Its columns must be independent, giving $q-2^i$ choices for each successive column. Each projected row has 2^d lifts. If $h < r-d$, a factor vanishes. The empty product is one. The total is $\Sigma_{R,S}\, m_L(R)\, M(R,S)\, m_R(S)$.

Integer factorization. In the interval $[T,U]$, let $Z_{R,W}$ be 1 if $R \le W$ and 0 otherwise. Möbius inversion gives a join matrix factorization [4].

$$M = Z\, \mathrm{diag}_{W \in [T,U]}(\mu(W,U) \cdot 2^{h \dim W})\, Z^{\top}. \quad (6)$$

Indeed, for W containing $R+S$, there are $2^{h \dim W}$ choices of hidden rows all in W. Sorting by dimension makes Z triangular with unit diagonal, hence invertible over $\mathbb{Z}$. The Smith factors are the powers $2^{h \dim W + c(c-1)/2}$ in increasing exponent order, with $c = r - \dim W$. Over $\mathbb{Q}$, M has full rank. When $h = 0$, only codimensions 0 and 1 have odd factors, so the rank modulo two is $2^{r-\dim T}$.

Computation with 256 private positions. The left block has 128 rows with first coordinate zero. The right block has 128 rows with second coordinate zero. The target rank is 6. Each of the first two columns has 2^{128+h} possibilities, including zero. Among nonzero pairs exactly $2^h - 1$ have equal columns, because their common support lies among the hidden rows. Choosing the other four columns by rank extension gives

$$((2^{128+h}-1)^2 - (2^h-1)) \prod_{i=2}^{5} (2^{256+h} - 2^i). \quad (7)$$

This independent formula agrees with both the join and (3) for $h = 0$ and $h = 2$. The exact counts have 1,280 and 1,292 bits respectively. The verifier checks 45 join matrices, 36 Smith forms from minors, and four shared values for non-coordinate observations.

7. Related work and conclusion

Prescribed support. The count in Theorem 2 is a case of the model of Lewis et al. [5, §5]. The pattern allows two entries in each row, at the endpoints of its edge, and sets all others to zero. Their Proposition 5.1 gives a congruence in powers of q−1, with modulus 1 at q = 2. Their formulas in Section 2 concern diagonal zeros. Congruence (4) concerns edge planes at q = 2, modulo four.

Critical problem. The kernels in (4) avoid the set of e_v and e_u+e_v for $uv \in E$. This is an instance of the critical problem of Crapo and Rota [8,9]. These vectors represent the graphic matroid of the cone over G. Its characteristic polynomial at 4 counts proper 4-colorings of the cone with the added vertex fixed to a particular color, giving $\chi_G(3)$. The kernel-to-coloring correspondence is thus classical. Theorem 2 relates it to the integer full-rank count reduced modulo four.

Modular counting. Classifications of modular CSPs and graph homomorphisms [10–12,18] provide the broader context. Theorem 90 of the arXiv version of Focke et al. [13] already implies $\oplus$P-completeness of list 3-coloring modulo two. Groenland et al. [17, §4.3 of the full version] study the same coloring orbits and encode lists using reference cliques. Here a single fixed triangle suffices by joining each vertex to its forbidden colors. Empty lists and isolated vertices are handled through their parity factors. These earlier results provide a route to the hardness of normalized colorings, then to bit 1 via (4). Section 4 gives a direct proof.

Matroids and parity. In the square case, full-rank selections are exactly common bases of a linear matroid and a partition matroid [3,14]. The Cauchy–Binet framework for common bases [15] covers parity over the binary field in this case. In the rectangular case several common bases can contribute to one tuple. Theorem 1 counts whole tuples.

The completeness result classifies the complexity of the second bit. A polynomial-time algorithm for this bit is equivalent to P = $\oplus$P. No separation of complexity classes is established here, and P versus NP remains open [7]. The targeted bibliography search establishes no priority.

The project was motivated by the AI-assisted work of Alpöge and Claude, which led to an announced proof of the Riemann hypothesis [6,16]. The theorems in this paper rely on finite algebraic and combinatorial arguments.

Declarations

Author. Zakaria Gharzouli conceived the research program and commissioned the revision. OpenAI Codex assisted with the mathematical arguments, computations, code and writing. Reproduction scripts are supplied and cross-checks have been rerun. The author personally validated the manuscript. He declares no funding and no conflicts of interest.

Status and license. This preprint has not been peer reviewed. The AI-assisted audit and tests do not constitute an independent human review. License CC BY 4.0.

References

[1] G.-C. Rota (1964). On the foundations of combinatorial theory I. Theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 2, 340–368. DOI · Text

[2] L. G. Valiant (1979). The complexity of computing the permanent. *Theoretical Computer Science* 8(2), 189–201. DOI

[3] J. Edmonds (1979). Matroid intersection. *Annals of Discrete Mathematics* 4, 39–49. DOI

[4] M. Mattila and P. Haukkanen (2013). Determinant and inverse of join matrices on two sets. *Linear Algebra and its Applications* 438(10), 3891–3904. DOI · arXiv:1110.4953

[5] J. B. Lewis, R. I. Liu, A. H. Morales, G. Panova, S. V. Sam and Y. X. Zhang (2011). Matrices with restricted entries and q-analogues of permutations. *Journal of Combinatorics* 2(3), 355–395. DOI · arXiv:1011.4539

[6] L. Alpöge (November 20, 2026). *Remark*, note accompanying the release of the Riemann hypothesis work with Claude. Author's note

[7] Clay Mathematics Institute. *P vs NP*. Accessed November 26, 2026. Problem page

[8] H. H. Crapo and G.-C. Rota (1970). *On the Foundations of Combinatorial Theory: Combinatorial Geometries*. MIT Press, chapter 16. Publisher

[9] J. P. S. Kung (1996). Critical problems. In *Matroid Theory*, Contemp. Math. 197, AMS, 1–127. Text

[10] J. Faben (2008). The complexity of counting solutions to Generalised Satisfiability Problems modulo k. arXiv:0809.1836

[11] J. Faben and M. Jerrum (2015). The Complexity of Parity Graph Homomorphism: An Initial Investigation. *Theory of Computing* 11(2), 35–57. DOI

[12] A. A. Bulatov and A. Kazeminia (2024). Complexity Classification of Counting Graph Homomorphisms Modulo a Prime Number. *SIAM Journal on Computing*, STOC22-182–223. DOI · arXiv:2106.04086

[13] J. Focke, L. A. Goldberg, M. Roth and S. Živný (2021). Counting Homomorphisms to K_4-Minor-Free Graphs, Modulo 2. *SIAM Journal on Discrete Mathematics* 35(4), 2749–2814. DOI · arXiv:2006.16632

[14] R. Rado (1942). A theorem on independence relations. *Quarterly Journal of Mathematics* os-13(1), 83–89. DOI

[15] K. Matoya and T. Oki (2022). Pfaffian Pairs and Parities: Counting on Linear Matroid Intersection and Parity Problems. *SIAM Journal on Discrete Mathematics* 36(3), 2121–2158. DOI

[16] M. Rubinstein (2026). Alpöge/Claude proof of the Generalized Riemann Hypothesis. Accessed November 26, 2026. Note and checks

[17] C. Groenland, I. Mannens, J. Nederlof and K. Szilágyi (2022). Tight Bounds for Counting Colorings and Connected Edge Sets Parameterized by Cutwidth. *STACS*, 36:1–36:20. DOI · arXiv:2110.02730

[18] A. A. Bulatov and A. Kazeminia (2026). Modular Counting over 3-Element and Conservative Domains. *STACS*, 22:1–22:20. DOI