

# SUBEXPONENTIAL WEIGHTED COUNTING AND EXACT OPTIMISATION

DEEP BHATTACHARJEE

**ABSTRACT.** *The deterministic algorithm of OpenAI that counts the satisfying assignments of a  $k$ -CNF in time  $2^{o(n)}$  is reproduced in full and extended from counting to sums of local weights with values in any commutative  $\mathbb{Q}$ -algebra. The extension gives deterministic  $2^{o(n)}$  algorithms for exact Max-Cut and Max- $k$ -SAT on sparse instances, for the independence polynomial, the maximum independent set and the knapsack problem with a conflict graph on every graph, for 0-1 integer programs whose rows involve boundedly many variables, for Vertex Cover in time  $2^{o(k)} + \text{poly}(n)$ , for the density of states of Ising and Potts models on bounded-degree graphs, for the Tutte polynomial on the hyperbolas  $(x-1)(y-1) = q$ , for the minimum distance and the maximum-likelihood decoding of codes with sparse parity checks, for Max-Cut and Max- $k$ -SAT at every average degree  $o(\sqrt{\log n / \log \log n})$ , for Max-Cut on dense graphs close to blow-ups of graphs with  $o(n / \log n)$  vertices, and for every problem in SNP.*

## 1. INTRODUCTION

**1.1. The theorem of OpenAI.** On 25 October 2026 OpenAI released a preprint [OAI26] with a deterministic algorithm that, for every fixed  $k \geq 2$ , counts the satisfying assignments of a  $k$ -CNF formula with  $n$  variables and  $m$  clauses in time

$$(n + m)^{O_k(1)} 2^{O_k(n \log \log n / \sqrt{\log n})} \quad (1.1)$$

and polynomial space. Since  $2^{O(n \log \log n / \sqrt{\log n})} = 2^{o(n)}$ , the exponential time hypothesis of Impagliazzo and Paturi [IP01], which asserted that 3-SAT needs time  $2^{\Omega(n)}$ , is false, and with it the strong hypothesis [IPZ01], the counting hypothesis #ETH [DHMTW14], and the gap and nondeterministic variants. The algorithm is algebraic. It writes the count as a sum over the Boolean cube of a product of local polynomials, replaces every variable occurrence by a linear form in a parameter  $T$  chosen so that at each of  $Q$  sample points the factors separate into  $r$  groups after a short list of integer measurements of the assignment has been fixed, recovers the Taylor coefficients of the resulting univariate polynomial at the sample points from Boolean substitutions, and interpolates back to  $T = 0$ . The present paper owes its existence to that preprint, and every step of the algorithm that is used below is OpenAI's.<sup>1</sup>

**1.2. What this paper adds.** The counting routine of [OAI26] is stated there for factors with rational coefficients, with the remark that it works over any commutative  $\mathbb{Q}$ -algebra. We take that remark seriously and run the algorithm over truncated polynomial rings  $\mathbb{Q}[W]/(W^N)$ . A factor may then be  $W^{w(x)}$  for an integer weight function  $w$  of a few variables, and the sum over

---

*Date:* 27 October 2026.

*2020 Mathematics Subject Classification.* Primary 68Q25; Secondary 68Q17, 68W40, 05C31, 05C85, 82B20.

*Key words and phrases.* exponential time hypothesis, weighted counting, independence polynomial, vertex cover.

<sup>1</sup>The preprint is unrefereed. For this reason the algorithm and its analysis are reproduced here in full, in Sections 3 and 4, so that no statement of this paper depends on a text outside it. The reproduction follows the preprint closely, with the notation of this paper, and the reader will find the corresponding statements of [OAI26] named in footnotes. The preprint also contains a Lean 4 formalisation of its main theorem; it is not used here.

the cube becomes a generating polynomial: its coefficient of  $W^j$  counts the assignments of total weight  $j$ . One number has turned into a profile, and the profile contains an optimum. The cost of the change is nil in the exponent, and polynomial in  $N$  elsewhere.

The technical statement is Theorem 5.3: for fixed  $k$  and  $c$ , a system of at most  $cn$  factors, each depending on at most  $k$  of the  $n$  Boolean variables and taking values in a commutative  $\mathbb{Q}$ -algebra  $\mathcal{A}$ , is summed over  $\{0, 1\}^n$  in

$$2^{O_{k,c}(\rho(n))} \text{poly}(n) \text{ operations in } \mathcal{A}, \quad \rho(n) = \frac{n(\log \log n)^2}{\sqrt{\log n}}. \quad (1.2)$$

The exponent is the one OpenAI obtain in their Sections 3–6 before the refinement that removes one logarithmic factor; we do not reproduce that refinement, because none of the consequences below needs it.

From Theorem 5.3 we deduce, with complete proofs:

- (i) for every fixed  $k$  and  $c$ , the *profile* of a constraint satisfaction instance with  $n$  Boolean variables and at most  $cn$  constraints of arity at most  $k$ , that is the number of assignments satisfying exactly  $j$  constraints for every  $j$ , in time  $2^{O_{k,c}(\rho(n))}$  (Theorem 6.1); in particular the exact optimum of Max- $k$ -SAT and of Max-Cut on sparse instances, together with the number of optimal solutions (Corollaries 6.2 and 6.4);
- (ii) the independence polynomial of *every* graph on  $n$  vertices, hence its independence number, clique number, vertex cover number and the number of maximum independent sets, in time  $2^{O(n(\log \log n)^2/(\log n)^{1/4})}$  (Theorem 6.7); the proof combines Theorem 5.3 with a branching rule on vertices of large degree whose cost we bound explicitly; the same branching carries any fixed number of vertex weights at once, which solves the knapsack problem with an arbitrary conflict graph and polynomially bounded weights, and its multidimensional form, exactly (Theorem 6.10, Corollary 6.11);
- (iii) Vertex Cover with parameter  $k$  in time  $2^{O(k(\log \log k)^2/(\log k)^{1/4})} + \text{poly}(n)$ , that is  $2^{o(k)}$  up to a polynomial, through the kernel of Nemhauser and Trotter (Theorem 6.13); the lower bound  $2^{\Omega(k)}$ , which was a standard consequence of the exponential time hypothesis [CFKL15, Theorem 14.6], therefore fails;
- (iv) for every fixed  $\Delta$ , the density of states of the Ising model with integer couplings and fields, and of the  $q$ -state Potts model, on graphs of maximum degree  $\Delta$  with  $n$  vertices, in time  $2^{O_{\Delta,q}(\rho(n))}$  (Theorem 6.18); every partition function and every ground state follows;
- (v) the minimum dominating set and the matching polynomial of a bounded-degree graph, the permanent of a sparse integer matrix, and the minimum hitting set and the maximum split of a sparse set system with sets of bounded size, in time  $2^{O(\rho(n))}$  (Theorems 6.20 and 6.21);
- (vi) for fixed domain size and arity, every sparse constraint system over a finite domain, hence every partition function  $Z_A(G)$  of a fixed symmetric rational matrix  $A$  and the number of homomorphisms into a fixed graph, on bounded-degree graphs, in time  $2^{O(\rho(n))}$  (Theorem 6.15, Corollary 6.16);
- (vii) the satisfying assignments of a sparse  $k$ -CNF formula of each Hamming weight, and those whose weight lies in a prescribed residue class, in time  $2^{O(\rho(n))}$  (Corollary 6.3, Proposition 6.23);
- (viii) further families worked out in Section 7: several linear statistics of the solutions of a sparse constraint system at once, with equations and congruences imposed (Theorem 7.1); Max- $k$ -Lin modulo  $p$ , Max- $q$ -Cut and balanced  $q$ -partitions (Corollary 7.4); on bounded-degree graphs,  $f$ -factors, even and odd subgraphs, orientations with prescribed indegrees, proper edge colourings and nowhere-zero flows by weight (Theorem 7.5), and the Tutte polynomial on every hyperbola  $(x-1)(y-1) = q$  with  $q$  a positive integer, which contains

the chromatic and flow polynomials at  $q$  and the  $q$ -state Potts model (Theorem 7.10); the weight enumerator, the minimum distance and maximum-likelihood decoding of a binary code with sparse parity checks (Theorem 7.12); matchings, transversals and colourings of bounded-degree hypergraphs (Theorem 7.14); the weighted output profile of a linear-size circuit of bounded fan-in (Theorem 7.15); and the optimum and the number of optimal solutions of a 0-1 or bounded integer program with  $O(n)$  rows, each involving a bounded number of variables with coefficients of any size, and a fixed number of further dense rows with polynomially bounded coefficients, which contains exact cover with bounded element frequency and the partially ordered knapsack problem (Theorem 7.17, Corollary 7.18), all in time  $2^{O(\rho(n))}$ ;

- (ix) the same for systems whose density grows with  $n$ : with parameters that follow the ratio  $c$  of factors to variables, the exponent is  $O(n g(\theta))$  with  $\theta = ck\sqrt{\log \log n / \log n}$  and  $g(x) = x \log^{3/2}(1/x)$  (Theorems 7.19 and 7.20), so that the maximum cut, the maximum bisection and the Ising and Potts ground states of a graph of average degree  $d(n)$ , the exact optimum of Max- $k$ -SAT with  $d(n)n$  clauses and that of an integer program with  $d(n)n$  sparse rows are found in time  $2^{o(n)}$  whenever  $d(n) = o(\sqrt{\log n / \log \log n})$  (Corollaries 7.21 and 7.26); and, through the complement, the cut profile of every graph whose complement has that many edges, a class that contains graphs with  $\Theta(n^2)$  edges (Corollary 7.22), and more generally the cut profile of every graph that differs in that many pairs from a blow-up of a graph on  $o(n/\log n)$  vertices, or from a weighted graph whose weights are a quadratic form in  $s$  vertex weights with  $s \log(n\Omega) = o(n)$  (Theorem 7.23 and Corollary 7.24); a count shows that these classes contain almost no graphs (Remark 7.25);
- (x)  $k$ -SAT lies in the class SE of problems solvable in time  $2^{\varepsilon n}$  poly for every  $\varepsilon > 0$ , by the sparsification lemma of Impagliazzo, Paturi and Zane [IPZ01], and therefore every problem of the class SNP of Papadimitriou and Yannakakis [PY91] lies in SE (Theorem 8.4).

Items (i)–(ix) concern problems for which no  $2^{o(n)}$  algorithm was known and for which such an algorithm was excluded under the exponential time hypothesis by standard reductions. The best exact algorithms for the maximum independent set run in time  $2^{0.1996n}$  [XN17], for Max-Cut in time  $2^{\omega n/3}$  where  $\omega$  is the exponent of matrix multiplication [Wil05], and for the permanent and the perfect matchings of sparse graphs in time  $2^{n-\Omega(n/\Delta)}$  [CP15]; each of these bounds is exceeded here asymptotically.

**1.3. What this paper does not claim.** The algorithm counts sums of products of *local* factors over a cube of dimension  $n$  and its running time is subexponential in  $n$ . Two features limit its reach, and they are stated here so that no reader expects more. First, the exponent is  $o(n)$  and not  $n^{o(1)}$ . A polynomial-time reduction from an NP-complete problem to 3-SAT multiplies the number of variables by a polynomial, so the algorithm says nothing about P versus NP, and nothing here bears on that question. Second, the number of factors and the total degree of their product must be  $O(n)$ , or  $o(n\sqrt{\log n / \log \log n})$  after Theorem 7.20, after a branching step that fixes  $O(n/\log n)$  variables; Remark 7.27 shows that the routine itself cannot go beyond total degree of order  $n \log n$ . Max-Cut on general graphs with  $\Theta(n^2)$  edges, the chromatic and Tutte polynomials as polynomials in  $q$  (as opposed to their values at a fixed positive integer  $q$ , which Theorem 7.10 does give on bounded-degree graphs), Orthogonal Vectors, 3SUM and All-Pairs Shortest Paths need either  $\Theta(n^2)$  constraints or reductions that multiply the variable count, and they are untouched. The paper contains no partial result on any of these problems.

**1.4. Organisation.** Section 2 fixes notation and the model of computation over an algebra. Sections 3 and 4 reproduce the summation routine and the construction of the replacement forms. Section 5 proves Theorem 5.3, including the branching on frequent variables and the

bit-complexity bound over  $\mathbb{Q}[W]/(W^N)$ . Section 6 proves items (i)–(vii), Section 7 proves items (viii) and (ix), Section 8 proves item (x), and Section 9 records what the method does not give. Appendix A describes exact-arithmetic checks of the extended algorithm on small instances and of the finite inequalities behind Section 7.9; no statement of the paper depends on them. The author’s earlier work on algorithmic structures in geometric and topological settings [Bha26a, Bha26b, Bha26c] is unrelated in method but shares the aim of making a computation explicit enough to be verified line by line.

A word on the constants. The parameters of Lemma 4.7 become admissible only when  $\sqrt{\log n}$  exceeds a large multiple of  $(\log \log n)^2$ , that is for  $n$  beyond any practical size; the same is true of OpenAI’s algorithm, and Figure 1 shows where the three costs of the routine settle; Theorem 7.19 asks for  $\log n \geq 2^{30} \log \log n$  at the least. The theorems are statements about asymptotic complexity, and the paper does not present them as anything else.

## 2. PRELIMINARIES

**2.1. Notation.** For an integer  $Q \geq 1$  we write  $[Q] = \{1, \dots, Q\}$ . Boolean assignments are vectors  $x \in \{0, 1\}^n$ , also read as vectors in  $\mathbb{Q}^n$ . For  $S \subseteq [n]$ ,  $x|_S$  is the restriction and, for  $\xi \in \mathbb{Q}^S$ ,  $\tilde{\xi} \in \mathbb{Q}^n$  is the extension of  $\xi$  by zero. The unit vector at  $v$  is  $\mathbf{e}_v$ , and  $\mathbf{1}_U$  is the indicator vector of  $U \subseteq [n]$ . Logarithms are to base 2 unless written  $\ln$ . For  $n \geq 20$  put

$$\lambda = \lceil \log n \rceil, \quad \nu = \lceil \log \lambda \rceil, \quad (2.1)$$

so that  $\lambda = \Theta(\log n)$  and  $\nu = \Theta(\log \log n)$ .

**2.2. Algebras and factors.** Throughout,  $\mathcal{A}$  is a commutative  $\mathbb{Q}$ -algebra with identity. The two cases of interest are  $\mathcal{A} = \mathbb{Q}$  and the truncated polynomial ring

$$\mathcal{A}_N = \mathbb{Q}[W]/(W^N), \quad N \geq 1, \quad (2.2)$$

whose elements are polynomials of degree less than  $N$  with rational coefficients, added coefficient-wise and multiplied with the terms of degree at least  $N$  discarded. An addition in  $\mathcal{A}_N$  costs  $N$  rational additions and a multiplication costs at most  $N^2$  rational multiplications and additions. More generally, for integers  $N_1, \dots, N_s \geq 1$  we use

$$\mathcal{A}_{N_1, \dots, N_s} = \mathbb{Q}[W_1, \dots, W_s]/(W_1^{N_1}, \dots, W_s^{N_s}), \quad (2.3)$$

a  $\mathbb{Q}$ -vector space of dimension  $N = N_1 \cdots N_s$  with the monomials  $W_1^{j_1} \cdots W_s^{j_s}$ ,  $0 \leq j_i < N_i$ , as a basis; its arithmetic costs the same as that of  $\mathcal{A}_N$  with this  $N$ . Both are commutative  $\mathbb{Q}$ -algebras, and so is the group algebra  $\mathbb{Q}[W]/(W^m - 1)$  of the cyclic group of order  $m$ , which appears in Section 6.

A *factor system* on  $n$  variables is a finite index set  $\mathcal{F}$  together with, for each  $f \in \mathcal{F}$ , a scope  $I_f \subseteq [n]$  and a polynomial  $\Phi_f \in \mathcal{A}[Y_v : v \in I_f]$ . We write

$$Z = \sum_{x \in \{0, 1\}^n} \prod_{f \in \mathcal{F}} \Phi_f((x_v)_{v \in I_f}) \in \mathcal{A} \quad (2.4)$$

for its sum over the cube, and

$$M = \sum_{f \in \mathcal{F}} |I_f|, \quad w = \max_f |I_f|, \quad b = \max_{v \in [n]} \#\{f : v \in I_f\}, \quad d = \sum_{f \in \mathcal{F}} \deg \Phi_f \quad (2.5)$$

for the number of incidences, the largest scope, the largest number of factors containing one variable, and the total degree. A function  $\phi : \{0, 1\}^I \rightarrow \mathcal{A}$  on at most  $k$  Boolean inputs is represented by its multilinear extension

$$\Phi(Y) = \sum_{U \subseteq I} \phi(\mathbf{1}_U) \prod_{v \in U} Y_v \prod_{v \in I \setminus U} (1 - Y_v), \quad (2.6)$$

a polynomial of degree at most  $|I| \leq k$  that agrees with  $\phi$  on  $\{0, 1\}^I$ . Every factor below is of this kind, so  $\deg \Phi_f \leq |I_f|$  and  $d \leq M$ .

**2.3. Weights.** Let  $\omega_f : \{0, 1\}^{I_f} \rightarrow \{0, 1, \dots\}$  be integer-valued and let  $N$  exceed  $\sum_f \max \omega_f$ . With  $\mathcal{A} = \mathcal{A}_N$  and  $\Phi_f$  the multilinear extension of  $W^{\omega_f}$ , the sum (2.4) is

$$Z = \sum_{x \in \{0, 1\}^n} W^{\sum_f \omega_f(x|_{I_f})} = \sum_{j \geq 0} \# \left\{ x : \sum_f \omega_f(x|_{I_f}) = j \right\} W^j, \quad (2.7)$$

and no term is lost to the truncation. We call the right-hand side the *profile* of the weight system. The largest  $j$  with a nonzero coefficient is the maximum of the total weight, the smallest is the minimum, and the coefficients count the assignments attaining each value.

**2.4. Model of computation.** Running times count bit operations on a deterministic random-access machine. Elements of  $\mathbb{Q}$  are stored as pairs of integers in lowest terms. When a bound is stated as a number of *operations in  $\mathcal{A}$* , each operation is an addition, a multiplication, or a multiplication by a rational scalar, and the bound on bit operations follows once the bit lengths of the operands are bounded. For  $\mathcal{A} = \mathcal{A}_N$  this is done in Lemma 5.5. The space bound “polynomial” means a polynomial in  $n$ ,  $|\mathcal{F}|$ ,  $N$  and the bit lengths of the input.

**2.5. Two classical tools.** We use Hoeffding’s inequality for a binomial variable [Hoe63]: if  $W \sim \text{Bin}(Q, p)$  and  $s > 0$  then

$$\Pr[|W - Qp| \geq s] \leq 2e^{-2s^2/Q}. \quad (2.8)$$

We use Hermite interpolation in the following form. Let  $m = QJ$  and let  $H \in \mathcal{A}[T]$  have degree less than  $m$ . The  $m$  quantities  $[z^j]H(t + z)$  for  $t \in [Q]$  and  $0 \leq j < J$  are rational linear combinations of the coefficients of  $H$ , through the  $m \times m$  matrix with entries  $\binom{i}{j}t^{i-j}$  (zero when  $i < j$ ) indexed by  $(t, j)$  and  $i$ . This matrix is invertible over  $\mathbb{Q}$ : a polynomial of degree less than  $QJ$  whose first  $J$  Taylor coefficients vanish at each of  $Q$  distinct points is divisible by  $\prod_t (T - t)^J$ , hence zero. Since the entries are rationals of  $O(QJ \log(QJ))$  bits, Gaussian elimination recovers the coefficients of  $H$  with  $O((QJ)^3)$  operations in  $\mathcal{A}$  and rational numbers of  $\text{poly}(QJ)$  bits.

### 3. THE SUMMATION ROUTINE

This section and the next reproduce the algorithm of OpenAI [OAI26, Sections 3–4] for a factor system with values in a commutative  $\mathbb{Q}$ -algebra  $\mathcal{A}$ . The algorithm is theirs; the statements are repeated with proofs so that the paper is self-contained.

**3.1. Replacement forms and the polynomial  $H$ .** Give each incidence  $(f, v)$ ,  $f \in \mathcal{F}$ ,  $v \in I_f$ , its own indeterminate  $Y_{f,v}$ , and suppose

$$\deg_{\{Y_{f,v}\}} \prod_{f \in \mathcal{F}} \Phi_f((Y_{f,v})_{v \in I_f}) \leq d. \quad (3.1)$$

With  $d$  as in (2.5) this holds. Choose for every incidence a row  $\vec{P}_{f,v}(T) \in \mathbb{Q}[T]^{1 \times n}$  with  $\vec{P}_{f,v}(0) = \mathbf{e}_v$  and  $\deg \vec{P}_{f,v} \leq R$ , and define

$$H(T) = \sum_{x \in \{0, 1\}^n} \prod_{f \in \mathcal{F}} \Phi_f(\langle \vec{P}_{f,v}(T), x \rangle_{v \in I_f}) \in \mathcal{A}[T]. \quad (3.2)$$

At  $T = 0$  every input is restored, so  $H(0) = Z$ , and each input has degree at most  $R$  in  $T$ , so

$$\deg_T H \leq d_H := dR. \quad (3.3)$$

The algorithm computes  $H$  at  $T = 0$  from information at the points  $T = 1, \dots, Q$ .

### 3.2. The design.

**Definition 3.1** (Design). Let  $r \geq 2$ ,  $D \geq 1$  and  $1 \leq R < Q$  be integers and  $0 < \tau \leq 1/2$ . A design for the scopes  $(I_f)_{f \in \mathcal{F}}$  consists of the following data.<sup>2</sup>

- (D1) For each  $t \in [Q]$  a partition  $[n] = S_{t,1} \sqcup \dots \sqcup S_{t,r}$ , parts possibly empty, with  $|S_{t,g}| \leq 2n/r$ .
- (D2) For each  $f \in \mathcal{F}$  a sequence  $\sigma_f \in [r]^Q$ ; the value  $\sigma_f(t)$  is the group of  $f$  at the point  $t$ .
- (D3) For each  $t \in [Q]$  and  $g \in [r]$  an integer matrix  $C_{t,g} \in \mathbb{Z}^{c_{t,g} \times S_{t,g}}$  whose rows have sum of absolute values at most  $D$ , with  $\sum_{g=1}^r c_{t,g} \leq 2\tau n$ . The entries of  $C_{t,g}$  are the *measurements* of  $x$  on the group  $g$  at  $t$ .
- (D4) For each incidence  $(f, v)$  a row  $\vec{P}_{f,v}(T) \in \mathbb{Q}[T]^{1 \times n}$  with  $\vec{P}_{f,v}(0) = \mathbf{e}_v$  and  $\deg \vec{P}_{f,v} \leq R$ , such that for all  $t \in [Q]$ ,  $g \neq \sigma_f(t)$  and  $\xi \in \mathbb{Q}^{S_{t,g}}$ ,

$$C_{t,g} \xi = 0 \implies \langle \vec{P}_{f,v}(t), \tilde{\xi} \rangle = 0. \quad (3.4)$$

- (D5) With  $d_{\text{den}} = (R-1)!Q!$ , every coefficient of  $d_{\text{den}} \vec{P}_{f,v}$  is an integer of  $O(Q \log Q + \log D)$  bits.

The kernel condition (3.4) has the following meaning. Fix  $t$ ,  $f$  and  $v \in I_f$ , put  $g = \sigma_f(t)$  and write  $x_h = x|_{S_{t,h}}$ . If  $x, x' \in \mathbb{Q}^n$  agree outside  $S_{t,h}$  for some  $h \neq g$  and  $C_{t,h}x_h = C_{t,h}x'_h$ , then  $x'_h - x_h \in \ker C_{t,h}$  and (3.4) gives  $\langle \vec{P}_{f,v}(t), x' \rangle = \langle \vec{P}_{f,v}(t), x \rangle$ . Applied one group at a time this shows that the tuple  $(\langle \vec{P}_{f,v}(t), x \rangle)_{v \in I_f}$  is determined by  $x_g$  together with the measurements  $(C_{t,h}x_h)_{h \neq g}$ . Once all measurements are fixed, the factor  $f$  at the point  $t$  is a function of  $x_g$  alone. The bound on the row sums in (D3) ensures that a measurement of a Boolean vector lies in  $[-D, D] \cap \mathbb{Z}$ ; a rank bound alone would not control the number of possible measurement values, as the single row  $(1, 2, 4, \dots, 2^{n-1})$  shows.

### 3.3. Separating the sum by fixing the measurements.

**Lemma 3.2** (Separation). Let  $[n] = S_1 \sqcup \dots \sqcup S_r$  be a partition, parts possibly empty, and for each  $g \in [r]$  let  $C_g$  be an integer matrix with  $c_g$  rows and columns indexed by  $S_g$ , each row with sum of absolute values at most  $D$ ; put  $c = \sum_g c_g$ . Let  $A_g : \mathbb{Q}^n \rightarrow \mathcal{A}$  satisfy

$$A_g(x + \tilde{\xi}) = A_g(x) \quad (x \in \mathbb{Q}^n, \xi \in \ker C_h, h \neq g). \quad (3.5)$$

For every  $E \subseteq [n]$  the sum

$$Z_E = \sum_{\substack{x \in \{0,1\}^n \\ x_i = 1 \ (i \in E)}} \prod_{g=1}^r A_g(x)$$

can be computed by examining at most  $2(2D+1)^c \sum_g 2^{|S_g|}$  Boolean vectors, with at most one evaluation of some  $A_g$ ,  $O(1)$  operations in  $\mathcal{A}$  and  $\text{poly}(n, r, c, \log(D+2))$  bit operations per vector, and with  $\text{poly}(n, r, c, \log(D+2))$  bits of storage beyond the evaluators' own and  $O(1)$  elements of  $\mathcal{A}$ .<sup>3</sup>

*Proof.* For  $y_g \in \mathbb{Z}^{c_g}$  let  $X_g(y_g) = \{u \in \{0,1\}^{S_g} : C_g u = y_g\}$  and let  $\mathcal{Y} = \prod_g \{-D, \dots, D\}^{c_g}$ , a set of  $(2D+1)^c$  tuples that contains the measurement tuple of every Boolean assignment. Fix  $\vec{y} = (y_1, \dots, y_r) \in \mathcal{Y}$  with every  $X_g(y_g)$  nonempty, choose representatives  $\tilde{x}_g \in X_g(y_g)$ , and put

<sup>2</sup>Definition 3.1 of [OAI26].

<sup>3</sup>Lemma 3.2 of [OAI26].

$X_{g,E}(y_g) = \{u \in X_g(y_g) : u_i = 1 \text{ } (i \in E \cap S_g)\}$ . If  $x_h \in X_h(y_h)$  for every  $h$ , then  $x_h - \bar{x}_h \in \ker C_h$ , and (3.5) applied for each  $h \neq g$  in turn gives

$$A_g\left(\sum_h \tilde{x}_h\right) = A_g\left(\tilde{x}_g + \sum_{h \neq g} \tilde{\tilde{x}}_h\right).$$

The assignments with measurement tuple  $\vec{y}$  and  $x_i = 1$  for  $i \in E$  are exactly the products  $\prod_g X_{g,E}(y_g)$ , so by distributivity they contribute  $\prod_g \Sigma_g(\vec{y})$  to  $Z_E$ , where

$$\Sigma_g(\vec{y}) = \sum_{x_g \in X_{g,E}(y_g)} A_g\left(\tilde{x}_g + \sum_{h \neq g} \tilde{\tilde{x}}_h\right), \quad (3.6)$$

an empty sum being zero. Every assignment has exactly one measurement tuple, hence

$$Z_E = \sum_{\vec{y} \in \mathcal{Y}} \prod_{g=1}^r \Sigma_g(\vec{y}), \quad (3.7)$$

where a product is zero when some  $X_g(y_g)$  is empty. The representatives need not satisfy the conditions indexed by  $E$ .

The algorithm runs through  $\vec{y} \in \mathcal{Y}$  with an accumulator  $Z \leftarrow 0$ . For each  $\vec{y}$  a first scan of each  $\{0, 1\}^{S_g}$  finds a representative  $\bar{x}_g$ , or shows that  $X_g(y_g)$  is empty, in which case the tuple is skipped. Otherwise it sets  $p \leftarrow 1$ , and for each  $g$  computes  $\Sigma_g(\vec{y})$  by a second scan of  $\{0, 1\}^{S_g}$ , adding  $A_g(\tilde{u} + \sum_{h \neq g} \tilde{\tilde{x}}_h)$  to a running sum  $s$  for each  $u$  with  $C_g u = y_g$  and  $u_i = 1$  on  $E \cap S_g$ , then sets  $p \leftarrow ps$ ; finally  $Z \leftarrow Z + p$ . By (3.7) the final value is  $Z_E$ . The two scans examine at most  $2 \sum_g 2^{|S_g|}$  vectors per tuple. Checking the measurements of a vector costs  $\text{poly}(n, r, c, \log(D+2))$  bit operations, since each row sum has  $O(\log(D+2))$  bits, and the retained state is  $\vec{y}$ , the representatives, the current  $g$  and  $u$ , and  $s, p, Z \in \mathcal{A}$ .  $\square$

To evaluate  $H(t)$  for a fixed  $t \in [Q]$ , group the transformed factors by  $\sigma_f(t)$ ,

$$A_g(x) = \prod_{f: \sigma_f(t)=g} \Phi_f(\langle \vec{P}_{f,v}(t), x \rangle)_{v \in I_f},$$

and apply the lemma with the partition  $(S_{t,g})_g$ , the matrices  $C_{t,g}$  and  $E = \emptyset$ ; the kernel condition (3.4) is the invariance (3.5). The two exponential costs are the number of measurement tuples and the number of assignments in one group.

**3.4. Taylor coefficients from Boolean substitutions.** The  $Q$  values  $H(1), \dots, H(Q)$  do not determine  $H$  when  $d_H \geq Q$ . Put

$$q = \lfloor d_H/Q \rfloor, \quad J = q + 1, \quad (3.8)$$

so that  $QJ > d_H$ . It suffices to compute the truncated Taylor expansions

$$H_t(z) = \sum_{j=0}^{J-1} ([z^j] H(t+z)) z^j \quad (t \in [Q]), \quad (3.9)$$

since Hermite interpolation (Section 2) recovers  $H$ , and hence  $Z = H(0)$ , from these  $QJ$  coefficients by rational linear combinations.

Fix  $t$ . The rows at  $t+z$  need not respect the design at  $t$ . Write  $\Delta \vec{P}_{f,v}(z) = \vec{P}_{f,v}(t+z) - \vec{P}_{f,v}(t)$ ; every entry is divisible by  $z$ . For a fixed  $x \in \mathbb{Q}^n$  introduce indeterminates  $\vec{X} = (X_1, \dots, X_n)$  and replace each input by  $\langle \vec{P}_{f,v}(t), x \rangle + \langle \Delta \vec{P}_{f,v}(z), \vec{X} \rangle$ . Let

$$G_x(\vec{X}, z) = \prod_{f \in \mathcal{F}} \Phi_f\left(\langle \vec{P}_{f,v}(t), x \rangle + \langle \Delta \vec{P}_{f,v}(z), \vec{X} \rangle\right)_{v \in I_f}. \quad (3.10)$$

Setting  $\vec{X} = x$  and summing over  $x \in \{0, 1\}^n$  gives  $H(t + z)$ . Holding  $\vec{X}$  fixed instead makes every correction independent of  $x$ , so the design at  $t$  still separates the factors; and every occurrence of  $\vec{X}$  carries a factor  $z$ , so

$$\deg_{\vec{X}}([z^j]G_x(\vec{X}, z)) \leq j. \quad (3.11)$$

**Lemma 3.3** (Boolean substitution). *Let  $q \geq 0$  and let  $\mathcal{A}$  be a commutative ring with identity. For each  $x \in \{0, 1\}^n$  let  $p_x \in \mathcal{A}[X_1, \dots, X_n]$  have degree at most  $q$ . Then*

$$\sum_{x \in \{0, 1\}^n} p_x(x) = \sum_{\substack{E \subseteq [n] \\ |E| \leq q}} \sum_{U \subseteq E} (-1)^{|E| - |U|} \sum_{\substack{x \in \{0, 1\}^n \\ x_i = 1 \ (i \in E)}} p_x(\mathbf{1}_U).^4 \quad (3.12)$$

*Proof.* Fix  $x$  and replace each positive power of  $X_i$  in  $p_x$  by  $X_i$ . Boolean evaluations are unchanged and the degree does not increase, so the result is a multilinear polynomial  $\tilde{p}_x = \sum_E c_{x,E} \prod_{i \in E} X_i$  of degree at most  $q$ . Evaluating at the indicator vectors and inverting on the Boolean lattice gives  $c_{x,E} = \sum_{U \subseteq E} (-1)^{|E| - |U|} p_x(\mathbf{1}_U)$ , and  $c_{x,E} = 0$  when  $|E| > q$ . Setting  $\vec{X} = x$  and summing over  $x$ , the monomial  $\prod_{i \in E} x_i$  selects exactly the assignments with  $x_i = 1$  for  $i \in E$ .  $\square$

Apply the lemma to  $[z^j]G_x$  for each  $j < J$ ; by (3.11) the degree bound  $q$  holds for all of them, so all coefficients use the same pairs  $U \subseteq E$ ,  $|E| \leq q$ , and may be computed together in the ring

$$\mathcal{R} = \mathcal{A}[z]/(z^J). \quad (3.13)$$

For a fixed pair  $(E, U)$  let  $\Phi_{f,U}(x; z)$  be the factor  $f$  in (3.10) after  $\vec{X} = \mathbf{1}_U$ , and group by assigned groups,

$$A_{g,U}(x; z) = \prod_{f: \sigma_f(t)=g} \Phi_{f,U}(x; z) \in \mathcal{R}. \quad (3.14)$$

Their product over  $g$  is  $G_x(\mathbf{1}_U, z)$ . For  $h \neq g$ , changing  $x$  by an element of  $\ker C_{t,h}$  changes neither the original inputs of the factors in  $A_{g,U}$  nor their fixed corrections, so Lemma 3.2 applies in  $\mathcal{R}$  and computes  $\sum_{x: x_i=1 \ (i \in E)} G_x(\mathbf{1}_U, z)$  by examining at most

$$2r(2D + 1)^{2\tau n} 2^{2n/r} \quad (3.15)$$

Boolean vectors. Summing over the pairs with the signs of (3.12),

$$H(t + z) \equiv \sum_{\substack{E \subseteq [n] \\ |E| \leq q}} \sum_{U \subseteq E} (-1)^{|E| - |U|} \sum_{\substack{x \in \{0, 1\}^n \\ x_i = 1 \ (i \in E)}} G_x(\mathbf{1}_U, z) \pmod{z^J}, \quad (3.16)$$

which is  $H_t(z)$ . The algorithm never stores a polynomial in the  $n$  indeterminates: it substitutes  $\vec{X} = \mathbf{1}_U$  first and keeps only  $J$  coefficients in  $\mathcal{A}$ .

**3.5. The three exponential costs.** For each  $t$ , the number of pairs  $U \subseteq E \subseteq [n]$  with  $|E| \leq q$  is

$$C_q = \sum_{s=0}^{\min(q,n)} 2^s \binom{n}{s}, \quad (3.17)$$

and for every  $0 < \eta \leq 1$ ,

$$\eta^q C_q \leq \sum_{s=0}^n (2\eta)^s \binom{n}{s} = (1 + 2\eta)^n, \quad (3.18)$$

---

<sup>4</sup>Lemma 3.3 of [OAI26].

because  $\eta^q \leq \eta^s$  for every retained  $s \leq q$ . The choice of  $\eta$  that minimises the right-hand side gives the bound that will be used when the ratio  $q/n$  is allowed to vary with  $n$ . Write  $H(u) = -u \log u - (1-u) \log(1-u)$  for the binary entropy, with  $H(0) = 0$ .

**Lemma 3.4** (Entropy bound). *For integers  $0 \leq q \leq n/2$ ,*

$$\log C_q \leq q + n H(q/n) \leq q \log \frac{2en}{q}, \quad (3.19)$$

*the last expression read as 0 when  $q = 0$ .*

*Proof.* For  $q = 0$  both sides vanish. For  $q \geq 1$  take  $\eta = q/(2(n-q))$ , which lies in  $(0, 1]$  since  $q \leq n/2$ . Then  $1 + 2\eta = n/(n-q)$ , and (3.18) gives

$$C_q \leq \eta^{-q} \left( \frac{n}{n-q} \right)^n = \frac{2^q n^n}{q^q (n-q)^{n-q}},$$

whose logarithm is  $q + n \log n - q \log q - (n-q) \log(n-q) = q + nH(q/n)$ . For the second inequality,  $-(1-u) \ln(1-u) \leq u$  for  $0 \leq u < 1$ , because the derivative of  $u + (1-u) \ln(1-u)$  is  $-\ln(1-u) \geq 0$ ; hence  $H(u) \leq u \log(e/u)$ , and with  $u = q/n$  this gives  $q + nH(q/n) \leq q + q \log(en/q)$ .  $\square$

**Proposition 3.5** (Summation routine). *Let  $\mathcal{A}$  be a commutative  $\mathbb{Q}$ -algebra, let the factor system satisfy (3.1), and let a design as in Definition 3.1 be given. With  $q$  and  $J$  as in (3.8), the routine above computes  $Z$  by examining at most*

$$2Qr C_q (2D+1)^{2\tau n} 2^{2n/r} \quad (3.20)$$

*Boolean vectors; for each vector it evaluates at most  $|\mathcal{F}|$  factors in  $\mathcal{A}[z]/(z^J)$ . The remaining operations per vector and the final interpolation are polynomially many operations in  $\mathcal{A}$  and over  $\mathbb{Q}$  in  $n, |\mathcal{F}|, D, Q, R, J, r$  and  $M$ . Apart from the factor evaluators' own storage the routine keeps polynomially many elements of  $\mathcal{A}$ , rationals and index bits.<sup>5</sup>*

*Proof.* There are  $Q$  points,  $C_q$  pairs per point, and at most (3.15) vectors per pair. Each group product uses at most  $|\mathcal{F}|$  factor evaluations, and addition and multiplication in  $\mathcal{A}[z]/(z^J)$  cost  $O(J^2)$  operations in  $\mathcal{A}$ . Correctness is (3.16) together with Hermite interpolation. The points, pairs, measurement tuples and local assignments are processed sequentially with their current indices and accumulators, which with the  $QJ$  interpolation data gives the storage bound.  $\square$

When  $d \leq \kappa n$  we have  $q \leq \kappa n R/Q$ , and (3.18) with  $\eta = R/Q$  gives

$$\log C_q \leq q \log \frac{Q}{R} + n \log \left( 1 + \frac{2R}{Q} \right) \leq \kappa n \frac{R}{Q} \log \frac{Q}{R} + \frac{2nR}{Q} \cdot \frac{1}{\ln 2}. \quad (3.21)$$

If factor evaluation and coefficient arithmetic cost polynomially many bit operations, the running time is a polynomial factor times

$$2^{O\left(n \left[ (\kappa+1) \frac{R}{Q} \log \frac{2Q}{R} + \tau \log(2D+1) + \frac{1}{\tau} \right] \right)}. \quad (3.22)$$

The three summands are the costs of recovering the Taylor coefficients, of enumerating the measurement values, and of enumerating the assignments of one group. Section 4 constructs a design for which all three are  $O(\nu^2/\sqrt{\lambda})$  up to constants depending on  $\kappa$  and on the logarithmic exponent bounding  $b$  and  $w$ .

<sup>5</sup>Proposition 3.4 of [OAI26].

## 4. CONSTRUCTING THE DESIGN

The routine needs one row  $\vec{P}_{f,v}(T)$  per incidence. This section constructs the rows and the matrices  $C_{t,g}$ , the partitions  $S_{t,g}$  and the group sequences  $\sigma_f$  of Definition 3.1 from the scopes alone, following [OAI26, Section 4]. The difficulty is to keep  $\deg \vec{P}_{f,v} \leq R$  with  $R$  much smaller than  $Q$ : the degree of  $H$  grows linearly in  $R$ , and a small  $R/Q$  reduces the number of Taylor coefficients needed at each point. Rows chosen independently at the  $Q$  points and interpolated would in general have degree  $Q$ . The construction instead produces, on a complete array of auxiliary coordinates, rows whose values at all points are controlled together, and then places each incidence in a copy of that array.

**4.1. Incidences and auxiliary coordinates.** Write  $\mathcal{I} = \{(f, v) : f \in \mathcal{F}, v \in I_f\}$ , so  $M = |\mathcal{I}|$ , and let  $w$  and  $b$  be as in (2.5). A factor with empty scope is a constant and is assigned to group one at every point. If necessary adjoin a constant-one factor with scope  $\{v\}$  for every  $v \in [n]$ ; this changes neither the product nor the sum, and raises  $b$  by at most one. With  $b$  the bound after this step we assume

$$n \geq 1, \quad b, w \geq 1, \quad n \leq M \leq bn. \quad (4.1)$$

Throughout,  $r \geq 2$  and  $1 \leq R < Q$  are integers and  $0 < \tau \leq 1/2$ .

At each point  $t$  the variable  $x_v$  belongs to one group while a factor containing it may be assigned to another. We shall choose

$$a_v \in [r]^Q \quad (v \in [n]), \quad \delta_f \in [r]^Q \quad (f \in \mathcal{F}),$$

where  $a_v(t)$  is the group of  $x_v$  and  $\delta_f(t)$  the group of  $f$  at the point  $t$ , and the incidence  $(f, v)$  will receive the label  $(a_v, \delta_f)$ . Before choosing them, take all possible labels,

$$\mathcal{B} = ([r]^Q)^2, \quad u = (a_u, \delta_u), \quad D = |\mathcal{B}| = r^{2Q}. \quad (4.2)$$

An auxiliary vector  $\zeta \in \mathbb{Q}^{\mathcal{B}}$  has one coordinate per label. For  $t \in [Q]$  and  $g \in [r]$  put

$$S_{t,g}^{\mathcal{B}} = \{u \in \mathcal{B} : a_u(t) = g\}, \quad (4.3)$$

which for fixed  $t$  partition  $\mathcal{B}$  into  $r$  sets of size  $D/r$ . We seek, for each  $u \in \mathcal{B}$ , a row  $\vec{P}_u(T) \in \mathbb{Q}[T]^{1 \times \mathcal{B}}$  of degree at most  $R - 1$  with  $\vec{P}_u(0) = \mathbf{e}_u$  such that the value  $\langle \vec{P}_u(t), \zeta \rangle$  is determined by  $\zeta|_{S_{t,\delta_u(t)}^{\mathcal{B}}}$  together with  $C^{\mathcal{B}}\zeta$ , where  $C^{\mathcal{B}}$  is one integer matrix with entries in  $\{0, 1, -1\}$  and few rows, the same for all  $u$  and  $t$ .

**4.2. The matrix  $M_0$  and the measurements  $C^{\mathcal{B}}$ .** We seek  $M_0 \in \mathbb{Q}^{\mathcal{B} \times \mathcal{B}}$  that fixes  $\ker C^{\mathcal{B}}$  and satisfies

$$M_0[u, y] \neq 0 \implies \#\{s \in [Q] : a_y(s) \neq \delta_u(s)\} \leq R - 1. \quad (4.4)$$

The identity

$$\zeta_u = ((I - M_0)\zeta)_u + \sum_{y \in \mathcal{B}} M_0[u, y]\zeta_y \quad (4.5)$$

explains both requirements: the first term is determined by  $C^{\mathcal{B}}\zeta$  and will be left independent of  $T$ , and each term of the sum will be multiplied by a polynomial equal to one at zero and vanishing at every  $t$  with  $a_y(t) \neq \delta_u(t)$ , which (4.4) allows with degree at most  $R - 1$ .

For  $s \in [Q]$  let  $\vartheta_s : \mathcal{B} \rightarrow \mathcal{B}$  exchange the entries  $a_u(s)$  and  $\delta_u(s)$  and leave the rest unchanged, and let  $\Pi_s$  be its permutation matrix,  $(\Pi_s \zeta)_u = \zeta_{\vartheta_s(u)}$ . The  $\Pi_s$  commute and  $\Pi_s^2 = I$ . Put

$$V_0 = \prod_{s=1}^Q \Pi_s, \quad W_0 = \frac{1}{2} \sum_{s=1}^Q (I - \Pi_s). \quad (4.6)$$

$V_0$  exchanges the two full sequences, so  $V_0[u, y] \neq 0$  forces  $a_y = \delta_u$ . On a common eigenvector  $\psi$  of the  $\Pi_s$  with  $\Pi_s \psi = -\psi$  for exactly  $j$  indices  $s$ ,

$$W_0 \psi = j\psi, \quad V_0 \psi = (-1)^j \psi, \quad V_0 p(W_0) \psi = (-1)^j p(j) \psi \quad (4.7)$$

for every polynomial  $p$ ; so  $V_0 p(W_0)$  fixes  $\psi$  whenever  $p(j) = (-1)^j$ . For an interval  $\mathcal{J} = \{A_0, \dots, A_0 + R - 1\} \subseteq \{0, \dots, Q\}$  the polynomial

$$f_{\mathcal{J}}(z) = (-1)^{A_0} \sum_{j=0}^{R-1} (-2)^j \binom{z - A_0}{j}, \quad \binom{z - A_0}{j} = \frac{1}{j!} \prod_{a=0}^{j-1} (z - A_0 - a), \quad (4.8)$$

has degree at most  $R - 1$  and satisfies  $f_{\mathcal{J}}(A_0 + i) = (-1)^{A_0} (1 - 2)^i = (-1)^{A_0+i}$  for  $0 \leq i < R$  by the binomial theorem. Set

$$M_0 = V_0 f_{\mathcal{J}}(W_0). \quad (4.9)$$

**Lemma 4.1.** *There is a matrix  $C^{\mathcal{B}} \in \{0, 1, -1\}^{c^{\mathcal{B}} \times \mathcal{B}}$  with  $c^{\mathcal{B}} \leq D \Pr[\mathbf{W} \notin \mathcal{J}]$ , where  $\mathbf{W} \sim \text{Bin}(Q, \frac{r-1}{2r})$ , such that  $M_0 \xi = \xi$  for every  $\xi \in \ker C^{\mathcal{B}}$ .<sup>6</sup>*

*Proof.* For  $(\alpha, \beta) \in [r]^2$  let  $e_{\alpha\beta}$  be the indicator function of that ordered pair. Exchanging the two entries of a pair has the orthogonal eigenbasis

$$e_{\alpha\alpha} \ (\alpha \in [r]), \quad e_{\alpha\beta} + e_{\beta\alpha}, \quad e_{\alpha\beta} - e_{\beta\alpha} \quad (1 \leq \alpha < \beta \leq r),$$

with eigenvalue  $+1$  for the first two kinds and  $-1$  for the third; there are  $r(r+1)/2$  vectors with eigenvalue  $+1$  and  $r(r-1)/2$  with  $-1$ . For each  $s \in [Q]$  choose one of these  $r^2$  functions,  $\psi_s : [r]^2 \rightarrow \{0, 1, -1\}$ , and put  $\psi(u) = \prod_s \psi_s(a_u(s), \delta_u(s))$ . The  $D = r^{2Q}$  choices give nonzero pairwise orthogonal vectors, since their inner products factor over  $s$ , so they form a basis of  $\mathbb{Q}^{\mathcal{B}}$  with entries in  $\{0, 1, -1\}$ . If exactly  $j$  of the chosen  $\psi_s$  have eigenvalue  $-1$ , then  $W_0 \psi = j\psi$  and  $V_0 \psi = (-1)^j \psi$  by (4.7), so  $M_0$  fixes  $\psi$  whenever  $j \in \mathcal{J}$ . Let the rows of  $C^{\mathcal{B}}$  be the transposes of the basis vectors with  $j \notin \mathcal{J}$ . A fraction  $(r-1)/(2r)$  of the choices at each position has eigenvalue  $-1$ , independently over the positions, so the fraction of basis vectors retained is  $\Pr[\mathbf{W} \notin \mathcal{J}]$ . If  $C^{\mathcal{B}} \xi = 0$  then  $\xi$  is orthogonal to the retained vectors, hence a combination of the others, all of which  $M_0$  fixes.  $\square$

For  $\epsilon > 0$  the condition

$$\Pr\left[\text{Bin}\left(Q, \frac{r-1}{2r}\right) \notin \mathcal{J}\right] \leq \epsilon \quad (4.10)$$

gives  $c^{\mathcal{B}} \leq \epsilon D$ ; we shall use  $\epsilon = \tau/(rb)$ .

**Lemma 4.2.** *If  $M_0[u, y] \neq 0$  then  $\#\{s \in [Q] : a_y(s) \neq \delta_u(s)\} \leq R - 1$ .<sup>7</sup>*

*Proof.* Expand  $f_{\mathcal{J}}(W_0)$  in the commuting matrices  $\Pi_s$ . Since  $\deg f_{\mathcal{J}} \leq R - 1$  and  $\Pi_s^2 = I$ , each monomial is a scalar multiple of  $\prod_{s \in K} \Pi_s$  for some  $K \subseteq [Q]$  with  $|K| \leq R - 1$ , and  $V_0 \prod_{s \in K} \Pi_s = \prod_{s \notin K} \Pi_s$ . A nonzero  $(u, y)$  entry of this permutation matrix requires  $a_y(s) = \delta_u(s)$  for every  $s \notin K$ , leaving at most  $R - 1$  positions where they may differ. If more than  $R - 1$  positions disagree, every summand has zero  $(u, y)$  entry.  $\square$

<sup>6</sup>Lemma 4.1 of [OAI26].

<sup>7</sup>Lemma 4.2 of [OAI26].

**4.3. The auxiliary rows.** For  $M_0[u, y] \neq 0$  put

$$E_{u,y} = \{s \in [Q] : a_y(s) \neq \delta_u(s)\}, \quad \varphi_{u,y}(T) = \prod_{s \in E_{u,y}} \left(1 - \frac{T}{s}\right), \quad (4.11)$$

the empty product being one; then  $\deg \varphi_{u,y} \leq R - 1$ ,  $\varphi_{u,y}(0) = 1$ , and  $\varphi_{u,y}(t) = 0$  whenever  $a_y(t) \neq \delta_u(t)$ . Define

$$\vec{P}_u(T) = \mathbf{e}_u^\top (I - M_0) + \sum_{\substack{y \in \mathcal{B} \\ M_0[u,y] \neq 0}} M_0[u, y] \varphi_{u,y}(T) \mathbf{e}_y^\top. \quad (4.12)$$

**Lemma 4.3.** *The rows (4.12) have degree at most  $R - 1$  and  $\vec{P}_u(0) = \mathbf{e}_u^\top$ . For  $t \in [Q]$  and  $g \in [r]$  let  $C_{t,g}^\mathcal{B}$  be the restriction of  $C^\mathcal{B}$  to the columns  $S_{t,g}^\mathcal{B}$ . For every  $u, t$  and  $g \neq \delta_u(t)$ , and every  $\xi \in \mathbb{Q}^{S_{t,g}^\mathcal{B}}$ ,*

$$C_{t,g}^\mathcal{B} \xi = 0 \implies \langle \vec{P}_u(t), \tilde{\xi} \rangle = 0.$$

*More generally,  $\langle \vec{P}_u(t), \zeta \rangle$  is determined by  $\zeta|_{S_{t,\delta_u(t)}^\mathcal{B}}$  and  $C^\mathcal{B} \zeta$ .<sup>8</sup>*

*Proof.* The degree bound is that of the  $\varphi_{u,y}$ , and  $\vec{P}_u(0) = \mathbf{e}_u^\top (I - M_0) + \mathbf{e}_u^\top M_0 = \mathbf{e}_u^\top$ . Applying (4.12) to  $\zeta$ ,

$$\langle \vec{P}_u(T), \zeta \rangle = ((I - M_0)\zeta)_u + \sum_{M_0[u,y] \neq 0} M_0[u, y] \varphi_{u,y}(T) \zeta_y. \quad (4.13)$$

At  $T = t$  every summand with  $a_y(t) \neq \delta_u(t)$  vanishes, so the remaining summands involve only  $y \in S_{t,\delta_u(t)}^\mathcal{B}$ ; if  $\tilde{\xi}$  is supported on  $S_{t,g}^\mathcal{B}$  with  $g \neq \delta_u(t)$  they are zero on  $\tilde{\xi}$ . The hypothesis  $C_{t,g}^\mathcal{B} \xi = 0$  gives  $C^\mathcal{B} \xi = 0$ , so  $(I - M_0)\tilde{\xi} = 0$  by Lemma 4.1. For the last assertion, two vectors with the same value of  $C^\mathcal{B} \zeta$  differ by an element of  $\ker C^\mathcal{B}$ , so  $(I - M_0)\zeta$  is determined by  $C^\mathcal{B} \zeta$ , and the other terms at  $T = t$  use only  $\zeta|_{S_{t,\delta_u(t)}^\mathcal{B}}$ .  $\square$

Fixing  $t$  and writing  $\zeta_g = \zeta|_{S_{t,g}^\mathcal{B}}$  we have  $C^\mathcal{B} \zeta = \sum_g C_{t,g}^\mathcal{B} \zeta_g$ , so fixing each of the  $r$  vectors  $C_{t,g}^\mathcal{B} \zeta_g$  fixes  $C^\mathcal{B} \zeta$  while imposing separate conditions on the  $r$  auxiliary groups; then  $\langle \vec{P}_u(t), \zeta \rangle$  depends on  $\zeta|_{S_{t,\delta_u(t)}^\mathcal{B}}$  alone.

**4.4. Choosing the groups.** We now choose  $(a_v)_v$  and  $(\delta_f)_f$  and set  $u_{f,v} = (a_v, \delta_f)$  for  $(f, v) \in \mathcal{I}$ . A label  $u$  occurring  $c_u$  times needs  $c_u$  positions in distinct copies of the auxiliary array, so the number of copies is

$$h = \max_{u \in \mathcal{B}} c_u, \quad c_u = \#\{(f, v) \in \mathcal{I} : u_{f,v} = u\}, \quad (4.14)$$

and the actual variable groups  $\{v : a_v(t) = g\}$  must be small.

**Lemma 4.4.** *Assume (4.1), let  $D = r^{2Q}$  and suppose*

$$2(b + w)D^2 \leq M, \quad 2Qr^2 \leq n. \quad (4.15)$$

*One can choose  $a_v, \delta_f \in [r]^Q$  deterministically, with  $\delta_f \equiv 1$  for factors of empty scope, such that*

$$\max_{u \in \mathcal{B}} \#\{(f, v) \in \mathcal{I} : (a_v, \delta_f) = u\} \leq \frac{2M}{D}, \quad \#\{v : a_v(t) = g\} \leq \frac{2n}{r} \quad (t \in [Q], g \in [r]),$$

*in  $\text{poly}(n, M, |\mathcal{F}|, D, Q, r, b, w)$  bit operations.<sup>9</sup>*

<sup>8</sup>Lemma 4.3 of [OAI26].

<sup>9</sup>Lemma 4.5 of [OAI26].

*Proof.* First let all  $a_v$ , and all  $\delta_f$  with nonempty  $I_f$ , be independent and uniform on  $[r]^Q$ ; then each label  $u_{f,v}$  is uniform on  $\mathcal{B}$ . Put  $n_{t,g} = \#\{v : a_v(t) = g\}$  and

$$\Phi = \sum_{u \in \mathcal{B}} \left(c_u - \frac{M}{D}\right)^2, \quad \Xi = \sum_{t=1}^Q \sum_{g=1}^r \left(n_{t,g} - \frac{n}{r}\right)^2.$$

Expanding the squares,

$$\Phi = \sum_{(f,v),(f',v') \in \mathcal{I}} \mathbf{1}\{(a_v, \delta_f) = (a_{v'}, \delta_{f'})\} - \frac{M^2}{D}, \quad \Xi = \sum_{t=1}^Q \sum_{v,v' \in [n]} \mathbf{1}\{a_v(t) = a_{v'}(t)\} - \frac{Qn^2}{r}.$$

Two labels are independent unless they share a factor or a variable, and each  $(f, v)$  shares one with at most  $b + w$  incidences including itself. Independent pairs are equal with probability  $1/D$ , the others with probability at most one, so  $\mathbb{E}\Phi \leq M(b + w)$ . For  $v \neq v'$  the values  $a_v(t), a_{v'}(t)$  are independent and uniform, so  $\mathbb{E}\Xi = Qn(1 - 1/r) \leq Qn$ .

To fix the choices deterministically, fix the  $a_v$  and the  $\delta_f$  with nonempty scope one at a time, each time choosing a value that does not increase the conditional expectation of  $\Phi/(M(b + w)) + \Xi/(Qn)$ ; such a value exists because the current expectation is the average over the possible values of the next choice. The final value is at most two, and both summands are nonnegative, so

$$c_u \leq \frac{M}{D} + \sqrt{2M(b + w)}, \quad n_{t,g} \leq \frac{n}{r} + \sqrt{2Qn}.$$

The inequalities (4.15) bound the square roots by  $M/D$  and  $n/r$ .

For the cost, put  $K = r^Q$ , so  $D = K^2$ . Each indicator in  $\Phi$  depends on at most four of the choices  $a_v, \delta_f, a_{v'}, \delta_{f'}$ , and each indicator in  $\Xi$  on at most two. After any partial fixing, the conditional probability of an indicator is computed by enumerating at most  $K^4 = D^2$  values of the unfixed choices it involves, and has denominator dividing  $K^4$ . Summing the  $M^2 + Qn^2$  indicators gives both expectations, with all rationals of  $O(\log(2D) + \log(2M) + \log(2n) + \log(2Q) + \log(2 + b + w))$  bits. There are at most  $n + |\mathcal{F}|$  choices to fix, each with  $K$  candidates.  $\square$

**4.5. Placing copies of the variables.** Assume (4.15) and apply Lemma 4.4. Index  $h$  copies of the auxiliary array by  $[h] \times \mathcal{B}$ , and assign to each  $(f, v) \in \mathcal{I}$  a copy  $c(f, v) \in [h]$  so that the positions  $(c(f, v), u_{f,v})$  are distinct; this is possible since no label occurs more than  $h$  times, and

$$hD \leq 2M \leq 2bn. \quad (4.16)$$

Define  $\iota : \mathbb{Q}^n \rightarrow \mathbb{Q}^{[h] \times \mathcal{B}}$  by  $(\iota x)_{c,u} = x_v$  if  $(c, u) = (c(f, v), u_{f,v})$  for some  $(f, v) \in \mathcal{I}$ , and 0 otherwise; the occupied positions are distinct, so this is well defined. Several positions may contain the same  $x_v$ , and unused positions are zero. Let  $\pi_c : \mathbb{Q}^{[h] \times \mathcal{B}} \rightarrow \mathbb{Q}^{\mathcal{B}}$  be the projection onto the copy  $c$ . Put

$$\vec{P}_{f,v}(T) = \vec{P}_{u_{f,v}}(T) \pi_{c(f,v)} \iota, \quad S_{t,g} = \{v : a_v(t) = g\}, \quad \sigma_f(t) = \delta_f(t). \quad (4.17)$$

At  $T = 0$  the row  $\vec{P}_{f,v}$  extracts the content of position  $(c(f, v), u_{f,v})$ , which is  $x_v$ ; so  $\vec{P}_{f,v}(0) = \mathbf{e}_v^\top$ , and its degree is at most  $R - 1$ . Lemma 4.4 gives  $|S_{t,g}| \leq 2n/r$ . Define

$$\tilde{C} = \text{diag}(C^{\mathcal{B}}, \dots, C^{\mathcal{B}}) \text{ (} h \text{ copies)}, \quad \hat{C} = \tilde{C} \iota, \quad C_{t,g} = \hat{C}|_{S_{t,g}}, \quad (4.18)$$

the restriction keeping the columns in  $S_{t,g}$  and all rows. Thus  $\hat{C}x$  records the auxiliary measurements in every copy after the copies have been filled from  $x$ . Each row of  $\tilde{C}$  has at most  $D$  nonzero entries, all  $\pm 1$ ; multiplication by  $\iota$  adds entries lying in a common column and discards

unused positions, so every row of  $C_{t,g}$  has sum of absolute values at most  $D$ , and its value on a Boolean vector lies in  $[-D, D] \cap \mathbb{Z}$ . If (4.10) holds, each  $C_{t,g}$  has  $hc^B$  rows, and

$$\sum_{g=1}^r c_{t,g} = rhc^B \leq rh\epsilon D \leq 2r\epsilon M \leq 2r\epsilon bn, \quad (4.19)$$

which with  $\epsilon = \tau/(rb)$  is at most  $2\tau n$ .

For the kernel condition fix  $t$ ,  $g \neq \sigma_f(t)$  and  $(f, v)$ , and let  $\xi \in \mathbb{Q}^{S_{t,g}}$  with  $C_{t,g}\xi = 0$ . Then  $\tilde{C}\iota\xi = 0$ , so  $C^B\pi_{c\iota}\tilde{\xi} = 0$  for every copy  $c$ . Every nonzero coordinate of  $\pi_{c\iota}\tilde{\xi}$  contains a variable  $v'$  with  $a_{v'}(t) = g$ , whose label has first sequence  $a_{v'}$ , so this vector is supported on  $S_{t,g}^B$ . Lemma 4.3 in the copy  $c(f, v)$ , where  $\delta_{u_{f,v}}(t) = \delta_f(t) = \sigma_f(t) \neq g$ , gives  $\langle \vec{P}_{f,v}(t), \tilde{\xi} \rangle = \langle \vec{P}_{u_{f,v}}(t), \pi_{c(f,v)}\iota\tilde{\xi} \rangle = 0$ , which is (3.4). All inputs of a factor use the group  $\delta_f(t)$ , and the sum still ranges over exactly the  $n$  Boolean variables.

#### 4.6. Coefficient bounds.

**Lemma 4.5.** *Let  $d_{\text{den}} = (R-1)!Q!$ . Then  $d_{\text{den}}\vec{P}_u(T) \in \mathbb{Z}[T]^{1 \times B}$  for every  $u$ , and the sum of the absolute values of all coefficients of  $d_{\text{den}}\vec{P}_u(T)$  is  $2^{O(Q \log Q)}$ . The same holds for  $d_{\text{den}}\vec{P}_{f,v}(T)$ . The matrices  $M_0$ ,  $C^B$  and the rows can be constructed in  $\text{poly}(D, Q, r)$  bit operations.<sup>10</sup>*

*Proof.* For  $0 \leq j < R$  the  $j$ th summand of  $f_{\mathcal{J}}(W_0)$  is

$$(-1)^{A_0}(-2)^j \binom{W_0 - A_0 I}{j} = \frac{(-1)^{A_0+j}}{j!} \prod_{a=0}^{j-1} (2W_0 - 2(A_0 + a)I),$$

and each factor on the right is an integer matrix because  $2W_0 = QI - \sum_s \Pi_s$ . Since  $j!$  divides  $(R-1)!$  and  $V_0$  is a permutation matrix,  $(R-1)!M_0$  is integral. The denominators of  $\varphi_{u,y}$  divide  $\prod_{s \in E_{u,y}} s$ , which divides  $Q!$ . This proves the common denominator. With  $\|A\|_{\infty} = \max_i \sum_j |A[i, j]|$  and  $\|\sum_j a_j T^j\|_1 = \sum_j |a_j|$ , we have  $\|W_0\|_{\infty} \leq Q$  and  $0 \leq A_0 + a \leq Q$ , so each factor  $2W_0 - 2(A_0 + a)I$  has norm at most  $4Q$ , whence

$$\|M_0\|_{\infty} \leq R(4Q)^{R-1}, \quad \|\varphi_{u,y}\|_1 \leq \prod_{s \in E_{u,y}} \left(1 + \frac{1}{s}\right) \leq 2^Q.$$

The sum of absolute values of the coefficients in the row  $d_{\text{den}}\vec{P}_u$  is at most  $d_{\text{den}}(1 + (1 + 2^Q)R(4Q)^{R-1})$ , whose logarithm is  $O(Q \log Q)$  since  $R < Q$ . Passing to  $\vec{P}_{f,v}$  by (4.17) adds coefficients in common columns and discards others, so the bound persists. Enumerating the eigenbasis, forming  $f_{\mathcal{J}}(W_0)$  and the  $\varphi_{u,y}$  takes  $\text{poly}(D, Q, r)$  arithmetic operations on rationals of  $O(Q \log Q)$  bits.  $\square$

**Lemma 4.6** (The completed design). *Let  $n, b, w \geq 1$  and let  $(I_f)_{f \in \mathcal{F}}$  be subsets of  $[n]$  of size at most  $w$  with each  $v$  in at most  $b$  of them,  $M = \sum_f |I_f|$  and  $n \leq M \leq bn$ . Let  $r \geq 2$ ,  $1 \leq R < Q$ ,  $0 < \tau \leq 1/2$  and  $D = r^{2Q}$  satisfy (4.15), and let an interval  $\mathcal{J} = \{A_0, \dots, A_0 + R - 1\} \subseteq \{0, \dots, Q\}$  satisfy*

$$\Pr\left[\text{Bin}\left(Q, \frac{r-1}{2r}\right) \notin \mathcal{J}\right] \leq \frac{\tau}{rb}. \quad (4.20)$$

*Then all the data of Definition 3.1 can be constructed deterministically with a number of bit operations polynomial in  $n, M, |\mathcal{F}|, D, Q, r, b$  and  $w$ .<sup>11</sup>*

<sup>10</sup>Lemma 4.6 of [OAI26].

<sup>11</sup>Lemma 4.7 of [OAI26].

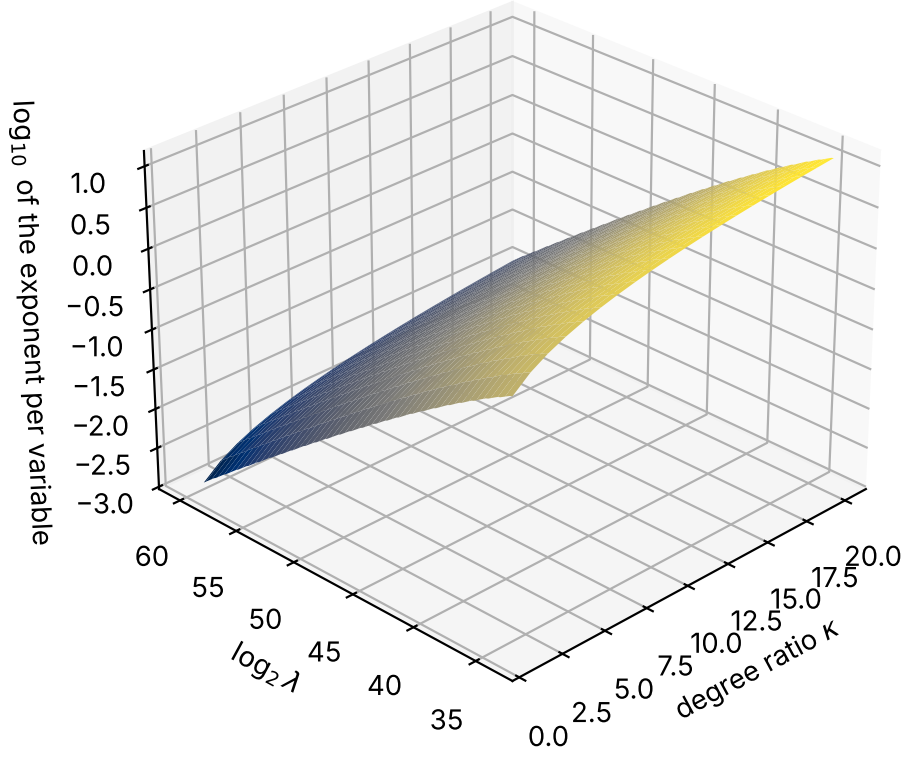


FIGURE 1. The sum of the three costs of (3.22) per variable, with the parameters (4.21) for  $\beta = 2$ , as a function of  $\lambda = \lceil \log n \rceil$  and of the degree ratio  $\kappa$ , drawn where the hypotheses  $r \geq 2$  and  $6R \leq Q$  of Lemma 4.7 hold. The surface falls like  $\nu^2/\sqrt{\lambda}$  in  $\lambda$  and grows linearly in  $\kappa$ .

*Proof.* Lemma 4.1 gives  $C^{\mathcal{B}}$  with  $c^{\mathcal{B}} \leq \tau D/(rb)$ ; Lemmas 4.2 and 4.3 give the auxiliary rows; Lemma 4.4 gives the groups with  $hD \leq 2M$  and  $|S_{t,g}| \leq 2n/r$ . The maps (4.17) and (4.18) give rows with  $\vec{P}_{f,v}(0) = \mathbf{e}_v^\top$ , degree at most  $R-1 \leq R$ , and the kernel condition; every row of  $C_{t,g}$  has absolute row sum at most  $D$ , and (4.19) gives  $\sum_g c_{t,g} \leq 2\tau n$ . Lemma 4.5 gives (D5).  $\square$

**4.7. Choosing the parameters.** The exponent of the routine depends on the three quantities of (3.22). The following choice bounds all three when  $b$  and  $w$  are bounded by a fixed power of  $\log n$ .

**Lemma 4.7** (Parameters). *Fix an integer  $\beta \geq 1$ . Let  $n \geq 20$ ,  $\lambda = \lceil \log n \rceil$ ,  $\nu = \lceil \log \lambda \rceil$ , and suppose  $b, w \leq \lambda^\beta$  and  $n \leq M \leq bn$ . For all sufficiently large  $n$  the parameters*

$$r = \left\lfloor \frac{\sqrt{\lambda}}{\nu^2} \right\rfloor, \quad Q = \left\lfloor \frac{\lambda}{128\nu} \right\rfloor, \quad R = (\beta + 3)\lceil \sqrt{\lambda} \rceil, \quad \tau = \frac{1}{r\lambda}, \quad \epsilon = \frac{1}{br^2\lambda}, \quad D = r^2Q, \quad (4.21)$$

*with the interval  $\mathcal{J}$  specified in the proof, satisfy the hypotheses of Lemma 4.6; moreover  $R \geq 32$ ,  $6R \leq Q$ ,  $D \leq 2^{\lambda/128}$ , and*

$$\frac{R}{Q} \log \frac{2Q}{R} + \tau \log(2D + 1) + \frac{1}{r} = O_\beta\left(\frac{\nu^2}{\sqrt{\lambda}}\right). \quad (4.22)$$

Given  $n, M, b, w$  the parameters can be computed and the hypotheses checked in  $\text{poly}_\beta(\log n)$  bit operations, and for fixed  $\beta$  only finitely many  $n$  fail the checks.<sup>12</sup>

*Proof.* For large  $n$ ,  $r = \Theta(\sqrt{\lambda}/\nu^2)$ ,  $Q = \Theta(\lambda/\nu)$  and  $R = \Theta_\beta(\sqrt{\lambda})$ ; so eventually  $r \geq 2$ ,  $R \geq 32$  and  $6R \leq Q$ . Since  $\log r \leq \nu/2$ ,  $\log D = 2Q \log r \leq \lambda/128$ . The balance inequalities (4.15) hold for large  $n$ :  $2(b+w)D^2 \leq 4\lambda^\beta 2^{\lambda/64} \leq n \leq M$  and  $2Qr^2 \leq n$ , because  $Q, r$  are bounded by fixed powers of  $\lambda$  while  $n > 2^{\lambda-1}$ . Put

$$\mu = \frac{Q(r-1)}{2r}, \quad A_0 = \left\lfloor \mu - \frac{R-1}{2} \right\rfloor, \quad \mathcal{J} = \{A_0, \dots, A_0 + R - 1\}.$$

Then  $Q/4 \leq \mu \leq Q/2$ , and since  $R \leq Q/6$  the interval lies in  $\{0, \dots, Q\}$ ; its centre is within one of  $\mu$ , and  $R \geq 32$  makes it contain every integer within  $R/4$  of  $\mu$ . For  $W \sim \text{Bin}(Q, \frac{r-1}{2r})$ , Hoeffding's inequality (2.8) with  $s = R/4$  gives

$$\Pr[W \notin \mathcal{J}] \leq 2e^{-R^2/(8Q)} \leq 2e^{-16(\beta+3)^2\nu} \leq 2^{-(\beta+2)\nu} \leq \lambda^{-(\beta+2)} \leq \frac{1}{br^2\lambda},$$

using  $R^2/(8Q) \geq (\beta+3)^2\lambda \cdot 128\nu/(8\lambda) = 16(\beta+3)^2\nu$ ,  $b \leq \lambda^\beta$  and  $r^2 \leq \lambda$ . This is (4.20) with  $\epsilon = \tau/(rb)$ , and  $0 < \tau \leq 1/2$  once  $r \geq 2$ . Finally  $R/Q = O_\beta(\nu/\sqrt{\lambda})$ ,  $\log(2Q/R) = O(\nu)$  and  $\tau \log(2D+1) = O(1/r)$ , while  $1/r = O(\nu^2/\sqrt{\lambda})$ , which proves (4.22).

For the computation,  $\lambda$  is the least integer with  $n \leq 2^\lambda$  and  $\nu$  the least with  $\lambda \leq 2^\nu$ ;  $r$  is the largest integer with  $r^2\nu^4 \leq \lambda$ ,  $Q$  is an integer division, and  $\lceil \sqrt{\lambda} \rceil$  comes from an integer square root. The algorithm reports failure if  $r < 2$  or  $Q \leq R$ ; otherwise it forms  $\tau, \epsilon, D$  and the endpoints of  $\mathcal{J}$  in rational arithmetic and checks  $\mathcal{J} \subseteq \{0, \dots, Q\}$ , (4.15),  $R \geq 32$ ,  $6R \leq Q$ , and (4.20), whose left side has denominator  $(2r)^Q$  and is an exact rational computation. All of this takes  $\text{poly}_\beta(\log n)$  bit operations, and the estimates above show that every check passes for large  $n$ .  $\square$

## 5. WEIGHTED COUNTING IN SUBEXPONENTIAL TIME

**5.1. Bounded occurrences.** We first state the outcome of Sections 3 and 4 for a factor system in which every variable occurs in few factors.

**Theorem 5.1** (Bounded occurrences). *Fix an integer  $\beta \geq 1$ . There are  $n_0 = n_0(\beta)$  and an absolute constant  $C_0$  with the following property. Let  $\mathcal{A}$  be a commutative  $\mathbb{Q}$ -algebra, let  $n \geq n_0$ , and let a factor system on  $n$  variables have largest scope  $w$ , largest occurrence number  $b$  and total degree  $d$  with*

$$w \leq \lambda^\beta, \quad b \leq \lambda^\beta, \quad d \leq \kappa n$$

*for some  $\kappa \geq 1$ . Then  $Z$  in (2.4) can be computed deterministically with at most*

$$|\mathcal{F}| \text{ poly}(n) 2^{C_0(\kappa+1)(\beta+3)n\nu^2/\sqrt{\lambda}} \quad (5.1)$$

*operations in  $\mathcal{A}$ , evaluations of factors at inputs in  $\mathbb{Q}[z]/(z^J)$  with  $J \leq \kappa n$ , and bit operations on rational numbers, using polynomially many elements of  $\mathcal{A}$  and polynomially many bits of storage in addition to the factors' own.*

*Proof.* Adjoin the constant-one factors of Section 4 so that (4.1) holds;  $b$  grows by at most one and  $d$  is unchanged, and we may take  $n_0$  so large that  $\lambda^\beta + 1 \leq \lambda^{\beta+1}$  is not needed: replace  $\beta$  by  $\beta+1$  in Lemma 4.7 if  $b$  was exactly  $\lambda^\beta$ , which changes only the constants. Choose the parameters (4.21) by Lemma 4.7, which for  $n \geq n_0(\beta)$  satisfy the hypotheses of Lemma 4.6; construct the design in polynomial time by that lemma; and run the routine of Proposition 3.5.

<sup>12</sup>Lemma 4.8 of [OAI26].

The degree bound (3.1) holds with  $d \leq \kappa n$ , so  $d_H \leq \kappa n R$ ,  $q \leq \kappa n R/Q$  and  $J \leq q + 1 \leq \kappa n$  because  $R < Q$ . By (3.21),

$$\log C_q \leq \kappa n \frac{R}{Q} \log \frac{Q}{R} + \frac{3nR}{Q} \leq (\kappa + 1) n \frac{R}{Q} \log \frac{2Q}{R} \cdot 3.$$

With the parameters (4.21) and  $n$  large,  $R/Q \leq 300(\beta + 3)\nu/\sqrt{\lambda}$  and  $\log(2Q/R) \leq \nu$ , while  $\tau \log(2D + 1) \leq 1/r$  and  $1/r \leq 2\nu^2/\sqrt{\lambda}$ . Hence the logarithm of the vector count (3.20) is at most

$$\log(2Qr) + 900(\kappa + 1)(\beta + 3) \frac{n\nu^2}{\sqrt{\lambda}} + \frac{4n\nu^2}{\sqrt{\lambda}} \leq \log(2Qr) + C_0(\kappa + 1)(\beta + 3) \frac{n\nu^2}{\sqrt{\lambda}}$$

with  $C_0 = 904$ , and each vector costs at most  $|\mathcal{F}|$  factor evaluations and polynomially many further operations. The storage bound is that of Proposition 3.5.  $\square$

**5.2. Branching on frequent variables.** A sparse system may still contain variables occurring in many factors. Fixing them removes the obstacle at a cost that is negligible against (5.1).

**Lemma 5.2** (Frequent variables). *Let a factor system on  $n \geq 20$  variables have  $|\mathcal{F}| \leq cn$  factors, each of scope at most  $k$ , and let  $V_{\text{hi}} = \{v \in [n] : \#\{f : v \in I_f\} > ck\lambda\}$ . Then  $|V_{\text{hi}}| < n/\lambda$ . For every assignment  $y \in \{0, 1\}^{V_{\text{hi}}}$  the system obtained by substituting  $x_v = y_v$  for  $v \in V_{\text{hi}}$  is a factor system on the  $n' = n - |V_{\text{hi}}|$  remaining variables, with at most  $cn$  factors, scopes at most  $k$ , occurrence numbers at most  $ck\lambda$  and total degree at most  $d$ , and*

$$Z = \sum_{y \in \{0, 1\}^{V_{\text{hi}}}} Z_y, \quad (5.2)$$

where  $Z_y$  is the sum (2.4) of the substituted system.

*Proof.* The incidences number  $M \leq k|\mathcal{F}| \leq ckn$ , so fewer than  $ckn/(ck\lambda) = n/\lambda$  variables occur more than  $ck\lambda$  times. Substituting values for the variables in  $V_{\text{hi}}$  restricts each  $\Phi_f$  to a polynomial in the remaining variables of its scope, of no larger degree; the scopes shrink and the occurrence numbers of the remaining variables are unchanged, hence at most  $ck\lambda$ . Splitting the sum over  $\{0, 1\}^n$  according to  $x|_{V_{\text{hi}}}$  gives (5.2).  $\square$

### 5.3. The weighted counting theorem.

**Theorem 5.3** (Weighted counting). *Fix integers  $k \geq 1$  and  $c \geq 1$ . There is a deterministic algorithm that, given a factor system on  $n \geq 20$  variables with at most  $cn$  factors, each of scope at most  $k$  and with values in a commutative  $\mathbb{Q}$ -algebra  $\mathcal{A}$ , computes*

$$Z = \sum_{x \in \{0, 1\}^n} \prod_{f \in \mathcal{F}} \Phi_f((x_v)_{v \in I_f})$$

with

$$\text{poly}(n) 2^{O_{k,c}(\rho(n))}, \quad \rho(n) = \frac{n(\log \log n)^2}{\sqrt{\log n}}, \quad (5.3)$$

operations in  $\mathcal{A}$  and evaluations of the factors at inputs in  $\mathbb{Q}[z]/(z^J)$ ,  $J \leq ckn$ , together with polynomially many bit operations on rationals, and polynomially many elements of  $\mathcal{A}$  and bits of storage. The implied constant depends only on  $k$  and  $c$ .

*Proof.* Let  $n_0 = n_0(2)$  be the threshold of Theorem 5.1 with  $\beta = 2$ , enlarged so that  $ck\lambda \leq (\lambda - 1)^2$  and  $k \leq (\lambda - 1)^2$  for  $n \geq n_0$ . If  $n < 2n_0$ , enumerate the cube directly:  $2^n \leq 2^{2n_0}$  is a constant depending on  $k, c$ . Otherwise  $n \geq 2n_0$ ; apply Lemma 5.2. For each of the fewer than  $2^{n/\lambda}$  assignments  $y$ , the substituted system has  $n' \geq n - n/\lambda \geq n/2 \geq n_0$  variables, scopes at most  $k \leq \lambda'^2$ , occurrences at most  $ck\lambda \leq \lambda'^2$ , where  $\lambda' = \lceil \log n' \rceil \geq \lambda - 1$ , and

total degree  $d \leq kcn \leq 2ckn'$ . Theorem 5.1 with  $\beta = 2$  and  $\kappa = 2ck$  computes  $Z_y$  with  $|\mathcal{F}| \text{poly}(n) 2^{5C_0(2ck+1)n'\nu'^2/\sqrt{\lambda'}}$  operations, and  $n'\nu'^2/\sqrt{\lambda'} \leq 2n\nu^2/\sqrt{\lambda}$  for large  $n$ . Summing the  $Z_y$  costs one operation each. The total is

$$2^{n/\lambda} \cdot cn \text{poly}(n) 2^{10C_0(2ck+1)n\nu^2/\sqrt{\lambda}} = \text{poly}(n) 2^{O_{k,c}(n\nu^2/\sqrt{\lambda})},$$

since  $n/\lambda \leq n\nu^2/\sqrt{\lambda}$ , and  $n\nu^2/\sqrt{\lambda} = \Theta(\rho(n))$ . The space bound follows from Theorem 5.1 because the branches are processed one after the other.  $\square$

*Remark 5.4.* For  $\mathcal{A} = \mathbb{Q}$  and  $\Phi_f$  the satisfaction indicator of a clause,  $Z$  is the number of satisfying assignments and the theorem is the sparse case of OpenAI's Theorem 1.1, with the exponent  $\rho(n)$  of their Proposition 6.4 in place of  $n \log \log n / \sqrt{\log n}$ . The dense case, and the better exponent, use their Sections 5, 6 and Appendices A and B, which are not reproduced here; the dense case is recovered in Section 8 from the published sparsification lemma, with a weaker time bound.

#### 5.4. Bit complexity over truncated polynomial rings.

**Lemma 5.5** (Bit complexity). *In the situation of Theorem 5.1, of Theorem 7.19 or of Theorem 5.3, with  $|\mathcal{F}|$  factors of scope at most  $k$  and total degree  $d$ , let  $\mathcal{A}$  be one of the algebras  $\mathcal{A}_N = \mathbb{Q}[W]/(W^N)$ ,  $\mathcal{A}_{N_1, \dots, N_s}$  of dimension  $N = N_1 \cdots N_s$ , or  $\mathbb{Q}[W]/(W^N - 1)$ , and suppose every coefficient of every  $\Phi_f$ , as a polynomial in the  $Y_v$  with coefficients in  $\mathcal{A}$ , is a polynomial in the generators with integer coefficients of absolute value at most  $2^L$ . Then every rational number that the algorithm forms has numerator and denominator of  $\text{poly}(n, |\mathcal{F}|, k, N, L)$  bits; hence the operation counts of those theorems turn into bounds on bit operations after multiplication by a polynomial in  $n, |\mathcal{F}|, k, N$  and  $L$ , and the space stays polynomial.*

*Proof.* All numbers arise in the following places. (a) The rows: by (D5),  $d_{\text{den}} \vec{P}_{f,v}$  has integer coefficients of  $O(Q \log Q + \log D) = O(\lambda)$  bits, and the values  $\langle \vec{P}_{f,v}(t), x \rangle$  for Boolean  $x$  and the coefficients of  $\langle \Delta \vec{P}_{f,v}(z), \mathbf{1}_U \rangle$  are sums of at most  $n$  such coefficients times powers  $t^i \leq Q^R$ , hence rationals with denominator  $d_{\text{den}}$  and numerators of  $O(\lambda + R \log Q + \log n) = O(\lambda)$  bits. (b) A factor evaluation:  $\Phi_f$  has degree at most  $k$ , so its value at inputs in  $\mathbb{Q}[z]/(z^J)$  with the bounds of (a) is an element of  $\mathcal{A}[z]/(z^J)$  with denominator  $d_{\text{den}}^k$  and numerators of  $O(L + k\lambda + k \log J)$  bits. (c) A group product is a product of at most  $|\mathcal{F}|$  factor values in  $\mathcal{A}[z]/(z^J)$ ; the number of monomials contributing to one coefficient is at most  $(NJ)^{|\mathcal{F}|}$ , so the product has denominator  $d_{\text{den}}^{k|\mathcal{F}|}$  and numerators of  $O(|\mathcal{F}|(L + k\lambda + k \log(NJ)))$  bits. (d) The sums (3.6), (3.7) and (3.16) add at most  $2^n \cdot 3^n \cdot (2D + 1)^{2\tau n}$  terms with a common denominator, adding  $O(n + n \log(2D + 1)) = O(n\lambda)$  bits to the numerators. (e) Hermite interpolation solves a  $QJ \times QJ$  rational system with entries of  $O(QJ \log(QJ))$  bits and right-hand sides as in (d); by Cramer's rule the solution has denominator and numerators of  $\text{poly}(QJ) \cdot O(n\lambda + |\mathcal{F}|(L + k\lambda + k \log(NJ)))$  bits. Since  $Q, \lambda = O(\log n)$  and  $J \leq d + 1 \leq k|\mathcal{F}| + 1$ , all bit lengths are  $\text{poly}(n, |\mathcal{F}|, k, N, L)$ , and each operation in  $\mathcal{A}$  costs  $O(N^2)$  operations on such rationals. The count of monomials in (c) used only that a product of two basis elements of  $\mathcal{A}$  is a basis element or zero, which holds in all three algebras. The case  $\mathcal{A} = \mathbb{Q}$  is  $N = 1$ .  $\square$

## 6. PROFILES, OPTIMA AND POLYNOMIALS

This section draws the consequences of Theorem 5.3 and of Theorem 5.1. The pattern is the same each time: a combinatorial quantity is written as the sum (2.4) of a factor system over a cube, the algebra is chosen so that the sum records what one wants to know, and the two theorems are applied. The only place where something has to be added to the summation routine

is the independence polynomial of a graph without any bound on the degrees, in Section 6.2, where a branching rule reduces the general graph to graphs of small degree; the cost of the rule is bounded by an explicit induction.

**6.1. Profiles of sparse constraint systems.** A *constraint system* on  $n$  Boolean variables consists of constraints  $C_1, \dots, C_m$ , where  $C_f : \{0, 1\}^{I_f} \rightarrow \{0, 1\}$  is given by its truth table on a scope  $I_f \subseteq [n]$ , together with nonnegative integer weights  $\omega_1, \dots, \omega_m$ . For an assignment  $x$  write

$$\text{sat}_\omega(x) = \sum_{f: C_f(x|_{I_f})=1} \omega_f$$

for the weight of the constraints it satisfies. The *profile* of the system is the sequence

$$p_j = \#\{x \in \{0, 1\}^n : \text{sat}_\omega(x) = j\}, \quad 0 \leq j \leq \Omega_{\text{tot}} = \sum_f \omega_f. \quad (6.1)$$

**Theorem 6.1** (Profiles). *Fix integers  $k \geq 1$  and  $c \geq 1$ . There is a deterministic algorithm that, given a constraint system on  $n \geq 20$  variables with  $m \leq cn$  constraints of scope at most  $k$  and weights  $\omega_f \leq \Omega$ , computes the whole profile (6.1) in time  $\text{poly}(n, \Omega) 2^{O_{k,c}(\rho(n))}$  and polynomial space.*

*Proof.* Let  $N = 1 + \Omega_{\text{tot}} \leq 1 + cn\Omega$  and  $\mathcal{A} = \mathcal{A}_N$ . For each  $f$  let  $\Phi_f$  be the multilinear extension (2.6) of the function  $\phi_f(x) = W^{\omega_f C_f(x)}$  on  $\{0, 1\}^{I_f}$ ; its coefficients are the polynomials  $\sum_{U \subseteq S} (-1)^{|S|-|U|} \phi_f(\mathbf{1}_U)$ , each a sum of at most  $2^k$  monomials  $\pm W^j$ , so they are polynomials in  $W$  with integer coefficients of absolute value at most  $2^k$ . The system  $(\Phi_f)$  is a factor system with at most  $cn$  factors of scope at most  $k$ , and by (2.7), since  $\sum_f \omega_f C_f(x) = \text{sat}_\omega(x) \leq \Omega_{\text{tot}} < N$ ,

$$Z = \sum_x \prod_f W^{\omega_f C_f(x|_{I_f})} = \sum_j p_j W^j$$

with no term lost to the truncation. Theorem 5.3 computes  $Z$  with  $\text{poly}(n) 2^{O_{k,c}(\rho(n))}$  operations in  $\mathcal{A}_N$  and factor evaluations, and Lemma 5.5 with  $L = k$  bounds the bit lengths by  $\text{poly}(n, N)$ , so the bit cost is  $\text{poly}(n, \Omega) 2^{O_{k,c}(\rho(n))}$ . The coefficients  $p_j$  are read off from  $Z$ .  $\square$

The profile contains, besides the optimum, every quantity that depends only on the distribution of the objective over the cube.

**Corollary 6.2** (Max- $k$ -SAT). *Fix  $k, c \geq 1$ . Given a  $k$ -CNF formula with  $n$  variables and  $m \leq cn$  clauses with integer weights in  $[0, \Omega]$ , the following are computed deterministically in time  $\text{poly}(n, \Omega) 2^{O_{k,c}(\rho(n))}$ : the maximum weight of simultaneously satisfiable clauses, the number of assignments attaining it, the minimum such weight, and for every  $j$  the number of assignments satisfying clauses of total weight exactly  $j$ . In particular Max- $k$ -SAT, Min- $k$ -SAT and weighted Max- $k$ -SAT with polynomially bounded weights are solved exactly in time  $2^{o(n)}$  on instances with  $O(n)$  clauses, and the number of optimal assignments is counted.*

*Proof.* A clause is a constraint whose truth table is the disjunction of its literals. Apply Theorem 6.1; the maximum is the largest  $j$  with  $p_j > 0$ , the minimum the smallest, and  $p_j$  is the required count.  $\square$

The Hamming weight of an assignment is itself a sum of local weights, one per variable, so the profile may be taken with respect to it.

**Corollary 6.3** (Satisfying assignments of each weight). *Fix  $k, c \geq 1$ . Given a  $k$ -CNF formula with  $n$  variables and  $m \leq cn$  clauses, the numbers*

$$s_i = \#\{x : x \text{ satisfies the formula and has exactly } i \text{ ones}\}, \quad 0 \leq i \leq n,$$

are computed deterministically in time  $\text{poly}(n)2^{O_{k,c}(\rho(n))}$ . Hence the satisfying assignment with the fewest, or the most, ones is found and counted within the same time.

*Proof.* Take  $\mathcal{A} = \mathcal{A}_{n+1}$ . For each clause  $f$  let  $\Phi_f$  be the multilinear extension of its satisfaction indicator, with coefficients in  $\{0, \pm 1\} \subseteq \mathbb{Q}$ , and for each variable  $v$  add the factor  $\Phi_v = 1 + (W - 1)Y_v$  of scope  $\{v\}$ , the multilinear extension of  $W^{x_v}$ . The system has at most  $(c+1)n$  factors of scope at most  $k$ , and its sum is  $Z = \sum_x [x \text{ satisfies the formula}] W^{|x|} = \sum_i s_i W^i$ , where  $|x|$  is the number of ones. Theorem 5.3 and Lemma 5.5 with  $N = n+1$ ,  $L = 1$  give the time bound. A satisfying assignment of a prescribed weight  $i$  with  $s_i > 0$  is then found by the usual self-reduction: fix the first variable to 0 and to 1, recompute the profiles of the two restricted formulas, which are again sparse  $k$ -CNF formulas, keep a value for which the coefficient of  $W^i$  or  $W^{i-1}$  is positive, and continue; this multiplies the time by at most  $2n$ .  $\square$

The next corollary uses an algebra in two variables, one recording the cut and one the size of a side.

**Corollary 6.4** (Max-Cut and bisections). *Fix  $c \geq 1$ . Given a graph  $G$  with  $n$  vertices and  $m \leq cn$  edges with integer weights in  $[0, \Omega]$ , the numbers*

$$q_{i,j} = \#\{S \subseteq V(G) : |S| = i, \text{ the edges between } S \text{ and } V \setminus S \text{ have total weight } j\}$$

*are computed for all  $i$  and  $j$  in deterministic time  $\text{poly}(n, \Omega)2^{O_c(\rho(n))}$ . In particular the maximum cut, the maximum and the minimum bisection, with the number of optimal solutions of each, are computed exactly in time  $2^{o(n)}$  on graphs with  $O(n)$  edges.*

*Proof.* Let  $N_1 = 1 + \sum_e \omega_e \leq 1 + cn\Omega$  and  $N_2 = n+1$ , and let  $\mathcal{A} = \mathcal{A}_{N_1, N_2} = \mathbb{Q}[W, U]/(W^{N_1}, U^{N_2})$ . Encode  $S$  by  $x \in \{0, 1\}^n$ . For an edge  $e = \{u, v\}$  take the multilinear extension of  $W^{\omega_e[x_u \neq x_v]}$ , which is  $1 + (W^{\omega_e} - 1)(Y_u + Y_v - 2Y_u Y_v)$ , and for each vertex  $v$  the factor  $1 + (U - 1)Y_v$ . There are at most  $(c+1)n$  factors of scope at most 2, with coefficients that are polynomials in  $W$  and  $U$  with integer coefficients of absolute value at most 2. The sum is  $Z = \sum_x W^{\text{cut}(x)} U^{|x|} = \sum_{i,j} q_{i,j} W^i U^j$ , and no term is truncated. Theorem 5.3 and Lemma 5.5 with  $N = N_1 N_2$  and  $L = 1$  give the bound. The maximum cut is the largest  $j$  with  $\sum_i q_{i,j} > 0$ , the maximum bisection the largest  $j$  with  $q_{\lfloor n/2 \rfloor, j} > 0$ , the minimum bisection the smallest such  $j$ , and the counts are the coefficients.  $\square$

*Remark 6.5.* The best exact algorithm for Max-Cut on general graphs runs in time  $2^{\omega n/3} \leq 2^{0.792n}$  [Wil05], and the exact algorithms for sparse Max-Cut run in time  $2^{\Theta(m)}$  [SS09]; on graphs with  $m = O(n)$  edges both are exponential in  $n$ . The proof of OpenAI's Corollary 1.2 contains a circuit encoding that, combined with their Theorem 1.1, also yields the exact optimum of a sparse Max- $k$ -SAT or Max-Cut instance in time  $2^{o(n)}$ , by a binary search over the threshold; that route handles the unweighted objective. The algebra  $\mathcal{A}_N$  handles integer weights up to  $\Omega$  at a cost polynomial in  $\Omega$ , which a circuit encoding cannot do without introducing  $\Theta(n \log \Omega)$  auxiliary variables.

**6.2. The independence polynomial of every graph.** Let  $G$  be a graph with vertex set  $V$ ,  $|V| = n$ , and let  $\omega : V \rightarrow \{0, 1, \dots\}$  be integer vertex weights. The weighted independence polynomial is

$$I_\omega(G; W) = \sum_{S \subseteq V \text{ independent}} W^{\omega(S)}, \quad \omega(S) = \sum_{v \in S} \omega_v, \quad (6.2)$$

and  $I(G; W) = I_1(G; W)$  is the independence polynomial. For a vertex  $v$  write  $N[v]$  for its closed neighbourhood and  $G - X$  for the graph induced on  $V \setminus X$ . Every independent set either avoids  $v$  or contains  $v$  and avoids  $N[v]$ , so

$$I_\omega(G; W) = I_\omega(G - v; W) + W^{\omega_v} I_\omega(G - N[v]; W). \quad (6.3)$$

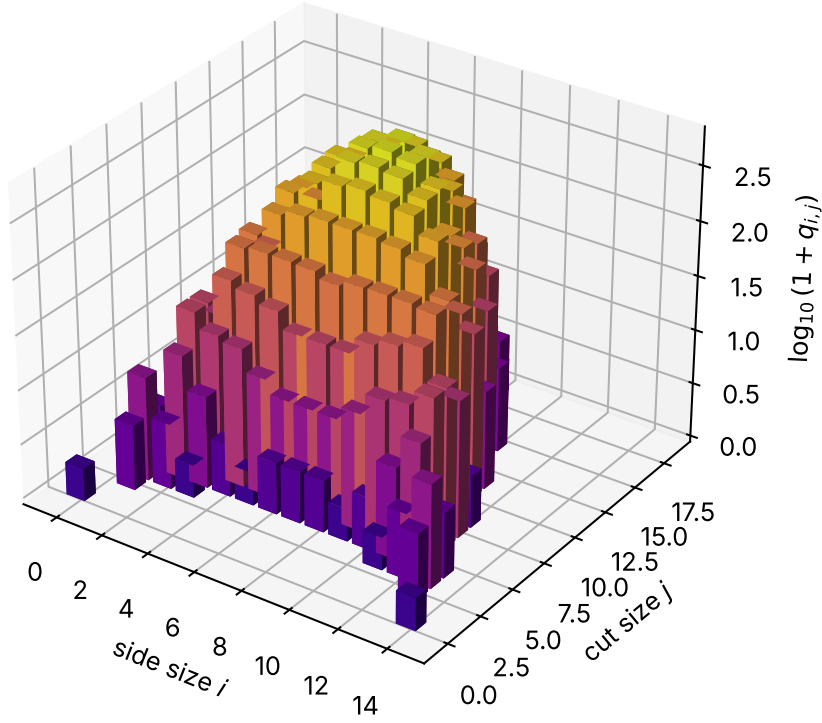


FIGURE 2. The array  $q_{i,j}$  of Corollary 6.4 for a graph with 14 vertices and 21 edges, a Hamilton cycle with seven chords, computed directly from the  $2^{14}$  subsets; the height is  $\log_{10}(1 + q_{i,j})$ . The algorithm returns the whole array as one element of  $\mathbb{Q}[W, U]/(W^{22}, U^{15})$ ; its maximum cut and its maximum bisection are both 18.

**Lemma 6.6** (Forcing branching). *Let  $b \geq 2$  be an integer and  $G$  a weighted graph with  $n$  vertices. There is a binary tree, computable in time  $\text{poly}(n)$  per node, whose leaves  $\ell$  carry induced subgraphs  $H_\ell$  of  $G$  of maximum degree at most  $b$  and integers  $s_\ell \geq 0$  such that*

$$I_\omega(G; W) = \sum_{\ell} W^{s_\ell} I_\omega(H_\ell; W), \quad (6.4)$$

and the number of leaves is at most

$$\left(1 + \frac{2 \ln(b+2)}{b+1}\right)^n. \quad (6.5)$$

*Proof.* Build the tree from the root  $(G, 0)$ . At a node  $(H, s)$  in which some vertex  $v$  has degree greater than  $b$  in  $H$ , attach the two children  $(H - v, s)$  and  $(H - N_H[v], s + \omega_v)$ ; by (6.3) the sum  $\sum W^s I_\omega(H; W)$  over the nodes of any antichain that cuts every root-to-leaf path is the same, and it equals  $I_\omega(G; W)$  at the root. A node in which every vertex has degree at most  $b$  is a leaf. This gives (6.4), and each node is processed in polynomial time.

Let  $L(n)$  be the largest number of leaves of such a tree over all weighted graphs with at most  $n$  vertices, so that  $L$  is nondecreasing and  $L(n) = 1$  for  $n \leq b + 1$ , because a graph with at most  $b + 1$  vertices has maximum degree at most  $b$ . If  $v$  has degree at least  $b + 1$  then  $H - v$  has  $n - 1$

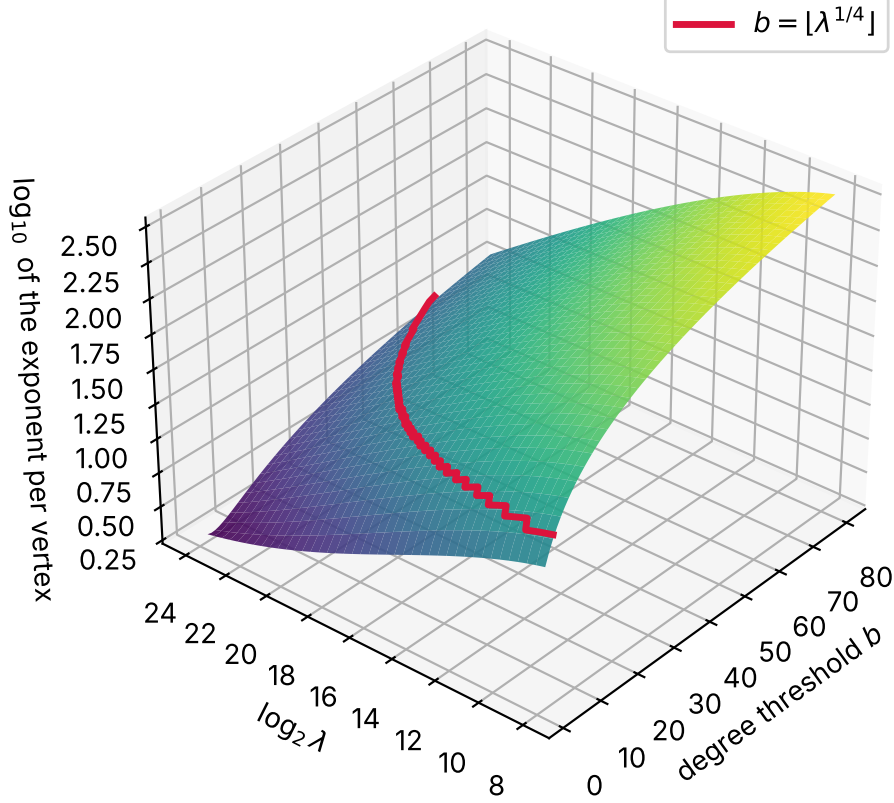


FIGURE 3. The two costs of the proof of Theorem 6.7 per vertex, in  $\log_{10}$ , with the constant  $C_0$  omitted: the leaf count  $2 \ln(b+2)/(b+1)$  of Lemma 6.6 and the leaf cost  $(b+2)\nu^2/\sqrt{\lambda}$  of Theorem 5.1, as a function of the degree threshold  $b$  and of  $\lambda$ ; the curve marks the choice  $b = \lfloor \lambda^{1/4} \rfloor$  of the proof.

vertices and  $H - N_H[v]$  at most  $n - b - 2$ , so

$$L(n) \leq L(n-1) + L(n-b-2) \quad (n \geq b+2),$$

with  $L(j) = 1$  for  $j \leq b+1$ , including  $j \leq 0$ . Put  $\varepsilon = 2 \ln(b+2)/(b+1)$ . We claim  $L(n) \leq (1+\varepsilon)^n$  for all  $n \geq 0$ . This holds for  $n \leq b+1$ . For  $n \geq b+2$ , by induction,

$$L(n) \leq (1+\varepsilon)^{n-1} + (1+\varepsilon)^{n-b-2} = (1+\varepsilon)^{n-b-2}((1+\varepsilon)^{b+1} + 1),$$

and the right side is at most  $(1+\varepsilon)^n$  exactly when  $\varepsilon(1+\varepsilon)^{b+1} \geq 1$ . Now  $b \geq 2$  gives  $2 \ln(b+2) \leq b+1$ , since this holds at  $b=2$  ( $2 \ln 4 < 3$ ) and the difference  $b+1 - 2 \ln(b+2)$  increases with  $b$ ; so  $\varepsilon \leq 1$ , and for  $0 \leq \varepsilon \leq 1$  one has  $\ln(1+\varepsilon) \geq \varepsilon - \varepsilon^2/2 \geq \varepsilon/2$ . Therefore

$$(1+\varepsilon)^{b+1} \geq e^{(b+1)\varepsilon/2} = e^{\ln(b+2)} = b+2, \quad \varepsilon(1+\varepsilon)^{b+1} \geq \frac{2 \ln(b+2)}{b+1}(b+2) \geq 2 \ln 4 > 1.$$

This proves the claim, and (6.5) follows.  $\square$

**Theorem 6.7** (Independence polynomial). *There is a deterministic algorithm that, given a graph  $G$  with  $n$  vertices and integer vertex weights in  $[0, \Omega]$ , computes the weighted independence*

polynomial  $I_\omega(G; W)$  in time

$$\text{poly}(n, \Omega) 2^{O(n(\log \log n)^2 / (\log n)^{1/4})}$$

and polynomial space. No condition is placed on the graph.

*Proof.* Let  $n_0 = n_0(1)$  be the threshold of Theorem 5.1 with  $\beta = 1$  and let  $n_1 \geq n_0^2$  be so large that for  $n \geq n_1$  the following hold, with  $\lambda = \lceil \log n \rceil$  and  $b = \lfloor \lambda^{1/4} \rfloor$ :  $b \geq 2$ ,  $\lambda^{1/4} + 1 \leq \lambda/2 - 1$ , and  $\lambda/2 - 1 \geq \lambda/3$ . For  $n < n_1$  compute  $I_\omega(G; W)$  by listing the  $2^n < 2^{n_1}$  subsets, a constant. For  $n \geq n_1$  apply Lemma 6.6 with this  $b$  and compute  $I_\omega(H_\ell; W)$  at each leaf as follows.

If  $H = H_\ell$  has  $n_H < \sqrt{n}$  vertices, list its subsets; this costs  $2^{\sqrt{n}} \text{poly}(n, \Omega)$ . Otherwise  $n_H \geq \sqrt{n}$ , so  $\lambda_H = \lceil \log n_H \rceil \geq \lambda/2 - 1 \geq \lambda/3$  and  $n_H \geq n_0$ . Take  $\mathcal{A} = \mathcal{A}_N$  with  $N = 1 + \sum_v \omega_v \leq 1 + n\Omega$ , and the factor system on the  $n_H$  variables  $x_v$ ,  $v \in V(H)$ , with the factor  $1 - Y_u Y_v$  for each edge  $uv$  of  $H$  and the factor  $1 + (W^{\omega_v} - 1)Y_v$  for each vertex  $v$ . The edge factor is the multilinear extension of the indicator that  $u$  and  $v$  are not both chosen, and the vertex factor that of  $W^{\omega_v x_v}$ , so the sum (2.4) is  $I_\omega(H; W)$ , with no term truncated. Each variable occurs in at most  $\deg_H(v) + 1 \leq b + 1 \leq \lambda^{1/4} + 1 \leq \lambda_H$  factors, the scopes are at most  $2 \leq \lambda_H$ , and the total degree is  $d = 2|E(H)| + n_H \leq (b + 1)n_H$ . Theorem 5.1 with  $\beta = 1$  and  $\kappa = b + 1$  applies and computes  $I_\omega(H; W)$  with

$$|\mathcal{F}| \text{poly}(n_H) 2^{4C_0(b+2)n_H\nu_H^2/\sqrt{\lambda_H}} \leq \text{poly}(n) 2^{4C_0(\lambda^{1/4}+2)n\nu^2\sqrt{3/\lambda}} = 2^{O(n\nu^2/\lambda^{1/4})}$$

operations in  $\mathcal{A}_N$ , using  $\nu_H \leq \nu$  and  $\lambda_H \geq \lambda/3$ ; Lemma 5.5 with  $|\mathcal{F}| \leq (b + 1)n$ ,  $k = 2$ ,  $L = 1$  turns this into  $\text{poly}(n, \Omega) 2^{O(n\nu^2/\lambda^{1/4})}$  bit operations.

By (6.5) the tree has at most  $(1 + \varepsilon)^n \leq e^{\varepsilon n}$  leaves with  $\varepsilon = 2 \ln(b + 2)/(b + 1) \leq 4 \ln \lambda / \lambda^{1/4} = O(\nu / \lambda^{1/4})$ , hence at most  $2^{O(n\nu/\lambda^{1/4})}$  leaves and fewer internal nodes, each processed in polynomial time, and the leaf polynomials are combined by (6.4) with one multiplication by a power of  $W$  and one addition in  $\mathcal{A}_N$  per leaf. The tree is traversed depth first, so only one root-to-leaf path of length at most  $n$  is stored at a time, and the space is polynomial. The total time is

$$2^{O(n\nu/\lambda^{1/4})} \cdot (2^{\sqrt{n}} + 2^{O(n\nu^2/\lambda^{1/4})}) \text{poly}(n, \Omega) = \text{poly}(n, \Omega) 2^{O(n\nu^2/\lambda^{1/4})},$$

and  $n\nu^2/\lambda^{1/4} = \Theta(n(\log \log n)^2 / (\log n)^{1/4})$ .  $\square$

**Corollary 6.8.** *Each of the following is computed deterministically, for every graph  $G$  on  $n$  vertices, in time  $2^{O(n(\log \log n)^2 / (\log n)^{1/4})}$ : the independence number  $\alpha(G)$  and the number of maximum independent sets; the number of independent sets of each size, hence the partition function of the hard-core model at every activity; the clique number  $\omega(G)$  and the number of maximum cliques; the vertex cover number  $\tau(G)$  and the number of minimum vertex covers; and, for integer weights in  $[0, \Omega]$ , the maximum weight of an independent set and the number of independent sets of that weight, with a factor  $\text{poly}(\Omega)$  in the time.*

*Proof.*  $\alpha(G)$  is the degree of  $I(G; W)$ , its leading coefficient counts the maximum independent sets, and the coefficient of  $W^i$  counts the independent sets of size  $i$ ; the hard-core partition function at activity  $z$  is  $I(G; z)$ . The cliques of  $G$  are the independent sets of the complement, to which Theorem 6.7 applies without restriction. The vertex covers are the complements of the independent sets, so  $\tau(G) = n - \alpha(G)$  and the minimum vertex covers are counted by the same leading coefficient. The weighted statements read the degree and the leading coefficient of  $I_\omega(G; W)$ .  $\square$

*Remark 6.9.* The fastest published exact algorithm for the maximum independent set runs in time  $2^{0.1996n}$  [XN17], and  $2^{\Omega(n)}$  was a known consequence of the exponential time hypothesis. OpenAI's Corollary 1.3 gives the number of independent sets of every graph in time  $2^{o(n)}$  through

their Theorem 1.1 in its general form, which handles formulas with  $\Theta(n^2)$  clauses; that form, combined with a linear-size threshold circuit for the number of chosen vertices, also gives  $\alpha(G)$  with the exponent  $n \log \log n / \sqrt{\log n}$ . The route of Theorem 6.7 is different: it uses only the sparse routine reproduced in Sections 3 and 4, reaches every graph through the branching of Lemma 6.6, and returns the whole weighted polynomial. Its exponent is larger, by the factor  $(\log n)^{1/4} \log \log n$ , which is the price of the branching.

The branching of Lemma 6.6 records the weight of the chosen vertices only through the monomial  $W^{\omega_v}$  that multiplies the second child, and its proof uses nothing else about that monomial. Several weights can therefore be carried at once, with one generator for each, and a budget can be placed on each of them. No restriction is placed on the graph.

**Theorem 6.10** (Independent sets under budgets). *Fix  $s \geq 0$ . There is a deterministic algorithm that, given a graph  $G$  with  $n$  vertices and weight vectors  $\omega^{(0)}, \dots, \omega^{(s)} \in \{0, 1, \dots, \Omega\}^V$ , computes the numbers*

$$a(j_0, \dots, j_s) = \#\{S \subseteq V \text{ independent} : \omega^{(i)}(S) = j_i \ (0 \leq i \leq s)\}$$

for all tuples  $(j_0, \dots, j_s)$  in time

$$\text{poly}(n, (1 + n\Omega)^{s+1}) 2^{O(n(\log \log n)^2 / (\log n)^{1/4})}$$

and polynomial space, the implied constant depending only on  $s$ .

*Proof.* Put  $N_i = 1 + \sum_v \omega_v^{(i)} \leq 1 + n\Omega$  and  $\mathcal{A} = \mathcal{A}_{N_0, \dots, N_s}$ , of dimension  $N = N_0 \cdots N_s \leq (1 + n\Omega)^{s+1}$ , and for a vertex  $v$  write  $\mathbf{W}^v = \prod_{i=0}^s W_i^{\omega_v^{(i)}}$ , a basis monomial of  $\mathcal{A}$  since  $\omega_v^{(i)} < N_i$ . The polynomial

$$I(G) = \sum_{S \text{ independent}} \prod_{v \in S} \mathbf{W}^v = \sum_{(j_0, \dots, j_s)} a(j_0, \dots, j_s) W_0^{j_0} \cdots W_s^{j_s}$$

loses no term to the truncation, because  $\omega^{(i)}(S) < N_i$  for every  $S$ . It satisfies  $I(G) = I(G - v) + \mathbf{W}^v I(G - N[v])$  for every vertex  $v$ , by the argument for (6.3). The construction of the tree in Lemma 6.6 uses only this recursion, with the second child carrying the product of the monomials  $\mathbf{W}^v$  of the vertices fixed to 1 on the way, and the bound (6.5) on the number of leaves uses only the numbers of vertices removed. So the lemma holds verbatim for  $I(G)$ , with leaf terms  $\mathbf{m}_\ell I(H_\ell)$  in which  $\mathbf{m}_\ell$  is a monomial.

At a leaf  $H = H_\ell$  with  $n_H \geq \sqrt{n}$  vertices take, as in the proof of Theorem 6.7, the factor  $1 - Y_u Y_v$  for each edge  $uv$  of  $H$  and the factor  $1 + (\mathbf{W}^v - 1)Y_v$  for each vertex  $v$ , the multilinear extension of  $\prod_i W_i^{\omega_v^{(i)}} x_v$ . The sum (2.4) of this system is  $I(H)$ . The occurrence numbers, the scopes and the total degree are those of the proof of Theorem 6.7, since only the coefficient of  $Y_v$  in the vertex factors has changed, and that coefficient is a polynomial in the generators with integer coefficients of absolute value at most 1. Theorem 5.1 with  $\beta = 1$  and  $\kappa = b + 1$  gives  $\text{poly}(n) 2^{O(nv^2/\lambda^{1/4})}$  operations in  $\mathcal{A}$ , and Lemma 5.5 with  $L = 1$  turns them into  $\text{poly}(n, N) 2^{O(nv^2/\lambda^{1/4})}$  bit operations. Leaves with fewer than  $\sqrt{n}$  vertices are summed directly, each of the at most  $2^{\sqrt{n}}$  subsets costing  $\text{poly}(n, N)$ . The count of leaves, the depth-first traversal and the final estimate are those of the proof of Theorem 6.7, with  $\text{poly}(n, \Omega)$  replaced by  $\text{poly}(n, N)$ .  $\square$

The knapsack problem with a conflict graph [PS09] asks for a set of items of largest value whose total weight is within a capacity and which contains no two items joined by an edge of a given graph. It contains the maximum independent set, and with several capacities it contains the multidimensional knapsack problem.

**Corollary 6.11** (Knapsack with an arbitrary conflict graph). *Fix  $s \geq 1$ . Given a graph  $G$  on  $n$  vertices, values  $\omega_v^{(0)}$  and weights  $\omega_v^{(1)}, \dots, \omega_v^{(s)}$  in  $[0, \Omega]$ , and capacities  $B_1, \dots, B_s$ , the largest value  $\omega^{(0)}(S)$  of an independent set  $S$  of  $G$  with  $\omega^{(i)}(S) \leq B_i$  for  $1 \leq i \leq s$ , and the number of independent sets attaining it, are computed deterministically in time  $\text{poly}(n, \Omega^{s+1}) 2^{O(n(\log \log n)^2 / (\log n)^{1/4})}$ . The same holds with any of the inequalities replaced by equalities or by lower bounds, and with independent sets replaced by cliques.*

*Proof.* By Theorem 6.10 compute all  $a(j_0, \dots, j_s)$ . The optimum is the largest  $j_0$  for which  $\sum a(j_0, j_1, \dots, j_s) > 0$ , the sum taken over the tuples with  $j_i \leq B_i$  for  $1 \leq i \leq s$ , and that sum is the number of optimal sets; it has at most  $(1 + n\Omega)^s$  terms. Other relations between  $j_i$  and  $B_i$  change only the range of the sum. Cliques of  $G$  are the independent sets of its complement.  $\square$

**6.3. Vertex Cover in time  $2^{o(k)}$ .** The parameterised Vertex Cover problem asks, for a graph  $G$  with  $n$  vertices and an integer  $k$ , whether  $G$  has a vertex cover with at most  $k$  vertices. We use the kernel of Chen, Kanj and Jia [CKJ01], which rests on the theorem of Nemhauser and Trotter [NT75] on the half-integral relaxation of the problem.

**Theorem 6.12** (Nemhauser and Trotter; Chen, Kanj and Jia). *There is a polynomial-time algorithm that, given a graph  $G$  and an integer  $k \geq 0$ , either reports correctly that  $G$  has no vertex cover with at most  $k$  vertices, or returns an induced subgraph  $G'$  of  $G$  with at most  $2k$  vertices and an integer  $k' \leq k$  such that  $G$  has a vertex cover with at most  $k$  vertices if and only if  $G'$  has one with at most  $k'$  vertices.<sup>13</sup>*

**Theorem 6.13** (Vertex Cover). *Parameterised Vertex Cover is decided deterministically in time*

$$2^{O(k(\log \log k)^2 / (\log k)^{1/4})} + \text{poly}(n),$$

*and the number of vertex covers of  $G'$  of each size is obtained within the same time.*

*Proof.* Run the algorithm of Theorem 6.12. If it reports that no cover of size  $k$  exists, answer no. Otherwise it returns  $G'$  on  $n' \leq 2k$  vertices and  $k'$ . Apply Theorem 6.7 to  $G'$  with unit weights and read  $\tau(G') = n' - \alpha(G')$  by Corollary 6.8; answer yes if and only if  $\tau(G') \leq k'$ . The time of the second step is  $2^{O(n'(\log \log n')^2 / (\log n')^{1/4})}$  with  $n' \leq 2k$ , which is  $2^{O(k(\log \log k)^2 / (\log k)^{1/4})}$ , and the first step is polynomial in  $n$ .  $\square$

*Remark 6.14.* Under the exponential time hypothesis, Vertex Cover has no algorithm with running time  $2^{o(k)} \text{poly}(n)$  [CFKL15, Theorem 14.6]; the hypothesis being false, the bound of Theorem 6.13 is the first of this shape. The same kernel argument applies to every parameterised problem with a kernel of  $O(k)$  vertices whose solution is read from an independence polynomial, but we state no further instance.

**6.4. Finite domains, homomorphisms and graphical models.** Variables with more than two values are encoded by blocks of Boolean variables, one block per variable, and the blocks are kept on the unit vectors by one factor each. The construction is used for Potts systems below; we record it in general, since it turns every sparse constraint system over a fixed finite domain into one over  $\{0, 1\}$ .

**Theorem 6.15** (Finite domains). *Fix integers  $D \geq 2$ ,  $k \geq 1$  and  $c \geq 1$ . Let  $n \geq 20$  variables take values in  $[D]$ , and let  $\mathcal{F}$  be a set of at most  $cn$  factors, each a function  $\phi_f : [D]^{I_f} \rightarrow \mathcal{A}$  with*

<sup>13</sup>In [CKJ01] the running time is  $O(kn + k^3)$ ; the half-integral optimum is found by a bipartite matching in the double cover of  $G$ , and the theorem of [NT75] says that the vertices valued 1 lie in some minimum vertex cover while those valued 0 lie outside it, so that only the vertices valued 1/2, at most  $2k$  of them when a cover of size  $k$  exists, remain.

$|I_f| \leq k$ , given by its table of  $D^{|I_f|}$  values, where  $\mathcal{A}$  is  $\mathbb{Q}$  or one of the algebras of Lemma 5.5 of dimension  $N$  and the table entries are polynomials in the generators with integer coefficients of at most  $L$  bits. Then

$$Z = \sum_{\sigma \in [D]^n} \prod_{f \in \mathcal{F}} \phi_f(\sigma|_{I_f})$$

is computed deterministically in time  $\text{poly}(n, N, L) 2^{O_{D,k,c}(\rho(n))}$  and polynomial space.

*Proof.* Introduce the  $Dn$  Boolean variables  $x_{v,a} = [\sigma_v = a]$ . For each  $v$  let  $\chi_v$  be the indicator on  $\{0, 1\}^D$  that exactly one of  $x_{v,1}, \dots, x_{v,D}$  equals 1, and for each  $f$  let  $\tilde{\phi}_f$  be the function on  $\{0, 1\}^{D|I_f|}$  that equals  $\phi_f(\sigma|_{I_f})$  when every block of its input is a unit vector, encoding  $\sigma|_{I_f}$ , and 0 otherwise. Take as factors the multilinear extensions of the  $n$  functions  $\chi_v$ , of scope  $D$ , and of the at most  $cn$  functions  $\tilde{\phi}_f$ , of scope at most  $kD$ . The product of the  $\chi_v$  vanishes unless  $x$  encodes some  $\sigma \in [D]^n$ , and then the product of all factors is  $\prod_f \phi_f(\sigma|_{I_f})$ ; so the sum (2.4) of this system is  $Z$ . The system has  $n' = Dn$  variables, at most  $(c+1)n \leq (c+1)n'$  factors of scope at most  $kD$ , and coefficients of at most  $L + kD$  bits, since a coefficient of a multilinear extension is a signed sum of at most  $2^{kD}$  table entries. Theorem 5.3 with the constants  $kD$  and  $c+1$  and Lemma 5.5 give time  $\text{poly}(n, N, L) 2^{O_{D,k,c}(\rho(Dn))}$ , and  $\rho(Dn) \leq 4D\rho(n)$  for  $n \geq D$ .  $\square$

**Corollary 6.16** (Homomorphisms and partition functions). *Fix  $\Delta \geq 1$  and  $q \geq 1$ . Let  $A$  be a symmetric  $q \times q$  matrix and  $\beta$  a vector of length  $q$ , both with rational entries of at most  $L$  bits. For a graph  $G$  with  $n$  vertices and maximum degree  $\Delta$ , the partition function*

$$Z_{A,\beta}(G) = \sum_{\sigma: V(G) \rightarrow [q]} \prod_{uv \in E(G)} A_{\sigma_u \sigma_v} \prod_{v \in V(G)} \beta_{\sigma_v}$$

is computed deterministically in time  $\text{poly}(n, L) 2^{O_{\Delta,q}(\rho(n))}$ . In particular the number of homomorphisms from  $G$  to a fixed graph  $H$  on  $q$  vertices, with loops allowed, is computed within that time, and so is the number of proper  $q$ -colourings of  $G$ .

*Proof.* Apply Theorem 6.15 with  $D = q$ ,  $\mathcal{A} = \mathbb{Q}$ , a factor of scope  $\{u, v\}$  with table  $A_{ab}$  for each edge and a factor of scope  $\{v\}$  with table  $\beta_a$  for each vertex: at most  $\Delta n/2 + n$  factors of scope at most 2. Homomorphisms to  $H$  are counted by the adjacency matrix of  $H$  and  $\beta = \mathbf{1}$ ; proper colourings by  $H = K_q$ .  $\square$

*Remark 6.17.* Dyer and Greenhill [DG00] and Bulatov and Grohe [BG05] showed that  $Z_A(G)$  is  $\#P$ -hard for every symmetric nonnegative  $A$  outside an explicit family of matrices of rank at most two in each block; on bounded-degree graphs every such partition function is now computed exactly in time  $2^{o(n)}$ . The same holds for the partition function of any discrete graphical model whose factors involve a bounded number of variables of bounded domain and whose factor count is linear: Theorem 6.15 computes it, and the marginal probability that a variable takes a value is the quotient of two such sums, the numerator with one indicator factor added.

**6.5. Spin systems on bounded-degree graphs.** Let  $G$  be a graph with  $n$  vertices and maximum degree  $\Delta$ . An Ising system on  $G$  is given by integer couplings  $J_{uv}$  on the edges and integer fields  $h_v$  on the vertices; a configuration is  $\sigma \in \{-1, 1\}^V$  and its energy is

$$E(\sigma) = - \sum_{uv \in E(G)} J_{uv} \sigma_u \sigma_v - \sum_v h_v \sigma_v. \quad (6.6)$$

The *density of states* is the function  $g(E) = \#\{\sigma : E(\sigma) = E\}$ , and the partition function at inverse temperature  $\beta$  is  $\mathcal{Z}(\beta) = \sum_{\sigma} e^{-\beta E(\sigma)} = \sum_E g(E) e^{-\beta E}$ . Put  $E_{\max} = \sum_{uv} |J_{uv}| + \sum_v |h_v|$ , so that  $|E(\sigma)| \leq E_{\max}$ .

**Theorem 6.18** (Ising and Potts density of states). *Fix  $\Delta \geq 1$  and  $q \geq 2$ .*

- (i) The density of states of an Ising system on a graph of maximum degree  $\Delta$  with  $n$  vertices is computed deterministically in time  $\text{poly}(n, E_{\max})2^{O_{\Delta}(\rho(n))}$ .
- (ii) For the  $q$ -state Potts system with energy  $E(\sigma) = -\sum_{uv} J_{uv}[\sigma_u = \sigma_v] - \sum_v h_v(\sigma_v)$ ,  $\sigma \in [q]^V$ , with integer couplings and integer fields  $h_v : [q] \rightarrow \mathbb{Z}$ , the density of states is computed in time  $\text{poly}(n, E_{\max})2^{O_{\Delta,q}(\rho(n))}$ , where now  $E_{\max} = \sum |J_{uv}| + \sum_v \max_s |h_v(s)|$ .

In both cases the ground-state energy and its degeneracy, and the partition function at every inverse temperature as the finite sum  $\sum_E g(E)e^{-\beta E}$ , follow at once.

*Proof.* (i) Encode  $\sigma_v = (-1)^{x_v}$  with  $x \in \{0, 1\}^n$ . The shifted energy  $E(\sigma) + E_{\max}$  is an integer in  $[0, 2E_{\max}]$  and is the sum of the local terms  $|J_{uv}| - J_{uv}\sigma_u\sigma_v \in \{0, 2|J_{uv}|\}$  over the edges and  $|h_v| - h_v\sigma_v \in \{0, 2|h_v|\}$  over the vertices, each nonnegative. Take  $\mathcal{A} = \mathcal{A}_N$  with  $N = 2E_{\max} + 1$  and the factor system consisting of the multilinear extension of  $W^{|J_{uv}| - J_{uv}\sigma_u\sigma_v}$  on the scope  $\{u, v\}$  for each edge and of  $W^{|h_v| - h_v\sigma_v}$  on  $\{v\}$  for each vertex. There are at most  $\Delta n/2 + n$  factors of scope at most 2, with coefficients that are polynomials in  $W$  with integer coefficients of absolute value at most 4, and the sum is  $\sum_{\sigma} W^{E(\sigma) + E_{\max}} = \sum_E g(E)W^{E + E_{\max}}$ . Theorem 5.3 with  $k = 2$ ,  $c = \Delta/2 + 1$  and Lemma 5.5 with  $L = 2$  give the bound.

(ii) This is an instance of Theorem 6.15; we write out the encoding. Encode  $\sigma_v \in [q]$  by the  $q$  Boolean variables  $x_{v,s} = [\sigma_v = s]$ , so that there are  $qn$  variables, and let  $\phi_v(x_{v,1}, \dots, x_{v,q})$  be the indicator that exactly one of them is 1, multiplied by  $W^{\max_t |h_v(t)| - h_v(s)}$  for that  $s$ ; this is a function on  $\{0, 1\}^q$  with values in  $\{0\} \cup \{W^j : 0 \leq j \leq 2 \max_t |h_v(t)|\}$ . For an edge  $uv$  let  $\phi_{uv}$  be the function on  $\{0, 1\}^{2q}$  that equals  $W^{|J_{uv}| - J_{uv}}$  when the two blocks are equal unit vectors,  $W^{|J_{uv}|}$  when they are distinct unit vectors, and 1 otherwise. Take  $\mathcal{A} = \mathcal{A}_N$  with  $N = 2E_{\max} + 1$  and the multilinear extensions of these functions as factors:  $n$  of scope  $q$  and at most  $\Delta n/2$  of scope  $2q$ , with coefficients of absolute value at most  $2^{2q}$ . The product of the vertex factors vanishes unless every block is a unit vector, in which case  $x$  encodes a configuration  $\sigma$  and the product of all factors is  $W^{E(\sigma) + E_{\max}}$ . So the sum is  $\sum_E g(E)W^{E + E_{\max}}$ . Theorem 5.3 on  $n' = qn$  variables with  $k = 2q$ ,  $c = \Delta/2 + 1$  and Lemma 5.5 with  $L = 2q$  give time  $\text{poly}(n, E_{\max})2^{O_{\Delta,q}(\rho(qn))}$ , and  $\rho(qn) \leq qn(\log \log n + 1)^2 / \sqrt{\log n} \leq 4q\rho(n)$  for  $n \geq q$ .  $\square$

*Remark 6.19.* The ground state of an Ising spin glass with couplings in  $\{-1, 0, 1\}$  on a planar graph with a field, or on a cubic lattice, is NP-hard to find [Bar82], and its density of states is #P-hard to compute, while the partition function of the ferromagnetic model admits a polynomial-time randomised approximation [JS93]; for bounded-degree graphs Theorem 6.18 computes both in time  $2^{o(n)}$  whenever the couplings are polynomially bounded. The  $q$ -colourings of a graph are the ground states of the antiferromagnetic Potts system with all  $J_{uv} = -1$  and no field, at energy  $E = 0$ ; their number for every graph is already in OpenAI's Corollary 1.3, and (ii) adds the whole density of states on bounded-degree graphs.

## 6.6. Bounded-degree graphs and sparse set systems.

**Theorem 6.20.** Fix  $\Delta \geq 1$ ,  $k \geq 1$  and  $c \geq 1$ . Each of the following is computed deterministically in time  $\text{poly}(n, \Omega)2^{O_{\Delta,k,c}(\rho(n))}$ .

- (i) For a graph with  $n$  vertices and maximum degree  $\Delta$ , with integer vertex weights in  $[0, \Omega]$ : the number of dominating sets of each weight, hence the minimum weight of a dominating set and the number of minimum dominating sets.
- (ii) For a graph with  $n$  vertices and maximum degree  $\Delta$ , with integer edge weights in  $[0, \Omega]$ : the number of matchings of each weight, hence the matching polynomial, the maximum weight of a matching, the number of maximum matchings and the number of perfect matchings.
- (iii) For a set system of  $m \leq cn$  sets of size at most  $k$  on  $n$  elements with integer element weights in  $[0, \Omega]$ : the number of hitting sets of each weight, hence the minimum hitting

set; and the number of 2-colourings of the elements under which exactly  $j$  sets are not monochromatic, for every  $j$ , hence the maximum number of sets split by a 2-colouring.

*Proof.* (i) Variables  $x_v = [v \in S]$ . For each vertex  $v$  take the multilinear extension of the function on the scope  $N[v]$  that equals  $W^{\omega_v x_v}$  if some  $x_u$ ,  $u \in N[v]$ , is 1 and 0 otherwise; it vanishes unless  $S$  dominates  $v$ . There are  $n$  factors of scope at most  $\Delta + 1$ , the sum is  $\sum_{S \text{ dominating}} W^{\omega(S)}$ , and Theorem 5.3 with  $k = \Delta + 1$ ,  $c = 1$ ,  $N = 1 + n\Omega$ ,  $L = \Delta + 1$  applies.

(ii) Discard isolated vertices, so that  $n \leq 2m$  where  $m \leq \Delta n/2$  is the number of edges. Variables  $y_e = [e \in M]$ , one per edge. For each vertex  $v$  take the multilinear extension of the indicator, on the scope of the edges at  $v$ , that at most one of them is chosen; for each edge the factor  $1 + (W^{\omega_e} - 1)Y_e$ . There are  $n + m \leq 3m$  factors of scope at most  $\Delta$  on  $m$  variables, the sum is  $\sum_{M \text{ matching}} W^{\omega(M)}$ , and Theorem 5.3 on  $m$  variables with  $k = \Delta$ ,  $c = 3$  applies;  $\rho(m) \leq \rho(\Delta n/2) = O_\Delta(\rho(n))$ . The coefficient of  $W^{n/2}$  for unit weights counts the perfect matchings.

(iii) Variables  $x_v = [v \in S]$ . For the hitting sets, take for each set  $A$  the multilinear extension of the indicator that some  $x_v$ ,  $v \in A$ , is 1, and for each element the factor  $1 + (W^{\omega_v} - 1)Y_v$ : at most  $(c + 1)n$  factors of scope at most  $k$ . For the splittings, take for each set  $A$  the multilinear extension of  $W^{[x|_A \text{ is not constant}]}$ ,  $m \leq cn$  factors of scope at most  $k$ , with  $N = m + 1$ . In both cases Theorem 5.3 and Lemma 5.5 with  $L = k$  give the bound.  $\square$

The permanent of a sparse matrix is a weighted count of perfect matchings in which the weights multiply rather than add; it needs no truncated algebra at all.

**Theorem 6.21** (Sparse permanents). *Fix  $\Delta \geq 1$ . The permanent of an  $n \times n$  integer matrix with at most  $\Delta$  nonzero entries in every row and every column, with entries of absolute value at most  $2^L$ , is computed deterministically in time  $\text{poly}(n, L)2^{O_\Delta(\rho(n))}$ .*

*Proof.* Let  $B$  be the bipartite graph on the rows and columns with an edge  $e = (i, j)$  for each nonzero entry  $a_{ij}$ ; it has  $n' \leq \Delta n$  edges, and  $\text{per}(A) = \sum_M \prod_{e \in M} a_e$  over the perfect matchings  $M$  of  $B$ . If some row or column has no entry the permanent is 0; otherwise  $n' \geq n$ . Take  $\mathcal{A} = \mathbb{Q}$ , one variable  $y_e$  per edge, for each of the  $2n$  vertices the multilinear extension of the indicator that exactly one incident edge is chosen, of scope at most  $\Delta$ , and for each edge the factor  $1 + (a_e - 1)Y_e$ . There are  $2n + n' \leq 3n'$  factors on  $n'$  variables, every coefficient is an integer of at most  $L + 1$  bits, and the sum (2.4) is  $\text{per}(A)$ . Theorem 5.3 with  $k = \Delta$ ,  $c = 3$  and Lemma 5.5 with  $N = 1$  give the bound, and  $\rho(n') = O_\Delta(\rho(n))$ .  $\square$

*Remark 6.22.* The permanent of a 0-1 matrix is #P-complete [Val79], and for sparse matrices the best bound known was  $2^{n - \Omega(n/\Delta)}$  [CP15]. OpenAI's Corollary 1.3 counts the perfect matchings of every graph in time  $2^{o(v+e)}$  by their general theorem; Theorem 6.21 is the sparse weighted version, with the weights absorbed into the factors over  $\mathbb{Q}$ .

**6.7. One global side constraint.** The factors of a factor system are local, but the algebra may carry one global statistic for free. Corollary 6.3 used this for the Hamming weight. The group algebra of a cyclic group records a residue.

**Proposition 6.23** (Counting under a congruence). *Fix  $k, c \geq 1$ . Given a  $k$ -CNF formula with  $n$  variables and  $m \leq cn$  clauses, integer weights  $\omega_v \in [0, \Omega]$  on the variables and a modulus  $M \geq 1$ , the numbers*

$$r_j = \#\left\{x : x \text{ satisfies the formula and } \sum_v \omega_v x_v \equiv j \pmod{M}\right\}, \quad 0 \leq j < M,$$

*are computed deterministically in time  $\text{poly}(n, M, \log \Omega)2^{O_{k,c}(\rho(n))}$ .*

*Proof.* Take  $\mathcal{A} = \mathbb{Q}[W]/(W^M - 1)$ , the clause indicators as factors, and for each variable the factor  $1 + (W^{\omega_v \bmod M} - 1)Y_v$ . The sum is  $\sum_x [x \text{ satisfies}] W^{\sum_v \omega_v x_v \bmod M} = \sum_j r_j W^j$ , since  $W^M = 1$ . Theorem 5.3 and Lemma 5.5 with  $N = M$ ,  $L = 1$  give the bound; the reduction of the exponents modulo  $M$  costs  $O(\log \Omega)$  bits per variable.  $\square$

Taken together, Corollary 6.3 and Proposition 6.23 count the solutions of a sparse  $k$ -CNF formula subject to one linear equation  $\sum_v \omega_v x_v = s$  or one linear congruence over the integers, constraints that no bounded number of local factors can express.

## 7. FURTHER WORKED APPLICATIONS

The applications of Section 6 used one Boolean variable per vertex or per clause variable, one algebra generator, and constraints of the simplest kind. This section works out eight further families in which one of these three choices is made differently: several generators record several linear statistics at once; the variables range over a finite domain and the objective is a cut into several classes; the variables sit on the edges of a bounded-degree graph and the constraints at the vertices; the algebra is  $\mathbb{Q}$  but the factor is the Fortuin–Kasteleyn weight of an edge, which reaches the Tutte polynomial on a hyperbola; the constraints are the parity checks of a linear code; the constraints are the incidences of a hypergraph; the constraints are the gates of a circuit; and the constraints are the rows of an integer program, linear inequalities with coefficients of arbitrary size. In each case the proof consists of naming the variables, the factors and the algebra, checking the three hypotheses of Theorem 5.3 or Theorem 6.15, and reading the answer off the resulting element of the algebra. We use throughout that  $\rho$  is increasing on  $[20, \infty)$ ,<sup>14</sup> so that  $\rho(m) \leq \rho(Cn) \leq 4C\rho(n)$  whenever  $20 \leq m \leq Cn$  and  $n \geq C$ , as in the proof of Theorem 6.15.

**7.1. Several linear side constraints.** Corollary 6.3 and Proposition 6.23 recorded one linear statistic of the assignment in the exponent of the single generator  $W$ . With the algebra  $\mathcal{A}_{N_1, \dots, N_s}$  of (2.3) one records  $s$  of them at once, and a fixed number of linear equations or congruences may be imposed on the solutions of a sparse constraint system.

**Theorem 7.1** (Linear statistics). *Fix integers  $k, c, s \geq 1$ . Let  $(\Phi_f)_{f \in \mathcal{F}}$  be a factor system on  $n \geq 20$  Boolean variables with at most  $cn$  factors of scope at most  $k$ , with values in  $\mathcal{A}_{N_0}$  for some  $N_0 \geq 1$  and coefficients that are polynomials in  $W_0$  with integer coefficients of at most  $L$  bits, and let  $\omega^{(1)}, \dots, \omega^{(s)} \in \{0, 1, \dots, \Omega\}^n$  be weight vectors.*

(i) *The elements*

$$r(j_1, \dots, j_s) = \sum_{x \in \{0,1\}^n: \sum_v \omega_v^{(i)} x_v = j_i \ (1 \leq i \leq s)} \prod_{f \in \mathcal{F}} \Phi_f(x) \in \mathcal{A}_{N_0}$$

*are computed, for all  $(j_1, \dots, j_s)$  at once, in time  $\text{poly}(n, L, N_0, (1 + n\Omega)^s) 2^{O_{k,c}(\rho(n))}$ .*

(ii) *For moduli  $M_1, \dots, M_s \geq 1$  the same holds for the sums over the  $x$  with  $\sum_v \omega_v^{(i)} x_v \equiv j_i \pmod{M_i}$  for all  $i$ , in time  $\text{poly}(n, L, N_0, M_1 \cdots M_s, \log \Omega) 2^{O_{k,c}(\rho(n))}$ ; equations and congruences may be mixed.*

*Proof.* (i) Put  $N_i = 1 + \sum_v \omega_v^{(i)} \leq 1 + n\Omega$  for  $1 \leq i \leq s$  and  $\mathcal{A} = \mathcal{A}_{N_0, N_1, \dots, N_s}$ , of dimension  $N = N_0 N_1 \cdots N_s$ . Keep the given factors, now read in  $\mathcal{A}$ , and add for each variable  $v$  the factor

$$\Psi_v = 1 + \left( \prod_{i=1}^s W_i^{\omega_v^{(i)}} - 1 \right) Y_v,$$

<sup>14</sup>With  $u = \log x$  and  $u' = 1/(x \ln 2)$  the logarithmic derivative of  $\rho(x) = x(\log u)^2 / \sqrt{u}$  is  $1/x + 2u'/(u \ln 2 \log u) - u'/(2u)$ , and the negative term is at most  $1/(2xu \ln 2) < 1/x$  for  $u \geq 1$ .

the multilinear extension of  $\prod_i W_i^{\omega_v^{(i)} x_v}$  on the scope  $\{v\}$ , of degree one in  $Y_v$  and with coefficients that are polynomials in the generators with integer coefficients of absolute value at most 1.<sup>15</sup> The new system has at most  $(c+1)n$  factors of scope at most  $k$ , and its sum is

$$\sum_x \prod_f \Phi_f(x) \prod_{i=1}^s W_i^{\sum_v \omega_v^{(i)} x_v} = \sum_{(j_1, \dots, j_s)} r(j_1, \dots, j_s) W_1^{j_1} \cdots W_s^{j_s},$$

with nothing lost to the truncation because every exponent of  $W_i$  is less than  $N_i$ . Theorem 5.3 computes the sum with  $\text{poly}(n)2^{O_{k,c}(\rho(n))}$  operations in  $\mathcal{A}$ , and Lemma 5.5, with  $|\mathcal{F}| \leq (c+1)n$  and the bound  $\max(L, 1)$  on the coefficient bits, turns this into  $\text{poly}(n, L, N)2^{O_{k,c}(\rho(n))}$  bit operations. (ii) Replace  $W_i^{N_i} = 0$  by  $W_i^{M_i} = 1$ , that is take  $\mathcal{A} = \mathcal{A}_{N_0} \otimes \mathbb{Q}[W_1, \dots, W_s]/(W_1^{M_1} - 1, \dots, W_s^{M_s} - 1)$ , with the exponents  $\omega_v^{(i)}$  reduced modulo  $M_i$  at a cost of  $O(\log \Omega)$  bits each; the sum is then  $\sum_j r(j)W^j$  with  $j$  running over the residues. The proof of Lemma 5.5 used of the algebra only that the product of two basis monomials is a basis monomial or zero, which holds here, so the bit bound follows with  $N = N_0 M_1 \cdots M_s$ . A mixture of equations and congruences uses  $W_i^{N_i} = 0$  for the former and  $W_i^{M_i} = 1$  for the latter.  $\square$

**Corollary 7.2.** *Fix  $k, c, s \geq 1$  and  $\Delta \geq 1$ . Each of the following is computed deterministically in time  $\text{poly}(n, \Omega^s)2^{O(\rho(n))}$ , the implied constant depending on  $k, c, s$  or  $\Delta, s$ .*

- (a) *For a  $k$ -CNF formula with  $n$  variables and at most  $cn$  clauses and weight vectors  $\omega^{(1)}, \dots, \omega^{(s)} \in [0, \Omega]^n$ : the number of satisfying assignments  $x$  with  $\sum_v \omega_v^{(i)} x_v = j_i$  for all  $i$ , for every tuple  $(j_1, \dots, j_s)$ ; in particular whether a satisfying assignment of prescribed weights exists.*
- (b) *For a constraint system as in Theorem 6.1: the number of assignments that satisfy constraints of total weight exactly  $j$  and have Hamming weight exactly  $i$ , for all  $(i, j)$ .*
- (c) *For a graph of maximum degree  $\Delta$  with vertex weights  $\omega^{(1)}, \omega^{(2)} \in [0, \Omega]^V$  and a budget  $B$ : the maximum of  $\omega^{(1)}(S)$  over the independent sets  $S$  with  $\omega^{(2)}(S) \leq B$ , and the number of independent sets attaining it.*

*Proof.* (a) is Theorem 7.1(i) with the clause indicators as factors and  $N_0 = 1$ . (b) is the theorem with the factors of the proof of Theorem 6.1, which take values in  $\mathcal{A}_{N_0}$  with  $N_0 = 1 + \sum_f \omega_f$  and have coefficients of at most  $k$  bits, and the single weight vector  $\omega^{(1)} = (1, \dots, 1)$ . (c) is the theorem with the edge factors  $1 - Y_u Y_v$  of the proof of Theorem 6.7, at most  $\Delta n/2$  factors of scope 2,  $N_0 = 1$ , and the two weight vectors: the maximum is the largest  $j_1$  with  $\sum_{j_2 \leq B} r(j_1, j_2) > 0$  and the count is that sum.  $\square$

**7.2. Finite domains: cuts into several classes.** Theorem 6.1 dealt with Boolean variables. Over a finite domain the same profile is obtained from Theorem 6.15, and the extra generators of Theorem 7.1 record the sizes of the classes of a partition.

**Theorem 7.3** (Profiles over a finite domain). *Fix  $D \geq 2$  and  $k, c \geq 1$ . Given  $n \geq 20$  variables with values in  $[D]$  and  $m \leq cn$  constraints  $C_f : [D]^{I_f} \rightarrow \{0, 1\}$  with  $|I_f| \leq k$ , given by their tables, with integer weights  $\omega_f \in [0, \Omega]$ , the numbers*

$$p_j = \#\left\{\sigma \in [D]^n : \sum_f \omega_f C_f(\sigma|_{I_f}) = j\right\}, \quad 0 \leq j \leq \sum_f \omega_f,$$

*are computed deterministically in time  $\text{poly}(n, \Omega)2^{O_{D,k,c}(\rho(n))}$ .*

<sup>15</sup>The product  $\prod_i (1 + (W_i^{\omega_v^{(i)}} - 1)Y_v)$  takes the same values at  $Y_v \in \{0, 1\}$ , but it has degree  $s$  in  $Y_v$  and differs from  $\Psi_v$  at the non-Boolean inputs  $\langle \vec{P}_{f,v}(t), x \rangle$  of the routine; at  $W_1 = 2, W_2 = 3, Y_v = \frac{1}{2}$  and unit weights it equals 3, while  $\Psi_v = \frac{7}{2}$ .

*Proof.* Apply Theorem 6.15 with  $\mathcal{A} = \mathcal{A}_N$ ,  $N = 1 + \sum_f \omega_f \leq 1 + cn\Omega$ , and the tables  $\phi_f(\sigma) = W^{\omega_f C_f(\sigma)}$ , whose entries are monomials with coefficient 1, so  $L = 1$ . The sum is  $\sum_{\sigma} W^{\sum_f \omega_f C_f(\sigma|_{I_f})} = \sum_j p_j W^j$ , with no term truncated.  $\square$

**Corollary 7.4** (Max- $k$ -Lin, Max- $q$ -Cut and balanced partitions). *Fix  $p, q, k, c \geq 2$ . The following are computed deterministically in time  $\text{poly}(n, \Omega)2^{O(\rho(n))}$ , the implied constant depending on the fixed parameters.*

- (i) *For a system of  $m \leq cn$  linear equations over  $\mathbb{Z}/p\mathbb{Z}$  in  $n$  unknowns, each equation involving at most  $k$  unknowns, with weights  $\omega_f \in [0, \Omega]$ : the number of assignments satisfying equations of total weight exactly  $j$ , for every  $j$ ; hence the maximum satisfiable weight and the number of optimal assignments.*
- (ii) *For a graph  $G$  with  $n$  vertices and  $m \leq cn$  edges with weights  $\omega_e \in [0, \Omega]$ : the number of maps  $\sigma : V \rightarrow [q]$  for which the edges  $uv$  with  $\sigma_u \neq \sigma_v$  have total weight exactly  $j$ , for every  $j$ ; hence the maximum  $q$ -cut, and the minimum, which is the least weight of edges left monochromatic by a  $q$ -colouring, with the number of optimal maps.*
- (iii) *In the situation of (ii), the number of maps  $\sigma$  with  $|\sigma^{-1}(a)| = i_a$  for  $a < q$  and cut weight  $j$ , for all  $(i_1, \dots, i_{q-1}, j)$ ; hence the maximum and the minimum  $q$ -cut over the partitions with prescribed class sizes, in particular over the balanced ones with every class of size  $\lfloor n/q \rfloor$  or  $\lceil n/q \rceil$ , and the number of optimal partitions.*

*Proof.* (i) is Theorem 7.3 with  $D = p$  and the equations as constraints. (ii) is the theorem with  $D = q$  and the constraint  $C_e(a, b) = [a \neq b]$  of weight  $\omega_e$  on the scope  $\{u, v\}$  of each edge. For (iii) apply Theorem 6.15 with  $D = q$  and  $\mathcal{A} = \mathcal{A}_{N_0, n+1, \dots, n+1}$ , with  $q-1$  copies of  $n+1$  and  $N_0 = 1 + \sum_e \omega_e$ , of dimension  $N = N_0(n+1)^{q-1}$ , which is polynomial in  $n$  and  $\Omega$  for fixed  $q$ . Take for each edge the table  $W^{\omega_e[a \neq b]}$  and for each vertex the table  $U_a$  for  $a < q$  and 1 for  $a = q$ ; all entries are monomials,  $L = 1$ , and there are at most  $(c+1)n$  factors of scope at most 2. The sum is  $\sum_{\sigma} W^{\text{cut}_{\omega}(\sigma)} \prod_{a < q} U_a^{|\sigma^{-1}(a)|}$  and nothing is truncated, since  $|\sigma^{-1}(a)| \leq n$ . The balanced optimum is read off from the coefficients with every  $i_a$ , and  $n - \sum_{a < q} i_a$ , in  $\{\lfloor n/q \rfloor, \lceil n/q \rceil\}$ .  $\square$

**7.3. Edge variables on bounded-degree graphs.** In Theorem 6.20(ii) the variables sat on the edges of a graph and the constraints at its vertices. On a graph of maximum degree  $\Delta$  every constraint of that shape has scope at most  $\Delta$ , so every counting problem whose feasibility is a condition at each vertex on its incident edges is within reach. We collect the classical ones.

**Theorem 7.5** (Vertex conditions on edge variables). *Fix  $\Delta \geq 1$  and  $q \geq 2$ . Let  $G$  be a graph with  $n$  vertices,  $m$  edges and maximum degree  $\Delta$ . Each of the following is computed deterministically in time  $\text{poly}(n, \Omega)2^{O_{\Delta, q}(\rho(n))}$ .*

- (i) *For sets  $F_v \subseteq \{0, 1, \dots, \Delta\}$ ,  $v \in V$ , and integer edge weights  $\omega_e \in [0, \Omega]$ : the number of spanning subgraphs  $H$  of  $G$  with  $\deg_H(v) \in F_v$  for every  $v$  and  $\omega(E(H)) = j$ , for every  $j$ . In particular the number of  $f$ -factors for any  $f : V \rightarrow \{0, \dots, \Delta\}$ , of  $k$ -regular spanning subgraphs for any  $k$ , of 2-factors, and of spanning subgraphs with all degrees even or all degrees odd, each by weight, and the maximum weight of each.*
- (ii) *For a function  $d : V \rightarrow \{0, \dots, \Delta\}$ : the number of orientations of  $G$  in which every vertex  $v$  has indegree  $d(v)$ ; for  $d = \deg/2$  this is the number of Eulerian orientations.*
- (iii) *The number of proper edge colourings of  $G$  with  $q$  colours, and more generally the number of maps  $E \rightarrow [q]$  that satisfy at every vertex  $v$  a prescribed condition, given by a table on  $[q]^{\deg v}$ , on the colours of the edges at  $v$ .*
- (iv) *For a fixed orientation of  $G$ : the number of nowhere-zero  $\mathbb{Z}/q\mathbb{Z}$ -flows, that is of maps  $y : E \rightarrow \{1, \dots, q-1\}$  with  $\sum_{e \text{ into } v} y_e \equiv \sum_{e \text{ out of } v} y_e \pmod{q}$  at every vertex; and the*

number of  $\mathbb{Z}/q\mathbb{Z}$ -flows with values in a prescribed subset of  $\mathbb{Z}/q\mathbb{Z}$  and prescribed residues  $b_v$  at the vertices.

*Proof.* An isolated vertex contributes the factor  $[0 \in F_v]$  in (i), the factor  $[d(v) = 0]$  in (ii), the factor 1 in (iii) and the factor  $[b_v = 0]$  in (iv); so we may delete the isolated vertices, after which  $n \leq 2m$  and  $m \leq \Delta n/2$ . If  $m < 20$  the count is obtained by listing the at most  $2^{19}$  or  $q^{19}$  assignments. Otherwise the variables are indexed by the edges, and  $\rho(m) \leq \rho(\Delta n/2) = O_\Delta(\rho(n))$ .

(i) Boolean variables  $y_e = [e \in E(H)]$ . For each vertex  $v$  take the multilinear extension of the indicator, on the scope  $E(v)$  of the edges at  $v$ , that  $\sum_{e \in E(v)} y_e \in F_v$ ; its coefficients are integers of absolute value at most  $2^\Delta$ . For each edge take  $1 + (W^{\omega_e} - 1)Y_e$ . There are  $n + m \leq 3m$  factors of scope at most  $\Delta$  on  $m$  variables, and the sum is  $\sum_H [\deg_H(v) \in F_v \ \forall v] W^{\omega(E(H))}$  in  $\mathcal{A}_N$ ,  $N = 1 + \sum_e \omega_e$ , with nothing truncated. Theorem 5.3 with  $k = \Delta$ ,  $c = 3$  and Lemma 5.5 with  $L = \Delta$  give the bound. The special cases are  $F_v = \{f(v)\}$ ,  $F_v = \{k\}$ ,  $F_v = \{2\}$ , and  $F_v$  the even or the odd numbers up to  $\Delta$ .

(ii) Fix a labelling of the vertices and let  $y_e = 1$  when  $e$  is oriented towards its endpoint of larger label. The indegree of  $v$  is  $\sum_{e \in E(v)} [e \text{ is oriented into } v]$ , a function of  $y|_{E(v)}$ . Take  $\mathcal{A} = \mathbb{Q}$  and for each vertex the multilinear extension of the indicator that this function equals  $d(v)$ ; the sum of the  $n \leq 2m$  factors of scope at most  $\Delta$  is the number of orientations with the prescribed indegrees.

(iii) Apply Theorem 6.15 on the  $m$  variables  $y_e \in [q]$  with the  $n \leq 2m$  factors given by the tables at the vertices, of scope at most  $\Delta$  and entries 0 or 1; for proper colourings the table at  $v$  is the indicator that the colours of the edges at  $v$  are distinct.

(iv) Apply Theorem 6.15 with  $D = q - 1$  for the nowhere-zero flows, the variable  $y_e \in [q - 1]$  being the flow value, or  $D = |S|$  for values in a subset  $S$ , and at each vertex the table of the condition  $\sum_{\text{in}} y_e - \sum_{\text{out}} y_e \equiv b_v \pmod{q}$  on  $[D]^{\deg v}$ .  $\square$

**Corollary 7.6** (Hafnians). *Fix  $\Delta \geq 1$ . The hafnian  $\text{haf}(A) = \sum_M \prod_{\{i,j\} \in M} a_{ij}$  of a symmetric  $2n \times 2n$  integer matrix  $A$  with at most  $\Delta$  nonzero entries in every row, the sum running over the perfect matchings  $M$  of the complete graph on  $[2n]$ , with entries of absolute value at most  $2^L$ , is computed deterministically in time  $\text{poly}(n, L) 2^{O_\Delta(\rho(n))}$ .*

*Proof.* Only matchings inside the support graph  $G$  of  $A$ , of maximum degree  $\Delta$ , contribute. Take  $\mathcal{A} = \mathbb{Q}$ , one Boolean variable per edge of  $G$ , for each vertex the multilinear extension of the indicator that exactly one incident edge is chosen, and for each edge the factor  $1 + (a_e - 1)Y_e$ : the sum is  $\text{haf}(A)$ , and the proof of Theorem 6.21 applies word for word.  $\square$

*Remark 7.7.* Counting Eulerian orientations is #P-complete [MW96], deciding whether a 3-regular graph has a proper 3-edge-colouring is NP-complete [Hol81], and the number of nowhere-zero  $\mathbb{Z}/q\mathbb{Z}$ -flows is the flow polynomial  $F(G; q)$  of Tutte [Tut54], whose evaluation at every fixed integer  $q \geq 3$  is #P-hard [JW90]. On graphs of bounded degree each of them is now computed in time  $2^{o(n)}$ . The flow polynomial is reached by a second route in the next subsection.

**7.4. The Tutte polynomial on the hyperbolas  $(x - 1)(y - 1) = q$ .** Let  $G = (V, E)$  be a graph, loops and multiple edges allowed, and for  $A \subseteq E$  let  $c(A)$  be the number of connected components of the spanning subgraph  $(V, A)$ . The Tutte polynomial is

$$T(G; x, y) = \sum_{A \subseteq E} (x - 1)^{c(A) - c(E)} (y - 1)^{c(A) + |A| - |V|}, \quad (7.1)$$

and the random-cluster, or multivariate Potts, partition function of Fortuin and Kasteleyn [FK72] is

$$Z(G; q, v) = \sum_{A \subseteq E} q^{c(A)} v^{|A|}. \quad (7.2)$$

Both are polynomials in their arguments, and a comparison of exponents gives, for  $x \neq 1$ ,  $y \neq 1$ ,

$$T(G; x, y) = (x - 1)^{-c(E)} (y - 1)^{-|V|} Z(G; (x - 1)(y - 1), y - 1), \quad (7.3)$$

since  $(x - 1)^{c(A)} (y - 1)^{c(A)} (y - 1)^{|A|}$  divided by  $(x - 1)^{c(E)} (y - 1)^{|V|}$  is the summand of (7.1). Thus the Tutte polynomial on the hyperbola  $H_q = \{(x, y) : (x - 1)(y - 1) = q\}$  is the function  $v \mapsto Z(G; q, v)$  up to the explicit prefactor, and for a positive integer  $q$  that function is a sum over colourings [FK72, Sok05].

**Lemma 7.8** (Fortuin–Kasteleyn). *For every integer  $q \geq 1$  and every  $v$ ,*

$$Z(G; q, v) = \sum_{\sigma: V \rightarrow [q]} \prod_{uv \in E} (1 + v[\sigma_u = \sigma_v]).$$

*Proof.* Expand the product: it is  $\sum_{A \subseteq E} v^{|A|} \prod_{uv \in A} [\sigma_u = \sigma_v]$ . For fixed  $A$  the product is 1 exactly when  $\sigma$  is constant on each component of  $(V, A)$ , and there are  $q^{c(A)}$  such  $\sigma$ . Summing over  $\sigma$  first gives (7.2).  $\square$

**Lemma 7.9** (Flows). *Fix an orientation of  $G$  and an integer  $q \geq 2$ . For  $A \subseteq E$  the number of  $\mathbb{Z}/q\mathbb{Z}$ -flows on  $(V, A)$ , that is of maps  $y : A \rightarrow \mathbb{Z}/q\mathbb{Z}$  with  $\sum_{e \in A \text{ into } v} y_e = \sum_{e \in A \text{ out of } v} y_e$  at every vertex, is  $q^{|A| - |V| + c(A)}$ . Consequently the number  $F(G; q)$  of nowhere-zero  $\mathbb{Z}/q\mathbb{Z}$ -flows on  $G$  satisfies*

$$F(G; q) = \sum_{A \subseteq E} (-1)^{|E \setminus A|} q^{|A| - |V| + c(A)} = (-1)^{|E|} q^{-|V|} Z(G; q, -q). \quad (7.4)$$

*Proof.* Let  $\partial : (\mathbb{Z}/q\mathbb{Z})^A \rightarrow (\mathbb{Z}/q\mathbb{Z})^V$  send  $y$  to the vector of net inflows  $(\partial y)_v = \sum_{\text{into } v} y_e - \sum_{\text{out of } v} y_e$ . The flows on  $(V, A)$  form  $\ker \partial$ . Each edge contributes  $+y_e$  at one endpoint and  $-y_e$  at the other, both in the same component of  $(V, A)$ , so the image of  $\partial$  lies in the group  $B$  of vectors whose sum over each component is zero, of order  $q^{|V| - c(A)}$ . Conversely every  $b \in B$  is attained: choose a spanning forest  $T$  of  $(V, A)$  and proceed by induction on  $|V|$ ; if  $|V| = 1$  then  $b = 0$  and  $y = 0$  serves; otherwise pick a leaf  $w$  of  $T$  with pendant edge  $e$  to  $w'$ , or an isolated vertex  $w$  of  $T$ , for which  $b_w = 0$ . In the first case set  $y_e = \pm b_w$  according to the orientation of  $e$ , so that  $(\partial y)_w = b_w$ , delete  $w$ , replace  $b_{w'}$  by  $b_{w'} + b_w$ , which keeps the component sums zero, and apply the induction to  $(V \setminus w, A \setminus e)$ ; in the second case delete  $w$ . Hence  $|\ker \partial| = q^{|A|} / q^{|V| - c(A)}$ .

For (7.4), the flows on  $(V, E)$  that vanish on a set  $S \subseteq E$  correspond to the flows on  $(V, E \setminus S)$ . By inclusion and exclusion over the set of edges at which a flow vanishes,  $F(G; q) = \sum_{S \subseteq E} (-1)^{|S|} q^{|E \setminus S| - |V| + c(E \setminus S)}$ , which is the first expression with  $A = E \setminus S$ ; and  $(-1)^{|E|} q^{-|V|} Z(G; q, -q) = \sum_A (-1)^{|E| - |A|} q^{|A| + c(A) - |V|}$  is the same sum.  $\square$

**Theorem 7.10** (Tutte polynomial on bounded-degree graphs). *Fix  $\Delta \geq 1$  and an integer  $q \geq 1$ . Let  $G$  be a graph with  $n$  vertices and maximum degree  $\Delta$ , and let  $v = a/b$  be a rational number with  $|a|, |b| \leq 2^L$ . Then  $Z(G; q, v)$  is computed deterministically in time  $\text{poly}(n, L) 2^{O_{\Delta, q}(\rho(n))}$ . Consequently  $T(G; x, y)$  is computed in that time at every rational point  $(x, y)$  with  $(x - 1)(y - 1) = q$  and  $y \neq 1$ , and in particular*

- (i) the chromatic polynomial  $P(G; q) = Z(G; q, -1)$ , the number of proper  $q$ -colourings;
- (ii) the flow polynomial  $F(G; q) = (-1)^{|E|} q^{-|V|} Z(G; q, -q)$ , the number of nowhere-zero  $\mathbb{Z}/q\mathbb{Z}$ -flows;

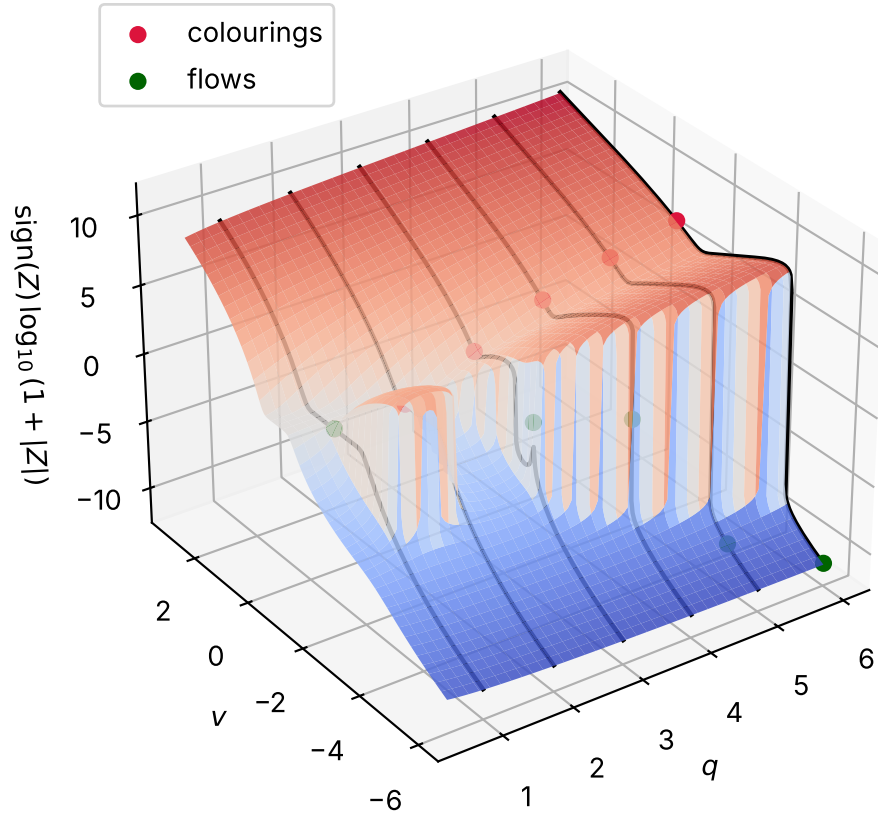


FIGURE 4. The random-cluster partition function  $Z(G; q, v)$  of the Petersen graph over the  $(q, v)$ -plane, as  $\text{sign}(Z) \log_{10}(1 + |Z|)$ , computed from the  $2^{15}$  edge subsets. Theorem 7.10 evaluates it along the lines  $q = 1, 2, 3, \dots$  drawn on the surface, which are the hyperbolas  $H_q$  of the Tutte plane; the marked points are the proper  $q$ -colourings ( $v = -1$ ) and the nowhere-zero flows ( $v = -q$ ).

- (iii) the partition function of the  $q$ -state Potts model with coupling  $J$  at inverse temperature  $\beta$ ,  $\sum_{\sigma} \exp(\beta J \sum_{uv} [\sigma_u = \sigma_v]) = Z(G; q, e^{\beta J} - 1)$ , whenever  $e^{\beta J}$  is rational, and more generally with any rational edge-dependent couplings  $v_e$  in place of  $v$ .

*Proof.* By Lemma 7.8,  $b^m Z(G; q, v) = \sum_{\sigma: V \rightarrow [q]} \prod_{uv \in E} (b + a[\sigma_u = \sigma_v])$ , where  $m \leq \Delta n/2$  is the number of edges. This is the sum of Theorem 6.15 with  $D = q$ ,  $\mathcal{A} = \mathbb{Q}$ , one factor of scope  $\{u, v\}$  per edge with the table  $(s, t) \mapsto b + a[s = t]$ , whose entries are integers of at most  $L + 1$  bits, and no vertex factors; a loop at  $u$  contributes the constant factor  $a + b$ . Theorem 6.15 with  $k = 2$  and  $c = \Delta/2 + 1$  gives  $b^m Z(G; q, v)$  in the stated time, and the division by  $b^m$  is polynomial. For  $T(G; x, y)$  use (7.3), with  $c(E)$  found by a graph search. For edge-dependent  $v_e = a_e/b_e$  the table of the edge  $e$  is  $b_e + a_e[s = t]$  and the sum is  $\prod_e b_e$  times  $\sum_{\sigma} \prod_e (1 + v_e[\sigma_u = \sigma_v])$ . Item (i) is Lemma 7.8 with  $v = -1$ , since  $1 - [\sigma_u = \sigma_v]$  is the indicator of a proper edge; item (ii) is Lemma 7.9; item (iii) is Lemma 7.8 with  $1 + v = e^{\beta J}$ , the factor  $e^{\beta J[\sigma_u = \sigma_v]}$  being  $1 + v[\sigma_u = \sigma_v]$ .  $\square$

*Remark 7.11.* Evaluating the Tutte polynomial is #P-hard at every point of the plane except on the hyperbola  $H_1$  and at eight special points [JWV90]; the points of  $H_q$  for positive integers  $q$

include the chromatic and flow polynomials at  $q$ , the Ising model ( $q = 2$ ) and the Potts models. On bounded-degree graphs Theorem 7.10 evaluates all of them in time  $2^{o(n)}$ . The theorem does not evaluate  $T$  off these hyperbolas: the number of acyclic orientations  $T(G; 2, 0)$ , which lies on  $H_{-1}$ , the number of forests  $T(G; 2, 1)$  on  $H_0$ , and the reliability polynomial on  $H_1$ , trivial, are not covered, and neither is the chromatic polynomial as a polynomial in  $q$ , which would need the values at  $n + 1$  integers  $q$  while the constant in the exponent grows with  $q$ .

**7.5. Linear codes with sparse parity checks.** A binary linear code of length  $n$  given by a parity-check matrix  $H \in \mathbb{F}_2^{m \times n}$  is  $\mathcal{C} = \{x \in \mathbb{F}_2^n : Hx = 0\}$ . Its weight enumerator is the sequence  $A_i = \#\{x \in \mathcal{C} : \text{wt}(x) = i\}$ ,  $0 \leq i \leq n$ , where  $\text{wt}$  is the Hamming weight, and its minimum distance is  $d(\mathcal{C}) = \min\{i \geq 1 : A_i > 0\}$ . For a received word  $y \in \mathbb{F}_2^n$  the distance distribution of  $y$  is  $A_i(y) = \#\{x \in \mathcal{C} : \text{wt}(x + y) = i\}$ , the weight enumerator of the coset  $y + \mathcal{C}$ ; a codeword at distance  $\min\{i : A_i(y) > 0\}$  from  $y$  is a maximum-likelihood decoding of  $y$  on the binary symmetric channel. The codes whose parity-check matrix has rows of bounded weight are the low-density parity-check codes of Gallager [Gal62].

**Theorem 7.12** (Sparse parity checks). *Fix  $k, c \geq 1$ . Let  $\mathcal{C} \subseteq \mathbb{F}_2^n$  be given by a parity-check matrix with  $m \leq cn$  rows, each of weight at most  $k$ .*

- (i) *The weight enumerator  $(A_i)_{0 \leq i \leq n}$ , and hence  $d(\mathcal{C})$  and the number of minimum-weight codewords, are computed deterministically in time  $\text{poly}(n)2^{O_{k,c}(\rho(n))}$ .*
- (ii) *For every  $y \in \mathbb{F}_2^n$  the distance distribution  $(A_i(y))_i$  is computed in the same time, and a codeword nearest to  $y$  is found in  $2n + 1$  times that time.*
- (iii) *For a code over a finite field  $\mathbb{F}_Q$ , or over any alphabet of fixed size  $Q$  with the checks given as tables on  $[Q]^k$ , given by  $m \leq cn$  checks each involving at most  $k$  coordinates, the Hamming weight enumerator and the distance distribution of every received word are computed in time  $\text{poly}(n)2^{O_{Q,k,c}(\rho(n))}$ ; so is the complete weight enumerator, the number of codewords with  $i_s$  coordinates equal to  $s$  for each symbol  $s$ , in time  $\text{poly}(n^Q)2^{O_{Q,k,c}(\rho(n))}$ .*

*Proof.* (i) Boolean variables  $x_1, \dots, x_n$  and  $\mathcal{A} = \mathcal{A}_{n+1}$ . For each row  $h$  of  $H$  take the multilinear extension of the indicator, on the scope  $\text{supp}(h)$  of size at most  $k$ , that  $\sum_{v \in \text{supp}(h)} x_v$  is even; its coefficients are integers of absolute value at most  $2^k$ . For each coordinate take  $1 + (W - 1)Y_v$ . There are at most  $(c + 1)n$  factors of scope at most  $k$ , and the sum is  $\sum_{x \in \mathcal{C}} W^{\text{wt}(x)} = \sum_i A_i W^i$ , with no truncation. Theorem 5.3 and Lemma 5.5 with  $N = n + 1$ ,  $L = k$  give the time;  $d(\mathcal{C})$  and  $A_d$  are read off.

(ii) Replace the coordinate factor by the multilinear extension of  $W^{[x_v \neq y_v]}$ , which is  $1 + (W - 1)Y_v$  if  $y_v = 0$  and  $W + (1 - W)Y_v$  if  $y_v = 1$ ; the sum is  $\sum_{x \in \mathcal{C}} W^{\text{wt}(x+y)} = \sum_i A_i(y) W^i$ . For the nearest codeword let  $i_0 = \min\{i : A_i(y) > 0\}$  and fix the coordinates one at a time: fixing  $x_1 = \epsilon$  replaces every row containing coordinate 1 by the indicator that the remaining coordinates have parity  $\epsilon h_1$ , a factor of smaller scope, and leaves the system sparse; compute the two distance distributions, keep a value  $\epsilon$  for which the coefficient of  $W^{i_0}$  is positive, and continue with the next coordinate. After  $n$  steps the assignment is a codeword at distance  $i_0$  from  $y$ , and  $2n + 1$  distributions have been computed.

(iii) Apply Theorem 6.15 with  $D = Q$ : the variables take values in the alphabet, each check is a table on at most  $k$  coordinates, and the coordinate factor has the table  $s \mapsto W^{[s \neq 0]}$ , or  $s \mapsto W^{[s \neq y_v]}$  for the distance to  $y$ , in  $\mathcal{A}_{n+1}$ ; for the complete weight enumerator the table is  $s \mapsto W_s$  in  $\mathcal{A}_{n+1, \dots, n+1}$  with one generator per symbol, of dimension  $(n + 1)^Q$ .  $\square$

**Remark 7.13.** Computing the minimum distance of a binary linear code is NP-hard [Var97], and so is maximum-likelihood decoding [BMT78]; both are posed with the code given by a generator or a parity-check matrix of unrestricted density. For the low-density codes of Gallager,

with  $m \leq n$  rows of bounded weight, Theorem 7.12 gives the exact minimum distance, the complete weight enumerator and exact maximum-likelihood decoding of every received word in deterministic time  $2^{o(n)}$ , where the previously known exact methods are exponential in the dimension or the redundancy of the code. No bound on the column weights is needed, by Lemma 5.2.

**7.6. Bounded-degree hypergraphs.** A hypergraph  $\mathcal{H} = (V, \mathcal{E})$  has vertex set  $V$ ,  $|V| = n$ , and a multiset  $\mathcal{E}$  of edges  $e \subseteq V$ . A matching is a set of pairwise disjoint edges, a perfect matching or exact cover is a matching whose union is  $V$ , a transversal is a set of vertices meeting every edge, an independent set is a set of vertices containing no edge, and a strong independent set is a set of vertices no two of which lie in a common edge. A  $q$ -colouring of  $V$  is proper if no edge of size at least two is monochromatic.

**Theorem 7.14** (Hypergraphs). *Fix  $k, \Delta, c \geq 1$  and  $q \geq 2$ , and let  $\mathcal{H}$  have  $n$  vertices and edges of size at most  $k$ .*

- (i) *If every vertex lies in at most  $\Delta$  edges and the edges have integer weights in  $[0, \Omega]$ : the number of matchings of each weight and the number of exact covers of each weight, hence the maximum weight of a packing and the number of exact covers, in time  $\text{poly}(n, \Omega) 2^{O_{k, \Delta}(\rho(n))}$ .*
- (ii) *If  $|\mathcal{E}| \leq cn$  and the vertices have integer weights in  $[0, \Omega]$ : the number of transversals, of independent sets and of strong independent sets of each weight, in time  $\text{poly}(n, \Omega) 2^{O_{k, c}(\rho(n))}$ .*
- (iii) *If  $|\mathcal{E}| \leq cn$ : the number of proper  $q$ -colourings, and the number of  $q$ -colourings with exactly  $j$  monochromatic edges for every  $j$ , in time  $\text{poly}(n) 2^{O_{q, k, c}(\rho(n))}$ .*

*Proof.* (i) A vertex in no edge lies in no exact cover and in every matching trivially; delete such vertices, so that  $n \leq k|\mathcal{E}|$ , and if  $|\mathcal{E}| < 20$  list the subsets. Boolean variables  $y_e$ ,  $e \in \mathcal{E}$ ; for each vertex  $v$  the multilinear extension on the scope  $\{e : v \in e\}$ , of size at most  $\Delta$ , of the indicator that at most one (for matchings) or exactly one (for exact covers) incident edge is chosen; for each edge  $1 + (W^{\omega_e} - 1)Y_e$ . There are  $n + |\mathcal{E}| \leq (k + 1)|\mathcal{E}|$  factors of scope at most  $\Delta$ , the sum is  $\sum_M W^{\omega(M)}$  over the matchings, or over the exact covers, in  $\mathcal{A}_N$  with  $N = 1 + \sum_e \omega_e$ , and Theorem 5.3 on  $|\mathcal{E}| \leq \Delta n$  variables with the constants  $\Delta$  and  $k + 1$ , with Lemma 5.5,  $L = \Delta$ , gives the bound;  $\rho(|\mathcal{E}|) = O_\Delta(\rho(n))$ .

(ii) Boolean variables  $x_v$ ; for each edge  $e$  the multilinear extension, on the scope  $e$ , of the indicator that some  $x_v$ ,  $v \in e$ , is 1 (transversals), that not all are 1 (independent sets), or that at most one is 1 (strong independent sets); for each vertex  $1 + (W^{\omega_v} - 1)Y_v$ . At most  $(c + 1)n$  factors of scope at most  $k$ ; Theorem 5.3 and Lemma 5.5 with  $L = k$ .

(iii) Theorem 7.3 with  $D = q$  and, for each edge of size at least two, the constraint that it is monochromatic, of weight 1: the coefficient  $p_j$  counts the colourings with  $j$  monochromatic edges and  $p_0$  the proper ones.  $\square$

**7.7. Circuits of linear size.** A Boolean circuit with  $n$  inputs is a sequence of gates  $g_1, \dots, g_s$ , each computing a function of at most  $k$  earlier gates or inputs given by its truth table, with  $t$  of the gates designated as outputs. Its size is  $s$ .

**Theorem 7.15** (Linear-size circuits of bounded fan-in). *Fix  $k, c \geq 1$ . Let  $C$  be a circuit with  $n \geq 20$  inputs,  $s \leq cn$  gates of fan-in at most  $k$ , and outputs  $o_1, \dots, o_t$  with integer weights  $\omega_1, \dots, \omega_t \in [0, \Omega]$ . The numbers*

$$N_{i,j} = \#\left\{x \in \{0, 1\}^n : \text{wt}(x) = i, \sum_{l=1}^t \omega_l o_l(x) = j\right\}$$

are computed deterministically, for all  $(i, j)$ , in time  $\text{poly}(n, \Omega) 2^{O_{k,c}(\rho(n))}$ . In particular the number of inputs accepted by a single-output circuit, and the maximum over the inputs of a weighted sum of its output bits with the number of inputs attaining it, are computed in that time.

*Proof.* Take the  $n + s$  Boolean variables  $x_1, \dots, x_n, g_1, \dots, g_s$  and  $\mathcal{A} = \mathcal{A}_{n+1, 1 + \sum_l \omega_l}$  with generators  $U$  and  $W$ . For each gate  $g$  take the multilinear extension, on the scope consisting of  $g$  and its at most  $k$  arguments, of the indicator that  $g$  equals the value of its function at the arguments; for each input  $1 + (U - 1)Y_{x_v}$ ; for each output  $o_l$  the factor  $1 + (W^{\omega_l} - 1)Y_{o_l}$ . There are  $s + n + t \leq (2c + 1)n \leq (2c + 1)(n + s)$  factors of scope at most  $k + 1$ , with coefficients of absolute value at most  $2^{k+1}$ . For each input  $x$  exactly one assignment of the gate variables makes every gate factor 1, the one in which each  $g$  carries its value at  $x$ , and every other assignment makes some gate factor 0; so the sum over the cube is  $\sum_x U^{\text{wt}(x)} W^{\sum_l \omega_l o_l(x)} = \sum_{i,j} N_{i,j} U^i W^j$ , with nothing truncated. Theorem 5.3 on  $n + s \leq (c + 1)n$  variables with the constants  $k + 1$  and  $2c + 1$ , and Lemma 5.5 with  $L = k + 1$ , give the time, since  $\rho(n + s) \leq \rho((c + 1)n) = O_c(\rho(n))$ .  $\square$

*Remark 7.16.* The count of the accepted inputs of a circuit with  $s$  gates is, through the Tseitin encoding used in the proof, the count of the satisfying assignments of a  $(k + 1)$ -CNF formula with  $n + s$  variables, so for every  $s$  it is contained in OpenAI's Theorem 1.1 with the exponent  $o(n + s)$ ; their Corollary 1.4 draws a circuit lower bound from exactly this fact. Theorem 7.15 adds the Hamming weight of the input and the weighted output profile for the sparse case, which turns any objective computable by a linear-size circuit of bounded fan-in on  $n$  Boolean variables into one that is maximised exactly in time  $2^{o(n)}$ , together with the number of maximisers. Theorems 6.1, 7.5 and 7.12 are instances, since their constraints are computed by such circuits; the direct encodings given there avoid the gate variables and have smaller constants.

**7.8. Integer programs with sparse rows.** A 0-1 integer program asks for a vector  $x \in \{0, 1\}^n$  satisfying a system of linear inequalities and maximising a linear objective. Its rows need not be clauses, cuts or parity checks: a row is an arbitrary linear form with integer coefficients, compared with an integer by  $\leq$ ,  $=$  or  $\geq$ . When every row involves at most  $k$  variables and there are  $O(n)$  of them, each row is a local constraint, and Theorem 7.1 adds a fixed number of dense rows and the objective as linear statistics.

**Theorem 7.17** (Sparse integer programs). *Fix integers  $k, c \geq 1$  and  $s \geq 0$ . Consider  $n \geq 20$  variables,  $m \leq cn$  sparse rows*

$$\sum_{v \in I_f} a_{f,v} x_v \bowtie_f b_f, \quad |I_f| \leq k, \quad \bowtie_f \in \{\leq, =, \geq\}, \quad a_{f,v}, b_f \in \mathbb{Z},$$

*with coefficients of arbitrary size, a cost vector  $\gamma$  and dense rows  $\alpha^{(1)}, \dots, \alpha^{(s)}$  in  $\{-\Omega, \dots, \Omega\}^n$ ,  $\Omega \geq 1$ , and let  $P \subseteq \{0, 1\}^n$  be the set of vectors satisfying the sparse rows.*

(i) *The numbers*

$$p(j_0, \dots, j_s) = \#\{x \in P : \gamma \cdot x = j_0, \alpha^{(i)} \cdot x = j_i \ (1 \leq i \leq s)\}$$

*are computed for all tuples in time  $\text{poly}(\ell, (1 + 2n\Omega)^{s+2}) 2^{O_{k,c,s}(\rho(n))}$ , where  $\ell$  is the bit length of the input. Hence the largest and the smallest value of  $\gamma \cdot x$  over the  $x \in P$  with  $\alpha^{(i)} \cdot x \bowtie_i \beta_i$  for prescribed relations  $\bowtie_i$  and integers  $\beta_i$ , and the number of vectors attaining each, are computed within the same time.*

(ii) *For a fixed  $D \geq 1$  the same holds for vectors  $x \in \{0, 1, \dots, D\}^n$ , with time  $\text{poly}(\ell, (1 + 4Dn\Omega)^{s+2}) 2^{O_{D,k,c,s}(\rho(n))}$ .*

*Proof.* (i) For each sparse row let  $\phi_f : \{0, 1\}^{I_f} \rightarrow \{0, 1\}$  be the indicator that the row holds, and let  $\Phi_f$  be its multilinear extension (2.6). The table of  $\phi_f$  has at most  $2^k$  entries, each found by

one evaluation of the row, at a cost of  $\text{poly}(\ell)$  bit operations; the coefficients of  $\Phi_f$  are integers of absolute value at most  $2^k$ . So  $(\Phi_f)$  is a factor system with at most  $cn$  factors of scope at most  $k$  and values in  $\mathbb{Q}$ , and its product at  $x \in \{0, 1\}^n$  is the indicator of  $P$ . To bring the statistics to nonnegative weights, put  $\omega_v^{(0)} = \gamma_v + \Omega$ ,  $\omega_v^{(i)} = \alpha_v^{(i)} + \Omega$  for  $1 \leq i \leq s$ , and  $\omega_v^{(s+1)} = 1$ , all in  $[0, 2\Omega]$ . For every  $x$ ,

$$\gamma \cdot x = \omega^{(0)} \cdot x - \Omega |x|, \quad \alpha^{(i)} \cdot x = \omega^{(i)} \cdot x - \Omega |x|, \quad |x| = \omega^{(s+1)} \cdot x.$$

Theorem 7.1(i) with  $N_0 = 1$ , the  $s + 2$  weight vectors  $\omega^{(0)}, \dots, \omega^{(s+1)}$  and the weight bound  $2\Omega$  computes the numbers  $r(j'_0, \dots, j'_{s+1})$  of  $x \in P$  with  $\omega^{(i)} \cdot x = j'_i$  for all  $i$ , in time  $\text{poly}(n, (1 + 2n\Omega)^{s+2})2^{O_{k,c,s}(\rho(n))}$ , to which the cost  $cn2^k \text{poly}(\ell)$  of the tables is added. Then

$$p(j_0, \dots, j_s) = \sum_{t=0}^n r(j_0 + \Omega t, j_1 + \Omega t, \dots, j_s + \Omega t, t),$$

because a vector with  $|x| = t$  has  $\omega^{(i)} \cdot x = \alpha^{(i)} \cdot x + \Omega t$ . The optimum over the  $x \in P$  with  $\alpha^{(i)} \cdot x \bowtie_i \beta_i$  is the largest, or smallest,  $j_0$  for which  $\sum p(j_0, j_1, \dots, j_s) > 0$ , the sum taken over the tuples with  $j_i \bowtie_i \beta_i$ , and that sum counts the optimal vectors.

(ii) Let  $T = \lceil \log_2(D + 1) \rceil$  and replace each  $x_v$  by  $T$  Boolean variables  $x_{v,0}, \dots, x_{v,T-1}$  with  $x_v = \sum_t 2^t x_{v,t}$ . A sparse row becomes a sparse row in the new variables, with the coefficients  $2^t a_{f,v}$  and at most  $kT$  variables; add for each  $v$  the row  $\sum_t 2^t x_{v,t} \leq D$ , which involves  $T$  variables. The cost and the dense rows become vectors with entries  $2^t \gamma_v$  and  $2^t \alpha_v^{(i)}$ , of absolute value at most  $2^T \Omega \leq 2(D + 1)\Omega \leq 4D\Omega$ . The new program has  $n' = Tn$  variables and at most  $(c + 1)n \leq (c + 1)n'$  sparse rows of at most  $kT$  variables, and its feasible vectors correspond bijectively to the vectors of  $\{0, \dots, D\}^n$  satisfying the original sparse rows, with the same values of  $\gamma \cdot x$  and  $\alpha^{(i)} \cdot x$ . Part (i) applies, with  $\rho(Tn) \leq 4T\rho(n)$  for  $n \geq T$ .  $\square$

The theorem places no bound on the coefficients of the sparse rows, only on their number of variables; the polynomial bound  $\Omega$  is needed only for the objective and the dense rows, which enter through the exponents of the generators. Three classical problems that are NP-hard in the strong sense fall under it.

**Corollary 7.18.** *Fix  $k, c, s \geq 1$ . Each of the following is computed deterministically in time  $\text{poly}(\ell, \Omega^{s+2})2^{O(\rho(n))}$ , where  $\ell$  is the length of the input, together with the number of optimal solutions, the implied constant depending on  $k, c, s$ .*

- (a) *For  $n$  sets over a universe of at most  $cn$  elements, each element lying in at most  $k$  of the sets, with costs in  $[0, \Omega]$ : the least cost of an exact cover, of a set cover and the largest cost of a set packing. Exact cover by 3-sets is NP-complete already when every element lies in exactly three sets [Gon85].*
- (b) *For  $n$  items with values and  $s$  weight vectors in  $[0, \Omega]$ , capacities  $B_1, \dots, B_s$  and a partial order given by at most  $cn$  relations  $u \prec v$ : the largest value of a set of items within the capacities that contains, with each item, every item below it. For  $s = 1$  this is the partially ordered knapsack problem, which is NP-complete in the strong sense [JN83].*
- (c) *For a system of at most  $cn$  linear equations  $\sum_{v \in I_f} a_{f,v} x_v = b_f$  with integer coefficients,  $|I_f| \leq k$ , and a cost vector in  $[-\Omega, \Omega]^n$ : the number of its 0-1 solutions of each cost, and the least and largest cost; in particular the number of solutions of a 1-in- $k$ -SAT instance with at most  $cn$  clauses, and its solutions with the fewest true variables.*

*Proof.* (a) Take a variable for each set and, for each element, the row  $\sum_{S \ni e} x_S = 1, \geq 1$  or  $\leq 1$ ; each row involves at most  $k$  variables, and there are at most  $cn$  rows. The cost is  $\gamma$ , with  $s = 0$ . (b) Take a variable for each item, the row  $x_v - x_u \leq 0$ , which says that  $x_v = 1$  forces  $x_u = 1$ ,

for each relation  $u \prec v$ , and the weight vectors as dense rows with  $\beta_i = B_i$  and  $\bowtie_i$  the relation  $\leq$ ; the downward closure of the chosen set under the given relations is equivalent to its closure under the order they generate. (c) The equations are the sparse rows; for 1-in- $k$ -SAT a clause on the literals  $\ell_1, \dots, \ell_k$  is the equation  $\sum_i \ell_i = 1$ , where a negated literal  $1 - x_v$  moves the constant 1 to the right side. All three are Theorem 7.17(i).  $\square$

**7.9. Growing density.** Every statement so far fixes the ratio  $c$  of factors to variables. The parameters of Lemma 4.7 do not depend on the degree ratio  $\kappa$ , and the exponent of Theorem 5.1 is linear in it. When  $\kappa$  grows with  $n$  it pays to let the parameters follow it: fewer groups, and so a smaller  $Q/R$  in the Taylor count, in exchange for a larger share of the cost spent on one group. With the entropy bound of Lemma 3.4 in place of (3.21), this extends the subexponential regime to every ratio  $\kappa = o(\sqrt{\log n / \log \log n})$ .

Throughout,  $\lambda = \lceil \log n \rceil$ ,  $\nu = \lceil \log \lambda \rceil$ , and

$$g(x) = x (\log(1/x))^{3/2} \quad (0 < x < 1),$$

which is increasing on  $(0, 2^{-3}]$  and tends to 0 with  $x$ .

**Theorem 7.19** (Bounded occurrences at a growing degree ratio). *There is an absolute constant  $\lambda_0$  with the following property. Let  $\mathcal{A}$  be a commutative  $\mathbb{Q}$ -algebra, and let a factor system on  $n$  variables with  $\lambda \geq \lambda_0$  have largest scope  $w \leq \lambda^2$ , largest occurrence number  $b \leq \lambda^2 - 1$  and total degree  $d \leq an$ , where  $a \geq 1$  is a real number with*

$$\theta := a \sqrt{\nu/\lambda} \leq 2^{-15}.$$

*Then  $Z$  in (2.4) is computed deterministically with at most*

$$|\mathcal{F}| \text{ poly}(n) 2^{790 n g(\theta)} \quad (7.5)$$

*operations in  $\mathcal{A}$ , evaluations of factors at inputs in  $\mathbb{Q}[z]/(z^J)$  with  $J \leq n$ , and bit operations on rational numbers, using polynomially many elements of  $\mathcal{A}$  and polynomially many bits of storage in addition to the factors' own. The polynomial is absolute.*

*Proof.* Adjoin the constant-one factors of Section 4, so that (4.1) holds with  $b \leq \lambda^2$  and  $d$  unchanged. Replacing  $a$  by  $\max(1, d/n)$ , which is rational, computed from the input and not larger than  $a$ , decreases  $\theta$  and, as  $g$  is increasing, the bound (7.5); so we may assume  $a = \max(1, d/n)$ . Put  $L = \log(1/\theta) \geq 15$  and choose

$$t = \lfloor L \rfloor, \quad r = 2^t, \quad Q = \left\lfloor \frac{\lambda}{128t} \right\rfloor, \quad R = 32 + \lceil 16\sqrt{Q\nu} \rceil, \quad \tau = \frac{1}{r\lambda}, \quad D = r^{2Q}.$$

Thus  $t$  is the largest integer with  $4^t a^2 \nu \leq \lambda$ , and  $1/(2\theta) < r \leq 1/\theta$ . Since  $a \geq 1$  and  $\lambda \leq 2^\nu$ ,  $L \leq \frac{1}{2} \log(\lambda/\nu) \leq \nu/2$ , so  $t \leq \nu/2$  and  $r^2 \leq 2^\nu \leq 2\lambda$ .

We check the hypotheses of Lemma 4.6. First,  $\lambda \geq 256t$ , so  $Q \geq \lambda/(256t)$ , and  $\log D = 2Qt \leq \lambda/64$ . Since  $Q\nu \geq 1$ ,  $R \leq 33 + 16\sqrt{Q\nu} \leq 49\sqrt{Q\nu}$ , hence

$$\frac{R}{Q} \leq 49 \sqrt{\frac{\nu}{Q}} \leq 784 \sqrt{\frac{t\nu}{\lambda}} \leq 784 \frac{\nu}{\sqrt{2\lambda}}, \quad (7.6)$$

which is at most  $1/6$  once  $\lambda \geq (6 \cdot 784)^2 \nu^2/2$ ; so  $R \geq 32$ ,  $6R \leq Q$  and  $R < Q$  for  $\lambda \geq \lambda_0$ . The balance conditions (4.15) hold:  $2(b+w)D^2 \leq 4\lambda^2 2^{\lambda/32} \leq 2^{\lambda-1} < n \leq M$  and  $2Qr^2 \leq (\lambda/64t) \cdot 2\lambda < n$  for  $\lambda \geq \lambda_0$ . For the interval take  $\mu = Q(r-1)/(2r)$ ,  $A_0 = \lfloor \mu - (R-1)/2 \rfloor$  and  $\mathcal{J} = \{A_0, \dots, A_0 + R-1\}$  as in the proof of Lemma 4.7; since  $R \leq Q/6$  it lies in  $\{0, \dots, Q\}$ ,

and since  $R \geq 32$  it contains every integer within  $R/4$  of  $\mu$ . Hoeffding's inequality (2.8) with  $s = R/4$  and  $R^2 \geq 256Q\nu$  gives

$$\Pr\left[\text{Bin}\left(Q, \frac{r-1}{2r}\right) \notin \mathcal{J}\right] \leq 2e^{-R^2/(8Q)} \leq 2e^{-32\nu} \leq 2^{-4\nu} \leq \frac{1}{r^2 b \lambda} = \frac{\tau}{rb},$$

because  $r^2 \leq 2^\nu$ ,  $b \leq \lambda^2 \leq 2^{2\nu}$  and  $\lambda \leq 2^\nu$ . This is (4.20), and  $0 < \tau \leq 1/2$ . Lemma 4.6 constructs the design in polynomial time.

Now the costs of Proposition 3.5. By (3.1) with  $d \leq an$ ,  $q = \lfloor dR/Q \rfloor \leq \chi n$  with  $\chi = aR/Q$ . Since  $\sqrt{\nu/\lambda} = \theta/a \leq \theta$  and  $t \leq L$ , (7.6) gives

$$\chi \leq 784a\sqrt{\frac{t\nu}{\lambda}} \leq 784\theta\sqrt{L} =: \bar{\chi},$$

and  $\bar{\chi} \leq 784 \cdot 2^{-15} \sqrt{15} < \frac{1}{10}$  because  $x\sqrt{\log(1/x)}$  increases on  $(0, 2^{-15}]$ . In particular  $q \leq n/2$  and  $J = q + 1 \leq n$ . Lemma 3.4 and the monotonicity of  $x \log(2en/x)$  in  $x \leq n$  give

$$\log C_q \leq \bar{\chi} n \log \frac{2e}{\bar{\chi}} = 784\theta\sqrt{L} n \left( L + \log \frac{2e}{784\sqrt{L}} \right) \leq 784 n \theta L^{3/2} = 784 n g(\theta).$$

The measurement count contributes  $2\tau n \log(2D + 1) \leq \frac{2n}{r\lambda} \left( \frac{\lambda}{64} + 2 \right) \leq \frac{n}{r}$  and one group  $\frac{2n}{r}$ , together  $\frac{3n}{r} < 6n\theta \leq 6n g(\theta)$ . The logarithm of (3.20) is therefore at most  $\log(2Qr) + 790 n g(\theta)$ . Each vector costs at most  $|\mathcal{F}|$  factor evaluations and a number of further operations polynomial in  $n$ ,  $|\mathcal{F}|$ ,  $D$ ,  $Q$ ,  $R$ ,  $J$ ,  $r$  and  $M \leq \lambda^2 n$ , all of which are bounded by powers of  $n$  and  $|\mathcal{F}|$  since  $D \leq 2^{\lambda/64} \leq 2n^{1/64}$ . The storage bound is that of Proposition 3.5.

The parameters are computed from  $n$  and  $d$  with integer square roots and exact rational comparisons, in  $\text{poly}(\log n, \log d)$  bit operations, and the conditions of Lemma 4.6 are checked exactly as in the proof of Lemma 4.7; the estimates above show that every check passes.  $\square$

For a fixed ratio  $a$  one has  $\theta = a\sqrt{\nu/\lambda}$  and  $\log(1/\theta) \leq \nu/2$ , so  $g(\theta) \leq a\nu^2/(2\sqrt{2\lambda})$ , and (7.5) is the bound of Theorem 5.1 with  $\beta = 2$  up to the constant. The gain is in the dependence on  $a$ : the exponent of Theorem 5.1 is  $an\nu^2/\sqrt{\lambda}$ , which is  $o(n)$  only for  $a = o(\sqrt{\lambda}/\nu^2)$ , while  $ng(\theta)$  is  $o(n)$  for every  $a = o(\sqrt{\lambda}/\nu)$ . Two changes produce the gain. The entropy bound (3.19) charges the Taylor count  $\log C_q \approx q \log(n/q)$ , whereas (3.21) charges  $q \log(Q/R)$ , larger by the factor  $\log(Q/R)/\log(Q/(aR))$ , which grows with  $a$ . And the number of groups  $r = 2^{\lfloor \log(1/\theta) \rfloor}$  shrinks as  $a$  grows, which makes  $Q$  larger and  $R/Q$  smaller at the price of the cost  $3n/r$  of the groups, balanced against the Taylor count.

**Theorem 7.20** (Growing density). *There are absolute constants  $n_1$  and  $\theta_1 > 0$  with the following property. Let  $n \geq n_1$ , let  $k \geq 1$  be an integer and  $c \geq 1$  a real number with*

$$\theta_n := ck \sqrt{\frac{\log \log n}{\log n}} \leq \theta_1,$$

*and let a factor system on  $n$  variables have at most  $cn$  factors, each of scope at most  $k$ , with values in a commutative  $\mathbb{Q}$ -algebra  $\mathcal{A}$ . Then  $Z$  in (2.4) is computed deterministically with at most*

$$\text{poly}(n) 2^{4800 n g(\theta_n)} \leq \text{poly}(n) 2^{1700 ck \rho(n)}$$

*operations in  $\mathcal{A}$ , evaluations of the factors at inputs in  $\mathbb{Q}[z]/(z^J)$ ,  $J \leq n$ , and bit operations on rationals, in polynomial space; the polynomial is absolute. In particular the bound is  $2^{o(n)}$  whenever  $ck = o(\sqrt{\log n / \log \log n})$ . If the coefficients of the factors are polynomials in the generators of  $\mathcal{A} = \mathcal{A}_N$  or  $\mathcal{A}_{N_1, \dots, N_s}$  with integer coefficients of at most  $L$  bits, the bound holds for bit operations after multiplication by a polynomial in  $n$ ,  $N$  and  $L$ .*

*Proof.* Put  $\theta_1 = 2^{-18}$  and  $\varepsilon_n = \sqrt{\log \log n / \log n}$ , so that  $\theta_n = ck \varepsilon_n \geq \varepsilon_n$ . Let  $n_1$  be so large that for  $n \geq n_1$  one has  $\lambda - 1 \geq \lambda_0$ , the absolute constant of Theorem 7.19, and  $\theta_1 \sqrt{\log n / \log \log n} \leq \lambda/2$ , so that  $ck \leq \lambda/2$ . Apply Lemma 5.2: fewer than  $2^{n/\lambda}$  assignments  $y$ , each leaving a system on  $n' \geq n - n/\lambda \geq n/2$  variables with  $\lambda' = \lceil \log n' \rceil \in \{\lambda - 1, \lambda\}$ ,  $\nu' \leq \nu$ , scopes at most  $k \leq \lambda'^2$ , occurrence numbers at most  $ck\lambda \leq \lambda^2/2 \leq \lambda'^2 - 1$ , and total degree at most  $kcn \leq 2ck n'$ . Theorem 7.19 applies with  $a = 2ck$  and

$$\theta' = 2ck \sqrt{\nu'/\lambda'} \leq 2ck \sqrt{2\nu/\lambda} \leq 6ck \varepsilon_n = 6\theta_n \leq 2^{-15},$$

where  $\nu \leq \log \log n + 2 \leq 3 \log \log n$  and  $\lambda \geq \log n$  were used for  $n \geq 20$ . As  $g$  is increasing on  $(0, 2^{-3}]$ ,  $g(\theta') \leq g(6\theta_n) \leq 6\theta_n \log^{3/2}(1/\theta_n) = 6g(\theta_n)$ . Each  $Z_y$  costs at most  $cn \text{poly}(n) 2^{4740ng(\theta_n)}$  operations, and  $cn \leq n^2$ . The branching adds the factor  $2^{n/\lambda} \leq 2^{n\varepsilon_n} \leq 2^{ng(\theta_n)}$ , since  $1/\lambda \leq 1/\log n \leq \varepsilon_n \leq \theta_n$  and  $\log(1/\theta_n) \geq 18$ . The total is at most  $\text{poly}(n) 2^{4741ng(\theta_n)}$ , which gives the first bound. For the second,  $\log(1/\theta_n) \leq \log(1/\varepsilon_n) \leq \frac{1}{2} \log \log n$ , so  $g(\theta_n) \leq ck \varepsilon_n (\log \log n)^{3/2} / 2^{3/2} = ck \rho(n) / (2^{3/2}n)$ , and  $4800/2^{3/2} < 1700$ . If  $ck = o(\sqrt{\log n / \log \log n})$  then  $\theta_n \rightarrow 0$  and  $g(\theta_n) \rightarrow 0$ . The space bound and  $J \leq n' \leq n$  come from Theorem 7.19, and the bit complexity is Lemma 5.5: its proof uses of the parameters only  $Q, \lambda = O(\log n)$ ,  $R < Q$  and  $D \leq 2^\lambda$ , which hold here, and its polynomial depends on  $|\mathcal{F}| \leq n^2$  and  $k \leq \lambda^2$  only through powers of  $n$ .  $\square$

The corollaries of Sections 6 and 7 that are proved by one application of Theorem 5.3 to a factor system whose number of factors is bounded by the number of constraints or edges of the input, and whose scopes are bounded, therefore extend to inputs of growing density. For a function  $d(n) \geq 1$  write

$$\delta_n = d(n) \sqrt{\frac{\log \log n}{\log n}};$$

the condition  $\delta_n \rightarrow 0$  means  $d(n) = o(\sqrt{\log n / \log \log n})$ . We record the two for which the question of the dense case was raised in the introduction.

**Corollary 7.21** (Cuts and spin systems at growing average degree). *Let  $d = d(n) \geq 1$  satisfy  $\delta_n \rightarrow 0$ , and let  $G$  be a graph with  $n$  vertices and at most  $d(n)n$  edges.*

- (i) *The cut profile  $(\#\{S \subseteq V : |S| = i, e(S, V \setminus S) = j\})_{i,j}$ , in particular the maximum cut, the maximum bisection and their numbers of optimal solutions, is computed in time  $\text{poly}(n) 2^{20000ng(\delta_n)} = 2^{o(n)}$ .*
- (ii) *The density of states of an Ising system on  $G$  with integer couplings and fields, and that of the  $q$ -state Potts system on  $G$  for fixed  $q$ , are computed in time  $\text{poly}(n, E_{\max}) 2^{O_q(n g(\delta_n))} = \text{poly}(E_{\max}) 2^{o(n)}$ .*

*In particular the maximum cut and the ground-state energy of an Ising system on a graph of average degree  $d(n)$  are found in time  $2^{o(n)}$  whenever  $d(n) = o(\sqrt{\log n / \log \log n})$ .*

*Proof.* (i) The factor system of Corollary 6.4 has one factor of scope 2 per edge and one of scope 1 per vertex, with values in  $\mathcal{A}_{n+1, m+1}$  and coefficients of absolute value at most 2; so it has at most  $(d(n) + 1)n \leq 2d(n)n$  factors of scope at most 2. Theorem 7.20 with  $k = 2$  and  $c = 2d(n)$  applies once  $4\delta_n \leq \theta_1$ , and gives  $\text{poly}(n) 2^{4800ng(4\delta_n)}$  operations;  $g(4\delta_n) \leq 4g(\delta_n)$ , and Lemma 5.5 with  $N = (n + 1)(m + 1) \leq n^3$  and  $L = 1$  turns the count into the time bound. The optimum and its multiplicity are read off the profile as in Corollary 6.4. (ii) The factor system in the proof of Theorem 6.18(i) has one factor of scope 2 per edge and one of scope 1 per vertex, with values in  $\mathcal{A}_N$ ,  $N = 2E_{\max} + 1$ , and coefficients of  $O(\log E_{\max})$  bits; it has at most  $2d(n)n$  factors, and Theorem 7.20 with  $k = 2$ ,  $c = 2d(n)$  and Lemma 5.5 give the bound

as in (i). The factor system in the proof of Theorem 6.18(ii) lives on  $n' = qn$  variables and has  $n$  factors of scope  $q$  and at most  $d(n)n$  of scope  $2q$ , with coefficients of  $O(q + \log E_{\max})$  bits; so it has at most  $(d(n) + 1)n \leq (d(n) + 1)n'$  factors of scope at most  $2q$ . Theorem 7.20 on  $n'$  variables with  $k = 2q$  and  $c = d(n) + 1$  has  $\theta_{n'} \leq 4q d(n) \sqrt{\log \log n' / \log n'} \leq 4q \delta_n$  for large  $n$ , because  $\log \log x / \log x$  decreases for  $x \geq 16$ ; it gives  $\text{poly}(n) 2^{4800 qn g(4q\delta_n)}$  operations, and  $g(4q\delta_n) \leq 4q g(\delta_n)$ . Lemma 5.5 turns the counts into time bounds, and the ground-state energy and the partition functions follow as in Theorem 6.18.  $\square$

At  $d(n) = \sqrt{\log n} / (\log \log n)^3$ , for instance,  $\delta_n = (\log \log n)^{-5/2}$  and the exponent in (i) is  $O(n / (\log \log n)^2)$ .

The cut function of a graph and that of its complement determine each other, so the corollary also reaches graphs with  $\Theta(n^2)$  edges whose complement is sparse.

**Corollary 7.22** (Graphs with sparse complements). *Let  $d = d(n) \geq 1$  satisfy  $\delta_n \rightarrow 0$  and let  $\Omega_0 \geq 1$  be an integer. Let  $H$  be a graph on a vertex set  $V$  of size  $n$  with at most  $d(n)n$  edges, and let  $\omega$  be a weighting of the pairs of distinct vertices with  $\omega_{uv} = \Omega_0$  for  $uv \notin E(H)$  and  $\omega_{uv} \in \{0, 1, \dots, \Omega_0\}$  for  $uv \in E(H)$ . Then the numbers*

$$p_{i,j} = \#\left\{S \subseteq V : |S| = i, \sum_{u \in S, v \notin S} \omega_{uv} = j\right\}$$

*are computed for all  $i, j$  in time  $\text{poly}(n, \Omega_0) 2^{O(n g(\delta_n))} = \text{poly}(\Omega_0) 2^{o(n)}$ . In particular, for every graph  $G$  on  $n$  vertices whose complement has at most  $d(n)n$  edges, the cut profile of  $G$ , the maximum cut, the maximum and the minimum bisection, and the numbers of optimal solutions, are computed in time  $2^{o(n)}$ .*

*Proof.* Put  $\eta_{uv} = \Omega_0 - \omega_{uv} \in \{0, \dots, \Omega_0\}$  for  $uv \in E(H)$ . Every pair  $\{u, v\}$  with  $u \in S$  and  $v \notin S$  carries the weight  $\Omega_0$ , less  $\eta_{uv}$  when it is an edge of  $H$ , and there are  $|S|(n - |S|)$  such pairs; hence

$$\sum_{u \in S, v \notin S} \omega_{uv} = \Omega_0 |S| (n - |S|) - \text{cut}_\eta(S), \quad (7.7)$$

where  $\text{cut}_\eta(S)$  is the total  $\eta$ -weight of the edges of  $H$  between  $S$  and  $V \setminus S$ . Corollary 6.4, with Theorem 7.20 in place of Theorem 5.3 as in the proof of Corollary 7.21(i), computes the numbers  $q_{i,j}$  of sets  $S$  of size  $i$  with  $\text{cut}_\eta(S) = j$  in the stated time; the algebra is  $\mathcal{A}_{n+1, \Omega_0 m + 1}$  with  $m = |E(H)|$ , of dimension polynomial in  $n$  and  $\Omega_0$ . By (7.7),  $p_{i,j} = q_{i, \Omega_0 i(n-i) - j}$ , read as 0 when the index is negative or exceeds  $\Omega_0 m$ . For a graph  $G$  take  $H$  its complement,  $\Omega_0 = 1$  and  $\omega = 0$  on  $E(H)$ ; then  $\omega$  is the indicator of  $E(G)$ , and the maximum cut and the bisections are read off the numbers  $p_{i,j}$ .  $\square$

Corollary 7.22 rests on the fact that the cut of the complete graph depends on  $|S|$  alone. More generally, the cut of a graph whose adjacency is a fixed quadratic expression in a few vertex weights depends only on the values of those weights on  $S$ , and these values are linear statistics in the sense of Theorem 7.1. A sparse perturbation of such a graph is then within reach.

**Theorem 7.23** (Sparse perturbations of graphs of bounded rank). *Let  $d = d(n) \geq 1$  satisfy  $\delta_n \rightarrow 0$ , and let  $s \geq 1$  and  $\Omega \geq 1$  be integers. Let  $V$  be a set of size  $n$ , let  $w_1, \dots, w_s : V \rightarrow \{0, 1, \dots, \Omega\}$  and let  $C = (c_{ab})$  be a symmetric  $s \times s$  integer matrix with entries in  $[-\Omega, \Omega]$ . Put*

$$\omega_{uv} = \sum_{a,b=1}^s c_{ab} w_a(u) w_b(v) \quad (u \neq v),$$

let  $H$  be a graph on  $V$  with at most  $d(n)n$  edges and let  $\eta : E(H) \rightarrow \{-\Omega, \dots, \Omega\}$ , extended by 0 to the other pairs. For  $S \subseteq V$  let  $\text{cut}(S) = \sum_{u \in S, v \notin S} (\omega_{uv} + \eta_{uv})$ . Then the numbers

$$p_{i,j} = \#\{S \subseteq V : |S| = i, \text{cut}(S) = j\} \quad (i \geq 0, j \in \mathbb{Z})$$

are computed for all  $i, j$  in deterministic time

$$\text{poly}(n, \Omega, (1 + n\Omega)^s) 2^{O(n g(\delta_n))},$$

which is  $2^{o(n)}$  whenever  $s \log(n\Omega) = o(n)$ .

*Proof.* Encode  $S$  by its indicator  $x \in \{0, 1\}^V$ , and put  $j_a = \sum_v w_a(v)x_v$  and  $\Sigma_a = \sum_v w_a(v)$ . Every ordered pair  $(u, v)$  with  $u \in S$  and  $v \notin S$  has  $u \neq v$ , so

$$\sum_{u \in S, v \notin S} \omega_{uv} = \sum_{a,b} c_{ab} \left( \sum_{u \in S} w_a(u) \right) \left( \sum_{v \notin S} w_b(v) \right) = \sum_{a,b} c_{ab} j_a (\Sigma_b - j_b) =: P(j_1, \dots, j_s), \quad (7.8)$$

and  $\text{cut}(S) = P(j_1, \dots, j_s) + \text{cut}_{\eta^+}(S) - \text{cut}_{\eta^-}(S)$ , where  $\eta^\pm = \max(\pm\eta, 0)$  and  $\text{cut}_{\eta^\pm}(S)$  is the  $\eta^\pm$ -weight of the edges of  $H$  between  $S$  and  $V \setminus S$ . Let  $N_\pm = 1 + \sum_e \eta_e^\pm \leq 1 + d(n)n\Omega$ ,  $N_a = 1 + \Sigma_a \leq 1 + n\Omega$ , and

$$\mathcal{A} = \mathbb{Q}[W_+, W_-, U, T_1, \dots, T_s] / (W_+^{N_+}, W_-^{N_-}, U^{N_+}, T_1^{N_1}, \dots, T_s^{N_s}),$$

of dimension  $N \leq (1 + d(n)n\Omega)^2(n+1)(1+n\Omega)^s$ . For an edge  $e = \{u, v\}$  of  $H$  take the factor  $1 + (W_+^{\eta_e^+} W_-^{\eta_e^-} - 1)(Y_u + Y_v - 2Y_u Y_v)$ , the multilinear extension of  $W_+^{\eta_e^+ [x_u \neq x_v]} W_-^{\eta_e^- [x_u \neq x_v]}$ , and for a vertex  $v$  the factor  $1 + (U T_1^{w_1(v)} \dots T_s^{w_s(v)} - 1)Y_v$ , of degree one in  $Y_v$  as in the proof of Theorem 7.1. There are at most  $(d(n)+1)n \leq 2d(n)n$  factors of scope at most 2, with coefficients that are polynomials in the generators with integer coefficients of absolute value at most 2, and their sum is

$$Z = \sum_x W_+^{\text{cut}_{\eta^+}(x)} W_-^{\text{cut}_{\eta^-}(x)} U^{|x|} T_1^{j_1} \dots T_s^{j_s} = \sum r(h_+, h_-, i, j_1, \dots, j_s) W_+^{h_+} W_-^{h_-} U^i T_1^{j_1} \dots T_s^{j_s},$$

with no term truncated. Theorem 7.20 with  $c = 2d(n)$  and  $k = 2$  applies once  $4\delta_n \leq \theta_1$  and computes  $Z$  with  $\text{poly}(n)2^{4800ng(4\delta_n)}$  operations in  $\mathcal{A}$ , and  $g(4\delta_n) \leq 4g(\delta_n)$ ; the bit bound of that theorem gives the stated time. Finally  $p_{i,j} = \sum r(h_+, h_-, i, j_1, \dots, j_s)$  over the tuples with  $h_+ - h_- + P(j_1, \dots, j_s) = j$ , a sum over at most  $N$  coefficients.  $\square$

**Corollary 7.24** (Blow-ups with few changes). *Let  $d = d(n)$  satisfy  $\delta_n \rightarrow 0$  and let  $s = s(n) = o(n/\log n)$ . Let  $F$  be a graph on  $\{1, \dots, s\}$ , loops allowed, and let  $V = V_1 \cup \dots \cup V_s$  be a partition of a set of size  $n$ . The blow-up  $F[V_1, \dots, V_s]$  joins  $u \in V_a$  and  $v \in V_b$ ,  $u \neq v$ , exactly when  $ab \in E(F)$ . For every graph  $G$  on  $V$  that differs from the blow-up in at most  $d(n)n$  pairs, the cut profile of  $G$ , its maximum cut, its maximum and minimum bisection, and the numbers of optimal solutions, are computed in time  $2^{o(n)}$ . This holds in particular for every graph that differs in at most  $d(n)n$  pairs from a complete multipartite graph with  $o(n/\log n)$  parts.*

*Proof.* Take  $\Omega = 1$ ,  $w_a$  the indicator of  $V_a$  and  $c_{ab} = 1$  for  $ab \in E(F)$  and 0 otherwise; then  $\omega_{uv}$  is the adjacency of the blow-up, since exactly one term of the sum is nonzero. Let  $H$  be the graph of the pairs where  $G$  and the blow-up differ, with  $\eta = +1$  on the pairs that are edges of  $G$  only and  $\eta = -1$  on the others; then  $\omega + \eta$  is the adjacency of  $G$ . Theorem 7.23 applies with  $s \log(n+1) = o(n)$ , and the optima are read off the numbers  $p_{i,j}$  as in Corollary 6.4.  $\square$

Corollary 7.22 is the case  $s = 1$ ,  $w_1 = 1$ ,  $c_{11} = \Omega_0$  of the theorem. Without the perturbation the cut is the polynomial  $P$  of the linear statistics  $j_a$ , and its maximisation needs no part of the routine: the  $s$ -tuples  $(j_a)$  that occur are counted in time  $\text{poly}(n\Omega)^s$  by dynamic programming over the vertices, and for a positive semidefinite objective of fixed rank Allemand, Fukuda,

Liebling and Steiner [AFLS01] solve unconstrained 0-1 quadratic optimisation in polynomial time. The perturbation  $\eta$  is what requires Theorem 7.20: for  $s = 1$  and  $\omega = 0$  it is Max-Cut on a graph of average degree  $d(n)$ .

*Remark 7.25.* The graphs reached by Theorem 7.23 in time  $2^{o(n)}$  form a vanishingly small part of all graphs. Fix  $\varepsilon > 0$  and parameters with  $s \log(n\Omega) \leq \varepsilon n$ . A graph  $G$  covered by the theorem is determined by  $w_1, \dots, w_s$ ,  $C$ , the edge set of  $H$  and the restriction of the adjacency of  $G$  to  $E(H)$ , since  $\eta$  is then the difference of that restriction and  $\omega$ . There are at most

$$(\Omega + 1)^{ns} (2\Omega + 1)^{s^2} n^{2d(n)n} 2^{d(n)n}$$

such choices. Because  $s \leq n$ , both  $ns \log(\Omega + 1)$  and  $s^2 \log(2\Omega + 1)$  are at most  $2n s \log(n\Omega) \leq 2\varepsilon n^2$ , and  $d(n)n(2 \log n + 1) = o(n^2)$ , so the count is at most  $2^{5\varepsilon n^2}$  for  $n$  large, while there are  $2^{n(n-1)/2}$  graphs on  $V$ . As  $\varepsilon$  is arbitrary, the proportion of graphs on  $V$  covered by the theorem with parameters for which its bound is  $2^{o(n)}$  tends to 0, and the random graph  $G(n, \frac{1}{2})$  is covered with probability tending to 0. The exact maximum cut of every graph in time  $2^{o(n)}$  is therefore not a consequence of the theorem, and it is not proved in this paper.

The bookkeeping of Corollary 7.22 applies as well to the profiles of Theorem 6.1 and to the programs of Theorem 7.17.

**Corollary 7.26** (Max- $k$ -SAT and integer programs at growing density). *Let  $d = d(n) \geq 1$  satisfy  $\delta_n \rightarrow 0$ , and fix  $k \geq 1$  and  $s \geq 0$ .*

- (i) *For a  $k$ -CNF formula with  $n$  variables and at most  $d(n)n$  clauses with integer weights in  $[0, \Omega]$ , the profile (6.1), and with it the maximum satisfiable weight and the number of assignments attaining it, is computed in time  $\text{poly}(n, \Omega) 2^{O_k(n g(\delta_n))} = \text{poly}(\Omega) 2^{o(n)}$ .*
- (ii) *For a program as in Theorem 7.17 with at most  $d(n)n$  sparse rows, each involving at most  $k$  variables, the numbers  $p(j_0, \dots, j_s)$  of Theorem 7.17(i) are computed in time  $\text{poly}(\ell, (1 + 2n\Omega)^{s+2}) 2^{O_k(n g(\delta_n))}$ .*

*Proof.* (i) The factor system of the proof of Theorem 6.1 has at most  $d(n)n$  factors of scope at most  $k$ , with values in  $\mathcal{A}_N$ ,  $N = 1 + \sum_f \omega_f \leq 1 + d(n)n\Omega$ , and coefficients of at most  $k$  bits. Theorem 7.20 with  $c = d(n)$  applies once  $k\delta_n \leq \theta_1$  and gives  $\text{poly}(n) 2^{4800 n g(k\delta_n)}$  operations, with  $g(k\delta_n) \leq k g(\delta_n)$ ; Lemma 5.5 gives the time bound. (ii) The factor system of the proof of Theorem 7.17, through that of Theorem 7.1, consists of the at most  $d(n)n$  indicator factors of the sparse rows and the  $n$  factors  $\Psi_v$  of scope 1, at most  $2d(n)n$  factors of scope at most  $k$ , with values in an algebra of dimension at most  $(1 + 2n\Omega)^{s+2}$  and coefficients of at most  $k + 1$  bits. Theorem 7.20 with  $c = 2d(n)$  and Lemma 5.5 compute its sum, and the numbers  $p(j_0, \dots, j_s)$  are read off as in the proof of Theorem 7.17.  $\square$

*Remark 7.27.* The regime of Theorem 7.20 ends at  $ck$  of order  $\sqrt{\log n / \log \log n}$ . Below that order the exponent is  $o(n)$ ; at that order  $\theta_n$  is bounded away from 0 and (7.5) is no better than enumeration. The routine of Proposition 3.5 cannot reach quadratic density with any design at all. The balance condition (4.15) and  $M \leq bn$  give  $D^2 \leq M/(2(b+w)) < n/2$ , and  $D = r^{2Q} \geq 2^{2Q}$ , so every admissible design has  $Q < \frac{1}{4} \log n$ . Consequently  $q = \lfloor dR/Q \rfloor \geq \lfloor 4d/\log n \rfloor$ , and  $q \geq n$  as soon as the total degree satisfies  $d \geq \frac{1}{4}(n+1) \log n$ ; then  $C_q = \sum_{s=0}^n 2^s \binom{n}{s} = 3^n$ , and the Taylor count (3.17) alone is exponential. For the cut factors  $1 + (W-1)(Y_u + Y_v - 2Y_u Y_v)$  of a graph with  $m$  edges the product over the edges has degree exactly  $2m$  in the incidence indeterminates, its top coefficient being  $(-2)^m (W-1)^m \neq 0$  in  $\mathcal{A}_{m+1}$ , so the degree bound (3.1) cannot be lowered and a graph with  $m \geq \frac{1}{8}(n+1) \log n$  edges already gives  $q \geq n$ . Branching on  $h$  vertices costs  $2^h$  and removes at most  $hn$  edges, so for a graph with  $\Theta(n^2)$  edges any subexponential

branching leaves  $\Theta(n^2)$  edges. This is a limitation of the routine, not a lower bound for Max-Cut. The theorem is nonetheless a bound of the form  $2^{o(n)}$  for the exact maximum cut on graphs whose average degree tends to infinity, and Corollary 7.22 and Theorem 7.23 give one for classes of graphs with  $\Theta(n^2)$  edges, though not for almost every graph (Remark 7.25); the algorithm of Williams [Wil05] runs in time  $2^{\omega n/3}$  on every graph, and under the exponential time hypothesis no bound  $2^{o(n)}$  was possible even at bounded degree.

## 8. DENSE FORMULAS AND THE CLASS SNP

**8.1. Subexponential time and sparsification.** Following Impagliazzo, Paturi and Zane [IPZ01], a decision problem is considered together with a *complexity parameter*  $\mu(x) \geq 1$ , a polynomial-time computable integer function of the instance  $x$ ; for a formula it is the number of variables, for a graph the number of vertices. The class SE consists of the parameterised problems that admit, for every  $\varepsilon > 0$ , a deterministic algorithm running in time  $2^{\varepsilon \mu(x)} \text{poly}(|x|)$  on every instance  $x$ . The exponential time hypothesis is the statement that 3-SAT, with the number of variables as parameter, is not in SE.

We use the sparsification lemma in the following form.

**Theorem 8.1** (Sparsification lemma, [IPZ01]). *For every integer  $k \geq 2$  and every  $\varepsilon > 0$  there is a constant  $C = C(k, \varepsilon)$  and a deterministic algorithm that, given a  $k$ -CNF formula  $\varphi$  with  $n$  variables and  $m$  clauses, outputs in time  $2^{\varepsilon n} \text{poly}(n + m)$  a list of at most  $2^{\varepsilon n}$   $k$ -CNF formulas  $\varphi_1, \dots, \varphi_t$  on the same  $n$  variables, each with at most  $Cn$  clauses, such that an assignment satisfies  $\varphi$  if and only if it satisfies at least one  $\varphi_i$ . Quantitative forms of the dependence of  $C$  on  $k$  and  $\varepsilon$  are given in [CIP06].*

**Theorem 8.2** ( $k$ -SAT is in SE). *For every  $k \geq 2$ ,  $k$ -SAT with the number of variables as parameter lies in SE. Explicitly, for every  $\varepsilon > 0$  there is  $n_\varepsilon$  such that the satisfiability of a  $k$ -CNF formula with  $n \geq n_\varepsilon$  variables and  $m$  clauses is decided deterministically in time  $2^{\varepsilon n} \text{poly}(n + m)$ .*

*Proof.* Given  $\varepsilon$ , apply Theorem 8.1 with  $\varepsilon/2$  and  $C = C(k, \varepsilon/2)$ . For each  $\varphi_i$  of the list, apply Theorem 5.3 with  $\mathcal{A} = \mathbb{Q}$ ,  $c = C$  and the satisfaction indicators of the clauses as factors: it returns the number of satisfying assignments of  $\varphi_i$  in time  $\text{poly}(n) 2^{O_{k,C}(\rho(n))}$ . Since  $\rho(n) = o(n)$ , there is  $n_\varepsilon$  with  $\text{poly}(n) 2^{O_{k,C}(\rho(n))} \leq 2^{\varepsilon n/2}$  for  $n \geq n_\varepsilon$ . The formula  $\varphi$  is satisfiable if and only if one of the counts is positive, and the total time is  $2^{\varepsilon n/2} \cdot 2^{\varepsilon n/2} \text{poly}(n + m)$ .  $\square$

*Remark 8.3.* The decision of a dense formula through sparsification and the sparse routine loses the exact count, because the lists of satisfying assignments of the  $\varphi_i$  overlap. OpenAI obtain the exact count of every  $k$ -CNF, with the exponent  $n \log \log n / \sqrt{\log n}$ , through the counting-preserving sparsification of Dell, Husfeldt, Marx, Taslaman and Wahlén [DHMTW14, Appendix A] and their Sections 5 and 6; that is their Theorem 1.1, and it is not reproduced here. Theorem 8.2 is the statement needed below.

**8.2. Strict NP.** Papadimitriou and Yannakakis [PY91] define the class SNP of properties of finite relational structures expressible by a sentence

$$\exists S_1 \cdots \exists S_t \forall y_1 \cdots \forall y_r \psi(y_1, \dots, y_r; S_1, \dots, S_t; G_1, \dots, G_s), \quad (8.1)$$

where the  $G_j$  are the input relations on a universe  $U$  of size  $n$ , the  $S_i$  are relation variables of arities  $a_i$ , the  $y_j$  range over  $U$ , and  $\psi$  is quantifier-free, built from atoms  $S_i(\bar{y})$ ,  $G_j(\bar{y})$  and equalities between the  $y_j$  by Boolean connectives. The problem associated with (8.1) asks whether a given structure satisfies the sentence. Following [IPZ01] we take as complexity

parameter the number

$$\mu = \sum_{i=1}^t n^{a_i} \quad (8.2)$$

of Boolean choices in the second-order variables, which is  $tn$  when every  $S_i$  is unary. For  $k$ -SAT with the clauses given as relations, for  $q$ -colourability, for independent set, clique and vertex cover with a threshold, and for  $k$ -set cover, all arities are one and  $\mu$  is a constant multiple of  $n$ , so that membership in SE with respect to  $\mu$  is membership with respect to the natural parameter.

**Theorem 8.4** (SNP  $\subseteq$  SE). *Every problem in SNP, with the parameter (8.2), lies in SE.*

*Proof.* Fix the sentence (8.1) and let  $k \geq 2$  be the number of atoms  $S_i(\bar{y})$  occurring in  $\psi$ , counted with their tuples of variables, or 2 if that number is smaller. Given the structure, introduce a Boolean variable  $z_{i,\bar{u}}$  for every  $i \leq t$  and every  $\bar{u} \in U^{a_i}$ ; there are  $\mu$  of them. For each  $\bar{v} \in U^r$  the formula  $\psi(\bar{v}; S; G)$  becomes, after the input atoms and equalities are evaluated to constants and each atom  $S_i(\bar{w})$  with  $\bar{w}$  a tuple of entries of  $\bar{v}$  is replaced by  $z_{i,\bar{w}}$ , a Boolean formula  $\psi_{\bar{v}}$  in at most  $k$  of the variables  $z$ . Write  $\psi_{\bar{v}}$  as a conjunction of at most  $2^k$  clauses of width at most  $k$  by listing the falsifying assignments of its at most  $k$  variables. The conjunction  $\Phi$  of all  $\psi_{\bar{v}}$  over  $\bar{v} \in U^r$  is a  $k$ -CNF with  $\mu$  variables and at most  $2^k n^r$  clauses, computable in polynomial time, and the structure satisfies (8.1) exactly when  $\Phi$  is satisfiable: a choice of relations  $S_i$  is the same thing as an assignment of the  $z$ , and  $\forall \bar{y} \psi$  holds exactly when every  $\psi_{\bar{v}}$  is true. Theorem 8.2 decides  $\Phi$  in time  $2^{\varepsilon \mu} \text{poly}(\mu + 2^k n^r) = 2^{\varepsilon \mu} \text{poly}(n)$  for every  $\varepsilon > 0$  once  $\mu \geq n_\varepsilon$ , and the finitely many structures with  $\mu < n_\varepsilon$  are decided in constant time.  $\square$

**Corollary 8.5.** *For every  $\varepsilon > 0$  and every fixed  $k, q \geq 2$ , the satisfiability of a  $k$ -CNF formula with  $n$  variables and any number of clauses, and the  $q$ -colourability of a graph with  $n$  vertices, are decided deterministically in time  $2^{\varepsilon n} \text{poly}$ . More generally, every property expressed by a sentence (8.1) whose second-order variables are all unary is decided in time  $2^{\varepsilon tn} \text{poly}(n)$  for every  $\varepsilon > 0$ .*

*Proof.* Both are in SNP with unary second-order variables [PY91, IPZ01]. For  $k$ -SAT take the variables as the universe and, for each of the  $2^k$  sign patterns  $\sigma \in \{0, 1\}^k$ , a  $k$ -ary input relation  $G_\sigma$  containing the tuples of variables of the clauses with that pattern; the sentence is  $\exists T \forall y_1 \cdots y_k \bigwedge_\sigma (G_\sigma(\bar{y}) \rightarrow \bigvee_i (T(y_i) \leftrightarrow \sigma_i = 1))$ , with  $T$  the set of true variables, and  $\mu = n$ . For  $q$ -colourability take the vertices as the universe, the edge relation  $E$ , and  $\exists C_1 \cdots \exists C_q \forall x \forall y (\bigvee_i C_i(x) \wedge (E(x, y) \rightarrow \bigwedge_i \neg(C_i(x) \wedge C_i(y))))$ , with  $\mu = qn$ . The general statement is Theorem 8.4 with  $\mu = tn$ .  $\square$

Independent set, clique and vertex cover with a given threshold are not expressed by (8.1) without an auxiliary ordering, and we do not pursue that encoding: Corollary 6.8 already decides them for every graph with the explicit exponent  $n(\log \log n)^2 / (\log n)^{1/4}$ , which is sharper than membership in SE.

## 9. WHAT THE METHOD DOES NOT GIVE

The summation routine needs three things of a factor system: scopes of size  $\lambda^{O(1)}$ , occurrence numbers of size  $\lambda^{O(1)}$  after the branching of Lemma 5.2, and total degree  $O(n)$ . The following problems fail one of them, and this paper proves nothing about them.

- (a) *Dense objectives.* Max-Cut on a graph with  $\Theta(n^2)$  edges has total degree  $\Theta(n^2)$  and  $\Theta(n)$  occurrences per variable; branching on frequent variables fixes  $\Theta(n)$  of them. Theorem 7.20 reaches average degree  $o(\sqrt{\log n / \log \log n})$ , and by Remark 7.27 no design for the routine reaches total degree  $\frac{1}{4}(n+1) \log n$ . Theorem 7.23 and Corollary 7.24 cover the dense graphs

that differ in  $o(n\sqrt{\log n / \log \log n})$  pairs from a graph whose adjacency is a quadratic form in  $o(n/\log n)$  vertex weights, a blow-up of a graph on  $o(n/\log n)$  vertices for instance; by Remark 7.25 these are almost no graphs, and  $G(n, \frac{1}{2})$  is not among them. The chromatic and Tutte polynomials as polynomials in  $q$  need connectivity, which is not a bounded sum of local factors; Theorem 7.10 gives their values at each fixed positive integer  $q$  on bounded-degree graphs, and nothing off the hyperbolas  $H_q$ . Section 6.2 shows that the independence polynomial escapes the density limit because a vertex of large degree forces many others; no such forcing is available for a cut.

- (b) *Problems in P with quadratic conjectured complexity.* The lower bounds for Orthogonal Vectors, 3SUM, All-Pairs Shortest Paths and their relatives that were derived from the strong exponential time hypothesis lose their foundation, as OpenAI note; but the reductions that gave them run from CNF-SAT on  $n$  variables to instances of size  $2^{n/2}$ , and an algorithm with exponent  $o(n)$  for the former gives nothing below  $N^{2-o(1)}$  for the latter.
- (c) *P versus NP.* A polynomial-time reduction from an arbitrary problem in NP to 3-SAT produces formulas with  $n^C$  variables for some  $C > 1$ , and  $2^{o(n^C)}$  is not polynomial. Nothing in this paper bears on the question.

Each of the results of Sections 6, 7 and 8 has been stated with the exact hypothesis under which it is proved. The constants in the exponents are explicit through Lemma 4.7, Theorem 5.1 and Theorem 7.19, and the dependence on the weight bound  $\Omega$  is polynomial in every statement.

#### APPENDIX A. AN EXACT-ARITHMETIC CHECK

No statement of the paper depends on a computation. The program described here was written to confirm, on instances small enough to be summed directly, that the algebraic identities of Sections 3 and 4 hold verbatim over a truncated polynomial ring, and it is supplied with the archive [Bha26d].

The program builds a design with the parameters  $r = 2$ ,  $Q = 3$ ,  $R = 2$ ,  $A_0 = 0$ , far below the asymptotic regime of Lemma 4.7, which is permissible because the identity  $H(0) = Z$  of Proposition 3.5 needs only the five conditions of Definition 3.1 and not the size bounds. The auxiliary set has  $D = r^{2Q} = 64$  elements; the measurement tuples are drawn from  $\{0, \dots, R-1\}^{\tau n}$  with  $\tau n = 2$ ; the labels of Lemma 4.4 are chosen at random from a fixed seed and the conditions (4.15) and (3.4) are asserted, as is the fixed-point property of  $M_0$  and the normalisation  $P_u(0) = \mathbf{e}_u$  of Lemma 4.3. Every number is a rational in lowest terms, every element of  $\mathcal{A}_N$  a list of  $N$  rationals, and every element of  $\mathcal{A}_N[z]/(z^J)$  a list of  $J$  such lists; the Hermite system is solved by exact Gaussian elimination.

Seven sums were computed on  $n = 6$  variables and compared with the direct sum over the 64 assignments: the number of satisfying assignments of a random 2-CNF with seven clauses over  $\mathcal{A} = \mathbb{Q}$ ; the Max-2-SAT profile of the same formula over  $\mathbb{Q}[W]/(W^8)$ , which came out as  $4W^3 + 8W^4 + 28W^5 + 16W^6 + 8W^7$ ; the independence polynomial  $1 + 6W + 8W^2 + 2W^3$  of the hexagon with one chord, over  $\mathbb{Q}[W]/(W^7)$ , from the factor system of Theorem 6.7; the cut profile  $2 + 12W^2 + 18W^3 + 18W^4 + 12W^5 + 2W^7$  of the same graph, from the factor system of Corollary 6.4 with  $U = 1$ ; the weight enumerator  $1 + 4W^3 + 3W^4$  of the binary code of length 6 with the parity checks  $\{1, 2, 3\}$ ,  $\{3, 4, 5\}$ ,  $\{5, 6, 1\}$ , from the factor system of Theorem 7.12; the number 2 of nowhere-zero  $\mathbb{Z}_3$ -flows of  $K_{2,3}$ , from the factor system of Theorem 7.5(iv) over  $\mathbb{Q}$ , which agrees with Lemma 7.9 since the flow polynomial of  $K_{2,3}$  is  $(q-1)(q-2)$ ; and the objective profile  $2W + W^2 + 4W^3 + 3W^4 + 2W^5 + 3W^6 + W^8$  of the 0-1 program with the rows  $2x_1 - x_2 + 3x_3 \leq 3$ ,  $x_3 + x_4 + x_5 = 1$ ,  $x_1 + x_5 - x_6 \geq 0$  and the cost vector  $(1, 2, 1, 3, 1, 2)$ , from the factor system of Theorem 7.17, over  $\mathbb{Q}[W]/(W^{11})$ . All seven agreed with the direct sums.

The run took about twenty minutes in exact arithmetic, almost all of it in the sums (3.16) over the pairs  $(E, U)$ , whose number is the quantity  $C_q$  of (3.17) with  $q$  between 8 and 13.

A second program, also in exact integer and rational arithmetic and supplied with the archive, checks the finite statements behind Section 7.9: the inequality  $C_q q^q (n - q)^{n-q} \leq 2^q n^n$  of Lemma 3.4 for  $1 \leq n \leq 200$  and  $0 \leq q \leq n/2$ ; the identity (7.7) and the resulting transfer of cut profiles for every graph on at most six vertices, 33868 graphs, and for 200 weighted instances; the inequalities of the proof of Theorem 7.19 at 32 pairs  $(\lambda, a)$  with  $\lambda$  between  $2^{40}$  and  $2^{128}$  and  $\theta \leq 2^{-15}$ , where the asymptotic estimates are replaced by exact comparisons of integers; the degree  $2m$  and the top coefficient  $(-2)^m (W - 1)^m$  of Remark 7.27 for  $m \leq 5$ ; the two values of the footnote to the proof of Theorem 7.1; and the identity (7.8) with the profile read off the generating function of Theorem 7.23, on 300 random weighted instances, together with the representation of Corollary 7.24 for 200 perturbed blow-ups. All eight checks passed in a few seconds.

## DECLARATIONS

**Data and code availability.** The sources of this paper and the exact-arithmetic program of Appendix A, with its recorded output, are available at <https://github.com/creelie/eCon>, release v1.4.0, archived at [doi:10.5281/zenodo.23718560](https://doi.org/10.5281/zenodo.23718560).

**Funding.** No funding was received for conducting this study.

**Declaration of competing interest.** The author has no competing interests to declare that are relevant to the content of this article.

**Declaration of generative AI and AI-assisted technologies in the manuscript preparation process.** During the preparation of this work the author used Claude (Anthropic) to write and run the program of Appendix A and to draft text. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

## REFERENCES

- [AFLS01] K. Allemand, K. Fukuda, T. M. Liebling and E. Steiner, *A polynomial case of unconstrained zero-one quadratic optimization*, Math. Program. **91** (2001), 49–52. [doi:10.1007/s101070100233](https://doi.org/10.1007/s101070100233)
- [Bar82] F. Barahona, *On the computational complexity of Ising spin glass models*, J. Phys. A **15** (1982), no. 10, 3241–3253. [doi:10.1088/0305-4470/15/10/028](https://doi.org/10.1088/0305-4470/15/10/028)
- [BG05] A. Bulatov and M. Grohe, *The complexity of partition functions*, Theoret. Comput. Sci. **348** (2005), no. 2–3, 148–186. [doi:10.1016/j.tcs.2005.09.011](https://doi.org/10.1016/j.tcs.2005.09.011)
- [Bha26a] S. Singha Roy, D. Bhattacharjee, S. Samal, S. Nandi, N. Thakur, S. Sadhu, U. Bhattacharya, S. Dey, L. Frederick, S. Ghora and K. Patra, *Quantum algorithms on symplectic state manifolds*, SSRN preprint (2026). [doi:10.2139/ssrn.6114107](https://doi.org/10.2139/ssrn.6114107)
- [Bha26b] D. Bhattacharjee and U. Bhattacharya, *Iterated nerve complexes for k-fold sensor coverage*, Preprints.org (2026). [doi:10.20944/preprints202606.1351.v2](https://doi.org/10.20944/preprints202606.1351.v2)
- [Bha26c] D. Bhattacharjee, *Neuromorphic Grothendieck topoi for grid-embedded circuit architectures*, SSRN preprint (2026). [doi:10.2139/ssrn.6194939](https://doi.org/10.2139/ssrn.6194939)
- [Bha26d] D. Bhattacharjee, *Source and verification code for the paper Subexponential Weighted Counting and Exact Optimisation*, GitHub repository <https://github.com/creelie/eCon>, release v1.4.0 (2026). [doi:10.5281/zenodo.23718560](https://doi.org/10.5281/zenodo.23718560).
- [BMT78] E. R. Berlekamp, R. J. McEliece and H. C. A. van Tilborg, *On the inherent intractability of certain coding problems*, IEEE Trans. Inform. Theory **24** (1978), no. 3, 384–386. [doi:10.1109/TIT.1978.1055873](https://doi.org/10.1109/TIT.1978.1055873)
- [CFKL15] M. Cygan, F. V. Fomin, Ł. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk and S. Saurabh, *Parameterized Algorithms*, Springer, Cham, 2015. [doi:10.1007/978-3-319-21275-3](https://doi.org/10.1007/978-3-319-21275-3)

- [CIP06] C. Calabro, R. Impagliazzo and R. Paturi, *A duality between clause width and clause density for SAT*, in: 21st Annual IEEE Conference on Computational Complexity (CCC 2006), IEEE, 2006, pp. 252–260. doi:10.1109/CCC.2006.6
- [CKJ01] J. Chen, I. A. Kanj and W. Jia, *Vertex cover: further observations and further improvements*, J. Algorithms **41** (2001), no. 2, 280–301. doi:10.1006/jagm.2001.1186
- [CP15] M. Cygan and M. Pilipczuk, *Faster exponential-time algorithms in graphs of bounded average degree*, Inform. and Comput. **243** (2015), 75–85. doi:10.1016/j.ic.2014.12.007
- [DG00] M. Dyer and C. Greenhill, *The complexity of counting graph homomorphisms*, Random Structures Algorithms **17** (2000), no. 3–4, 260–289. doi:10.1002/1098-2418(200010/12)17:3/4<260::AID-RSA5>3.0.CO;2-W
- [DHMTW14] H. Dell, T. Husfeldt, D. Marx, N. Taslaman and M. Wahlén, *Exponential time complexity of the permanent and the Tutte polynomial*, ACM Trans. Algorithms **10** (2014), no. 4, Art. 21, 1–32. doi:10.1145/2635812
- [FK72] C. M. Fortuin and P. W. Kasteleyn, *On the random-cluster model. I. Introduction and relation to other models*, Physica **57** (1972), no. 4, 536–564. doi:10.1016/0031-8914(72)90045-6
- [Gal62] R. G. Gallager, *Low-density parity-check codes*, IRE Trans. Inform. Theory **8** (1962), no. 1, 21–28. doi:10.1109/TIT.1962.1057683
- [Gon85] T. F. Gonzalez, *Clustering to minimize the maximum intercluster distance*, Theoret. Comput. Sci. **38** (1985), 293–306. doi:10.1016/0304-3975(85)90224-5
- [Hoe63] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), no. 301, 13–30. doi:10.1080/01621459.1963.10500830
- [Hol81] I. Holyer, *The NP-completeness of edge-coloring*, SIAM J. Comput. **10** (1981), no. 4, 718–720. doi:10.1137/0210055
- [IP01] R. Impagliazzo and R. Paturi, *On the complexity of  $k$ -SAT*, J. Comput. System Sci. **62** (2001), no. 2, 367–375. doi:10.1006/jcss.2000.1727
- [IPZ01] R. Impagliazzo, R. Paturi and F. Zane, *Which problems have strongly exponential complexity?*, J. Comput. System Sci. **63** (2001), no. 4, 512–530. doi:10.1006/jcss.2001.1774
- [JN83] D. S. Johnson and K. A. Niemi, *On knapsacks, partitions, and a new dynamic programming technique for trees*, Math. Oper. Res. **8** (1983), no. 1, 1–14. doi:10.1287/moor.8.1.1
- [JS93] M. Jerrum and A. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, SIAM J. Comput. **22** (1993), no. 5, 1087–1116. doi:10.1137/0222066
- [JVV90] F. Jaeger, D. L. Vertigan and D. J. A. Welsh, *On the computational complexity of the Jones and Tutte polynomials*, Math. Proc. Cambridge Philos. Soc. **108** (1990), 35–53. doi:10.1017/S03050004100068936
- [MW96] M. Mihail and P. Winkler, *On the number of Eulerian orientations of a graph*, Algorithmica **16** (1996), no. 4–5, 402–414. doi:10.1007/BF01940872
- [NT75] G. L. Nemhauser and L. E. Trotter, Jr., *Vertex packings: structural properties and algorithms*, Math. Programming **8** (1975), 232–248. doi:10.1007/BF01580444
- [OAI26] OpenAI, *The exponential time hypothesis is false*, preprint, 25 October 2026, <https://github.com/openai/math/blob/main/preprints/The-Exponential-Time-Hypothesis-is-False-October-25-2026/paper.pdf>.
- [PS09] U. Pferschy and J. Schauer, *The knapsack problem with conflict graphs*, J. Graph Algorithms Appl. **13** (2009), no. 2, 233–249. doi:10.7155/jgaa.00186
- [PY91] C. H. Papadimitriou and M. Yannakakis, *Optimization, approximation, and complexity classes*, J. Comput. System Sci. **43** (1991), no. 3, 425–440. doi:10.1016/0022-0000(91)90023-X
- [Sok05] A. D. Sokal, *The multivariate Tutte polynomial (alias Potts model) for graphs and matroids*, in: Surveys in Combinatorics 2005, London Math. Soc. Lecture Note Ser. **327**, Cambridge Univ. Press, 2005, pp. 173–226. doi:10.1017/CBO9780511734885.009
- [SS09] A. D. Scott and G. B. Sorkin, *Polynomial constraint satisfaction problems, graph bisection, and the Ising partition function*, ACM Trans. Algorithms **5** (2009), no. 4, Art. 45. doi:10.1145/1597036.1597049
- [Tut54] W. T. Tutte, *A contribution to the theory of chromatic polynomials*, Canad. J. Math. **6** (1954), 80–91. doi:10.4153/CJM-1954-010-9
- [Val79] L. G. Valiant, *The complexity of enumeration and reliability problems*, SIAM J. Comput. **8** (1979), no. 3, 410–421. doi:10.1137/0208032
- [Var97] A. Vardy, *The intractability of computing the minimum distance of a code*, IEEE Trans. Inform. Theory **43** (1997), no. 6, 1757–1766. doi:10.1109/18.641542
- [Wil05] R. Williams, *A new algorithm for optimal 2-constraint satisfaction and its implications*, Theoret. Comput. Sci. **348** (2005), no. 2–3, 357–365. doi:10.1016/j.tcs.2005.09.023

- [XN17] M. Xiao and H. Nagamochi, *Exact algorithms for maximum independent set*, Inform. and Comput. **255** (2017), 126–146. doi:[10.1016/j.ic.2017.06.001](https://doi.org/10.1016/j.ic.2017.06.001)

FORMERLY, ELECTRO-GRAVITATIONAL SPACE PROPULSION LABORATORY (EGSPL), BHUBANESWAR, ODISHA 751030, INDIA

FORMERLY, UNIVERSITY OF CALCUTTA (HERAMBA CHANDRA COLLEGE), 23/49 GARIAHAT ROAD, GOLPARK, KOLKATA 700029, WEST BENGAL, INDIA

*Email address:* [d.bhattacharjee@erl-forschung.de](mailto:d.bhattacharjee@erl-forschung.de)

*Email address:* [itsdeep@live.com](mailto:itsdeep@live.com)

*Email address:* [egspl@proton.me](mailto:egspl@proton.me)

*URL:* <https://orcid.org/0000-0003-0466-750X>