Path Decompositions of Sparse Graphs via Gaussian Sweeps and Faster Exponential Algorithms

Serge Gaspers
UNSW Sydney, Australia
serge.gaspers@unsw.edu.au

Abstract

Using the bisection bound of Monien and Preis, Fomin and Høie proved in 2006 that every graph of maximum degree 3 on n vertices has pathwidth at most $n/6 + o(n)$. We prove that the pathwidth, bisection width and cutwidth of these graphs are at most $c_3\, n + o(n)$, where $c_3 = \frac{3}{4\pi} \arccos \frac{5}{6} \approx 0.1398$. For bisection width this improves the bound $\frac{n}{6} + o(n)$ of Monien and Preis (2006) and extends to all such graphs the bound that Lichev and Mitsche (2023) obtained for random cubic graphs. For maximum degree 4, we prove that bisection width and cutwidth are at most $\frac{n}{3} + o(n)$, improving the bound $\frac{2n}{5} + o(n)$ of Monien and Preis for 4-regular graphs, and that pathwidth is at most $0.2458n$, improving the bound $\frac{n}{3}$ of Fomin, Gaspers, Saurabh and Stepanov (2009). We show that every graph with m edges has pathwidth at most $\frac{7}{48}m + o(n)$, improving the bound $\frac{13}{75}m$ of Scott and Sorkin (2007), and we give several further bounds, none best for all graphs, of the form $\sum_{d \geq 3} a_d\, n_d + o(n)$ with coefficients $a_d < 1$, where n_d is the number of vertices of degree d. The path decompositions, and linear orderings of the stated cutwidth, can be computed in polynomial time.

Instead of deriving pathwidth bounds from bisection width, we give each vertex a Gaussian label, a weighted sum of independent standard normal variables, one for each vertex of a ball around it, so that adjacent vertices have strongly correlated labels. Sweeping a threshold through the labels gives a path decomposition whose bags are frontiers: the vertices above the threshold that have a neighbor below it. We bound the probability that a vertex lies in a frontier from above by comparison with an explicit Gaussian model of a vertex and its neighbors; a refinement of this sweep gives the width $c_3\, n + o(n)$ for maximum degree 3.

Faster exact algorithms follow. For graphs of bounded maximum or average degree, they include algorithms for feedback vertex set and the traveling salesperson problem. For general graphs, a branching algorithm that finishes sparse instances by dynamic programming becomes faster for minimum weight dominating set, and the bound in m gives faster algorithms for Max 2-CSP and for counting subgraphs with k edges.

Code and formal proofs: [Gas26], https://serggasp.github.io/pathwidth-site/.

Contents

1 Introduction **1**

2 Technical overview **5**

I Width bounds **12**

3 Short proof of a pathwidth bound below $\frac{n}{6}$ for cubic graphs **12**

4 Preliminaries **21**

5 Gaussian labels and edge correlations **27**

6 Star bounds **33**

7 The plain sweep: maximum degree 4 and 5 **38**

8 The median-switched sweep: maximum degree 3 and 4 **45**

9 Pathwidth sequences and the bootstrap **56**

10 Computing the decompositions **68**

II Algorithmic consequences **72**

11 Dynamic programming over the decompositions **73**

12 Separate, Measure and Conquer in polynomial space **80**

13 Counting dominating sets **80**

14 Connectivity problems **84**

15 Conclusion and open problems **86**

AI methodology statement **87**

A Deferred proofs **88**

A.1 Proof of the reduction lemma 88
A.2 The star-bound certificate 89
A.3 Maximum degree 4: proof of Theorem 8.2 and its certificates 98
A.4 Pathwidth sequences: proofs, and comparison with previous bounds 106
A.5 Proof of the algorithmic theorem 111
A.6 Computing the median-switched decompositions 133
A.7 Proofs for Part II 136

B Formal verification **143**

C Separate, Measure and Conquer: the analyses **147**

D Counting dominating sets: the analysis **172**

References **181**

1 Introduction

How large can the pathwidth of a sparse graph be? For graphs of maximum degree 3 the best upper bound has stood for twenty years: Fomin and Høie [FH06] proved that they have pathwidth at most $\frac{1}{6}n + o(n)$, where n is the number of vertices. We prove that graphs of maximum degree 3 have pathwidth at most

$$c_3\, n + o(n), \qquad c_3 = \frac{3}{4\pi} \arccos \frac{5}{6} \approx 0.13982,$$

with corresponding improvements at every degree, per edge, and in the running times of the exact algorithms built on these bounds.

The proof of the old bound cuts the graph into two halves along a bisection with about $\frac{n}{6}$ edges, which exists by the theorem of Monien and Preis [MP06, Theorem 1], and then decomposes each half with a prescribed boundary. Each of the two steps needs width about $\frac{n}{6}$. The per-degree bounds of Fomin, Gaspers, Saurabh and Stepanov [FGSS09] and of Edwards and McDermid [EM15] extend the bound to all graphs by deleting vertices of large degree, and the per-edge bound $\frac{13}{75}m$ of Scott and Sorkin [SS07], as well as the bound $\frac{m}{5.769}$ of Kneis, Mölle, Richter and Rossmanith [KMRR09], also follows from these per-degree bounds. Many exact algorithms for sparse instances run a dynamic program over a decomposition of such a width, and their running times are stated through these constants. We replace bisection by a different construction: a threshold sweep through a correlated Gaussian labeling of the vertices. Counting only the edges that cross a threshold of this sweep already gives pathwidth at most $\left(\frac{3}{2\pi} \arcsin \frac{1}{3}\right) n + o(n) \approx 0.1623n$, below $\frac{n}{6}$, with a short complete proof whose probabilistic tools are the random-hyperplane lemma of Goemans and Williamson [GW95] and Chebyshev's inequality (Theorem 3.1).

1.1 Width bounds

Throughout, graphs are finite and simple, $n = |V(G)|$, $m = |E(G)|$, $n_d = n_d(G)$ is the number of vertices of degree d, and $\Delta(G)$ is the maximum degree. Every remainder $o(n)$ (or $o(m)$) stands for a function of n (or m) alone, the same for all graphs in the stated class. Since every path decomposition is a tree decomposition, all bounds also hold for treewidth.

Theorem 1.1 (bounded degree)**.** *Let G be a graph with n vertices.*

(a) *If $\Delta(G) \le 3$, then $\mathrm{pw}(G) \le c_3\, n + o(n)$, where $c_3 = \frac{3}{4\pi} \arccos \frac{5}{6} \approx 0.13982$.*

(b) *If $\Delta(G) \le 4$, then $\mathrm{pw}(G) \le F_4\, n + o(n)$, where $F_4 \approx 0.24577$ is the maximum over the threshold of an explicit Gaussian integral, the ideal 4-star function* (2.2)*.*

Part (a) is Theorem 8.1, proved in Section 8, and part (b) is Theorem 7.1(a), proved in Section 7. The constant c_3 also bounds two other width parameters. The *bisection width* $\mathrm{bw}(G)$ is the minimum number of edges between the two sides of a partition of $V(G)$ into parts whose sizes differ by at most one, and the *cutwidth* $\mathrm{cw}(G)$ is the minimum, over all linear orderings of $V(G)$, of the maximum number of edges joining a vertex before a gap of the ordering to a vertex after it.

Corollary 1.2 (bisection width and cutwidth)**.** *Every graph G with $\Delta(G) \le 3$ satisfies $\mathrm{bw}(G) \le \mathrm{cw}(G) \le c_3\, n + o(n)$.*

The corollary follows from Theorem 1.1(a) through the chain $\mathrm{bw}(G) \le \mathrm{cw}(G) \le \mathrm{pw}(G) + 2$ for maximum degree 3: bisection width is at most cutwidth [DPS02, Lemma 2.4], cutwidth equals the search number for maximum degree 3 [MS89, Corollary 3.3], the search number is at most the vertex

separation number plus 2 [EST94, Theorem 2.1] (see also [Bod98]), and vertex separation equals pathwidth [Kin92, Theorem 3.1].

For bisection width, Corollary 1.2 improves the bound $\frac{n}{6} + o(n)$ that Monien and Preis [MP06, Theorem 1] proved for cubic graphs. For cutwidth, the previous bound $\frac{n}{6} + o(n)$ followed from the theorem of Fomin and Høie by the same relations. Corollary 10.2 computes, in deterministic time $n^{2+o(1)}$, an ordering of cutwidth at most $c_3 n + o(n)$ and a bisection with at most that many edges between its sides.

The constant c_3 is not new as a bisection constant: Lichev and Mitsche [LM23] proved that *random* cubic graphs have bisection width at most $c_3 n + o(n)$ asymptotically almost surely (stated there as $0.139822n$), and at least $0.103295n$. What is new is that the same constant holds for every graph of maximum degree 3, and for pathwidth. The lower bound for random cubic graphs gives pathwidth at least $0.103295n - 2$. We conjecture (Conjecture 15.1) the upper bound $\mathrm{pw}(G) \le (c_3 - 3p_{\mathrm{ch}})n + o(n) \approx 0.1322n + o(n)$ for cubic graphs.

The simplest form of our method, which counts the edges that cross a threshold, also bounds the cutwidth for larger maximum degree. Theorem 3.6 in Section 3 shows that every graph G with $\Delta(G) \le p$ has cutwidth at most $\gamma_p m + o(n)$, where $\gamma_p = \frac{1}{\pi} \arccos \frac{2\sqrt{p-1}}{p}$. For $\Delta(G) \le 4$ this is $\frac{m}{6} + o(n) \le \frac{n}{3} + o(n)$, which improves the second bound of Monien and Preis [MP06, Theorem 2], $\frac{2n}{5} + o(n)$ for the bisection width of 4-regular graphs; for $\Delta(G) \le 5$ it is at most $0.5121n + o(n)$. For $p \in \{4, 5\}$ the orderings can be computed in deterministic polynomial time (Theorem 10.3). The bound also gives the edge isoperimetric number (Section 3.6).

For graphs with vertices of all degrees we prove one bound per vertex, with a closed-form coefficient for every degree, and one bound per edge.

Theorem 1.3 (all graphs)**.** *Every graph G satisfies* $\mathrm{pw}(G) \le \sum_{d \ge 3} a_d\, n_d + o(n)$ *for the sequence*

$$a_d = \begin{cases} \dfrac{3d-4}{32} & \text{for } 3 \le d \le 6, \\ 1 - \dfrac{111}{28\,(d+1)} & \text{for } 7 \le d \le 10, \\ 1 - \dfrac{4}{d+1} & \text{for } d \ge 11. \end{cases}$$

The first coefficients are $(a_3, a_4, a_5, a_6) = (\frac{5}{32}, \frac{1}{4}, \frac{11}{32}, \frac{7}{16})$; rounded up to four decimals, $(a_3, \ldots, a_{11})$ is (0.1563, 0.2500, 0.3438, 0.4375, 0.5045, 0.5596, 0.6036, 0.6397, 0.6667).

Corollary 1.4 (edges)**.** *Every graph G satisfies* $\mathrm{pw}(G) \le \frac{7}{48}\, m + o(n)$.

Both are proved in Section 9. Table 1 compares the coefficients with the previous ones. The relative improvement is largest at degree 4: from $\frac{1}{3}$ to $\frac{1}{4}$ per vertex of degree 4, and to $F_4 \approx 0.2458$ for the whole class $\Delta \le 4$. The per-edge bound comes from the degree d that maximizes $\frac{2a_d}{d}$. For the previous coefficients this is $d = 5$, with $\frac{13}{75} \approx 0.1733$; for ours it is $d = 6$, with $\frac{2a_6}{6} = \frac{7}{48} \approx 0.1458$. For large degrees the coefficients of Theorem 1.3 approach 1 as $1 - \frac{4}{d+1}$, against $1 - \frac{3.5}{d+1} + O(d^{-3})$ for the coefficients $g_{1/6}(d)$ of Edwards and McDermid [EM15].

Theorem 1.3 is one of several bounds of this form. Section 9 states all coefficient vectors of the paper as *pathwidth sequences*, among them the pair $(c_3, 0.2627)$ for maximum degree 4 (Theorem 8.2), and extends each of them to all degrees by a bootstrap lemma. It also shows that the bootstrap of $(\frac{5}{32}, \frac{1}{4}, \frac{11}{32})$, which Theorem 1.3 rounds up to closed forms, has the smallest coefficient at every degree $d \ge 5$.

Table 1: Upper bounds on pathwidth, without the lower-order terms: per vertex for the classes $\Delta \le 3$ and $\Delta \le 4$ (Theorem 1.1), per vertex of degree d (Theorem 1.3) and per edge (Corollary 1.4). Previous bounds: [FH06] for $\Delta \le 3$; [FGSS09] for $\Delta \le 4$ and for degrees 3 to 6; the coefficients $g_{1/6}(d)$ of [EM15] for $d \ge 7$; and $\frac{13}{75}m$ from [SS07] ([KMRR09] state $\frac{m}{5.769}$). Decimals are rounded to four places. In Theorem 1.3 the coefficient of degree 3 exceeds c_3: the triple behind it trades a larger coefficient at degree 3 for smaller ones at degrees 4 and 5, and for $\Delta \le 3$ Theorem 1.1(a) applies.

	$\Delta \le 3$	$\Delta \le 4$	$d = 3$	$d = 4$	$d = 5$	$d = 6$	$d = 7$	$d \ge 11$	per edge
previous	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{13}{30}$	$\frac{23}{45}$	0.5698	$1 - \frac{3.5}{d+1} + O(d^{-3})$	$\frac{13}{75}$
this paper	0.1398	0.2458	$\frac{5}{32}$	$\frac{1}{4}$	$\frac{11}{32}$	$\frac{7}{16}$	0.5045	$1 - \frac{4}{d+1}$	$\frac{7}{48}$

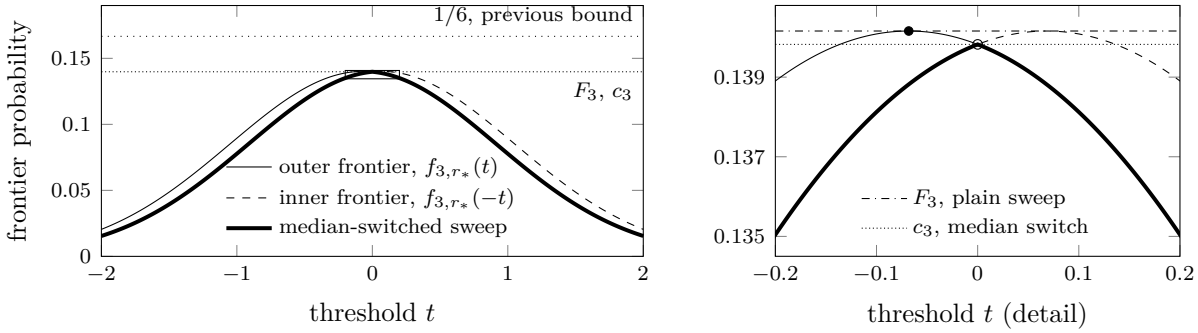


Figure 1: The ideal 3-star function f_{3,r_*} at the cubic correlation $r_* = \frac{2\sqrt{2}}{3}$, computed from the ideal star function (2.2) of Section 2, and its reflection. The right panel enlarges the box drawn in the left panel, and the dotted line at 1/6 is the previous bound for maximum degree 3 [FH06]. The plain sweep has width at most $F_3 n + o(n) \approx 0.1402n$, where F_3 is the maximum of the thin solid curve, attained at $t \approx -0.068$ (filled dot). The median-switched sweep keeps inner frontiers for $t < 0$ and outer frontiers for $t > 0$. It has width at most $c_3 n + o(n) \approx 0.1398n$, where c_3 is the maximum of the thick envelope, attained at $t = 0$ (open dot): by the switch lemma (Section 2.4), the passage from inner to outer frontiers at $t = 0$ adds nothing at leading order.

The method in one paragraph. We give every vertex a Gaussian label, a weighted sum of independent standard normal variables over a ball around the vertex with critically decaying weights, and sweep a threshold through the labels. The vertex separation layout turns the sweep into a path decomposition whose width is the size of the largest *frontier*, the set of vertices on one side of the threshold with a neighbor on the other side. Whether a vertex lies in a frontier depends only on the labels of its closed neighborhood, a *star*. We bound that probability from above by the one for an explicit *ideal* star, using that the regular simplex minimizes Gaussian orthant probabilities. A threshold has two frontiers: its *outer* frontier is the set of vertices above the threshold with a neighbor at or below it, and its *inner* frontier is the set of vertices at or below the threshold with a neighbor above it. For maximum degree 3, outer frontiers at every threshold give width about $0.1402n$; inner frontiers for thresholds below the median and outer frontiers above it lower the width to $c_3 n + o(n)$. Figure 1 shows the two frontier curves of the ideal star and the envelope that bounds the median-switched sweep; Section 2 develops these ideas.

1.2 Algorithms and consequences

The proofs of the width bounds are probabilistic. Section 10 makes them algorithmic: it approximates the Gaussian variables by a finite rational probability distribution and derandomizes by the method of conditional expectations. Theorem 10.1 there is the precise form of the following theorem and states the computational model.

Theorem 1.5 (computing the decompositions)**.** *Let $\beta(G)$ be any of the bounds of Theorems 1.1, 1.3, 7.1, 7.2, 8.2 and 9.2 and Corollaries 1.4 and 9.5 (for Theorem 7.1*(a)*, $F_4 n$), without its remainder term. For every graph G in the class of the bound, a path decomposition of G of width at most $\beta(G) + o(n)$ can be computed in deterministic time $n^{2+o(1)} + O(n(n+m))$, and by a Las Vegas algorithm in expected time $n^{1+o(1)} + O(n(n+m))$.*

Part II follows the new constants through the literature of exact algorithms. The five most striking consequences are the following, and Table 2 lists forty-five rows of bounds that, as far as we know, two recent preprints would not supersede (see the paragraph on the Exponential Time Hypothesis below). To our knowledge, each bound of Theorem 1.6 and Table 2 is the fastest known for its problem and graph class (for a deterministic bound, among deterministic algorithms), subject to the table notes.

Theorem 1.6 (faster exact algorithms)**.** *Polynomial factors are omitted; r is the domain size, n the number of variables or vertices, m the number of constraints or edges and k the pattern size. The following problems can be solved in the stated time.*

(i) *Minimum Weight Dominating Set and #Minimum Weight Dominating Set, for vertex weights with polynomially many distinct subset sums, in time $O(1.4815^n)$.*

(ii) *Max $(r,2)$-CSP in time $r^{7m/48+o(m)}$.*

(iii) *The number of subgraphs of a graph that are isomorphic to a given pattern with k edges, in time $k^{O(k)} n^{7k/48+o(k)}$.*

(iv) *Permutation Pattern Matching, in time $2^{O(k)} n^{253k/1215+o(k)}$, and k-opt Optimization for the traveling salesperson problem, in time $O(n^{(253/1215+\varepsilon_k)k})$ with $\varepsilon_k \to 0$ as $k \to \infty$.*

(v) *On graphs of maximum degree 4, Feedback Vertex Set by a Monte Carlo algorithm in time $O(1.2600^n)$ and deterministically in time $O(1.4853^n)$, and the traveling salesperson problem deterministically in time $O(1.4214^n)$; on graphs of maximum degree 3, Hamiltonian Cycle and Connected Dominating Set by Monte Carlo algorithms in time $O(1.1312^n)$ and $O(1.1661^n)$.*

Items (i) to (v) are Theorem 13.1, Theorem 11.3(2), Theorem 11.5, Theorem 11.8 and Theorem 14.1. Table 2 gives, for each bound, the previous bound, the work that proves it, and the theorem that proves the new one. Polynomial factors are omitted, and every exponent carries an additive $o(\cdot)$ term. The notation is that of Theorem 1.6; in addition, $\bar{d}$ is the average degree, and for Co-Path Packing and k-Co-Path Set, k is the number of deleted vertices or edges. The algorithms of Theorem 14.1 for connectivity and path problems are Monte Carlo algorithms unless marked det.

The Exponential Time Hypothesis. Two recent preprints, by OpenAI [Ope26a] and Anthropic [Cla26], claim to refute the Exponential Time Hypothesis (ETH) [IP01] by giving a subexponential-time algorithm for #3-SAT. As far as we know, they do not affect the bounds of Table 2. Part II contains seventeen further bounds that they would supersede; its opening lists them (twelve in Table 5 and five in polynomial space) and explains how the preprints reach them.

1.3 Evidence and verification

Theorem 8.1, which is Theorem 1.1(a), and Corollary 1.2 use no computation, and Theorem 8.1 has a complete formal proof in the Lean proof assistant, without hypotheses. Their written proofs use the simplex theorem for three Gaussian variables, which is a case of Mulgund's theorem [Mul26], a preprint whose Lean formalization we checked (Appendix B); for three variables it also has an elementary proof (Remark 6.6), which the formal proof of Theorem 8.1 uses. The other width bounds also use the cases of four and five variables, which rest on the same preprint; for four variables the theorem also follows from the four-cap theorem of Fejes Tóth [FT50, Bör04]. They rest in addition on computer-assisted certificates in ball arithmetic and exact rational arithmetic (Propositions 6.9 and 7.5, and two certificates for Theorem 8.2). Theorem 1.1(b), Theorems 1.3, 7.1, 7.2 and 8.2, Corollary 1.4 and Theorem 9.2 for its seeds other than C also have formal proofs in Lean, conditional on these inputs; for the seed C the formal proof is unconditional. Appendix B contains further details on the formal verification and states which result needs which input.

Throughout, "the accompanying code" refers to the programs that perform the computer-assisted steps; they and the Lean proofs are published on the companion website of this paper [Gas26].

1.4 Related work

Width bounds. Fomin, Grandoni and Kratsch [FGK05, §4.3.1] asked for the best possible upper bound on the pathwidth and treewidth of graphs of maximum degree 3, and Fomin and Høie [FH06, §5] and Fomin and Kratsch [FK10, notes to Chapter 5, p. 100] asked to reduce the gap between the upper bound $\frac{n}{6}$ of Fomin and Høie, the best previous one for maximum degree 3, and a lower bound $0.082n$. The question of the largest pathwidth of an n-vertex cubic graph is listed among the unsolved problems of Wikipedia [Wik26b, Wik26a, Wik26c]. Van Rooij asked at a Dagstuhl seminar for stronger bounds on the treewidth of graphs of bounded degree, with associated tree decompositions computable in polynomial time [FIK$^+$08]. He stated the cubic case as an open problem: to prove that cubic graphs have treewidth at most $c_t n + o(n)$ for some constant $c_t < \frac{1}{6}$, with tree decompositions of this width computable in polynomial time [vR11, Open Problem 14.2]. Theorems 1.1 and 1.5 solve Open Problem 14.2 with $c_t = c_3$, even for path decompositions.

Gaussian methods on sparse graphs. Thresholding Gaussian waves is a classical local algorithm on random regular graphs and graphs of large girth [CGHV15, Lyo17, HV15]. The cut density $\frac{3}{2\pi} \arccos \frac{2\sqrt{2}}{3}$ of a single threshold appears in Lyons [Lyo17], and the correlation sequence of our labels on the 3-regular tree is the extremal one of Backhausz, Szegedy and Virág [BSV15] for factor of i.i.d. processes. Lichev and Mitsche [LM23] used such waves, with a selection of helpful vertices, for the bisection width of random cubic graphs. Their analysis uses that random cubic graphs converge locally to the 3-regular tree, where the correlations of the wave are explicit. In an arbitrary graph, short cycles leave the correlation of two neighbors of a vertex unknown; the star comparison of Section 2.3 takes the place of the tree structure. Orderings by a Gaussian projection with bounds on prefix boundaries appear in the approximation algorithms of Feige [Fei00] and of Louis, Raghavendra and Vempala [LRV13], with non-constant-factor losses. We are not aware of any previous work that uses a threshold-by-threshold Gaussian layout to bound a width parameter of all graphs in a class, nor of the star comparison.

2 Technical overview

Our decompositions are of a different kind from the bisection-based ones of Section 1. Every vertex receives a real label, a threshold sweeps through the labels, and each bag is an outer frontier, the

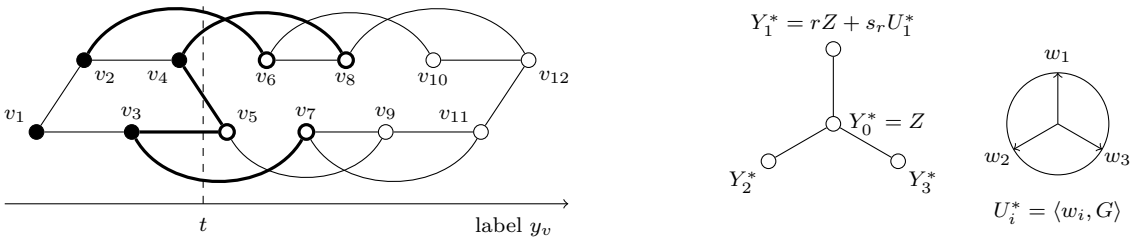


Figure 2: Left: a sweep. The vertices are placed at their labels; the dashed line is the threshold t, the filled vertices form S_t, the thick edges cross the threshold, and the vertices with a thick border form $\partial^+ S_t = \{v_5, v_6, v_7, v_8\}$, while $\partial^- S_t = \{v_2, v_3, v_4\}$. Right: the ideal 3-star of Section 2.3. The center Z and the planar standard Gaussian vector G are independent, w_1, w_2, w_3 are unit vectors at mutual angles $\frac{2\pi}{3}$, and $s_r = \sqrt{1 - r^2}$; the neighbors have correlation r with the center and sum to $3rZ$.

set of vertices above the threshold with a neighbor at or below it, together with the vertex being introduced. The analysis of such a sweep splits into a local question, how likely a given vertex is to lie on the frontier of the current threshold, and a global step: the frontier sizes concentrate around their means. This section explains the ideas without proofs, in the order of Part I, and points to the exact statements. Sections 2.1 and 2.2 describe the sweep and the labels, Section 2.3 bounds the probability that a vertex lies on a frontier, Section 2.4 improves the layout for maximum degree 3, Section 2.5 extends the bounds to mixed and large degrees, and Section 2.6 makes them algorithmic and explains how the widths become running times.

2.1 Sweeps and the layout lemma

Let $y\colon V(G) \to \mathbb{R}$ be an injective labeling. For a threshold $t \in \mathbb{R}$, the sublevel set $S_t = \{v : y_v \le t\}$ has an *outer frontier* $\partial^+ S_t$ and an *inner frontier* $\partial^- S_t$,

$$\partial^+ S_t = \{v :\ y_v > t,\ \exists u \in N(v) :\ y_u \le t\}, \qquad \partial^- S_t = \{u :\ y_u \le t,\ \exists v \in N(u) :\ y_v > t\}.$$

Every path decomposition in Part I is built from the layouts of the following lemma.

Lemma (layout; Lemma 4.1)**.** *For every injective labeling y of a graph G,* $\mathrm{pw}(G) \le \max_t |\partial^+ S_t|$ *and* $\mathrm{pw}(G) \le \max_t |\partial^- S_t|$.

The decomposition behind the lemma is the vertex separation layout of Kinnersley [Kin92], shown in Fig. 2: with $v_1, \dots, v_n$ the vertices in increasing order of their labels and $T_i = \{v_1, \dots, v_i\}$, its bags are $W_i = \{v_i\} \cup \partial^+ T_i$. So we need labels for which every frontier is small at every threshold. For random labels, $\mathbb{E}|\partial^+ S_t| = \sum_v \mathbb{P}(v \in \partial^+ S_t)$; if each label depends only on a bounded neighborhood, the count concentrates (Lemma 4.10), and a finite grid of thresholds covers all thresholds (Lemma 7.8). Everything then rests on the probability that one vertex lies on the frontier.

2.2 Gaussian labels with critical decay

A vertex lies on a frontier only if one of its edges crosses the threshold, that is, only if the two ends of that edge fall on different sides of it. If two neighbors had equal labels, their edge would never cross. For two standard normal labels with correlation ρ, the probability that they fall on different sides of zero is $\arccos(\rho)/\pi$ (Sheppard's formula, Lemma 4.9). So the labels should be as correlated

as possible across every edge, while each label keeps unit variance and depends only on a bounded neighborhood, so that labels of far-apart vertices are independent. We take independent standard normal variables $(Z_w)_{w\in V(G)}$ and, for a radius R, with $B_R(v) = \{w : \operatorname{dist}(v,w) \le R\}$ the ball of radius R around v,

$$X_v = \sum_{w\in B_R(v)} 2^{-\operatorname{dist}(v,w)/2}\, Z_w, \qquad Y_v = X_v/\sqrt{\operatorname{Var} X_v} \tag{2.1}$$

(Definition 5.2). For maximum degree p the base 2 is replaced by $p-1$.

Why this choice of weights? A label is a weighted sum of the variables around the vertex. Normalizing does not change correlations, so it suffices to look at the sums X_u, X_v of (2.1) for two neighbors u, v. They share the variables Z_w of every vertex w in both balls, and the two weights of such a Z_w differ by a factor at most $\sqrt{2}$, because $\operatorname{dist}(u,w)$ and $\operatorname{dist}(v,w)$ differ by at most 1. Weights with such a bounded ratio give the shared part a correlation at least $\operatorname{sech}(\frac{1}{2}\ln 2) = \frac{2\sqrt{2}}{3}$ (Lemma 5.7); with base $p-1$ the factor is $\sqrt{p-1}$ and the bound $\frac{2\sqrt{p-1}}{p}$. The only loss comes from the outer shell, the vertices in the ball of u but not in that of v, and the decay rate decides how much squared mass lies there. If the coefficients decay faster than $2^{-\operatorname{dist}/2}$, the squared mass sits near the center, each label is close to Z_v alone, and neighboring labels are nearly independent. If they decay more slowly, the outermost sphere of the ball dominates, and that is where the balls of u and v differ. At the *critical* rate $2^{-\operatorname{dist}/2}$, every sphere of a tree carries the same squared mass, so the squared mass is spread evenly over the R layers of the ball, and the outer shell carries only a fraction $O(1/R)$ of the variance (Eq. (5.2)). In a graph that is not a tree the spheres are smaller, which only helps; nothing is assumed about girth, regularity or connectivity.

Lemma (edge correlations; Lemma 5.4)**.** *Let G be a graph with $\Delta(G) \le 3$ and let (Y_v) be the labels (2.1) at radius R. Then every edge uv satisfies $\operatorname{Corr}(Y_u, Y_v) \ge r_*(1-\frac{1}{R})$, where $r_* = \frac{2\sqrt{2}}{3} \approx 0.9428$. For maximum degree p and base $p-1$ the bound is $\frac{2\sqrt{p-1}}{p}(1-\frac{1}{R})$.*

The constant is best possible for labels whose weights depend only on the distance: on the infinite p-regular tree, the local limit of p-regular graphs of large girth, no such *radial* Gaussian labeling has a neighbor correlation above $\frac{2\sqrt{p-1}}{p}$, and the labels approximate the *Gaussian wave*, which attains this value in the limit (Remark 5.1).

Vertices of degree at most 2 can be removed by series–parallel reductions at a cost of $O(\log n)$ in the pathwidth (Lemma 4.5), so it suffices to bound graphs with $\delta(G) \ge 3$. For such graphs with vertices of several degrees we use *mixed labels*: a step away from v through a vertex x multiplies the coefficient by $(d(x)-1)^{-1/2}$, so that the decay is critical at every vertex at once. An edge whose ends have degrees i and j then has correlation at least $r_{ij} = \operatorname{sech}(\frac{1}{4}\ln((i-1)(j-1)))$ up to a factor $1-o(1)$ (Lemma 5.5); $r_{33} = r_*$, $r_{44} = \frac{\sqrt{3}}{2}$ and $r_{55} = \frac{4}{5}$. For both kinds of labels, Y_v depends only on the variables in a ball of bounded radius around v, the labels are distinct with probability 1, and frontier counts have variance $O(n)$ (Lemmas 4.10 and 5.6). The radius R grows slowly with n, so that the loss $\frac{1}{R}$ vanishes while the dependence stays local.

2.3 The star comparison and the simplex theorem

This subsection bounds the probability that a single vertex lies on an outer frontier, using only the correlation bounds of Section 2.2. Summed over the vertices, these bounds give the plain sweep bounds at the end of the subsection, and the median switch and the charging certificate of the next two subsections build on them.

Fix a vertex v of degree d with neighbors $u_1, \dots, u_d$. The vector $(Y_v, Y_{u_1}, \dots, Y_{u_d})$ is a centered Gaussian vector with unit variances whose *incident* correlations $\mathrm{Corr}(Y_v, Y_{u_i})$, between the center and a neighbor, are at least known lower bounds r_i; we call it a *d-star with lower bounds r* (Definition 6.1), with center $Y_0 = Y_v$ and neighbors $Y_i = Y_{u_i}$. These lower bounds are all that the labels guarantee. The correlations $\mathrm{Corr}(Y_i, Y_j)$ between two neighbors are not controlled: short cycles through v may make them large or small. The vertex lies in $\partial^+ S_\tau$ exactly when the *frontier event* $\{Y_0 > \tau,\ \min_i Y_i \le \tau\}$ occurs, and we bound its probability from above in terms of r alone, in three steps.

Reduction to the center (Lemma 6.2). The first step removes the correlations between neighbors from the problem. The event $\{Y_0 > \tau\}$ is the disjoint union of the frontier event and the event that the whole star lies above τ. Let Φ be the standard normal distribution function. Since Y_0 is standard normal and its distribution is symmetric about 0, $\mathbb{P}(Y_0 > \tau) = 1 - \Phi(\tau) = \Phi(-\tau)$ for every $\tau \in \mathbb{R}$, positive or negative, and so

$$\mathbb{P}\big(Y_0 > \tau,\ \min_i Y_i \le \tau\big) = \Phi(-\tau) - \mathbb{P}\big(Y_k > \tau \text{ for } 0 \le k \le d\big).$$

It therefore suffices to bound the probability that the whole star lies above τ from below. Positively correlated variables tend to exceed a threshold together, so this probability should be smallest when the correlations are smallest. Slepian's inequality [Sle62] (Lemma 4.8) makes this precise: if two centered Gaussian vectors have the same variances and every covariance of the first is at least the corresponding covariance of the second, then all coordinates of the first exceed τ with at least the probability that all coordinates of the second do. So we lower each incident correlation $\rho_i = \mathrm{Corr}(Y_0, Y_i)$ to exactly r_i, without raising any other correlation. With fresh independent standard normal variables $\xi_1, \dots, \xi_d$, put

$$Y_0' = Y_0, \qquad Y_i' = Y_i - (\rho_i - r_i)\, Y_0 + \sqrt{\rho_i^2 - r_i^2}\ \xi_i \qquad (1 \le i \le d).$$

Subtracting a multiple of the center removes the excess correlation with the center, and the noise ξ_i restores unit variance. A direct computation gives $\mathrm{Var}\, Y_i' = 1$, $\mathrm{Corr}(Y_0', Y_i') = r_i$ and $\mathrm{Cov}(Y_i', Y_j') = \mathrm{Cov}(Y_i, Y_j) - \rho_i\rho_j + r_i r_j \le \mathrm{Cov}(Y_i, Y_j)$ for $i \ne j$. So the correlations between neighbors go down as well, whatever they were, and by Slepian's inequality the new star has at least the frontier probability of the old one.

The new star has a simple structure. Each neighbor is a multiple of the center plus an independent remainder: $Y_i' = r_i Y_0 + s_i U_i$, with $s_i = \sqrt{1 - r_i^2}$ and a centered Gaussian vector $U = (U_1, \dots, U_d)$ with unit variances, independent of Y_0. The incident correlations are now known exactly, and everything unknown about the star sits in the correlation matrix of U. Given $Y_0 = z$, the neighbor Y_i' lies above τ exactly when $U_i > a_i(z) = (\tau - r_i z)/s_i$. These thresholds decrease as z grows: a high center pulls its neighbors up. Integrating over the values $z > \tau$ of the center against its density φ, the standard normal density, and taking for each z the least favorable correlation matrix of U, gives

$$\mathbb{P}\big(Y_0 > \tau,\ \min_i Y_i \le \tau\big) \ \le\ \Phi(-\tau) - \int_\tau^\infty \varphi(z)\, P_{\min}\big(a(z)\big)\, dz, \qquad a_i(z) = \frac{\tau - r_i z}{s_i},$$

where $P_{\min}(a)$ is the infimum of $\mathbb{P}(U_i > a_i\ \forall i)$ over all correlation matrices of U. The right-hand side depends only on τ and r, and the next step evaluates $P_{\min}$.

The simplex theorem (Lemma 6.5). For equal thresholds $a_1 = \cdots = a_d = a$ and $d \le 5$ the infimum is attained by the most spread-out configuration, the *regular simplex vector* U^*: d standard normal variables with all pairwise correlations $-\frac{1}{d-1}$, the most negative common value possible, at which their sum is 0. This is a case of a recent theorem of Mulgund [Mul26], which proves a

conjecture of Balitskiy, Karasev and Tsigler [BKT17]; for three coordinates it also has an elementary proof, and for four it also follows from the four-cap theorem of L. Fejes Tóth [FT50, Bör04]. Write $S_d(a) = \mathbb{P}(U_i^* > a \ \forall i)$.

Ideal stars (Definition 6.7). The first two steps combine into one explicit worst case. Suppose that all lower bounds are equal to some r, as on a d-regular graph. Then all thresholds $a_i(z)$ are equal to $(\tau - rz)/s_r$, the simplex theorem gives $P_{\min}(a(z)) = S_d((\tau - rz)/s_r)$, and the bound of the first step becomes an explicit integral. This integral is not only a bound: it is the frontier probability of an actual star. Let Z be standard normal and U^* the regular simplex vector in d variables, independent of Z. The *ideal d-star* is $Y_0^* = Z$, $Y_i^* = rZ + s_r U_i^*$ (Fig. 2, right). Its incident correlations are exactly r, and its neighbors are spread out as far as these correlations allow: the remainders U_i^* sum to 0, so the neighbors sum to drZ, and any two neighbors have correlation $r^2 - s_r^2/(d-1) = (dr^2 - 1)/(d-1)$, the smallest common value possible. Given $Z = z$, all neighbors lie above τ with probability exactly $S_d((\tau - rz)/s_r)$, so the computation of the first step holds with equality for this star. Its frontier probability, the *ideal star function*, is therefore

$$f_{d,r}(\tau) = \mathbb{P}(Y_0^* > \tau,\ \min_i Y_i^* \le \tau) = \Phi(-\tau) - \int_\tau^\infty \varphi(z)\, S_d\Big(\frac{\tau - rz}{s_r}\Big) dz. \tag{2.2}$$

For $3 \le d \le 5$, every d-star whose lower bounds are all at least r has frontier probability at most $f_{d,r}(\tau)$ at every threshold τ, and so does every star with fewer than d neighbors and such lower bounds, because adding neighbors can only enlarge the frontier event (Corollary 6.8). As the ideal star is itself a d-star with lower bounds r, no better bound follows from the lower bounds alone.

The layout lemma takes the largest frontier over all thresholds, so what counts is the worst threshold. The suprema

$$F_3 = \sup_\tau f_{3,r_{33}}(\tau) \approx 0.1401583, \qquad F_4 = \sup_\tau f_{4,r_{44}}(\tau) \approx 0.2457720$$

are enclosed rigorously in ball arithmetic (Proposition 6.9); the first is attained near $\tau \approx -0.068$, the second near $\tau \approx -0.164$. For a graph of maximum degree 4, the labels with base 3 give every edge correlation at least $r_{44}(1 - \frac{1}{R})$, so every vertex lies in $\partial^+ S_\tau$ with probability at most $f_{4,r_{44}}(\tau) + o(1) \le F_4 + o(1)$. By linearity of expectation, $\mathbb{E}|\partial^+ S_\tau| \le F_4 n + o(n)$ at every threshold, and concentration, the finite grid of thresholds and the layout lemma turn this into a bound on the pathwidth. For maximum degree 3 the same argument with the labels (2.1) gives F_3. This is the *plain sweep*, the layout lemma applied to the Gaussian labels.

Theorem (plain sweep; Theorem 7.1(a) and the case $n_4 = 0$ of Theorem 7.1(b))**.** *Every graph with $\Delta(G) \le 4$ satisfies $\mathrm{pw}(G) \le F_4 n + o(n)$, and every graph with $\Delta(G) \le 3$ satisfies $\mathrm{pw}(G) \le F_3 n + o(n)$.*

This is where the plain sweep stops. For maximum degree 3, the coefficient F_3 is the best that any analysis of the outer frontiers can give, because it is essentially attained: on cubic graphs of large girth the stars of the labels are nearly ideal, so the outer frontier at the threshold $\tau \approx -0.068$ has about $F_3 n$ vertices. Improving the bound requires changing the layout, not the analysis.

2.4 The median switch

The change is suggested by the shape of $f_{3,r}$ (Fig. 1). At the median threshold 0 its value has a closed form. Since $Y_1^* + Y_2^* + Y_3^* = 3rY_0^*$, all three neighbors can be positive only when the center is. So $f_{3,r}(0)$ is one half minus the probability that three standard normal variables with pairwise correlation $b(r) = \frac{3r^2-1}{2}$ are all positive, and Sheppard's formula gives

$$f_{3,r}(0) = \kappa(r) := \frac{3}{4\pi} \arccos \frac{3r^2 - 1}{2}, \qquad \kappa(r_*) = \frac{3}{4\pi} \arccos \frac{5}{6} = c_3 \tag{2.3}$$

(Lemma 8.6). Moreover $f_{3,r}(t) \le f_{3,r}(0)$ for every $t \ge 0$ (Lemma 8.7). So $f_{3,r}$ exceeds its value at the median only below the median, and there we use the inner frontier instead. Since $-Y$ has the same distribution as Y (*sign symmetry*), the event $v \in \partial^- S_t$ for Y is the event $v \in \partial^+ S_{-t}$ for $-Y$, whose probability is at most $f_{3,r}(-t) \le \kappa(r)$ when $t \le 0$ (Proposition 8.8). A layout with inner frontiers below zero and outer frontiers above zero therefore has expected frontier size at most $\kappa(r)n$ at every nonzero threshold.

The difficulty lies at threshold zero, where the layout must pass from the inner frontier $U = \partial^- S_0$ to the outer frontier $W = \partial^+ S_0$ while every edge between S_0 and its complement stays covered. Putting $U \cup W$ into one bag would give a bag of about $2\kappa(r)n$ vertices. The *connector* does better: it adds the vertices of W one at a time, in an order given by a purely combinatorial lemma (Lemma 8.3), and drops a vertex of U as soon as all its neighbors above zero are present. Only the vertices whose three neighbors all lie on the other side of zero obstruct this exchange; let D be their set.

Lemma (median-switched layout; Lemma 8.4)**.** *Let $\Delta(G) \le 3$ and let y be an injective labeling. If $|\partial^- S_t| \le w$ for all $t < 0$, $|\partial^+ S_t| \le w$ for all $t > 0$, and $\max(|U| + |D \cap W|,\ |W| + |D \cap U|) + 1 \le w$, then $\mathrm{pw}(G) \le w$.*

For a vertex v of degree $d \le 3$, the events $v \in U$ and $v \in D \cap W$ are $\{Y_v \le 0 < \max_i Y_{u_i}\}$ and $\{d = 3,\ Y_v > 0 \ge \max_i Y_{u_i}\}$. The connector adds nothing at leading order because of the following lemma, whose right side is the same bound $\kappa(r)$ as at nonzero thresholds.

Lemma (switch lemma; Lemma 8.9)**.** *Let $(Y_0, Y_1, \dots, Y_d)$ be a d-star with all lower bounds equal to $r \in (0,1)$ and $d \le 3$. Then*

$$\mathbb{P}(Y_0 \le 0,\ \exists i :\ Y_i > 0) + \mathbf{1}_{d=3}\mathbb{P}(Y_0 > 0,\ Y_i \le 0 \text{ for all } i) \ \le\ \kappa(r).$$

At the ideal star the second event is impossible, since the neighbor labels sum to $3r$ times the center label, and the first has probability exactly $\kappa(r)$; a star that makes the second event possible loses at least as much in the first. After conditioning on the center, the proof reduces to showing that the probability that at least two of three standard normal residuals exceed a level $h > 0$ is smallest for the equilateral matrix of U^* (Lemma 8.11). This is not an orthant event, so Slepian's inequality does not apply; a deformation monitored by Plackett's identity (Lemma 4.7) reduces it to the planar case. Concentration and $R \to \infty$ then give the bound with constant $\kappa(r_*) = c_3$.

Theorem (Theorem 8.1)**.** *Every graph G with $\Delta(G) \le 3$ satisfies $\mathrm{pw}(G) \le c_3\, n + o(n)$.*

The gain over the plain sweep is small, $F_3 - c_3 \approx 3.4 \cdot 10^{-4}$. But the median switch replaces a numerically enclosed maximum by a closed form, and its proof uses no computation. The value c_3 is exactly what the median-switched sweep gives on cubic graphs of large girth, so any further improvement must change the rule at threshold zero (Section 15). With mixed labels, the same layout gives Theorem 8.2.

2.5 Mixed degrees, the charging certificate and the bootstrap

Degrees 3 to 5. With mixed labels, the lower bounds $r_{d(v)d(u)}$ of the star of v depend on the degrees of its neighbors. So the bound of the reduction to the center depends on the *type* $T = (d; k_3, k_4, k_5)$ of v, where d is its degree and k_j counts its neighbors of degree j. We write this bound $C_T(\tau)$; there are 46 types for degrees 3 to 5 (Definition 7.3). A vertex of degree 3 next to vertices of higher degree has weaker correlations than one in a cubic region, and its bound exceeds

the coefficient a_3 that we aim for. We move the excess along the edges: a vertex of degree i sends an amount t_{ij} along each edge to a neighbor of degree $j > i$. The transfers sum to zero over the graph (Lemma 7.4). Hence $\mathbb{E}|\partial^+ S_\tau| \le \sum_d a_d n_d$ whenever (a_3, a_4, a_5) is *admissible*, that is, whenever for every $\tau \in [-3, 3]$ some transfer vector (t_{34}, t_{35}, t_{45}) brings the transferred bound of every type below the coefficient of its degree. (Beyond $|\tau| = 3$ the frontier probabilities are below $5\Phi(-3) < 0.007$.)

We verify admissibility by a computer-assisted proof (Proposition 7.5). The thresholds are partitioned into cells. Each cell has one transfer vector and, for every type, a rational lower bound on the integral in $C_T(\tau)$, obtained in ball arithmetic. The final inequalities are checked in exact rational arithmetic with positive slack. This certifies the pathwidth sequences

$$\mathsf{P} = (F_3,\, 0.2575,\, 0.3667), \qquad \mathsf{R} = (0.1455,\, 0.2530,\, 0.3600), \qquad \mathsf{S} = (\tfrac{5}{32},\, \tfrac{1}{4},\, \tfrac{11}{32})$$

of Theorem 7.2; the bounds $(F_3, 0.2575)$ and $(\frac{5}{32}, \frac{1}{4})$ for maximum degree 4 (Theorem 7.1(b),(c)) are their restrictions to graphs without vertices of degree 5. The triple P is best when few vertices have degree 4 or 5, S when many do. All coefficient vectors of the paper are collected in Theorem 9.2.

All graphs. Beyond degree 5 no probability is involved. A combinatorial induction in the spirit of Fomin, Gaspers, Saurabh and Stepanov [FGSS09] and of Edwards and McDermid [EM15] deletes vertices of large degree one at a time (Proposition 9.10). Putting a deleted vertex back into every bag adds at most one to the width, and we compare this increase with the decrease of the potential $\sum_v b_{d(v)}$. Deleting a vertex w of maximum degree d removes the term b_d of w itself. It also lowers the degree of each neighbor by one, which lowers the potential by an increment $h_{d'} = b_{d'} - b_{d'-1}$, where d' is the degree of the neighbor; this increment is at least the running minimum $H_d = \min_{3 \le i \le d} h_i$. With $d-1$ neighbors of degree d and one of smaller degree, the decrease balances the increase exactly when $1 - b_d = (d-1)H_d + H_{d-1}$, and the two worst cases give the recurrence

$$b_d = \max\Big\{ \frac{1 + (d-1)\, b_{d-1} - H_{d-1}}{d},\; 1 - d\, H_{d-1} \Big\} \qquad (d > k) \tag{2.4}$$

(Eq. (9.1)). Vertices of degree at most 2 are removed by the classical series–parallel reductions after every deletion, and all reductions together add $O(\log n)$ to the pathwidth (Lemma 4.5).

The bootstrap lemma (Lemma 9.3) gives the coefficients b_d for $d > k$ from a base of maximum degree k by (2.4). When the increments are nonincreasing, the first term of the maximum is the larger one, and (2.4) is the recurrence $d\, b_d = (d-2)\, b_{d-1} + b_{d-2} + 1$ of Edwards and McDermid (Remark 9.4). The recurrence also explains why a better cubic constant alone improves nothing at higher degrees: with $k = 3$ and $b_3 = c_3 < \frac{1}{6}$ the second term is the larger one, and $b_4 = 1 - 4c_3 \approx 0.4407$, worse than the known $\frac{1}{3}$. Theorem 1.3 is therefore bootstrapped from the degree-5 base S: its coefficients are $\frac{3d-4}{32}$ up to degree 6, then $1 - \frac{111}{28(d+1)}$ up to degree 10 and $1 - \frac{4}{d+1}$ from degree 11 on. The edge bound follows because $2a_d/d \le \frac{7}{48}$ for every d, with equality only at $d = 6$, where $a_6 = \frac{7}{16}$ (Corollary 1.4); applying this to every component separately also gives a remainder $o(m)$ (Remark 9.14).

2.6 Algorithms and running times

Section 10 turns the existence proofs into algorithms. It replaces the Gaussian variables by a finite rational distribution coupled to the normal distribution, which changes every per-vertex probability of Part I by $o(1)$ uniformly in the threshold (Lemmas A.23 and A.24), and then fixes the variables one at a time by the method of conditional expectations [ES73, Rag88], applied to finitely many local counts (Lemma A.27). The radius grows only like $\frac{1}{8} \log_2 \log_2 n$, so each conditional expectation is computable by local enumeration in $n^{o(1)}$ time.

Part II turns the widths into running times in three ways. *Substitution*: a dynamic program with $c^{w+1}\,\mathrm{poly}(n)$ steps over a path decomposition of width $W(G)+o(n)$ runs in time $c^{W(G)+o(n)}\,\mathrm{poly}(n)$ (Lemma 11.1); this gives most bounds of Section 11 and, through Cut&Count, the rank-based approach and algorithms over orderings of small cutwidth, those of Section 14. *Rerunning a measure-and-conquer analysis*: in polynomial space the widths enter Separate, Measure and Conquer [GS17] through balanced separators of size cn, and the separator constant enters the mathematical program of the analysis in one or two places (Lemma C.5, Theorem C.3). *Moving a switch point*: with coefficient $\frac{1}{4}$ in place of $\frac{1}{3}$ at degree 4, the algorithm of Nederlof, van Rooij and van Dijk [NvRvD14] for counting dominating sets can hand the sets of size 4 to its dynamic program, which lowers its running time from 1.5002^n to 1.4815^n (Theorem 13.1).

Part I
Width bounds

3 Short proof of a pathwidth bound below $\frac{n}{6}$ for cubic graphs

This section proves, without using the rest of the paper, that a short argument improves on the bound $\frac{n}{6}$ for the pathwidth, the cutwidth and the bisection width of graphs of maximum degree 3. The algorithm behind it is simple: give every vertex a random real label, sort the vertices by label, and use the sorted order as a layout. The analysis counts, for every threshold t, the edges with one end labelled at most t and the other above t. The main results (Section 1) are stronger; the point here is a short proof. The reader is assumed to know graphs, not probability: the facts about random variables that we use are collected in Section 3.2, and the only estimate on the graph is a count of the vertices that the balls around the two ends of an edge share. Up to the end of Section 3.4 the section is self-contained; Sections 3.5 and 3.6 connect the argument to the rest of the paper. The later sections of Part I refine each step: the layout lemma, Sheppard's formula and the concentration bound of Section 4 are the general forms of Lemma 3.2, Lemma 3.3(a) and Lemma 3.5, and Section 5 generalizes the labels and Lemma 3.4.

Graphs are finite and simple. For a graph G, let $n=|V(G)|$ and $m=|E(G)|$, let $\Delta(G)$ be its maximum degree, $N(v)$ the set of neighbors of a vertex v and $N(X)=\bigcup_{v\in X}N(v)$, let $\mathrm{dist}(v,w)$ be the distance in G, let $B_R(v)=\{w:\mathrm{dist}(v,w)\le R\}$ be the ball of radius R around v and $\{w:\mathrm{dist}(v,w)=k\}$ its *sphere* of radius k, and for $X\subseteq V(G)$ let $e(X)$ be the number of edges with exactly one end in X. A *path decomposition* of G is a sequence of sets (*bags*) $W_1,\dots,W_s\subseteq V(G)$ such that every edge has both ends in some bag and every vertex lies in a nonempty set of consecutive bags; its width is $\max_i|W_i|-1$, and the *pathwidth* $\mathrm{pw}(G)$ is the minimum width of a path decomposition of G. The *cutwidth* $\mathrm{cw}(G)$ is the minimum, over all orderings $v_1,\dots,v_n$ of $V(G)$, of $\max_i e(\{v_1,\dots,v_i\})$, and the *bisection width* is the minimum of $e(X)$ over all $X\subseteq V(G)$ with $|X|=\lfloor n/2\rfloor$. Put

$$c_*=\frac{3}{2\pi}\arcsin\frac{1}{3}=\frac{3}{2\pi}\arccos\frac{2\sqrt{2}}{3}\approx 0.16226;$$

the two forms agree because an angle with cosine $\frac{2\sqrt{2}}{3}$ has sine $\sqrt{1-\frac{8}{9}}=\frac{1}{3}$. Numerically c_* is below $\frac{1}{6}$ by about 0.0044, and this can be checked exactly. The inequality $c_*<\frac{1}{6}$ means $\arcsin\frac{1}{3}<\frac{\pi}{9}$, that is, $\frac{1}{3}<s$ for $s=\sin\frac{\pi}{9}$. By the triple-angle formula, $3s-4s^3=\sin\frac{\pi}{3}=\frac{\sqrt{3}}{2}$. The function $h(x)=3x-4x^3$ is increasing on $[0,\frac{1}{2}]$, which contains $\frac{1}{3}$ and $s<\sin\frac{\pi}{6}$, and $h(\frac{1}{3})=\frac{23}{27}<\frac{\sqrt{3}}{2}$ because

Table 2: Running-time bounds of this paper that, as far as we know, the preprints [Ope26a, Cla26] that claim to refute ETH do not affect (Section 1.2). Letters refer to the notes that follow the table.

Problem	Previous	New	References
Feedback vertex sets			
Feedback Vertex Set, $\Delta \leq 4$	1.443^n	$2^{n/3} < 1.2600^n$	[CNP+11][f], Thm. 14.1
Feedback Vertex Set, $\Delta \leq 4$, det.	1.7117^n	$5^{F_4 n} < 1.4853^n$	[GL17], Thm. 14.1
Feedback Vertex Set, $\Delta \leq 5$	1.61^n	$2^{0.5121n} < 1.4262^n$	[CNP+11][f], Thm. 14.1
Feedback Vertex Set, $\Delta \leq 6$	1.6297^n	$3^{7n/16} < 1.6172^n$	[LN19][f], Thm. 14.1
Connected FVS and others[e], $\Delta \leq 3$	1.26^n	$4^{c_3 n} < 1.2139^n$	[CNP+11][f], Thm. 14.1
Connected FVS and others[e], $\Delta \leq 4$	1.588^n	$4^{F_4 n} < 1.4060^n$	[CNP+11][f], Thm. 14.1
Connected FVS and others[e], $\Delta \leq 5$	1.824^n	$4^{11n/32} < 1.6105^n$	[CNP+11][f], Thm. 14.1
Domination			
(#)Min. Weight Dominating Set[c]	1.5002^n	1.4815^n	[NvRvD14][a], Thm. 13.1
(#)Min. Weight Dominating Set, $\Delta \leq 3$	$3^{n/6} < 1.2010^n$	$3^{c_3 n} < 1.1661^n$	[FH06], Thm. 11.4
Connected Dominating Set, $\Delta \leq 3$	1.26^n	$3^{c_3 n} < 1.1661^n$	[CNP+11][f], Thm. 14.1
Connected Dominating Set, $\Delta \leq 4$	1.588^n	$4^{F_4 n} < 1.4060^n$	[CNP+11][f], Thm. 14.1
Connected Dominating Set, $\Delta \leq 5$	1.8139^n	$4^{11n/32} < 1.6105^n$	[BF12], Thm. 14.1
Independence and vertex covers			
(#)Max. Weight Indep. Set, $\Delta \leq 3$	$2^{n/6} < 1.1225^n$	$2^{c_3 n} < 1.1018^n$	[FH06], Thm. 11.4
(#)Max. Weight Indep. Set, $\Delta \leq 4$	$2^{n/3} < 1.2600^n$	$2^{F_4 n} < 1.1858^n$	[FGSS09], Thm. 11.4
Connected Vertex Cover, $\Delta \leq 4$	1.443^n	$2^{n/3} < 1.2600^n$	[CNP+11][f], Thm. 14.1
Connected Vertex Cover, $\Delta \leq 5$	1.61^n	$2^{0.5121n} < 1.4262^n$	[CNP+11][f], Thm. 14.1
Max 2-CSP			
Max $(r,2)$-CSP, all instances	$r^{13m/75} < r^{m/5.769}$	$r^{7m/48} < r^{m/6.857}$	[SS07], Thm. 11.3
Max $(r,2)$-CSP, $\Delta \leq 3$	$r^{n/6}$	$r^{c_3 n} < r^{0.13983n}$	[FH06], Thm. 11.3
Max $(r,2)$-CSP, $\Delta \leq 4$	$r^{n/3}$, $r^{m/6}$	$r^{F_4 n}$, $r^{\frac{3F_4}{2+16F_4}m}$	[FGSS09], [SS07], Thm. 11.3
Max $(r,2)$-CSP[b], avg. degree $\bar{d} \geq 7$	$r^{n(1-3.5/(\bar{d}+1)+O(\bar{d}^{-3}))}$	$r^{n(1-3.964/(\bar{d}+1))}$	[EM15], Thm. 11.3
Max $(r,2)$-CSP[b], avg. degree $\bar{d} \geq 11$	$r^{n(1-3.5/(\bar{d}+1)+O(\bar{d}^{-3}))}$	$r^{n(1-4/(\bar{d}+1))}$	[EM15], Thm. 11.3
Patterns and packings			
Counting k-edge subgraphs[d]	$n^{0.174k}$	$n^{(7/48)k} < n^{0.1459k}$	[CDM17], Thm. 11.5
Counting k-paths, k-cycles[d]	$n^{0.1704k}$	$n^{0.1459k}$	[BIJK19a], Thm. 11.5
Permutation Pattern Matching	$n^{k/4}$	$2^{O(k)} n^{(253/1215)k}$	[BKM21], Thm. 11.8
k-opt for TSP, general graphs	$n^{(1/4+\varepsilon_k)k}$	$n^{(253/1215+\varepsilon_k)k}$	[CKS19], Thm. 11.8
k-opt for TSP, bounded degree	$n^{(23/135+\varepsilon_k)k}$	$n^{(7/48+\varepsilon_k)k}$	[BIJK19b], Thm. 11.8
Co-Path Packing	2.9241^k	2.8192^k	[LX26], Thm. 14.1
k-Co-Path Set	1.588^k	1.5423^k	[SvdP16], Thm. 14.1
Cycles and paths			
Hamiltonian Cycle, $\Delta \leq 3$	1.1583^n	1.1312^n	[CKN12][f], Thm. 14.1
Hamiltonian Cycle, $\Delta \leq 4$	1.5058^n	1.3416^n	[CKN12][f,g], Thm. 14.1
Hamiltonian Cycle, $\Delta \leq 5$	1.6246^n	1.5252^n	[BKKZ15][f], Thm. 14.1
Hamiltonian Cycle, avg. degree d	1.12^{dn}	1.0937^{dn}	[KM20], Thm. 14.1
TSP, $\Delta \leq 3$, det.	1.2128^n	1.1757^n	[BCKN15], Thm. 14.1
TSP, $\Delta \leq 4$, det.	1.6110^n	1.4214^n	[BCKN15][g], Thm. 14.1
TSP, $\Delta \leq 5$, det.	1.8588^n	1.6352^n	[BCKN15][g], Thm. 14.1
TSP, $\Delta \leq 6$, det.	1.9672^n	1.8699^n	[BHKK12], Thm. 14.1
TSP, avg. degree d, det.	1.1321^{dn}	1.1100^{dn}	[KM20], Thm. 14.1
Longest Path, $\Delta \leq 3$	1.26^n	1.1874^n	[CNP+11][f], Thm. 14.1
Longest Path, $\Delta \leq 4$	1.5875^n	1.3523^n	[BKKZ15][f], Thm. 14.1
Longest Path, $\Delta \leq 5$	1.6246^n	1.5252^n	[BKKZ15][f], Thm. 14.1
Other connectivity problems			
Steiner Tree, $\Delta \leq 3$	1.201^n	$3^{c_3 n} < 1.1661^n$	[CNP+11][f], Thm. 14.1
Steiner Tree, $\Delta \leq 4$	1.3533^n	$3^{F_4 n} < 1.3100^n$	[FGK+13], Thm. 14.1
Max Internal Spanning Tree, $\Delta \leq 6$	1.8423^n	$4^{7n/16} < 1.8341^n$	[BKKZ15][f], Thm. 14.1
Directed problems			
Directed problems[h], $\Delta \leq 3$	1.349^n	$6^{c_3 n} < 1.2848^n$	[CNP+11][f], Thm. 14.1
Directed problems[h], $\Delta \leq 4$	1.818^n	$6^{F_4 n} < 1.5533^n$	[CNP+11][f], Thm. 14.1

Notes to Table 2. A letter after a citation concerns the previous bound, and a letter after a problem concerns the whole row. [a] A gap in the analysis of this algorithm allowed us only to reconstruct a running time of $O(1.5004^n)$ (Appendix D.5). [b] A record only for moderate $\bar{d}$; for large $\bar{d}$, the algorithms of Williams [Wil05], for integer scores of polynomial size, and of Golovnev and Kutzkov [GK14], for Boolean domains, are faster (Theorem 11.3). [c] When the set of possible weight sums has polynomial size, as in [NvRvD14, Cor. 32] (minimum weight) and [vR11, Cor. 8.14] (counting). [d] Both bounds carry a further factor $k^{O(k)}$. [e] The problems of Theorem 14.1(b) other than Connected Dominating Set: Connected Feedback Vertex Set, Connected Odd Cycle Transversal, Exact k-Leaf Spanning Tree, Maximum Full Degree Spanning Tree, Graph Metric Traveling Salesperson, and undirected Min Cycle Cover and Longest Cycle; Maximum Leaf Spanning Tree also has the bounds of Connected Dominating Set. [f] By a Monte Carlo algorithm. [g] The cited algorithm, run with the pathwidth bounds of [FGSS09], and for TSP with $\omega \leq \frac{9}{4}$. [h] The directed versions of Min Cycle Cover, Longest Path and Longest Cycle, and Exact k-Leaf Outbranching; Δ is the maximum degree of the underlying undirected graph.
The bounds for TSP, previous and new, are evaluated with the matrix multiplication exponent $\omega \leq \frac{9}{4}$ [Ope26b, AvdBZ26] (Section 14).

$46^2 = 2116 < 2187 = 3 \cdot 27^2$. Hence $\frac{1}{3} < s$, and $c_* < \frac{1}{6}$.

Theorem 3.1 (cut sweep)**.** *Every graph G with $\Delta(G) \leq 3$ satisfies* $\mathrm{pw}(G) \leq \mathrm{cw}(G) \leq c_* n + o(n)$.

By considering a set X formed by the first $\lfloor n/2 \rfloor$ vertices of an ordering, we get that the bisection width is at most $c_* n + o(n)$ as well; for all three parameters the previous bound was $\frac{n}{6} + o(n)$ [MP06, Theorem 1][FH06]. Theorem 1.1 and Corollary 1.2 lower all three to $c_3 n + o(n)$ with $c_3 \approx 0.1398$.

3.1 The idea

Every vertex w draws a real number Z_w from the standard normal distribution $N(0,1)$, with mean 0 and variance 1, independently of all other vertices, and every vertex v receives the label Y_v, a weighted sum of the numbers Z_w over the ball of radius R around v, with coefficients that decrease with $\mathrm{dist}(v,w)$. We order the vertices by label. Every prefix of this order is a set $S_t = \{v : Y_v \leq t\}$ for some threshold t, so the cutwidth of the order is the largest number of edges that *cross* a threshold, that is, have one end labelled at most t and the other above t.

The two ends of an edge uv see almost the same ball, so Y_u and Y_v are sums of almost the same noise. Their *correlation* $\rho = \mathrm{Corr}(Y_u, Y_v)$ measures this: for such sums it is the overlap of the two coefficient sequences, normalized so that identical sequences have $\rho = 1$ (Section 3.2). So the ends of an edge have ρ close to 1. Two standard normal variables with correlation ρ lie on opposite sides of a threshold with probability at most $\arccos(\rho)/\pi$, which is small when ρ is close to 1 (about $\sqrt{2(1-\rho)}/\pi$). Summing over the at most $\frac{3}{2}n$ edges, the expected number of edges crossing any fixed threshold is at most $\frac{3}{2\pi}\arccos(\rho)\, n$. The bound $\frac{n}{6}$ corresponds to $\arccos\rho = \frac{\pi}{9}$, since $\frac{3}{2\pi} \cdot \frac{\pi}{9} = \frac{1}{6}$. The coefficients of Section 3.3, for a radius R, give every edge of every graph of maximum degree 3 correlation at least $\frac{2\sqrt{2}}{3}(1 - \frac{1}{R})$, and $\arccos\frac{2\sqrt{2}}{3} = \arcsin\frac{1}{3} < \frac{\pi}{9}$.

A bound on the expected number of crossing edges at each threshold is not yet a bound for one order. We need a single outcome of the random numbers in which *every* threshold is crossed by few edges. Since every label depends only on the numbers Z_w in a bounded ball, the count at a fixed threshold is a sum of many terms that are mostly independent of each other, so it stays close to its expectation with probability close to 1. We apply this to finitely many thresholds on a grid, which control all the others, and conclude that some outcome is good at all of them at once.

One may think of Z_w as a fair coin flip of ± 1; for large balls the labels are then approximately normal, and the normal variables make the probability calculation exact.

The proof has four steps.

(1) An order of the vertices gives a path decomposition whose width is at most the largest number of edges leaving a prefix (Lemma 3.2).

(2) Each edge crosses each threshold with probability at most $\arccos(\rho)/\pi$ (Lemma 3.3), where $\rho \geq \frac{2\sqrt{2}}{3}(1-\frac{1}{R})$ (Lemma 3.4). So the expected number of crossing edges is at most $(c_* + O(\frac{1}{R}))n$ at every threshold, which is below $\frac{n}{6}$ for R large.

(3) Each count is close to its expectation with high probability (Lemma 3.5 and Chebyshev's inequality), and a grid of thresholds reduces all thresholds to finitely many counts.

(4) The probability that some count on the grid is far from its expectation is less than 1, so some outcome of the random numbers gives an order of small cutwidth.

3.2 Facts from probability

This section lists everything about random variables that the proof uses. All random variables in the proof are functions of finitely many independent standard normal variables $(Z_w)_{w \in W}$; in Section 3.4, $W = V(G)$. For a random variable X with expectation $\mathbb{E}X$, the *variance* is $\operatorname{Var} X = \mathbb{E}(X - \mathbb{E}X)^2$. For two random variables the *covariance* is $\operatorname{Cov}(X, X') = \mathbb{E}[(X - \mathbb{E}X)(X' - \mathbb{E}X')]$. If both variances are positive, the *correlation* is $\operatorname{Corr}(X, X') = \operatorname{Cov}(X, X')/\sqrt{\operatorname{Var} X \operatorname{Var} X'}$, a number in $[-1, 1]$.

(P1) *Expectations and variances.* $\mathbb{E}(X + X') = \mathbb{E}X + \mathbb{E}X'$ for all random variables, independent or not (linearity of expectation). Also $\operatorname{Var}(\sum_i X_i) = \sum_{i,j} \operatorname{Cov}(X_i, X_j)$, and $|\operatorname{Cov}(X, X')| \leq \sqrt{\operatorname{Var} X \operatorname{Var} X'}$ (the Cauchy–Schwarz inequality). Functions of disjoint sets of the Z_w are independent, and independent variables have covariance 0.

(P2) *Linear sums.* For real coefficients (a_w), the sum $\sum_w a_w Z_w$ is normal with mean 0 and variance $\sum_w a_w^2$. If this variance s^2 is positive, the sum has the density $\frac{1}{s\sqrt{2\pi}} e^{-x^2/(2s^2)}$, which is symmetric and decreasing in $|x|$; in particular the sum takes any given value with probability 0. The covariance of $\sum_w a_w Z_w$ and $\sum_w b_w Z_w$ is $\sum_w a_w b_w$.

(P3) *Orthogonal sums are independent.* If $\sum_w a_w b_w = 0$, then $\sum_w a_w Z_w$ and $\sum_w b_w Z_w$ are independent. This is special to sums of independent normal variables. For general random variables, covariance 0 does not imply independence.

(P4) *Rotation invariance.* If G_1, G_2 are independent and standard normal, then the direction of the point (G_1, G_2) of the plane is uniformly distributed on the unit circle, because the joint density $\frac{1}{2\pi} e^{-(x^2+y^2)/2}$ depends only on the distance from the origin.

(P5) *Chebyshev's inequality.* $\mathbb{P}(|X - \mathbb{E}X| \geq \lambda) \leq \operatorname{Var} X/\lambda^2$ for every $\lambda > 0$.

(P6) *Union bound.* The probability that at least one of several events happens is at most the sum of their probabilities. So if these probabilities add up to less than 1, some outcome of the random numbers avoids all of the events (the probabilistic method).

We write $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ for the standard normal density and $\Phi(x) = \int_{-\infty}^{x} \varphi$ for the standard normal distribution function. Since $\Phi(-3) < 0.0014$, a standard normal variable lies outside $[-3, 3]$ with probability less than 0.003.

3.3 Four lemmas

The first lemma turns an ordering into a path decomposition; it is the easy direction of Kinnersley's theorem that pathwidth equals vertex separation [Kin92].

Lemma 3.2 (orderings)**.** *Let G be a graph, let $v_1, \dots, v_n$ be an ordering of $V(G)$, and let $T_i = \{v_1, \dots, v_i\}$. Then* $\mathrm{pw}(G) \le \max_i e(T_i)$*. In particular* $\mathrm{pw}(G) \le \mathrm{cw}(G)$*.*

Proof. Let $W_i = \{v_i\} \cup (N(T_i) \setminus T_i)$ for $1 \le i \le n$, and let f_j be the index of the first neighbor of v_j among $v_1, \dots, v_{j-1}$, or $f_j = j$ if there is none; that is, $f_j = \min(\{j\} \cup \{k < j : v_k \in N(v_j)\})$. For $i < j$, $v_j \in W_i$ if and only if v_j has a neighbor in T_i, that is, $f_j \le i$; also $v_j \in W_j$, and $v_j \notin W_i$ for $i > j$ because then $v_j \in T_i \setminus \{v_i\}$. So v_j lies exactly in the consecutive bags $W_{f_j}, \dots, W_j$. An edge $v_a v_b$ with $a < b$ has both ends in W_a, because $v_b \notin T_a$ has the neighbor $v_a \in T_a$. So the bags form a path decomposition. Its width is $\max_i |N(T_i) \setminus T_i|$, since $v_i \in T_i$, and $|N(T_i) \setminus T_i| \le e(T_i)$, since every vertex of $N(T_i) \setminus T_i$ is the end of an edge leaving T_i. Applied to an ordering of minimum cutwidth, this gives $\mathrm{pw}(G) \le \mathrm{cw}(G)$. □

The second lemma is the only fact about normal variables that we need beyond Section 3.2. Part (a) is Sheppard's formula [She99], known in algorithms as the random-hyperplane lemma of Goemans and Williamson [GW95]. Part (b) says that the median 0 is the worst threshold: the two variables straddle t most often where their density is highest.

Lemma 3.3 (two Gaussians)**.** *Let $(Z_w)_{w \in W}$ be independent standard normal variables, let $U = \sum_w a_w Z_w$ and $V = \sum_w b_w Z_w$ with $\sum_w a_w^2 = \sum_w b_w^2 = 1$, and let $\rho = \sum_w a_w b_w$, which is the correlation of U and V by* (P2)*. If $\rho > -1$, then*

(a) $\mathbb{P}(UV \le 0) = \arccos(\rho)/\pi$*;*

(b) $\mathbb{P}\big(\min(U,V) \le t < \max(U,V)\big) \le \arccos(\rho)/\pi$ *for every* $t \in \mathbb{R}$*.*

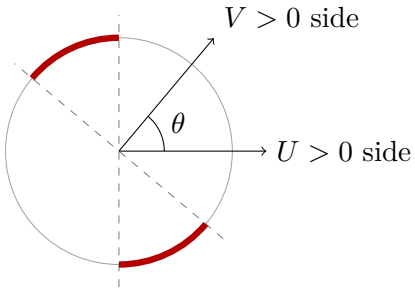


Figure 3: The proof of Lemma 3.3(a). The point (G_1, G_2) has a uniformly random direction. $U > 0$ when this direction lies in the half-circle to the right of the vertical dashed line, and $V > 0$ when it lies in the same half-circle rotated by $\theta = \arccos \rho$. The signs differ exactly on the two thick arcs, each of length θ.

Proof. If $\rho = 1$, then $\sum_w (a_w - b_w)^2 = 2 - 2\rho = 0$, so $U = V$, $\mathbb{P}(UV \le 0) = \mathbb{P}(U = 0) = 0$ by (P2), and the event of (b) is empty. Let $\rho < 1$ and $\theta = \arccos \rho \in (0, \pi)$.

(a) We write V in terms of U and a second, independent normal variable. Put $a'_w = (b_w - \rho a_w)/\sin\theta$. Then $\sum_w a_w a'_w = 0$ and $\sum_w a'^2_w = (1 - 2\rho^2 + \rho^2)/\sin^2\theta = 1$, so $G_1 = U$ and $G_2 = \sum_w a'_w Z_w$ are standard normal (P2) and independent (P3), and $V = \cos\theta\, G_1 + \sin\theta\, G_2$. Write the point (G_1, G_2) as $r(\cos\phi, \sin\phi)$ with $r \ge 0$. By (P4) the angle ϕ is uniform on the circle, and $U = r\cos\phi$ and $V = r\cos(\phi - \theta)$. Ignoring the outcomes where $U = 0$ or $V = 0$, which have probability 0 by (P2), $U > 0$ exactly when ϕ lies in the open half-circle centred at angle 0, and $V > 0$ exactly when ϕ lies in the half-circle centred at angle θ, which is the first one rotated by θ.

So $UV < 0$ exactly when ϕ lies in one of the two half-circles but not the other. As Fig. 3 shows,[1] this set consists of two arcs of length θ, so $\mathbb{P}(UV \le 0) = 2\theta/(2\pi) = \arccos(\rho)/\pi$.

(b) We compare the threshold t with the threshold 0, in four steps.

Step 1. Let $S = \frac{U+V}{2}$ and $D = \frac{U-V}{2}$. These are sums of the Z_w with coefficients $\frac{a_w+b_w}{2}$ and $\frac{a_w-b_w}{2}$, and $\sum_w \frac{(a_w+b_w)(a_w-b_w)}{4} = \frac{1}{4}(\sum_w a_w^2 - \sum_w b_w^2) = 0$. So S and D are independent by (P3). Also $\operatorname{Var} S = \frac{1+\rho}{2} > 0$, so S has a density f that is symmetric and decreasing in $|s|$ (P2).

Step 2. The event of (b) implies $(U-t)(V-t) \le 0$, and $(U-t)(V-t) = (S-t)^2 - D^2$. So its probability is at most $\mathbb{P}(|S-t| \le |D|)$.

Step 3. An interval of given length has the largest mass under f when it is centred at 0: for all $c \in \mathbb{R}$ and $w \ge 0$, $\int_{c-w}^{c+w} f \le \int_{-w}^{w} f$. Indeed, for $c \ge 0$ the difference of the two sides is $\int_0^c \big(f(u+w) - f(u-w)\big)\,du \le 0$, as $|u+w| \ge |u-w|$ for $u \ge 0$, and the case $c < 0$ follows by symmetry.

Step 4. Fix a value of D. Since S is independent of D, the probability that $|S-t| \le |D|$ is then the mass of the interval $[t-|D|, t+|D|]$ under f, which by Step 3 is at most the mass of $[-|D|, |D|]$. Averaging over the values of D gives $\mathbb{P}(|S-t| \le |D|) \le \mathbb{P}(|S| \le |D|) = \mathbb{P}(UV \le 0)$, since $UV = S^2 - D^2$. Now (a) finishes the proof. □

Fix an integer $R \ge 1$. For vertices v, w define the *coefficient* $c_v(w)$ and, for vertices u, v, the *overlap* $\sigma(u,v)$ by

$$c_v(w) = \begin{cases} 2^{-\operatorname{dist}(v,w)/2} & \text{if } w \in B_R(v), \\ 0 & \text{otherwise,} \end{cases} \qquad \sigma(u,v) = \sum_{w \in V(G)} c_u(w)\, c_v(w).$$

We call $c_v(w)^2 = 2^{-\operatorname{dist}(v,w)}$ the *weight* of w seen from v. So $\sigma(u,v)$ is a weighted count of the vertices of $B_R(u) \cap B_R(v)$, and $\sigma(v,v) = \sum_{k=0}^{R} 2^{-k}\,|\{w : \operatorname{dist}(v,w) = k\}|$ is the total weight of the ball around v. In Section 3.4 the label of v will be $Y_v = \sigma(v,v)^{-1/2} \sum_w c_v(w) Z_w$. By (P2), the correlation of the labels of u and v will then be $\sigma(u,v)/\sqrt{\sigma(u,u)\sigma(v,v)}$. The decay $2^{-k/2}$ of the coefficients balances the spheres: in the 3-regular tree the sphere of radius $k \ge 1$ around v has $3 \cdot 2^{k-1}$ vertices, so every sphere carries the same total weight $\frac{3}{2}$, and passing from v to a neighbor loses only part of the outermost of the R spheres. The third lemma shows that this holds in every graph of maximum degree 3, with no assumption on girth, regularity or connectivity. The same coefficients $(d-1)^{-k/2}$ on spheres are the test functions in Nilli's proof of the Alon–Boppana bound on the second eigenvalue of d-regular graphs [Nil91], and the constant $\frac{2\sqrt{2}}{3}$ below is the norm $2\sqrt{2}$ of the adjacency operator of the 3-regular tree divided by the degree.

Lemma 3.4 (short edges)**.** *Let G be a graph with $\Delta(G) \le 3$ and let $R \ge 1$ be an integer. Then every edge uv of G satisfies $\sigma(u,v) \ge r_R \sqrt{\sigma(u,u)\,\sigma(v,v)}$, where $r_R = \frac{2\sqrt{2}}{3}\big(1 - \frac{1}{R}\big)$.*

Proof. The proof has two steps: every shared vertex contributes to $\sigma(u,v)$ nearly its full weight from both ends, and the vertices that are not shared carry at most a $\frac{1}{R}$ fraction of the weight. Write $a_w = c_u(w)$, $b_w = c_v(w)$ and $I = B_R(u) \cap B_R(v)$. The coefficients are nonnegative and vanish outside the balls, so $\sigma(u,v) = \sum_{w \in I} a_w b_w$. For $w \in I$ the distances from u and from v to w differ by at most 1, so $\lambda = a_w/b_w \in [2^{-1/2}, 2^{1/2}]$; hence $\lambda + \frac{1}{\lambda} \le \frac{3}{\sqrt{2}}$, that is, $a_w^2 + b_w^2 \le \frac{3}{\sqrt{2}} a_w b_w$. Summing

[1] Part (a) also has a proof by calculus alone. For $-1 < \rho < 1$ let $P(\rho) = \mathbb{P}(U > 0, V > 0)$. By Plackett's identity [Pla54] for two variables, $P'(\rho)$ is the joint density of (U,V) at $(0,0)$, which is $1/(2\pi\sqrt{1-\rho^2})$. Since $P(0) = \frac{1}{4}$, integration gives $P(\rho) = \frac{1}{4} + \frac{\arcsin \rho}{2\pi}$, and by symmetry $\mathbb{P}(UV \le 0) = 1 - 2P(\rho) = \frac{\arccos \rho}{\pi}$.

over I, with $A = \sum_{w\in I} a_w^2$ and $B = \sum_{w\in I} b_w^2$ the weights of the shared vertices seen from u and from v, and using $A + B \ge 2\sqrt{AB}$,

$$\sigma(u,v) \;\ge\; \frac{\sqrt{2}}{3}(A+B) \;\ge\; \frac{2\sqrt{2}}{3}\sqrt{AB}.$$

It remains to show that the shared vertices carry almost all the weight, that is, $A \ge (1-\frac{1}{R})\sigma(u,u)$ and $B \ge (1-\frac{1}{R})\sigma(v,v)$. Let s_j be the number of vertices at distance j from u. For $j \ge 1$, a vertex at distance j has a neighbor at distance $j-1$ and hence at most two neighbors at distance $j+1$, so $s_{j+1} \le 2s_j$, and the sphere weights $s_j 2^{-j}$ are nonincreasing in $j \ge 1$. A vertex $w \in B_R(u) \setminus I$ has $\operatorname{dist}(v,w) > R$ and hence $\operatorname{dist}(u,w) \ge \operatorname{dist}(v,w) - 1 \ge R$, so it lies on the sphere of radius R. Since $\sigma(u,u) = 1 + \sum_{j=1}^{R} s_j 2^{-j}$,

$$\sigma(u,u) - A \;\le\; s_R 2^{-R} \;\le\; \frac{1}{R}\sum_{j=1}^{R} s_j 2^{-j} \;\le\; \frac{1}{R}\,\sigma(u,u),$$

and likewise $\sigma(v,v) - B \le \frac{1}{R}\sigma(v,v)$. Hence $\sigma(u,v) \ge r_R\sqrt{\sigma(u,u)\,\sigma(v,v)}$. □

The last lemma is the variance bound for sums of locally dependent terms.

Lemma 3.5 (local sums)**.** *Let G be a graph with $\Delta(G) \le 3$, let $(Z_w)_{w\in V(G)}$ be independent random variables, let $\ell \ge 0$ be an integer and $\alpha > 0$, and let $Q = \sum_{v\in V(G)} I_v$, where each I_v takes values in $[0,\alpha]$ and is a function of $(Z_w)_{w\in B_\ell(v)}$. Then* $\operatorname{Var} Q \le \frac{3}{4}\alpha^2\, 4^\ell\, n$.

Proof. We use (P1) throughout. We have $\operatorname{Var} Q = \sum_{u,v} \operatorname{Cov}(I_u, I_v)$. If $B_\ell(u) \cap B_\ell(v) = \emptyset$, then I_u and I_v are functions of disjoint sets of independent variables, hence independent, and $\operatorname{Cov}(I_u, I_v) = 0$. Otherwise $\operatorname{dist}(u,v) \le 2\ell$ and $|\operatorname{Cov}(I_u, I_v)| \le \sqrt{\operatorname{Var} I_u \operatorname{Var} I_v} \le \frac{\alpha^2}{4}$, because a variable X with values in $[0,\alpha]$ has $\operatorname{Var} X \le \mathbb{E}(X - \frac{\alpha}{2})^2 \le \frac{\alpha^2}{4}$. Since $\Delta(G) \le 3$, every ball satisfies $|B_{2\ell}(u)| \le 1 + \sum_{j=1}^{2\ell} 3\cdot 2^{j-1} < 3\cdot 4^\ell$, so at most $3\cdot 4^\ell n$ pairs (u,v) contribute. □

3.4 Proof of Theorem 3.1

Let G be a graph with $\Delta(G) \le 3$, so that $m \le \frac{3}{2}n$, and fix integers $R, K \ge 1$. Let $Z = (Z_w)_{w\in V(G)}$ have independent standard normal coordinates and give every vertex v the *label*

$$Y_v = \frac{1}{\sqrt{\sigma(v,v)}} \sum_{w\in V(G)} c_v(w)\, Z_w,$$

a weighted sum of the independent normal numbers in the ball $B_R(v)$. By (P2), Y_v is standard normal and $\operatorname{Corr}(Y_u, Y_v) = \sigma(u,v)/\sqrt{\sigma(u,u)\sigma(v,v)}$, and every pair (Y_u, Y_v) has the form of Lemma 3.3 with $a_w = c_u(w)/\sqrt{\sigma(u,u)}$ and $b_w = c_v(w)/\sqrt{\sigma(v,v)}$. For $t \in \mathbb{R}$ let $S_t = \{v : Y_v \le t\}$ and $e_t = e(S_t)$; an edge uv *crosses* t if exactly one of its ends lies in S_t, that is, if $\min(Y_u, Y_v) \le t < \max(Y_u, Y_v)$.

Distinct labels. We need the labels to be distinct, so that sorting by label gives an order in which every prefix is a set S_t. Let $u \ne v$. Among the coefficients of the Z_w in Y_u the largest is the one of Z_u, uniquely, and in Y_v it is the one of Z_v; so the two coefficient sequences differ. Hence $Y_u - Y_v$ is a normal variable with positive variance, so $\mathbb{P}(Y_u = Y_v) = 0$ by (P2). So all labels are distinct with probability 1.

One threshold. By Lemma 3.4 and Lemma 3.3(b), and since arccos is decreasing, every edge crosses every threshold with probability at most $\arccos(r_R)/\pi$. Summing over the $m \le \frac{3}{2}n$ edges (linearity of expectation, (P1)) gives

$$\mathbb{E}e_t \le \beta_R n \quad \text{for every } t \in \mathbb{R}, \qquad \beta_R = \frac{3}{2\pi}\arccos r_R. \tag{3.1}$$

Since $|\frac{d}{dx}\arccos x| = (1-x^2)^{-1/2} \le 3$ for $0 \le x \le \frac{2\sqrt{2}}{3}$ and $\frac{2\sqrt{2}}{3} - r_R = \frac{2\sqrt{2}}{3R}$, the mean value theorem gives $\beta_R \le c_* + \frac{3}{2\pi} \cdot 3 \cdot \frac{2\sqrt{2}}{3R} = c_* + \frac{3\sqrt{2}}{\pi R}$.

From one threshold to all thresholds. The bound (3.1) holds for each threshold separately, on average over the random numbers. We need one outcome in which all thresholds are good at the same time, and there are infinitely many thresholds. So we first reduce them to finitely many counts on a grid, then show that each count is close to its mean with high probability, and finally use the union bound (P6).

A grid of thresholds. Labels outside $[-3,3]$ are rare (Section 3.2), so the grid only needs to cover $[-3,3]$. Let $\tau_j = -3 + j/K$ for $0 \le j \le 6K$, and for an outcome of the labels let $N_j = |\{v : \tau_j < Y_v \le \tau_{j+1}\}|$, $C_- = |\{v : Y_v \le -3\}|$ and $C_+ = |\{v : Y_v > 3\}|$. Let $\tau_j \le \tau < \tau_{j+1}$. An edge ab with $Y_a \le \tau < Y_b$ either has $Y_a \le \tau_j$, and then it is counted by e_{τ_j}, or has $\tau_j < Y_a \le \tau < \tau_{j+1}$, and then a is one of the N_j counted vertices, each the end of at most three edges. Hence $e_\tau \le e_{\tau_j} + 3N_j$. In the same way $e_\tau \le 3C_-$ for $\tau < -3$, where the lower end of a crossing edge has label at most $\tau < -3$, and $e_\tau \le 3C_+$ for $\tau \ge 3$, where the upper end has label greater than 3.

Concentration. Consider the $12K+2$ counts e_{τ_j} and N_j for $0 \le j < 6K$, C_- and C_+. Each label Y_u is a function of $(Z_w)_{w \in B_R(u)}$. So $e_\tau = \sum_v |\{u \in N(v) : Y_v \le \tau < Y_u\}|$, which counts every crossing edge at its lower end, is a sum over v of $[0,3]$-valued functions of $(Z_w)_{w \in B_{R+1}(v)}$, and so are N_j, C_- and C_+. By Lemma 3.5 with $\alpha = 3$ and $\ell = R+1$, each count has variance at most $27 \cdot 4^R n$, so by Chebyshev's inequality (P5) it deviates from its mean by at least $n^{7/8}$ with probability at most $27 \cdot 4^R n^{-3/4}$. If

$$(12K+2) \cdot 27 \cdot 4^R < n^{3/4}, \tag{3.2}$$

then the probability that some count deviates from its mean by $n^{7/8}$ or more is less than 1. So by (P6) some outcome of the numbers Z_w has distinct labels and all $12K+2$ counts within $n^{7/8}$ of their means; fix such an outcome. Each Y_v is standard normal, and its density is at most $\varphi(0)$, so it lies in an interval of length $\frac{1}{K}$ with probability at most $\varphi(0)/K$. Hence $\mathbb{E}N_j \le \varphi(0)n/K$, and $\mathbb{E}C_\pm = \Phi(-3)n$. By (3.1) and the grid, every $\tau \in [-3,3)$ satisfies

$$e_\tau \;\le\; \Big(\beta_R + \frac{3\varphi(0)}{K}\Big)n + 4n^{7/8},$$

and every $\tau \notin [-3,3)$ satisfies $e_\tau \le 3\Phi(-3)n + 3n^{7/8}$, which is smaller, because $3\Phi(-3) < 0.005 < c_* \le \beta_R$.

Conclusion. Order $V(G)$ by increasing label in the fixed outcome. Every prefix of this ordering is a set S_t, so by Lemma 3.2

$$\mathrm{pw}(G) \;\le\; \mathrm{cw}(G) \;\le\; \max_t e_t \;\le\; \Big(\beta_R + \frac{3\varphi(0)}{K}\Big)n + 4n^{7/8}.$$

Take $K = R = \lfloor \frac{1}{4}\log_2 n \rfloor$. Then $4^R \le \sqrt{n}$ and $12K + 2 \le 3\log_2 n + 2$, so (3.2) follows from $27(3\log_2 n + 2) < n^{1/4}$, which holds for all $n \ge 2^{48}$ (at $n = 2^{48}$ it reads $3942 < 4096$). For $2^{46} \le n < 2^{48}$ we have $K = R = 11$, and (3.2) reads $134 \cdot 27 \cdot 4^{11} < n^{3/4}$, which holds since $134 \cdot 27 < 2^{12}$ and $2^{12} \cdot 2^{22} = 2^{34} < 2^{34.5} \le n^{3/4}$. So (3.2) holds for all $n \ge 2^{46}$. For these n, since $\frac{3\sqrt{2}}{\pi} + 3\varphi(0) < 2.55$,

$$\mathrm{pw}(G) \;\le\; \mathrm{cw}(G) \;\le\; c_* n + \frac{2.55}{R}n + 4n^{7/8} \;=\; c_* n + o(n).$$

For $n < 2^{46}$ the bound $\mathrm{cw}(G) \le m \le \frac{3}{2}n$ is absorbed by the remainder. □

3.5 Larger maximum degree

The cut sweep extends to every maximum degree. For an integer $p \geq 3$ put

$$r_{pp} = \frac{2\sqrt{p-1}}{p} \qquad \text{and} \qquad \gamma_p = \frac{\arccos r_{pp}}{\pi},$$

so that $r_{33} = \frac{2\sqrt{2}}{3}$, $r_{44} = \frac{\sqrt{3}}{2}$, $r_{55} = \frac{4}{5}$, and $\gamma_3 = \frac{2}{3}c_* \approx 0.10817$, $\gamma_4 = \frac{1}{6}$ and $\gamma_5 < 0.20484$.

Theorem 3.6 (cut sweep, maximum degree p)**.** *For every integer $p \geq 3$ there is a function $\eta_p(n) = o(n)$ such that every graph G with $\Delta(G) \leq p$ satisfies*

$$\mathrm{pw}(G) \;\leq\; \mathrm{cw}(G) \;\leq\; \gamma_p\, m + \eta_p(n).$$

In particular, the bisection width and the cutwidth of every graph G with $\Delta(G) \leq 4$ are at most $\frac{m}{6} + o(n) \leq \frac{n}{3} + o(n)$, and those of every graph G with $\Delta(G) \leq 5$ are at most $\gamma_5 m + o(n) \leq \frac{5}{2}\gamma_5 n + o(n) < 0.5121n + o(n)$.

Proof. We follow Section 3.4 and list the changes. Here and below, $O_p(\cdot)$ hides constants that depend on p. Let $c_v(w) = (p-1)^{-\operatorname{dist}(v,w)/2}$ for $w \in B_R(v)$, and $c_v(w) = 0$ otherwise, and define Y_v from c_v as in Section 3.4. Then every edge uv satisfies $\mathrm{Corr}(Y_u, Y_v) \geq r_{pp}(1 - \frac{1}{R})$: the proof of Lemma 3.4 goes through with 2 replaced by $p-1$, the ratio $\lambda \in [(p-1)^{-1/2}, (p-1)^{1/2}]$, $\lambda + \frac{1}{\lambda} \leq \frac{2}{r_{pp}}$, and $s_{j+1} \leq (p-1)s_j$. (These are the padded labels of Definition 5.2(a), and the bound is Lemma 5.4.) The labels are distinct with probability 1, as before.

One threshold. By Lemma 3.3(b), every edge crosses every threshold with probability at most $\arccos(r_{pp}(1-\frac{1}{R}))/\pi$, so $\mathbb{E}e_t \leq \frac{m}{\pi}\arccos(r_{pp}(1-\frac{1}{R}))$ for every t. On $[0, r_{pp}]$ the derivative of arccos has absolute value at most $(1-r_{pp}^2)^{-1/2} = \frac{p}{p-2}$, and $m \leq \frac{p}{2}n$, so

$$\mathbb{E}e_t \;\leq\; \gamma_p\, m + \frac{p^2 r_{pp}}{2\pi(p-2)} \cdot \frac{n}{R} \qquad \text{for every } t \in \mathbb{R}.$$

Grid and tails. With the grid of Section 3.4, an edge crossing $\tau \in [\tau_j, \tau_{j+1})$ either crosses τ_j or has its lower end among the N_j counted vertices, each the end of at most p edges, so $e_\tau \leq e_{\tau_j} + pN_j$. For the tails we weight by degree, so that the bound stays proportional to m: let $C'_- = \sum_{v:\, Y_v \leq -3} d(v)$ and $C'_+ = \sum_{v:\, Y_v > 3} d(v)$. Then $e_\tau \leq C'_-$ for $\tau < -3$ and $e_\tau \leq C'_+$ for $\tau \geq 3$, and $\mathbb{E}C'_\pm = 2m\Phi(-3)$.

Concentration. The counts e_{τ_j}, N_j and $C'_\pm$ are sums over v of $[0,p]$-valued functions of $(Z_w)_{w \in B_{R+1}(v)}$. The proof of Lemma 3.5 applies with $|B_{2\ell}(u)| \leq 1 + \sum_{j=1}^{2\ell} p(p-1)^{j-1} \leq \frac{p}{p-2}(p-1)^{2\ell}$, so each count has variance at most $\frac{p^3}{4(p-2)}(p-1)^{2R+2}n$. Take $K = R = \lfloor \frac{1}{8}\log_{p-1} n \rfloor$, so that $(p-1)^{2R} \leq n^{1/4}$. By Chebyshev's inequality and a union bound over the $12K+2$ counts, with probability at least $1 - O_p(n^{-1/2}\log n)$ all of them lie within $n^{7/8}$ of their means, and then every $\tau \in [-3, 3)$ satisfies

$$e_\tau \;\leq\; \gamma_p m + \Big(\frac{p^2 r_{pp}}{2\pi(p-2)} + p\,\varphi(0)\Big)\frac{n}{R} + (p+1)n^{7/8},$$

while every $\tau \notin [-3,3)$ satisfies $e_\tau \leq 2\Phi(-3)m + n^{7/8}$, which is smaller because $2\Phi(-3) < 0.003 < \gamma_3 \leq \gamma_p$. As in Section 3.4, ordering the vertices by their labels gives $\mathrm{pw}(G) \leq \mathrm{cw}(G) \leq \max_t e_t$, and the error terms are $O_p(n/R) + (p+1)n^{7/8} = o(n)$, since $R \to \infty$; for the finitely many n below the threshold where the probability is positive, $\mathrm{cw}(G) \leq m \leq \frac{p}{2}n$ is absorbed into η_p.

For the bisection width, the first $\lfloor n/2 \rfloor$ vertices of the ordering form a set X with $e(X) \leq \mathrm{cw}(G)$. For $p = 4$ we have $r_{44} = \cos\frac{\pi}{6}$, so $\gamma_4 = \frac{1}{6}$, and $m \leq 2n$; for $p = 5$, $m \leq \frac{5}{2}n$ and $\frac{5}{2}\gamma_5 < 0.5121$. $\square$

The proof is a randomized polynomial-time algorithm. The labels need the balls of radius $R = O(\log n)$, which have polynomial size, and the cut values of an ordering can be computed in polynomial time and compared with the explicit bound in the display above. So repeating with fresh Gaussians until the bound holds computes, in expected polynomial time, an ordering of cutwidth at most $\gamma_p m + o(n)$, for every $p \geq 3$. Since the algorithm checks the bound on the ordering it outputs, its answer is correct whatever numbers it samples; sampling the Gaussians to finite precision can affect only the success probability of a round, which we do not analyze. For $p \in \{4, 5\}$, Theorem 10.3 in Section 10.2 derandomizes it with the method of conditional expectations and computes an ordering of cutwidth at most $\gamma_p m + o(n)$, and a bisection with at most that many edges between its sides, in deterministic time $n^{2+o(1)}$.

For $p = 3$, Theorem 3.6 gives $\gamma_3 m \leq c_* n$, which Theorem 3.1 states, and the stronger Corollary 1.2 gives $c_3 n$. For $p = 4$ the pathwidth bound of Theorem 3.6, $\frac{n}{3}$, is weaker than $F_4 n$ (Theorem 7.1), but the bounds on bisection width and cutwidth are new: Monien and Preis [MP06, Theorem 2] proved that 4-regular graphs have bisection width at most $\frac{2n}{5} + o(n)$, and for maximum degree 4 the relations between cutwidth and pathwidth [MS89, Theorem 3.2][EST94, Theorem 2.1] give only $\mathrm{cw}(G) \leq 2\,\mathrm{pw}(G) + O(1) \leq 2F_4 n + o(n)$.

3.6 Remarks

The argument loses in two places. First, the angle: on cubic graphs of girth greater than $2R + 3$, where every ball $B_{R+1}(v)$ induces a tree, no choice of coefficients $c_v(w)$ supported on balls of radius R and depending only on $\mathrm{dist}(v, w)$ does better than $\arcsin \frac{1}{3}$. There the edge correlations of such labels are those of the 3-regular tree, at most $\frac{2\sqrt{2}}{3}$ because the adjacency operator of that tree has norm $2\sqrt{2}$ (Remark 5.1); so at least $c_* n$ edges cross the median in expectation, and counting crossing edges cannot give a smaller constant with such labels. Second, the count: a vertex of $N(S_t) \setminus S_t$ is counted once for every neighbor in S_t. Counting each such vertex once needs the joint distribution of a label and all its neighbor labels, in which the correlations between two neighbors are unknown; the star comparison of Section 6 handles them and gives the plain-sweep constant $F_3 \approx 0.14016$ (Theorem 7.1(b) with $n_4 = 0$), and switching from inner to outer frontiers at threshold 0 gives c_3 (Theorem 8.1). Finally, for graphs that converge locally to the 3-regular tree (for example, cubic graphs of large girth), c_* is the density of the sign cut (threshold 0) of the Gaussian wave [Lyo17].

The edge isoperimetric number. For a d-regular graph the bisection of Theorem 3.6 gives the edge isoperimetric number $i(G) = \min\{e(S)/|S| : 0 < |S| \leq \frac{n}{2}\}$, where $e(S)$ is the number of edges with exactly one end in S, at most $d\gamma_d + o(1) \leq \frac{d}{2} - \frac{2}{\pi}\sqrt{d-1} + o(1)$. Alon's upper bound has the form $\frac{d}{2} - c\sqrt{d}$ for $n > n_0(d)$, with $c = \frac{3}{16\sqrt{2}} < 0.133$ [Alo97, Theorem 1.1 and (11)]; for $d \geq 3$ and n large enough, our bound gives $\frac{2}{\pi}\sqrt{d-1} > 0.51\sqrt{d}$ in place of $c\sqrt{d}$. For cubic graphs, Corollary 1.2 gives $i(G) \leq 2c_3 + o(1) < 0.2797$, in place of the bound $\frac{1}{3}$ that follows from the theorem of Monien and Preis [MP06, Theorem 1], and 4-regular graphs have $i(G) \leq \frac{2}{3} + o(1)$.

4 Preliminaries

This section fixes notation and collects the tools that all later sections share. Two of them shape the whole of Part I. The layout lemma turns a labeling of the vertices into a path decomposition; it is the only way in which path decompositions are built in Part I (Section 4.2). The reductions on vertices of degree at most two let the main proofs assume minimum degree at least three, at an additive loss of $O(\log n)$ in the pathwidth (Section 4.3).

Of the remaining subsections, Section 4.1 fixes the notation for graphs, decompositions and remainders $o(n)$, Section 4.4 the Gaussian notation and comparison tools, and Section 4.5 proves a variance bound for sums of locally dependent variables, which makes frontier sizes concentrate.

4.1 Graphs, decompositions and remainders

Graphs are finite and simple. We follow Diestel's notation: $V(G)$ and $E(G)$ are the vertex and edge sets, $uv \in E(G)$ is an edge, $d_G(v)$ or $d(v)$ is the degree, $N(v)$ and $N[v] = N(v) \cup \{v\}$ are the open and closed neighborhoods, $N(S) = \bigcup_{v \in S} N(v)$ for $S \subseteq V(G)$, and $\delta(G)$, $\Delta(G)$ are the minimum and maximum degree. $G[U]$ is the subgraph induced by U and $G - U$ is obtained by deleting U. We write $n = |V(G)|$, $m = |E(G)|$ and $n_d = n_d(G)$ for the number of vertices of degree d. The graph distance $\operatorname{dist}(v, w) \in \mathbb{N} \cup \{\infty\}$ is the number of edges of a shortest v–w path, and $B_k(v) = \{w : \operatorname{dist}(v, w) \le k\}$ is the ball of radius k. The sizes of these balls enter the variance bound of Lemma 4.10 and the analysis of the algorithms. If $\Delta(G) \le \Delta$ with $\Delta \ge 2$, then

$$|B_k(v)| \le 1 + \sum_{i=1}^{k} \Delta(\Delta - 1)^{i-1} \le \sum_{i=0}^{k} \Delta^i \le 2\Delta^k, \tag{4.1}$$

because a vertex at distance $i \ge 1$ from v has at most $\Delta - 1$ neighbors at distance $i + 1$, and $\sum_{i=0}^{k} \Delta^i = (\Delta^{k+1} - 1)/(\Delta - 1) \le 2\Delta^k$ for $\Delta \ge 2$. For the maximum degrees 3, 4 and 5, which occur most often below, evaluating the middle sum gives the sharper bounds

$$\begin{aligned} |B_k(v)| &\le 3 \cdot 2^k - 2 && \text{if } \Delta = 3, \\ |B_k(v)| &\le 2 \cdot 3^k - 1 && \text{if } \Delta = 4, \\ |B_k(v)| &\le \tfrac{1}{3}\bigl(5 \cdot 4^k - 2\bigr) \le 2 \cdot 4^k && \text{if } \Delta = 5. \end{aligned}$$

Decompositions. A *tree decomposition* of G is a pair $(T, (W_t)_{t \in V(T)})$ of a tree T and bags $W_t \subseteq V(G)$ such that every edge of G lies in some bag and, for every vertex x, the nodes t with $x \in W_t$ induce a nonempty subtree T_x of T. Its width is $\max_t |W_t| - 1$, and the treewidth $\operatorname{tw}(G)$ is the minimum width. Path decompositions, pathwidth and cutwidth are as in Section 3; a path decomposition is a tree decomposition whose tree is a path. The empty graph has $\operatorname{pw} = \operatorname{tw} = -1$. We use the following elementary facts without further mention: $\operatorname{tw}(G) \le \operatorname{pw}(G)$; both widths are monotone under taking subgraphs; the pathwidth of a disjoint union is the maximum over its parts (concatenate the decompositions); and adding a set X of vertices to every bag of a path decomposition of $G - X$ gives a path decomposition of G, so $\operatorname{pw}(G) \le \operatorname{pw}(G - X) + |X|$. A path decomposition is given in *interval form* if it is described by the number s of bags and, for each vertex, the indices of the first and the last bag containing it; the algorithms of Section 10 output their decompositions in this form.

Remainders. Every remainder $o(n)$ in a statement about a class $\mathcal{G}$ of graphs means the following: there is a function $\rho\colon \mathbb{N} \to \mathbb{R}$ with $\rho(n)/n \to 0$ such that the statement holds with $o(n)$ replaced by $\rho(|V(G)|)$ for every $G \in \mathcal{G}$. So the remainder depends on n only, and not on the graph. The same convention applies to $o(m)$. The *running maximum* of a function $\rho\colon \mathbb{N} \to \mathbb{R}$ is

$$\bar{\rho}(x) := \max\bigl(\{0\} \cup \{\rho(k) : k \in \mathbb{N},\ k \le x\}\bigr) \qquad (x \ge 0). \tag{4.2}$$

It is nonnegative, nondecreasing and at least ρ on $\mathbb{N}$. If $\rho(n) = o(n)$, then $\bar{\rho}(x) = o(x)$: for $\varepsilon > 0$ there is N with $\rho(k) \le \varepsilon k$ for $k \ge N$, so $\bar{\rho}(x) \le \max\{\bar{\rho}(N), \varepsilon x\}$ for $x \ge N$, which is at most εx for all large x. Hence every remainder may be assumed nonnegative and nondecreasing.

4.2 Layouts from labelings

A *labeling* of G is a map $y\colon V(G) \to \mathbb{R}$, $v \mapsto y_v$; it is *injective* if all labels are distinct. For $t \in \mathbb{R}$ let $S_t = \{v : y_v \le t\}$ be the sublevel set, and define its outer and inner frontiers

$$\partial^+ S_t = N(S_t) \setminus S_t = \{v :\ y_v > t,\ \exists u \in N(v) :\ y_u \le t\}, \qquad \partial^- S_t = \{u \in S_t :\ N(u) \not\subseteq S_t\}.$$

The same notation $\partial^+ S = N(S) \setminus S$ and $\partial^- S = \{u \in S : N(u) \not\subseteq S\}$ is used for arbitrary sets $S \subseteq V(G)$. Every later bound on pathwidth comes from the following lemma, applied to a random labeling whose frontiers are small at every threshold. The maxima over t in the lemma are attained, since S_t takes at most $n+1$ values.

Lemma 4.1 (layout)**.** *Let G be a nonempty graph and y an injective labeling of G, let $v_1, \dots, v_n$ be the vertices in increasing order of their labels, and let $T_i = \{v_1, \dots, v_i\}$. Then the bags $W_i = \{v_i\} \cup \partial^+ T_i$ $(1 \le i \le n)$ form a path decomposition of G of width $\max_i |\partial^+ T_i|$. Moreover, $\mathrm{pw}(G) \le \max_{t \in \mathbb{R}} |\partial^+ S_t|$ and $\mathrm{pw}(G) \le \max_{t \in \mathbb{R}} |\partial^- S_t|$.*

The quantity $\max_i |\partial^- T_i|$, minimized over all orderings, is the *vertex separation number* of G, and Kinnersley proved that it equals the pathwidth [Kin92, Theorem 3.1]. The bound by inner frontiers is the easy direction of that theorem, and the bound by outer frontiers is the same statement for the reversed order. The algorithms of Section 10 use the explicit bags.

Proof. The proof of Lemma 3.2 shows that the bags W_i form a path decomposition of width $\max_i |\partial^+ T_i|$, since $\partial^+ T_i = N(T_i) \setminus T_i$; it uses neither the labels nor a degree bound. By injectivity $T_i = S_{y_{v_i}}$, so $\max_i |\partial^+ T_i| \le \max_t |\partial^+ S_t|$; in fact a bound $|\partial^+ S_t| \le w$ for all t in any set of thresholds that contains all labels already gives $\mathrm{pw}(G) \le w$.

For the bound by inner frontiers, apply the bound by outer frontiers to the injective labeling $-y$. Its sublevel sets are the sets $\{v : y_v \ge -t\} = V(G) \setminus A$, where $A = \{v : y_v < -t\}$ is either empty or equal to S_s for $s = \max\{y_v : y_v < -t\}$. For every $A \subseteq V(G)$ we have $\partial^+(V(G) \setminus A) = \{a \in A : N(a) \not\subseteq A\} = \partial^- A$, and $\partial^- \emptyset = \emptyset$. Hence the outer frontiers for $-y$ are inner frontiers for y, which gives the claim. □

4.3 Reductions on vertices of degree at most two

The pathwidth bounds of this paper are proved for graphs of minimum degree at least three and then transferred to all graphs. The transfer uses the classical series–parallel reductions, which remove vertices of degree at most two. Reduction rules of this kind are used in the pathwidth bounds of Kneis, Mölle, Richter and Rossmanith [KMRR09] and of Edwards and McDermid [EM15]. The reduction theorem of Edwards and McDermid [EM15, Theorem 3.2] interleaves the reductions with vertex deletions and bounds the effect of all reductions on the pathwidth once, at the end. Edwards and McDermid take two facts from [KMRR09]: that reductions commute with deletions, and that undoing all reductions increases the pathwidth by at most $O(\log n)$. Lemma 4.5 below replaces both facts by a single self-contained statement, which needs no uniqueness of the reduced graph. We first define the reductions and prove two auxiliary facts about treewidth, then prove Lemma 4.5 and derive the transfer from minimum degree three to all graphs (Corollary 4.6).

Definition 4.2 (reduction operations)**.** The *reduction operations* on a graph H are:

(R1) delete an isolated vertex;

(R2) delete a vertex of degree 1;

(R3) for a vertex v of degree 2 whose neighbors x, y are not adjacent, delete v and add the edge xy (*suppression*);

(R4) delete a vertex of degree 2 whose two neighbors are adjacent.

A sequence of reduction operations is *maximal* if no reduction operation applies to the final graph.

We use three properties of the reduction operations throughout.

(RP1) Each operation removes one vertex, produces a simple graph, and does not increase the degree of any remaining vertex: (R1), (R2) and (R4) only delete a vertex, and in (R3) the vertices x and y each lose the neighbor v and gain one new neighbor.

(RP2) Whether an operation applies at a vertex v, and its effect, depend only on v, its neighbors and the adjacency among them. So operations inside one component of a graph are operations on the whole graph.

(RP3) By (RP1), every sequence of operations on H has at most $|V(H)|$ steps, so maximal sequences exist. The final graph of a maximal sequence is empty or has minimum degree at least 3, since a vertex of degree 0, 1 or 2 admits (R1), (R2), or one of (R3) and (R4).

The proof of Lemma 4.5 needs two standard facts about graphs of bounded treewidth [Bod98].

Lemma 4.3. *Let $G = G_1 \cup G_2$ where $V(G_1) \cap V(G_2) = K$ and K is either a single vertex or the two ends of an edge that belongs to both G_1 and G_2. Then* $\mathrm{tw}(G) \le \max\{\mathrm{tw}(G_1), \mathrm{tw}(G_2)\}$.

Proof. Join tree decompositions of G_1 and G_2 of minimum width by an edge between a bag of each that contains K. □

Lemma 4.4. *Let $k \ge 0$ and let H be a graph with $N \ge 1$ vertices and* $\mathrm{tw}(H) \le k$. *Then* $\mathrm{pw}(H) \le (k+1)\lfloor \log_2 N \rfloor$.

Proof. Some bag X of a tree decomposition of width at most k leaves only components with at most $N/2$ vertices (the standard balanced-separator argument). Decompose these components by induction, concatenate the decompositions and add X to every bag. □

The next lemma is the bookkeeping device of the paper. Suppose that an arbitrary interleaving of reduction operations and vertex deletions is applied to a graph. Then the pathwidth of the original graph exceeds that of the final graph by at most the number of deleted vertices plus $O(\log n)$, however many reductions were made. The lemma is used for the passage from minimum degree three to all graphs (Corollary 4.6) and in the deletion process of Section 9. Its part (b) describes how $F - D$ attaches to the final graph; Appendix A.5.6 uses it to compute the decomposition.

Lemma 4.5 (undoing reductions)**.** *Let F be a nonempty graph. Apply to F any finite sequence of steps, each of which is a reduction operation or the deletion of a vertex of the current graph. Let D be the set of vertices removed by deletions and R the final graph. Then:*

(a) $\mathrm{pw}(F) \le |D| + \mathrm{pw}(R) + 3\log_2 |V(F)| + 1$;

(b) *every edge of $F - D$ with both ends in $V(R)$ is an edge of R, and for every component C of $(F - D) - V(R)$, the set $N_{F-D}(C)$ is empty, a single vertex, or the two ends of an edge of R, and* $\mathrm{tw}(F[C]) \le 2$.

The proof is in Appendix A.1.

In the simplest application no vertex is deleted and the sequence of reduction operations is maximal, so the final graph is empty or has minimum degree at least three by (RP3). For coefficients that are nondecreasing in the degree, this transfers every bound for minimum degree three to all graphs.

Corollary 4.6. *Let $k \in \{3, 4, \dots\} \cup \{\infty\}$, let $0 = a_0 = a_1 = a_2 \le a_3 \le a_4 \le \cdots$ (up to a_k), and let $\rho\colon \mathbb{N} \to \mathbb{R}_{\ge 0}$ be nondecreasing. Suppose that every graph J with $\delta(J) \ge 3$ and $\Delta(J) \le k$ satisfies $\mathrm{pw}(J) \le \sum_d a_d n_d(J) + \rho(|V(J)|)$. Then every nonempty graph G with $\Delta(G) \le k$ satisfies*

$$\mathrm{pw}(G) \le \sum_d a_d n_d(G) + \rho(n) + 3\log_2 n + 1.$$

Proof. Apply a maximal sequence of reduction operations to G and let R be the final graph. Degrees do not increase, so $\Delta(R) \le k$, and either R is empty or $\delta(R) \ge 3$. Since the a_d are nonnegative and nondecreasing in d, and every vertex of R has degree in R at most its degree in G,

$$\sum_d a_d n_d(R) = \sum_{v \in V(R)} a_{d_R(v)} \le \sum_{v \in V(R)} a_{d_G(v)} \le \sum_{v \in V(G)} a_{d_G(v)} = \sum_d a_d n_d(G).$$

If R is nonempty, the hypothesis and the monotonicity of ρ give $\mathrm{pw}(R) \le \sum_d a_d n_d(G) + \rho(n)$; if R is empty, $\mathrm{pw}(R) = -1$ satisfies the same bound. Lemma 4.5 with $D = \emptyset$ completes the proof. □

4.4 Gaussian notation and comparison tools

The star bounds of Section 6 and the switch lemma of Section 8 compare *orthant probabilities*, probabilities of the form $\mathbb{P}(Y_i > t_i\ \forall i)$, or in the switch lemma combinations of them, of Gaussian vectors Y with different correlation matrices. This subsection fixes the Gaussian notation and states the three classical tools that these comparisons use: Plackett's identity, Slepian's inequality and Sheppard's formula.

We write φ and Φ for the standard normal density and distribution function, $N(\mu, \sigma^2)$ for the normal distribution, and

$$\varphi_2(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\Big(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\Big)$$

for the standard bivariate normal density with correlation $\rho \in (-1, 1)$. Note that $\varphi_2(x, y; \rho) \le (2\pi\sqrt{1-\rho^2})^{-1}$.

A random vector is *centered Gaussian* if every linear combination of its coordinates is a centered normal variable (possibly degenerate); its distribution $N(0, \Sigma)$ is determined by its covariance matrix Σ, which may be singular. A *correlation matrix* is a positive semidefinite matrix with unit diagonal. Every $k \times k$ correlation matrix C is the Gram matrix $C_{ij} = \langle v_i, v_j \rangle$ of unit vectors $v_1, \dots, v_k \in \mathbb{R}^k$ (for instance the columns of $C^{1/2}$), and then $U_i = \langle v_i, G \rangle$ with $G \sim N(0, I_k)$ defines a centered Gaussian vector with correlation matrix C.

For $r \in [-1, 1]$ we write $s_r = \sqrt{1 - r^2}$. In Part I the letter r always denotes a correlation level. If $(Y_1, \dots, Y_k)$ is centered Gaussian with $\operatorname{Var} Y_i > 0$ for all i, then $\mathbb{P}(Y_i = t_i) = 0$ for every i and every $t_i \in \mathbb{R}$, so strict and weak inequalities between coordinates and constants may be exchanged in all probabilities below.

The comparison tool used throughout is Slepian's inequality: raising the correlations of a Gaussian vector, with its variances fixed, raises its orthant probabilities. We state it with arbitrary

thresholds and singular covariance matrices allowed, because both cases occur (the ideal stars of Section 6 have singular covariance). It is derived from Plackett's identity, which also bounds how fast an orthant probability can change with the correlations.

Lemma 4.7 (Plackett's identity [Pla54])**.** *Let $k \geq 2$ and $t \in \mathbb{R}^k$, and for a positive definite $k \times k$ correlation matrix Σ let $P(\Sigma) = \mathbb{P}(Y_i > t_i \ \forall i)$ with $Y \sim N(0, \Sigma)$. Then P is continuously differentiable in the off-diagonal entries $\Sigma_{ij} = \Sigma_{ji}$ $(i < j)$ on the open set of positive definite correlation matrices, and*

$$\frac{\partial P}{\partial \Sigma_{ij}} = \varphi_2(t_i, t_j; \Sigma_{ij}) \cdot \mathbb{P}\big(Y_l > t_l \ \forall l \notin \{i, j\} \mid Y_i = t_i,\ Y_j = t_j\big),$$

where the conditional probability refers to the nondegenerate Gaussian conditional distribution and equals 1 if $k = 2$. In particular $0 \leq \partial P / \partial \Sigma_{ij} \leq (2\pi\sqrt{1 - \Sigma_{ij}^2})^{-1}$.

The identity is classical [Pla54]; for a proof see [AT07, Section 2.2] and the survey [LS01]. It gives Slepian's inequality, and it is used in the lemma on two exceedances in the proof of the switch lemma (Lemma 8.11).

For $k = 3$, $\{i, j, l\} = \{1, 2, 3\}$ and equal thresholds $t_i = t_j = t_l = h$, the conditional distribution in Lemma 4.7 is explicit: the distribution of Y_l given $Y_i = Y_j = h$ is normal with

$$\text{mean} \ \ h\,\frac{\Sigma_{il} + \Sigma_{jl}}{1 + \Sigma_{ij}} \qquad \text{and variance} \ \ \frac{\det \Sigma}{1 - \Sigma_{ij}^2} > 0. \tag{4.3}$$

Indeed, the Gaussian conditioning formula gives the conditional mean $(\Sigma_{il}, \Sigma_{jl}) \left(\begin{smallmatrix} 1 & \Sigma_{ij} \\ \Sigma_{ij} & 1 \end{smallmatrix}\right)^{-1} \left(\begin{smallmatrix} h \\ h \end{smallmatrix}\right)$, and $(1, 1)^{\mathsf{T}}$ is an eigenvector of the 2×2 matrix with eigenvalue $1 + \Sigma_{ij}$; the conditional variance is the Schur complement of that matrix in Σ, which is $\det \Sigma / (1 - \Sigma_{ij}^2)$ and positive because Σ is positive definite. Section 8 uses this formula in the preparation of Lemma 8.11.

Lemma 4.8 (Slepian's inequality [Sle62])**.** *Let $Y = (Y_1, \dots, Y_k)$ and $Y' = (Y_1', \dots, Y_k')$ be centered Gaussian vectors with $\operatorname{Var} Y_i = \operatorname{Var} Y_i' > 0$ for all i and $\operatorname{Cov}(Y_i, Y_j) \geq \operatorname{Cov}(Y_i', Y_j')$ for all $i \neq j$. Then, for every $t \in \mathbb{R}^k$,*

$$\mathbb{P}(Y_i > t_i \ \forall i) \geq \mathbb{P}(Y_i' > t_i \ \forall i) \qquad \text{and} \qquad \mathbb{P}(Y_i \leq t_i \ \forall i) \geq \mathbb{P}(Y_i' \leq t_i \ \forall i).$$

Proof. For positive definite covariance matrices this is Slepian's inequality [Sle62] [AT07, Section 2.2]: after scaling to unit variances, integrate Lemma 4.7 along the segment from the correlation matrix of Y' to that of Y. For singular covariance matrices, apply it to $\sqrt{1-\eta}\,Y + \sqrt{\eta}\,\xi$ and $\sqrt{1-\eta}\,Y' + \sqrt{\eta}\,\xi'$ with independent standard Gaussian vectors ξ, ξ', and let $\eta \to 0$; the probabilities converge by dominated convergence, since $\mathbb{P}(Y_i = t_i) = 0$. The second inequality is the first one for $-Y$, $-Y'$ and $-t$. □

Sheppard's formula evaluates orthant probabilities at threshold zero. We use it for two and three coordinates.

Lemma 4.9 (Sheppard's formula [She99])**.** *Let (X_1, X_2, X_3) be centered Gaussian with unit variances and put $\vartheta_{ij} = \arccos \operatorname{Corr}(X_i, X_j) \in [0, \pi]$. Then*

$$\mathbb{P}(\operatorname{sgn} X_1 \neq \operatorname{sgn} X_2) = \frac{\vartheta_{12}}{\pi}, \qquad \mathbb{P}(X_1 > 0,\ X_2 > 0) = \frac{1}{2} - \frac{\vartheta_{12}}{2\pi} = \frac{1}{4} + \frac{\arcsin \operatorname{Corr}(X_1, X_2)}{2\pi},$$

and

$$\mathbb{P}(X_1 > 0,\ X_2 > 0,\ X_3 > 0) = \frac{1}{2} - \frac{\vartheta_{12} + \vartheta_{13} + \vartheta_{23}}{4\pi}.$$

Proof. Each X_i is nonzero with probability 1. The first formula is Lemma 3.3(a), whose proof applies to any centered Gaussian pair with unit variances (write it as a pair of unit-norm sums of independent standard normals; the case $\operatorname{Corr}(X_1, X_2) = -1$ is trivial). Since X and $-X$ have the same distribution, the two coordinates are both positive and both negative with equal probabilities, which gives the second; the last expression follows from $\arcsin x = \frac{\pi}{2} - \arccos x$.

For three coordinates, three nonzero signs have either 0 or 2 disagreeing pairs. Hence $\mathbb{E}\#\{\text{disagreeing pairs}\} = 2\mathbb{P}(\text{not all signs equal})$, so by the first formula $\mathbb{P}(\text{all signs equal}) = 1 - (\vartheta_{12} + \vartheta_{13} + \vartheta_{23})/(2\pi)$, and symmetry under $X \mapsto -X$ halves this. □

4.5 Local dependence and concentration

The labels of Section 5 are functions of independent variables attached to the vertices, and each label depends only on the variables in a ball of bounded radius. Counts such as frontier sizes are then sums of locally dependent indicators, whose variance is linear in n. The plain sweep (Section 7), the cubic sweep (Section 8) and the algorithms (Section 10) combine the following bound with Chebyshev's inequality.

Lemma 4.10 (local variance)**.** *Let G be a graph with n vertices, let $(Z_w)_{w \in V(G)}$ be independent random variables, and for every vertex v let $A_v \subseteq V(G)$ and let I_v be a random variable with values in $[0, a]$ that is a measurable function of $(Z_w)_{w \in A_v}$. Put $W = \sum_v I_v$. Then*

$$\operatorname{Var} W \le \frac{a^2}{4} \cdot \#\{(u, v) \in V(G)^2 : \ A_u \cap A_v \neq \emptyset\}.$$

If moreover $\Delta(G) \le \Delta$ for some $\Delta \ge 2$ and $A_v \subseteq B_\rho(v)$ for every v, then $\operatorname{Var} W \le \frac{a^2}{4}\, n \max_v |B_{2\rho}(v)| \le \frac{a^2}{2}\, n\Delta^{2\rho}$ and $\mathbb{P}(|W - \mathbb{E}W| \ge \lambda) \le a^2 n \Delta^{2\rho}/(2\lambda^2)$ for every $\lambda > 0$.

Proof. The first bound is proved as Lemma 3.5: $\operatorname{Cov}(I_u, I_v) = 0$ if $A_u \cap A_v = \emptyset$, and $|\operatorname{Cov}(I_u, I_v)| \le a^2/4$ otherwise. If $A_v \subseteq B_\rho(v)$ for all v, then $A_u \cap A_v \neq \emptyset$ implies $v \in B_{2\rho}(u)$, so there are at most $n \max_u |B_{2\rho}(u)| \le 2n\Delta^{2\rho}$ such pairs by (4.1). The last claim is Chebyshev's inequality. □

5 Gaussian labels and edge correlations

All sweeps of Part I order the vertices by random Gaussian labels and bound the sizes of the frontiers of the resulting order from above with the layout lemma (Lemma 4.1). This section constructs the labels and proves the two properties of them that the rest of Part I uses.

First, the labels of a vertex and of its neighbors form a star. For every vertex v, the vector $(Y_v; Y_u)_{u \in N(v)}$ is a centered Gaussian vector with unit variances (Lemma 5.6), and its incident correlations $\operatorname{Corr}(Y_v, Y_u)$ are bounded below by explicit constants that depend only on the degrees (Lemmas 5.4 and 5.5). When these constants are positive and v is not isolated, the vector is therefore a $d(v)$-star in the sense of Definition 6.1. The star bounds of Section 6 use nothing else about the labels. They allow arbitrary correlations between two neighbors of v, so short cycles and triangles through v need no separate treatment.

Second, the labels are local and distinct. Each label depends only on the independent variables in a ball of bounded radius, which makes frontier sizes concentrate (Lemma 4.10). The labels are distinct with probability 1, as Lemma 4.1 requires. Both facts are part of Lemma 5.6.

Section 5.1 defines the labels and states the three lemmas, and Sections 5.2 to 5.5 prove them. The technical overview (Section 2.2) explains why the chosen decay rate is the right one.

Remark 5.1 (optimality for radial labels). On the infinite p-regular tree, the local limit of p-regular graphs of large girth, the adjacency operator has norm $2\sqrt{p-1}$ (Kesten [Kes59]; [Woe00, Lemma 1.24]). Hence no Gaussian field $X_v = \sum_w f(\operatorname{dist}(v,w))Z_w$ with radial weights has a neighbor correlation above $\frac{2\sqrt{p-1}}{p} = r_{pp}$, since the p neighbor covariances of v sum to $\langle c_v, Ac_v\rangle$, where c_v is the coefficient vector of X_v. The value r_{pp} is attained in the limit by the Gaussian wave [CGHV15, HV15, BSV15], and the padded labels are its finite-radius approximations: their coefficients $(p-1)^{-k/2}$ on spheres are the test functions of Nilli's proof of the Alon–Boppana bound [Nil91]. So Lemma 5.4 is asymptotically sharp for radial labels. For mixed degrees the limiting edge correlation can lie slightly above r_{ij}. On the $(3,4)$-biregular tree, every generation of the two sides of an edge uv with $d(u)=3$ and $d(v)=4$ carries the same squared mass, and a ball contains $N_R + O(1)$ generations of each side, with $N_R = 4R/\ln 6$; summing over generations gives $\sigma_u^2 = \frac{3}{2}N_R + O(1)$, $\sigma_v^2 = \frac{4}{3}N_R + O(1)$ and $\operatorname{Cov}(X_u, X_v) = (2^{-1/2} + 3^{-1/2})N_R + O(1)$, so the correlation tends to $(2^{-1/2}+3^{-1/2})/\sqrt{2} = \frac{1}{2} + \frac{1}{\sqrt{6}} \approx 0.9082$, above $r_{34} \approx 0.9074$.

5.1 The labels

A label is a normalized weighted sum of independent standard normal variables, one for each vertex, with weights that decay exponentially in a weighted distance. An *arc* is an ordered pair (x,y) with $xy \in E(G)$, written $x \to y$.

Definition 5.2 (Gaussian labels)**.** Let G be a graph, let $(Z_w)_{w\in V(G)}$ be independent $N(0,1)$ variables, let every arc $x \to y$ have a length $\ell(x \to y) > 0$, and let $\hat{R} > 0$. The *directed distance* $D(v,w)$ is the minimum of $\sum_{i<k} \ell(x_i \to x_{i+1})$ over all walks $v = x_0, x_1, \dots, x_k = w$, with $D(v,v) = 0$ and $D(v,w) = \infty$ if there is no such walk. Put $B(v) = \{w : D(v,w) \le \hat{R}\}$ and

$$X_v = \sum_{w\in B(v)} e^{-D(v,w)} Z_w, \qquad \sigma_v^2 = \sum_{w\in B(v)} e^{-2D(v,w)}, \qquad Y_v = X_v/\sigma_v.$$

(a) *Padded labels of degree p*, for integers $p \ge 3$ and $R \ge 1$ and graphs with $\Delta(G) \le p$: every arc has length $\frac{1}{2}\ln(p-1)$, and $\hat{R} = \frac{R}{2}\ln(p-1)$.

(b) *Mixed labels*, for a real $R > 0$ and graphs with $\delta(G) \ge 3$: $\ell(x \to y) = \frac{1}{2}\ln\bigl(d(x)-1\bigr)$, and $\hat{R} = R$.

By independence, $\sigma_v^2 = \operatorname{Var} X_v$. For padded labels, $D(v,w) = \frac{1}{2}\ln(p-1)\operatorname{dist}(v,w)$ and $B(v) = B_R(v)$, so

$$X_v = \sum_{w\in B_R(v)} (p-1)^{-\operatorname{dist}(v,w)/2} Z_w;$$

for $p = 3$ we call them the *geometric labels*, with coefficients $2^{-\operatorname{dist}(v,w)/2}$. Padded labels are used with $p = 3$ for graphs of maximum degree 3 (Theorem 8.1) and with $p = 4$ for Theorem 7.1(a). In mixed labels the length of an arc depends on its tail only; they are used for graphs with $\delta(G) \ge 3$ and $\Delta(G) \le 5$ (Theorem 7.1(b),(c) and Theorem 7.2).

A walk of minimum length is a path, because all lengths are positive; since G is finite, the minimum in the definition of D is attained. The directed distance satisfies the triangle inequality $D(u,w) \le D(u,x) + D(x,w)$, and in particular

$$D(u,w) \le \ell(u \to v) + D(v,w) \qquad \text{for every edge } uv \text{ and every } w. \tag{5.1}$$

Table 3: The edge correlation bounds r_{ij} of (5.3) for end degrees $3 \le i \le j \le 5$.

(i,j)	$(i-1)(j-1)$	exact value	decimal
$(3,3)$	4	$\frac{2\sqrt{2}}{3}$	0.942809
$(3,4)$	6	$\frac{2\cdot 6^{1/4}}{1+\sqrt{6}}$	0.907430
$(3,5)$	8	$\frac{2\cdot 8^{1/4}}{1+2\sqrt{2}}$	0.878582
$(4,4)$	9	$\frac{\sqrt{3}}{2}$	0.866025
$(4,5)$	12	$\frac{2\cdot 12^{1/4}}{1+2\sqrt{3}}$	0.833856
$(5,5)$	16	$\frac{4}{5}$	0.800000

In the mixed case D is not symmetric. Its arc lengths lie between $\ell_{\min} = \frac{1}{2}\ln 2$ and $\frac{1}{2}\ln(\Delta(G)-1)$. The decay rate is chosen so that in both regimes the arcs leaving a vertex, other than one, carry total squared weight at most one:

$$\sum_{y \in N(x)\setminus\{y_0\}} e^{-2\ell(x\to y)} \le 1 \qquad \text{for every vertex } x \text{ and every } y_0 \in N(x). \tag{5.2}$$

Indeed the sum has $d(x)-1$ terms, each equal to $1/(p-1)$ in case (a), where $d(x) \le p$, and to $1/(d(x)-1)$ in case (b). In the mixed case, and in the padded case at vertices of degree p, (5.2) holds with equality. We call this *critical decay*; Section 2.2 explains its role.

Remark 5.3 (path products). The algorithms of Section 10 describe the same labels by vertex weights instead of arc lengths. Put $q(x) = p-1$ in case (a) and $q(x) = d(x)-1$ in case (b), and let $P(v,w)$ be the minimum of $\prod_{i<k} q(x_i)$ over all v–w paths $v = x_0, \dots, x_k = w$, with $P(v,v) = 1$. Since every arc $x \to y$ has length $\frac{1}{2}\ln q(x)$, the length of a walk is $\frac{1}{2}\ln$ of its product, so

$$e^{-D(v,w)} = P(v,w)^{-1/2} \qquad \text{and} \qquad X_v = \sum_{w \in B(v)} P(v,w)^{-1/2} Z_w,$$

with $B(v) = \{w : P(v,w) \le e^{2\hat{R}}\}$: the hop ball $B_R(v)$ in case (a), and $\{w : P(v,w) \le 2^h\}$ in case (b) when $R = \frac{h}{2}\ln 2$. The two descriptions define identical labels.

The correlation bounds are expressed through the following constants. For integers $i, j \ge 3$ put

$$r_{ij} = \operatorname{sech}\Big(\tfrac{1}{4}\ln\big((i-1)(j-1)\big)\Big) = \frac{2\,g_{ij}}{1+g_{ij}^2}, \qquad g_{ij} = \big((i-1)(j-1)\big)^{1/4}. \tag{5.3}$$

They are symmetric and decrease in each index. On the diagonal, $r_{pp} = \operatorname{sech}(\frac{1}{2}\ln(p-1)) = 2\sqrt{p-1}/p$; so $r_{33} = r_* = \frac{2\sqrt{2}}{3}$, $r_{44} = \frac{\sqrt{3}}{2}$ and $r_{55} = \frac{4}{5}$. Table 3 lists the values for $3 \le i \le j \le 5$.

The main statements of this section are the following two lemmas. The first covers the padded labels, and in particular the geometric labels of the cubic case.

Lemma 5.4 (edge correlations, padded labels)**.** *Let $p \ge 3$ and $R \ge 1$ be integers, let G be a graph with $\Delta(G) \le p$, and let (Y_v) be the padded labels of degree p (Definition 5.2*(a)*). Then every edge uv satisfies* $\operatorname{Corr}(Y_u, Y_v) \ge r_{pp}(1 - \frac{1}{R})$.

The second covers the mixed labels, where the bound depends on the degrees of the two ends of the edge. For $\Delta_0 \ge 3$ put $\ell_{\max} = \frac{1}{2}\ln(\Delta_0 - 1)$, so that all arc lengths of a graph with $\delta(G) \ge 3$ and $\Delta(G) \le \Delta_0$ lie in $[\ell_{\min}, \ell_{\max}]$, where $\ell_{\min} = \frac{1}{2}\ln 2$ as above. For $R > 2\ell_{\max}$ put

$$\varepsilon_R = \frac{2\ell_{\max}^2}{\ell_{\min}\,(R - 2\ell_{\max})}. \tag{5.4}$$

Lemma 5.5 (edge correlations, mixed labels)**.** *Let $\Delta_0 \geq 3$ and $R > 2\ell_{\max}$, let G be a graph with $\delta(G) \geq 3$ and $\Delta(G) \leq \Delta_0$, and let (Y_v) be the mixed labels (Definition 5.2*(b)*). Then every edge uv satisfies* $\operatorname{Corr}(Y_u, Y_v) \geq r_{d(u)d(v)}(1 - \varepsilon_R)$.

For $\Delta_0 = 5$ we have $\ell_{\max} = \ln 2$ and $2\ell_{\max}^2/\ell_{\min} = 4\ln 2$, so $\varepsilon_R = \frac{4\ln 2}{R - 2\ln 2}$; with $R = \frac{h}{2}\ln 2$ for an integer $h > 4$, also $R - 2\ell_{\max} = \frac{h-4}{2}\ln 2$ and $\varepsilon_R = \frac{8}{h-4}$. Since $\varepsilon_R \to 0$ as $R \to \infty$, every edge correlation is asymptotically at least the value r_{ij} determined by the degrees of its ends.

Lemma 5.6 (distinctness and locality)**.** *Let (Y_v) be labels as in Definition 5.2, in either regime. Put $\rho = R$ for padded labels and $\rho = \lfloor 2R/\ln 2 \rfloor$ for mixed labels. Then:*

(i) *$(Y_v)_{v \in V(G)}$ is a centered Gaussian vector,* $\operatorname{Var} Y_v = 1$ *and $\sigma_v \geq 1$ for every v;*

(ii) *the labels are pairwise distinct with probability* 1*;*

(iii) *$B(v) \subseteq B_\rho(v)$, so Y_v is a function of $(Z_w)_{w \in B_\rho(v)}$;*

(iv) *for every $t \in \mathbb{R}$, the events $\{v \in \partial^+ S_t\}$ and $\{v \in \partial^- S_t\}$ for the labeling Y are determined by $(Z_w)_{w \in B_{\rho+1}(v)}$.*

So every count of the form $\sum_v \mathbf{1}\{v \in \partial^+ S_t\}$ satisfies the hypothesis of Lemma 4.10 with $A_v = B_{\rho+1}(v)$. The radius ρ is fixed for fixed R; the sweeps let R grow slowly with n, so that $\varepsilon_R \to 0$ while the dependence radius stays small enough for concentration.

5.2 Distinctness and locality

Proof of Lemma 5.6. (i) Each X_v is a finite linear combination of the independent standard normal variables Z_w, so (X_v) and (Y_v) are centered Gaussian vectors. Since $v \in B(v)$ with coefficient $e^{-D(v,v)} = 1$, $\sigma_v^2 \geq 1$, and $\operatorname{Var} Y_v = 1$ by construction.

(ii) Let $u \neq v$. The coefficient of Z_u in $Y_u - Y_v$ is $1/\sigma_u - \mathbf{1}\{u \in B(v)\}e^{-D(v,u)}/\sigma_v$, and the coefficient of Z_v is $\mathbf{1}\{v \in B(u)\}e^{-D(u,v)}/\sigma_u - 1/\sigma_v$. If both vanish, then $u \in B(v)$ and $\sigma_v = \sigma_u e^{-D(v,u)}$, and $v \in B(u)$ and $\sigma_v = \sigma_u e^{D(u,v)}$. Hence $D(u,v) + D(v,u) = 0$, which is impossible because $u \neq v$ and all arc lengths are positive. So $Y_u - Y_v$ is a normal variable with positive variance, and $\mathbb{P}(Y_u = Y_v) = 0$. A union bound over the finitely many pairs gives the claim.

(iii) For padded labels $B(v) = B_R(v)$. For mixed labels, every arc has length at least $\ell_{\min} = \frac{1}{2}\ln 2$ because $\delta(G) \geq 3$, so a walk of length at most R has at most $R/\ell_{\min}$ arcs, and $D(v,w) \leq R$ implies $\operatorname{dist}(v,w) \leq \lfloor R/\ell_{\min} \rfloor = \lfloor 2R/\ln 2 \rfloor$.

(iv) By definition, $v \in \partial^+ S_t$ if and only if $Y_v > t$ and $Y_u \leq t$ for some $u \in N(v)$, and $v \in \partial^- S_t$ if and only if $Y_v \leq t$ and $Y_u > t$ for some $u \in N(v)$. Both are determined by $(Y_x)_{x \in N[v]}$, and by (iii) each such Y_x is a function of Z restricted to $B_\rho(x) \subseteq B_{\rho+1}(v)$. □

5.3 A ratio inequality

Both correlation lemmas reduce to the following inequality for the shared part of two labels: the weights a_w and b_w that the two labels give to the same variable Z_w.

Lemma 5.7. *Let $\alpha, \beta > 0$, let I be a finite set, and let $a_w, b_w > 0$ $(w \in I)$ satisfy $b_w/\alpha \leq a_w \leq \beta b_w$. Then, with $A = \sum_{w \in I} a_w^2$ and $B = \sum_{w \in I} b_w^2$,*

$$\sum_{w \in I} a_w b_w \;\geq\; \operatorname{sech}\Big(\tfrac{1}{2}\ln(\alpha\beta)\Big)\sqrt{AB}.$$

Proof. Since $1/\alpha \le a_w/b_w \le \beta$, the inequality of Diaz and Metcalf [DM63, Theorem 1] gives the first inequality below (it is the sum of the inequalities $(a_w - b_w/\alpha)(\beta b_w - a_w) \ge 0$), and the second is the inequality of arithmetic and geometric means:

$$\Big(\beta + \frac{1}{\alpha}\Big) \sum_{w \in I} a_w b_w \;\ge\; A + \frac{\beta}{\alpha} B \;\ge\; 2\sqrt{\frac{\beta}{\alpha}}\,\sqrt{AB}.$$

Finally $2\sqrt{\beta/\alpha}/(\beta + 1/\alpha) = 2\sqrt{\alpha\beta}/(\alpha\beta + 1) = 2/(\sqrt{\alpha\beta} + 1/\sqrt{\alpha\beta}) = \operatorname{sech}(\frac{1}{2}\ln(\alpha\beta))$. □

For an edge uv we apply the lemma to $a_w = e^{-D(u,w)}$ and $b_w = e^{-D(v,w)}$ on the common support $I = B(u) \cap B(v)$. By (5.1) and its counterpart with u and v exchanged, $a_w \ge e^{-\ell(u \to v)} b_w$ and $b_w \ge e^{-\ell(v \to u)} a_w$; so the hypothesis holds with $\alpha = e^{\ell(u \to v)}$ and $\beta = e^{\ell(v \to u)}$, and $\frac{1}{2}\ln(\alpha\beta) = \frac{1}{2} L_{uv}$ with $L_{uv} = \ell(u \to v) + \ell(v \to u)$. Because coefficients outside I are paired with zero coefficients,

$$\operatorname{Cov}(X_u, X_v) = \sum_{w \in I} a_w b_w \;\ge\; \operatorname{sech}(\tfrac{1}{2} L_{uv}) \sqrt{AB}, \qquad A = \sum_{w \in I} a_w^2, \quad B = \sum_{w \in I} b_w^2. \tag{5.5}$$

For padded labels $L_{uv} = \ln(p-1)$ and $\operatorname{sech}(\frac{1}{2}L_{uv}) = r_{pp}$. For mixed labels $L_{uv} = \frac{1}{2}\ln((i-1)(j-1))$ with $i = d(u)$, $j = d(v)$, and $\operatorname{sech}(\frac{1}{2}L_{uv}) = r_{ij}$. It remains to show that A and B are almost all of σ_u^2 and σ_v^2: the outer shells $B(u) \setminus I$ and $B(v) \setminus I$ carry little squared mass.

5.4 Padded labels: proof of Lemma 5.4

For $p = 3$ these are the geometric labels, and the lemma is Lemma 3.4; the proof of that lemma applies to every p with the base 2 replaced by $p - 1$.

Proof of Lemma 5.4. Write $q = p - 1 \ge 2$ and let uv be an edge. As in the proof of Lemma 3.4, a vertex at distance $j \ge 1$ from u has at most q neighbors at distance $j + 1$, so the sphere masses $s_j q^{-j}$ are nonincreasing in $j \ge 1$, and the outermost one carries at most a fraction $\frac{1}{R}$ of $\sigma_u^2 = 1 + \sum_{j=1}^{R} s_j q^{-j}$. Every vertex of $B_R(u) \setminus B_R(v)$ lies on the sphere of radius R around u. Hence $A \ge (1 - \frac{1}{R})\sigma_u^2$ in (5.5), and likewise $B \ge (1 - \frac{1}{R})\sigma_v^2$. Since $\operatorname{sech}(\frac{1}{2}\ln q) = 2\sqrt{q}/(q+1) = r_{pp}$, (5.5) gives $\operatorname{Corr}(Y_u, Y_v) \ge r_{pp}(1 - \frac{1}{R})$. □

No step uses regularity, girth, connectivity or a lower bound on the degrees: vertices of degree less than p only make the sphere bound $s_{j+1} \le q s_j$ slack, and vertices in different components have disjoint supports.

5.5 Mixed labels: proof of Lemma 5.5

For mixed labels the balls $B(v)$ are defined by the directed distance, whose spheres are not layers of the graph. The sphere argument is therefore replaced by a cut argument in a shortest-path tree rooted at one end u of the edge uv. Step 1 locates the outer shell in a layer of width $\ell_{\max}$ below the boundary of the ball. Steps 2 and 3 show, using the critical decay (5.2), that the squared mass of the cut at directed distance t from u is nonincreasing in t. Step 4 integrates the cut mass over t and compares the shell with the whole ball.

Proof of Lemma 5.5. Let uv be an edge with $d(u) = i$ and $d(v) = j$. All arc lengths lie in $[\ell_{\min}, \ell_{\max}]$, because $3 \le d(x) \le \Delta_0$ for every x. By (5.5), it suffices to show

$$\sum_{w \in B(u) \setminus B(v)} e^{-2D(u,w)} \le \varepsilon_R\, \sigma_u^2 \tag{5.6}$$

and the same statement with u and v exchanged. Indeed, then $A \ge (1-\varepsilon_R)\sigma_u^2$ and $B \ge (1-\varepsilon_R)\sigma_v^2$, and (5.5) gives $\mathrm{Corr}(Y_u, Y_v) \ge r_{ij}(1-\varepsilon_R)$ if $\varepsilon_R \le 1$; if $\varepsilon_R > 1$ the claim holds because $\mathrm{Cov}(X_u, X_v) \ge 0$, all coefficients being nonnegative. We prove (5.6); the other statement is symmetric, using $\ell(u \to v) \le \ell_{\max}$. Throughout, write $D(w) = D(u, w)$.

Step 1: the shell is thin. If $w \in B(u) \setminus B(v)$, then $D(v, w) > R$, and $D(v, w) \le \ell(v \to u) + D(w)$ by (5.1); hence

$$R - \ell_{\max} < D(w) \le R. \tag{5.7}$$

Step 2: a shortest-path tree. Let V_u be the set of vertices in the component of u. For $w \in V_u \setminus \{u\}$ choose a shortest u–w path and let $p(w)$ be the vertex before w on it. Its initial segment is a walk from u to $p(w)$, so $D(p(w)) \le D(w) - \ell(p(w) \to w)$, and the triangle inequality gives the reverse; hence

$$D(w) = D(p(w)) + \ell(p(w) \to w), \qquad D(p(w)) < D(w).$$

Following p from any w strictly decreases D and therefore reaches u. So the pairs $(p(w), w)$ form a tree $\mathcal{T}$ on V_u rooted at u, in which $p(w)$ is the parent of w. Write $\mathrm{ch}(x) = \{w : p(w) = x\}$ for the children of x. For $x \ne u$, the children of x are neighbors of x with larger D, so $\mathrm{ch}(x) \subseteq N(x) \setminus \{p(x)\}$, and by (5.2)

$$\sum_{c \in \mathrm{ch}(x)} e^{-2D(c)} = e^{-2D(x)} \sum_{c \in \mathrm{ch}(x)} e^{-2\ell(x \to c)} \le e^{-2D(x)} \qquad (x \ne u). \tag{5.8}$$

Call a set $\mathcal{A}$ of vertices an *antichain* if no element of $\mathcal{A}$ is a proper descendant of another in $\mathcal{T}$. We claim: for every $x \ne u$ and every antichain $\mathcal{A}$ of descendants of x (including x itself),

$$\sum_{w \in \mathcal{A}} e^{-2D(w)} \le e^{-2D(x)}. \tag{5.9}$$

We induct on the height of the subtree of x. If $x \in \mathcal{A}$, then $\mathcal{A} = \{x\}$. Otherwise $\mathcal{A}$ is the disjoint union, over $c \in \mathrm{ch}(x)$, of antichains of descendants of c, and by induction and (5.8) its mass is at most $\sum_c e^{-2D(c)} \le e^{-2D(x)}$.

Step 3: moving cuts. For $t > 0$ let

$$\mathcal{A}_t = \{w \in V_u \setminus \{u\} : \ D(p(w)) < t \le D(w)\}, \qquad m(t) = \sum_{w \in \mathcal{A}_t} e^{-2D(w)}.$$

We show that m is nonincreasing on $(0, \infty)$. Let $0 < t < t'$ and $w \in \mathcal{A}_{t'}$. Along the path $u = x_0, x_1, \dots, x_k = w$ in $\mathcal{T}$, D increases strictly from $0 < t$ to $D(w) \ge t' > t$, so there is exactly one index $s \ge 1$ with $D(x_{s-1}) < t \le D(x_s)$; that is, w has exactly one ancestor-or-self in $\mathcal{A}_t$. For $x \in \mathcal{A}_t$, the set of $w \in \mathcal{A}_{t'}$ whose ancestor-or-self in $\mathcal{A}_t$ is x is an antichain of descendants of x: if $w' \in \mathcal{A}_{t'}$ were a proper descendant of $w \in \mathcal{A}_{t'}$, then $D(p(w')) \ge D(w) \ge t'$, contradicting $w' \in \mathcal{A}_{t'}$. By (5.9),

$$m(t') = \sum_{x \in \mathcal{A}_t} \sum_{\substack{w \in \mathcal{A}_{t'} \\ w \text{ below } x}} e^{-2D(w)} \le \sum_{x \in \mathcal{A}_t} e^{-2D(x)} = m(t).$$

Step 4: integrating the cut. Each $w \in V_u \setminus \{u\}$ belongs to $\mathcal{A}_t$ exactly for t in the interval $J_w = (D(p(w)), D(w)]$, whose length $\ell(p(w) \to w)$ lies in $[\ell_{\min}, \ell_{\max}]$. Since the sums are finite, for $0 \le a < b$

$$\int_a^b m(t)\, dt = \sum_{w \in V_u \setminus \{u\}} e^{-2D(w)} \, \big|J_w \cap (a, b]\big|. \tag{5.10}$$

Put $T = R - 2\ell_{\max} > 0$.

- *Upper bound on $m(T)$.* A vertex w contributes to (5.10) with $(a,b] = (0,T]$ only if $D(p(w)) < T$, and then $D(w) < T + \ell_{\max} \le R$, so $w \in B(u) \setminus \{u\}$; its contribution is at most $\ell_{\max} e^{-2D(w)}$. Since m is nonincreasing,

$$T\, m(T) \le \int_0^T m(t)\, dt \le \ell_{\max} \sum_{w \in B(u)\setminus\{u\}} e^{-2D(w)} \le \ell_{\max}\, \sigma_u^2.$$

- *The shell as an integral.* Every w in the shell $B(u) \setminus B(v)$ satisfies (5.7), so $w \ne u$ and $D(p(w)) \ge D(w) - \ell_{\max} > R - 2\ell_{\max} = T$. Hence $J_w \subseteq (T,R]$, and w contributes $e^{-2D(w)}|J_w| \ge \ell_{\min} e^{-2D(w)}$ to (5.10) with $(a,b] = (T,R]$. All other contributions are nonnegative, and $m(t) \le m(T)$ for $t > T$, so

$$\ell_{\min} \sum_{w \in B(u)\setminus B(v)} e^{-2D(w)} \le \int_T^R m(t)\, dt \le 2\ell_{\max}\, m(T).$$

Combining the two bounds,

$$\sum_{w \in B(u)\setminus B(v)} e^{-2D(w)} \le \frac{2\ell_{\max}}{\ell_{\min}} \cdot \frac{\ell_{\max}\sigma_u^2}{R - 2\ell_{\max}} = \varepsilon_R\, \sigma_u^2$$

by the definition (5.4) of ε_R, which is (5.6). □

The proof uses the degrees only through the critical decay (5.2) and the bounds $\ell_{\min} \le \ell \le \ell_{\max}$ on the arc lengths. It therefore applies to any lengths with these two properties.

6 Star bounds

A vertex v lies in the outer frontier $\partial^+ S_\tau$ of the sweep exactly when $Y_v > \tau$ and some neighbor u of v has $Y_u \le \tau$. This section bounds the probability of this event from above in terms of lower bounds on the correlations between Y_v and the labels Y_u of the neighbors alone, which is all that the labels of Section 5 provide. These bounds for a single vertex are the input of the plain sweep (Section 7), of the cubic sweep (Section 8) and of the certificate (Appendix A.2). The technical overview explains the route (Section 2.3).

Section 6.1 defines stars and bounds the frontier probability by the center bound (Lemma 6.2), an integral over the label of the center whose integrand involves the infimum $P_{\min}$ of an orthant probability over all correlation matrices. Section 6.2 shows that for equal thresholds and three to five coordinates this infimum is attained by the regular simplex vector (the simplex theorem, Lemma 6.5), and Section 6.3 defines the ideal stars and encloses F_3 and F_4 (Definition 6.7 and Proposition 6.9). The bounds on $P_{\min}$ for unequal thresholds (Lemmas A.1 and A.2), the simplex recursion (Lemma A.3) and the proof of Proposition 6.9 are in Appendices A.2.1 and A.2.2.

The notation and conventions of Section 4.4 are in force; in particular $s_r = \sqrt{1-r^2}$. All coordinates below are standard normal, so strict and weak inequalities between a coordinate and a constant may be exchanged.

6.1 Stars and the reduction to the center

Definition 6.1 (stars)**.** Let $d \ge 1$ and $r = (r_1, \dots, r_d) \in (0,1)^d$. A *$d$-star with lower bounds r* is a centered Gaussian vector $(Y_0, Y_1, \dots, Y_d)$ with $\operatorname{Var} Y_k = 1$ for $0 \le k \le d$ and $\operatorname{Corr}(Y_0, Y_i) \ge r_i$ for $1 \le i \le d$. The coordinate Y_0 is the *center*. For a threshold $\tau \in \mathbb{R}$, the *frontier event* of the star is $\{Y_0 > \tau,\ \min_{1\le i\le d} Y_i \le \tau\}$.

No assumption is made on the correlations $\mathrm{Corr}(Y_i, Y_j)$ between two non-central coordinates; they may be negative, and the covariance matrix may be singular. In the sweep, Y_0 is the label of a vertex and $Y_1, \dots, Y_d$ are the labels of its neighbors; the frontier event is the event that the vertex lies in $\partial^+ S_\tau$.

Since $\{Y_0 > \tau\}$ is the disjoint union of the frontier event and the event $\{Y_k > \tau \text{ for } 0 \le k \le d\}$,

$$\mathbb{P}\big(Y_0 > \tau,\ \min_i Y_i \le \tau\big) = \Phi(-\tau) - \mathbb{P}\big(Y_k > \tau \text{ for } 0 \le k \le d\big). \tag{6.1}$$

So an upper bound on the frontier probability is the same as a lower bound on the probability that the whole star lies above τ. We now fix notation for the quantities that the lower bound will involve. For lower bounds $r \in (0,1)^d$, a threshold τ and $z \in \mathbb{R}$, put

$$s_i = \sqrt{1 - r_i^2}, \qquad a_i(z) = a_i(z; \tau, r) = \frac{\tau - r_i z}{s_i} \qquad (1 \le i \le d).$$

Each $a_i(z)$ is strictly decreasing in z, because $r_i > 0$. For $a \in \mathbb{R}^d$ let

$$P_{\min}(a) = P^{(d)}_{\min}(a) := \inf\Big\{\mathbb{P}\big(U_i > a_i \text{ for } 1 \le i \le d\big)\Big\},$$

where the infimum is over all centered Gaussian vectors $U = (U_1, \dots, U_d)$ with unit variances (equivalently, over all $d \times d$ correlation matrices, singular ones included). Finally, the *center bound* is

$$C_d(\tau; r) := \Phi(-\tau) - \int_\tau^\infty \varphi(z)\, P_{\min}\big(a(z; \tau, r)\big)\, dz. \tag{6.2}$$

By Lemma 6.3(a) below, $P_{\min}$ is continuous, so the integrand is continuous in z and the integral is well defined.

Lemma 6.2 (reduction to the center)**.** *Let $(Y_0, \dots, Y_d)$ be a d-star with lower bounds $r \in (0,1)^d$, and let $\tau \in \mathbb{R}$. There is a centered Gaussian vector $U = (U_1, \dots, U_d)$ with unit variances such that*

$$\mathbb{P}\big(Y_k > \tau \text{ for } 0 \le k \le d\big) \ \ge\ \int_\tau^\infty \varphi(z)\, \mathbb{P}\big(U_i > a_i(z; \tau, r) \text{ for } 1 \le i \le d\big)\, dz. \tag{6.3}$$

Consequently

$$\mathbb{P}\big(Y_0 > \tau,\ \min_{1 \le i \le d} Y_i \le \tau\big) \le C_d(\tau; r).$$

The proof has three steps. A coupling replaces each neighbor label by one whose correlation with the center equals its lower bound; this lowers the correlations between neighbors as well. By Slepian's inequality, the probability that the whole star lies above τ can then only decrease. Conditioning on the center finally gives (6.3). A separate lower bound for each neighbor is needed for graphs with several degrees, where the edges at one vertex have different correlation bounds.

Proof. Put $Z = Y_0$ and $\rho_i = \mathrm{Corr}(Y_0, Y_i)$, so $0 < r_i \le \rho_i \le 1$.

Coupling. Take independent standard normal variables $\xi_1, \dots, \xi_d$, independent of Y, and put

$$Y_0' = Z, \qquad Y_i' = Y_i - (\rho_i - r_i)\, Z + \sqrt{\rho_i^2 - r_i^2}\, \xi_i \qquad (1 \le i \le d).$$

The square root is real because $\rho_i \ge r_i > 0$. The vector Y' is centered Gaussian. A direct computation gives

$$\begin{aligned}
\mathrm{Var}\, Y_i' &= 1 - 2(\rho_i - r_i)\rho_i + (\rho_i - r_i)^2 + (\rho_i^2 - r_i^2) = 1, \\
\mathrm{Cov}(Y_0', Y_i') &= \rho_i - (\rho_i - r_i) = r_i \le \rho_i = \mathrm{Cov}(Y_0, Y_i), \\
\mathrm{Cov}(Y_i', Y_j') &= \mathrm{Cov}(Y_i, Y_j) - \rho_i\rho_j + r_i r_j \le \mathrm{Cov}(Y_i, Y_j) \qquad (i \ne j),
\end{aligned}$$

the last inequality because $0 < r_i \le \rho_i$ and $0 < r_j \le \rho_j$. The correlations $\mathrm{Cov}(Y_i, Y_j)$ between neighbors, which may be negative, enter only through the common term.

Comparison. The vectors Y and Y' have the same variances, and every off-diagonal covariance of Y is at least the corresponding one of Y'. By Slepian's inequality (Lemma 4.8; singular covariance matrices are allowed),

$$\mathbb{P}(Y_k > \tau \ \forall k) \ \ge \ \mathbb{P}(Y'_k > \tau \ \forall k).$$

Conditioning on the center. Put $U_i = (Y'_i - r_i Z)/s_i$. Since $Y'_i - r_i Z = Y_i - \rho_i Z + \sqrt{\rho_i^2 - r_i^2}\,\xi_i$ has variance $(1-\rho_i^2) + (\rho_i^2 - r_i^2) = s_i^2$ and is uncorrelated with Z, the vector $U = (U_1, \dots, U_d)$ is centered Gaussian with unit variances, (Z, U) is jointly Gaussian, and U is independent of Z. As $Y'_i = r_i Z + s_i U_i$, Fubini's theorem gives

$$\mathbb{P}(Y'_k > \tau \ \forall k) = \int_\tau^\infty \varphi(z)\, \mathbb{P}\big(r_i z + s_i U_i > \tau \ \forall i\big)\, dz = \int_\tau^\infty \varphi(z)\, \mathbb{P}\big(U_i > a_i(z;\tau,r) \ \forall i\big)\, dz.$$

This proves (6.3). The integrand is at least $\varphi(z) P_{\min}(a(z;\tau,r))$ by the definition of $P_{\min}$, and the second claim follows from (6.1). □

The center bound depends continuously on the lower bounds, with an explicit modulus. The labels of Section 5 give the lower bounds only up to a factor $1 - \varepsilon_R$ (Lemma 5.5), and the following estimate is how this factor is removed in the limit $R \to \infty$.

Lemma 6.3 (continuity of the center bound)**.** *Let $d \ge 1$ and write $P_{\min} = P_{\min}^{(d)}$. Then:*

(a) *$P_{\min}$ is nonincreasing in each coordinate, and $|P_{\min}(a) - P_{\min}(a')| \le \varphi(0) \sum_{i=1}^d |a_i - a'_i|$ for all $a, a' \in \mathbb{R}^d$.*

(b) *Let $\bar r \in (0,1)$, let $r, r' \in (0, \bar r]^d$ with $|r_i - r'_i| \le \varepsilon$ for all i, and let $\tau \in \mathbb{R}$. Then*

$$|C_d(\tau; r) - C_d(\tau; r')| \le d\, \varphi(0)\, (1-\bar r^2)^{-3/2} \Big(|\tau| + \sqrt{2/\pi} \Big)\, \varepsilon.$$

In particular, if $d \le 5$, $\bar r = r_{33} = \frac{2\sqrt2}{3}$ and $|\tau| \le 3$, then $|C_d(\tau; r) - C_d(\tau; r')| \le 205\, \varepsilon$.

Proof. (a) Fix a centered Gaussian vector U with unit variances. Raising a threshold shrinks the event $\{U_i > a_i \ \forall i\}$, so its probability is nonincreasing in each a_i; and the two events for a and a' differ only on $\bigcup_i \{U_i \text{ lies between } a_i \text{ and } a'_i\}$, which has probability at most $\sum_i |\Phi(a_i) - \Phi(a'_i)| \le \varphi(0) \sum_i |a_i - a'_i|$. An infimum of functions that are all nonincreasing in each coordinate and all satisfy the same Lipschitz bound has the same two properties.

(b) For $\rho \in (0, \bar r]$ and fixed τ, z, the map $\rho \mapsto (\tau - \rho z)/\sqrt{1-\rho^2}$ has derivative $(\rho\tau - z)/(1-\rho^2)^{3/2}$, whose absolute value is at most $(1-\bar r^2)^{-3/2}(|\tau| + |z|)$. By the mean value theorem, $|a_i(z;\tau,r) - a_i(z;\tau,r')| \le (1-\bar r^2)^{-3/2}(|\tau| + |z|)\, \varepsilon$ for each i. By part (a),

$$\begin{aligned} |C_d(\tau; r) - C_d(\tau; r')| &\le \int_\tau^\infty \varphi(z)\, d\, \varphi(0)\, (1-\bar r^2)^{-3/2}(|\tau| + |z|)\, \varepsilon\, dz \\ &\le d\, \varphi(0)\, (1-\bar r^2)^{-3/2}(|\tau| + \mathbb{E}|N|)\varepsilon, \end{aligned}$$

where N is standard normal and $\mathbb{E}|N| = \sqrt{2/\pi}$. For $\bar r = 2\sqrt2/3$ we have $1 - \bar r^2 = \frac19$ and $(1-\bar r^2)^{-3/2} = 27$. With $d \le 5$ and $|\tau| \le 3$ the constant is at most $5 \cdot 0.39895 \cdot 27 \cdot (3 + 0.79789) < 204.6$. □

6.2 The simplex theorem

The infimum $P_{\min}$ in the center bound is over all correlation matrices. For equal thresholds it is attained by the most "spread out" matrix, that of the regular simplex. This subsection defines the regular simplex vector and states this claim for three to five coordinates (Lemma 6.5), as cases of a theorem of Mulgund [Mul26].

For $d \geq 2$, let $w_1, \dots, w_d \in \mathbb{R}^{d-1}$ be unit vectors with $\langle w_i, w_j \rangle = -\frac{1}{d-1}$ for $i \neq j$ (the vertices of a regular simplex inscribed in the unit sphere S^{d-2}), so that $\sum_i w_i = 0$. Let G be a standard Gaussian vector in $\mathbb{R}^{d-1}$. The *regular simplex vector* is $U^* = (U_1^*, \dots, U_d^*)$ with $U_i^* = \langle w_i, G \rangle$. It is centered, with unit variances and all pairwise correlations $-\frac{1}{d-1}$, and $\sum_i U_i^* = \langle \sum_i w_i, G \rangle = 0$. Its *orthant function* and the reflection of the latter are

$$S_d(a) := \mathbb{P}\big(U_i^* > a \text{ for } 1 \leq i \leq d\big), \qquad T_d(h) := S_d(-h) = \mathbb{P}\big(\max_i U_i^* < h\big) \qquad (a, h \in \mathbb{R}),$$

the last identity because $-U^*$ has the same distribution as U^*. Since the coordinates of U^* sum to zero, they cannot all exceed a number $a \geq 0$; hence

$$S_d(a) = 0 \quad \text{for } a \geq 0, \qquad T_d(h) = 0 \quad \text{for } h \leq 0. \tag{6.4}$$

The bound of Lemma A.2 for three unequal thresholds starts by lowering all correlations until the correlation matrix is singular.

Lemma 6.4 (rank reduction)**.** *Let $d \geq 2$ and let C be a nonsingular $d \times d$ correlation matrix. There is a singular $d \times d$ correlation matrix C' such that, for $U \sim N(0, C)$ and $U' \sim N(0, C')$, $\mathbb{P}(U_i > a_i \ \forall i) \geq \mathbb{P}(U_i' > a_i \ \forall i)$ for every $a \in \mathbb{R}^d$.*

Proof. Let J be the all-ones matrix and $C_t = C - t(J - I)$ for $t \geq 0$; it has unit diagonal and its off-diagonal entries are those of C minus t. The set $\{t \geq 0 : C_t \succeq 0\}$ is closed and convex, contains a neighborhood of 0 in $[0, \infty)$ (as $C \succ 0$), and is bounded, because $\mathbf{1}^{\mathsf{T}} C_t \mathbf{1} = \mathbf{1}^{\mathsf{T}} C \mathbf{1} - d(d-1)t$. So it is an interval $[0, t^*]$ with $t^* > 0$. The matrix $C' := C_{t^*}$ is positive semidefinite, and it is singular by the maximality of t^*: the least eigenvalue of C_t is continuous in t, nonnegative at t^* and negative for $t > t^*$. The off-diagonal entries of C' are at most those of C, so Slepian's inequality (Lemma 4.8) gives the claim. □

Lemma 6.5 (simplex theorem for $d \leq 5$)**.** *Let $d \in \{3, 4, 5\}$ and let $U = (U_1, \dots, U_d)$ be a centered Gaussian vector with unit variances. Then $\mathbb{P}(U_i > a \text{ for } 1 \leq i \leq d) \geq S_d(a)$ for every $a \in \mathbb{R}$; consequently $P_{\min}^{(d)}(a, \dots, a) = S_d(a)$.*

Proof. The second claim follows from the first, since U^* itself is admissible in the infimum.

Mulgund [Mul26, Theorem 1.1] proves: for every $n \geq 2$, every $n \times n$ correlation matrix Γ (positive semidefinite with unit diagonal, singular matrices included) and every $t \in \mathbb{R}$,

$$\mathbb{P}_{X \sim N(0, \Sigma_n)}\big(\max_i X_i \leq t\big) \leq \mathbb{P}_{X \sim N(0, \Gamma)}\big(\max_i X_i \leq t\big),$$

where Σ_n has unit diagonal and off-diagonal entries $-\frac{1}{n-1}$. Apply this with $n = d$, $X = -U$ (whose correlation matrix is that of U) and $t = -a$. Since $-U^*$ has the distribution $N(0, \Sigma_d)$, this gives $\mathbb{P}(U_i \geq a \ \forall i) \geq \mathbb{P}(U_i^* \geq a \ \forall i)$, and since every coordinate has a continuous distribution, the weak inequalities may be replaced by strict ones. □

Remark 6.6 (status of the simplex theorem). For $d = 3$ the inequality also has an elementary proof: condition on one coordinate, apply the Fréchet bound $\mathbb{P}(A \cap B) \geq \mathbb{P}A + \mathbb{P}B - 1$ to the other two, note that the bound is attained by a planar configuration with the same conditional marginals, and compare planar configurations by the convexity of the probability in the arcs between the directions. For $d = 4$ the inequality also follows from the four-cap theorem of L. Fejes Tóth [FT50] ([Bör04, Theorem 5.6.2]) by slicing the Gaussian measure along spheres centered at the origin, after lowering the correlations until the correlation matrix is singular (Lemma 6.4); see [BKT17, §§2–3]. For $d = 5$ no other proof is known to us. The paper [Mul26] is a preprint (version 2), whose author discloses substantial AI assistance. The three-variable case enters the proof of Theorem 8.1 through Proposition 8.8, the four-variable case only through vertices of degree at least 4, and the five-variable case only through the triples $\mathsf{P}, \mathsf{R}, \mathsf{S}$ of Theorem 7.2, and through S also Theorem 1.3, Corollary 1.4 and the results derived from them.

6.3 Ideal stars and the constants F_3 and F_4

When all lower bounds of a star are equal, the center bound is the frontier probability of an explicit star, and this star is the worst case at every threshold.

Definition 6.7 (ideal stars)**.** Let $d \in \{3, 4, 5\}$ and $r \in (0, 1)$. Let Z be standard normal and U^* the regular simplex vector in d variables, independent of Z. The *ideal d-star at correlation r* is

$$Y_0^* = Z, \qquad Y_i^* = rZ + s_r U_i^* \qquad (1 \leq i \leq d).$$

Its *ideal star function* is

$$f_{d,r}(\tau) := \mathbb{P}\big(Y_0^* > \tau,\ \min_i Y_i^* \leq \tau\big) \qquad (\tau \in \mathbb{R}).$$

Put

$$F_3 := \sup_{\tau \in \mathbb{R}} f_{3,r_{33}}(\tau), \qquad F_4 := \sup_{\tau \in \mathbb{R}} f_{4,r_{44}}(\tau), \qquad r_{33} = \tfrac{2\sqrt{2}}{3}, \quad r_{44} = \tfrac{\sqrt{3}}{2}.$$

The constants F_3 and F_4 are those of Theorems 7.1 and 7.2. The ideal star is a d-star with all lower bounds equal to r: its incident correlations are r and its other correlations are $r^2 - s_r^2/(d-1) = (dr^2 - 1)/(d-1)$. Since U^* is independent of Z and $Y_i^* > \tau$ means $U_i^* > (\tau - rZ)/s_r$, Fubini's theorem and (6.1) give

$$f_{d,r}(\tau) = \Phi(-\tau) - \int_\tau^\infty \varphi(z)\, S_d\Big(\frac{\tau - rz}{s_r}\Big) dz. \tag{6.5}$$

Comparing this with (6.2) and using $P_{\min}(a, \ldots, a) = S_d(a)$ (Lemma 6.5),

$$C_d(\tau; (r, \ldots, r)) = f_{d,r}(\tau) \qquad (d \in \{3, 4, 5\},\ r \in (0, 1),\ \tau \in \mathbb{R}). \tag{6.6}$$

In the cubic case, $r_{33} = r_* = \frac{2\sqrt{2}}{3}$ is the edge correlation bound $r_*(1 - \frac{1}{R})$ of the geometric labels in the limit $R \to \infty$ (Lemma 5.4), and $f_{3,r_*}(0)$ is the constant c_3 of Theorem 8.1 (Lemma 8.6).

Corollary 6.8 (ideal star bound)**.** *Let $d \in \{3, 4, 5\}$, $r \in (0, 1)$ and $0 \leq k \leq d$, and let $(Y_0, Y_1, \ldots, Y_k)$ be a centered Gaussian vector with unit variances and $\mathrm{Corr}(Y_0, Y_i) \geq r$ for $1 \leq i \leq k$. Then for every $\tau \in \mathbb{R}$,*

$$\mathbb{P}\big(Y_0 > \tau,\ Y_i \leq \tau \text{ for some } 1 \leq i \leq k\big) \leq f_{d,r}(\tau).$$

Proof. If $k < d$, add coordinates $Y_j = rY_0 + s_r\eta_j$ for $k < j \le d$, with fresh independent standard normal η_j. The extended vector is a d-star with lower bounds $(r, \dots, r)$, and the event can only grow. Now apply Lemma 6.2 and (6.6). □

In particular $f_{d,r} \le f_{d,r'}$ pointwise whenever $0 < r' \le r < 1$: apply Corollary 6.8 with lower bound r' to the ideal d-star at correlation r.

The bound $f_{4,r}$ is used with $r = r_{44}(1 - 1/R)$ for the padded degree-4 labels (Lemma 5.4); by (6.6) and Lemma 6.3(b), it exceeds $f_{4,r_{44}}$ by at most $178/R$ at every threshold in $[-3, 3]$.

Proposition 6.9 (enclosures of F_3 and F_4)**.** *The suprema F_3 and F_4 of Definition 6.7 satisfy*

$$0.140158252331452163693940514017 2 \;\le\; F_3 \;\le\; 0.1401582524,$$
$$0.2457719746988 \;\le\; F_4 \;\le\; 0.24577197475.$$

The proof is computer-assisted, in ball arithmetic with the Arb library [Joh17], from one-dimensional integral formulas for the ideal star functions; it is in Appendix A.2.2.

Remark 6.10. The numbers in Proposition 6.9 are used as follows: the lower bounds of F_3 and F_4 in tail estimates and in the exact check of the certificate (Appendix A.2), and the upper bound of F_3 in the certificate for the triples R and S, whose coefficient at degree 3 exceeds F_3. Numerically, F_3 is attained at $\tau \approx -0.06814$ and F_4 at $\tau \approx -0.16367$; both maximizers are negative, so in the plain sweep the binding thresholds lie slightly below the median.

7 The plain sweep: maximum degree 4 and 5

This section bounds the pathwidth of graphs of maximum degree 4 and 5 by a weighted count of their vertices of degree 3, 4 and 5, up to a term $o(n)$. The two theorems below are its results. In them, n_d is the number of vertices of degree d, and F_3 and F_4 are the maxima of the ideal star functions $f_{3,r_{33}}$ and $f_{4,r_{44}}$ (Definition 6.7).

Theorem 7.1 (maximum degree 4)**.** *Every graph G with $\Delta(G) \le 4$ satisfies* $\mathrm{pw}(G) \le a_3\, n_3 + a_4\, n_4 + o(n)$ *for each of the pairs*

(a) $(a_3, a_4) = (F_4,\, F_4)$*, and in particular* $\mathrm{pw}(G) \le F_4\, n + o(n)$*;*

(b) $(a_3, a_4) = (F_3,\, 0.2575)$*;*

(c) $(a_3, a_4) = (\frac{5}{32},\, \frac{1}{4})$.

Theorem 7.2 (maximum degree 5)**.** *Every graph G with $\Delta(G) \le 5$ satisfies* $\mathrm{pw}(G) \le a_3\, n_3 + a_4\, n_4 + a_5\, n_5 + o(n)$ *for each of the triples*

P $(a_3, a_4, a_5) = (F_3,\, 0.2575,\, 0.3667)$*;*

R $(a_3, a_4, a_5) = (0.1455,\, 0.2530,\, 0.3600)$*;*

S $(a_3, a_4, a_5) = (\frac{5}{32},\, \frac{1}{4},\, \frac{11}{32})$.

Parts (b) and (c) of Theorem 7.1 are the cases P and S of Theorem 7.2 for graphs without vertices of degree 5; we state them separately because later sections use them in this form. With the bootstrap of Section 9, the two theorems give the sequences F, P, R and S of Theorem 9.2.

The method is the *plain sweep*: draw the Gaussian labels of Section 5, order the vertices by label, and take the layout of Lemma 4.1, whose width is the largest outer frontier $|\partial^+ S_\tau|$ over all thresholds τ. It therefore suffices to find one realization of the labels in which every outer frontier has at most $\sum_d a_d n_d + o(n)$ vertices. The proof has three parts.

(1) *Expected frontier at a fixed threshold.* A vertex lies in $\partial^+ S_\tau$ exactly when the frontier event of its star occurs. So $\mathbb{E}|\partial^+ S_\tau|$ is a sum of star probabilities, each bounded by the center bound of Lemma 6.2. The center bound depends on the degree of the vertex and on the degrees of its neighbors, which we record as its *type*. A vertex of degree 3 next to vertices of higher degree has weaker correlation bounds than one in a cubic region, and its center bound exceeds a_3. The excess is moved along the edges to the neighbors of higher degree (Lemma 7.4). Checking that after these transfers the bound at every vertex is at most the coefficient of its degree, at every threshold, is a finite but large computation: the certificate of Proposition 7.5.

(2) *Removing the correlation loss.* The labels give the correlation lower bounds only up to a factor $1 - \varepsilon_R$, where R is the parameter of the labels. The explicit continuity of the center bound (Lemma 6.3) turns this factor into an additive error that vanishes as $R \to \infty$, uniformly in the threshold.

(3) *Concentration.* Frontier sizes at finitely many thresholds are sums of local indicators, so they concentrate (Lemma 7.9). A grid argument (Lemma 7.8) passes from these thresholds to all thresholds.

The proof of Theorem 7.1(a) has the same three parts, with the padded labels of degree 4 instead of the mixed labels used for the other bounds (Definition 5.2). There all vertices are treated alike: the frontier probability of every vertex is at most the ideal star function $f_{4,r}$ (Corollary 6.8), so no charging and no certificate are needed.

Setting. In Sections 7.1 to 7.3, H is a graph with $\delta(H) \geq 3$ and $\Delta(H) \leq 5$, and $Y = (Y_v)_{v \in V(H)}$ are the mixed labels of Definition 5.2(b) with parameter R. We use two facts from Section 5.

(L1) For every edge uv of H with $d(u) = i$ and $d(v) = j$, $\operatorname{Corr}(Y_u, Y_v) \geq r_{ij}(1 - \varepsilon_R)$, where $r_{ij} = \operatorname{sech}(\frac{1}{4}\ln((i-1)(j-1)))$ and $\varepsilon_R = \frac{4\ln 2}{R - 2\ln 2}$ for $R > 2\ln 2$ (Lemma 5.5 with $\Delta_0 = 5$); in particular $\varepsilon_R \to 0$ as $R \to \infty$.

(L2) The labels are pairwise distinct with probability 1, and each Y_v is a function of $(Z_w)_{w \in B_{\varrho_R}(v)}$, where $\varrho_R := \lfloor 2R/\ln 2 \rfloor$ is the locality radius ρ of Lemma 5.6 for mixed labels.

Since sech is decreasing on $[0, \infty)$, the largest of the numbers r_{ij} with $3 \leq i, j \leq 5$ is $r_{33} = \frac{2\sqrt{2}}{3}$.

7.1 Vertex types and the charging scheme

The center bound of a vertex depends only on its degree and on the degrees of its neighbors. We record this dependence as a type, define the redistribution of the excess, and isolate the inequalities that a coefficient vector must satisfy.

Definition 7.3 (types, reference bounds and admissible coefficients)**.** Let D be $\{3, 4\}$ or $\{3, 4, 5\}$.

(i) A *D-type* is a tuple $T = (d; k_3, k_4, k_5)$ of nonnegative integers with $d \in D$, $k_3 + k_4 + k_5 = d$, and $k_j = 0$ for $j \notin D$. In a graph whose degrees all lie in D, the *type* of a vertex v is $T(v) = (d(v); k_3, k_4, k_5)$, where k_j is the number of neighbors of v of degree j.

(ii) The *reference bounds* of T are the vector $r_T \in (0,1)^d$ consisting of k_3 copies of r_{d3}, k_4 copies of r_{d4} and k_5 copies of r_{d5}, and the *reference star bound* of T is $C_T(\tau) := C_d(\tau; r_T)$, the center bound (6.2).

(iii) For a *transfer vector* $t = (t_{34}, t_{35}, t_{45}) \in \mathbb{R}^3$, the *transfer term* of T is

$$\lambda_T(t) = \begin{cases} -k_4 t_{34} - k_5 t_{35}, & d = 3, \\ k_3 t_{34} - k_5 t_{45}, & d = 4, \\ k_3 t_{35} + k_4 t_{45}, & d = 5. \end{cases}$$

(iv) Coefficients $(a_d)_{d \in D}$ are *admissible for* D if for every $\tau \in [-3, 3]$ there is a transfer vector t with

$$C_T(\tau) + \lambda_T(t) \le a_d \qquad \text{for every } D\text{-type } T = (d; k_3, k_4, k_5).$$

Thus a vertex of degree i sends the amount t_{ij} along each edge to a neighbor of degree $j > i$. Admissibility says that, at every threshold in $[-3, 3]$, some choice of these transfers makes the reference star bound plus the transfer term of every type at most the coefficient of its degree. There are 10 types of degree 3, 15 of degree 4 and 21 of degree 5, so 46 types for $D = \{3, 4, 5\}$ and 9 for $D = \{3, 4\}$. The *pure cubic* type $T_0 = (3; 3, 0, 0)$ has $C_{T_0} = f_{3,r_{33}} \le F_3$ by (6.6) and Definition 6.7, and $\lambda_{T_0} = 0$.

Lemma 7.4 (charging)**.** *Let D be as in Definition 7.3, let H be a graph all of whose degrees lie in D, and let $t \in \mathbb{R}^3$. Then $\sum_{v \in V(H)} \lambda_{T(v)}(t) = 0$. Consequently, if real numbers x_v ($v \in V(H)$) and $e \ge 0$ satisfy $x_v + \lambda_{T(v)}(t) \le a_{d(v)} + e$ for every v, then*

$$\sum_{v \in V(H)} x_v \;\le\; \sum_{d \in D} a_d n_d(H) + e\,|V(H)|.$$

Proof. Write $\lambda_{T(v)}(t)$ as a sum over the edges at v: by the definition, an edge vu with $d(v) = i$ and $d(u) = j$ contributes $-t_{ij}$ if $i < j$, $+t_{ji}$ if $i > j$, and 0 if $i = j$. So each edge joining degrees $i < j$ contributes $-t_{ij}$ at its end of degree i and $+t_{ij}$ at its end of degree j, and each edge joining equal degrees contributes nothing. Summing over all vertices therefore gives 0. Summing the inequalities $x_v \le a_{d(v)} + e - \lambda_{T(v)}(t)$ over v gives the second claim. □

7.2 The certified coefficients

Admissibility is a statement about all thresholds in $[-3, 3]$ and about 46 types at once, each with a reference star bound that is an integral of an infimum over correlation matrices. We verify it for three coefficient vectors by a computer-assisted proof in ball arithmetic and exact rational arithmetic.

Proposition 7.5 (certified admissibility)**.** *Each of the triples*

$$\mathsf{P} = (F_3,\, 0.2575,\, 0.3667), \qquad \mathsf{R} = (0.1455,\, 0.2530,\, 0.3600), \qquad \mathsf{S} = (\tfrac{5}{32},\, \tfrac{1}{4},\, \tfrac{11}{32})$$

is admissible for $D = \{3, 4, 5\}$.

If (a_3, a_4, a_5) is admissible for $\{3, 4, 5\}$, then (a_3, a_4) is admissible for $\{3, 4\}$: every $\{3, 4\}$-type is a $\{3, 4, 5\}$-type with the same reference star bound, and its transfer term, $-k_4 t_{34}$ for $d = 3$ and $k_3 t_{34}$ for $d = 4$, is the same function of the same transfer vector. In particular $(F_3, 0.2575)$ and $(\frac{5}{32}, \frac{1}{4})$ are admissible for $\{3, 4\}$.

What is certified. For each of the three triples, the accompanying code partitions $[-3, 3]$ into finitely many cells $[t_0, t_1]$. For each cell it stores one transfer vector t with $|t_{ij}| \le 1$ and, for each type T, a rational number $I_T^{\mathrm{lo}}(t_1)$. It then verifies, in exact rational arithmetic, that

$$\widehat{\Phi}(-t_0) - I_T^{\mathrm{lo}}(t_1) + \lambda_T(t) < a_d \qquad \text{for every type } T \ne T_0, \tag{7.1}$$

where $\widehat{\Phi}(-t_0)$ is a rational upper bound for $\Phi(-t_0)$. The numbers $I_T^{\mathrm{lo}}(t_1)$ are lower bounds for a Riemann sum of the integral in (6.2) at $\tau = t_1$, in which $P_{\min}$ is replaced by a lower bound: the bound of Lemma A.1, or, for the seven degree-3 types with unequal reference bounds, the larger of that bound and the arc bound of Lemma A.2. Appendix A.2 proves that (7.1) implies $C_T(\tau) + \lambda_T(t) \le a_d$ for every $\tau \in [t_0, t_1]$ (Lemmas A.6 and A.7). The pure cubic type T_0 is not checked: $C_{T_0} \le F_3 \le a_3$ for all three triples, and $\lambda_{T_0} = 0$.

How. Every value of φ, Φ, S_d and of the arc bound that enters I^{lo} is a lower or upper bound obtained from an enclosure in Arb ball arithmetic [Joh17], rounded outward. The arc bound is evaluated through a closed form in Owen's T function [Owe56] (Lemma A.8). A separate program rechecks (7.1) for all cells in exact rational arithmetic. Floating point is used only to make choices, such as the transfer vectors; no floating-point value enters (7.1). For P, whose coefficient at degree 3 is F_3 itself, the recheck uses the lower bound of F_3 in Proposition 6.9 in place of a_3. For R and S, the upper bound of F_3 there gives $F_3 \le a_3$. Appendices A.2.4 and A.2.5 give the details, the numbers of cells and the smallest slacks.

Proof of Proposition 7.5. Fix one of the triples and $\tau \in [-3, 3]$; let $[t_0, t_1]$ be a cell of the certificate containing τ (the cells tile $[-3, 3]$, which the code checks), and t its stored transfer vector. For $T = T_0$, $C_{T_0}(\tau) + \lambda_{T_0}(t) \le F_3 \le a_3$. For every other type T of degree d, Lemmas A.6 and A.7 and the construction of I_T^{lo} described in Appendix A.2 give $C_T(\tau) \le \Phi(-t_0) - I_T^{\mathrm{lo}}(t_1)$, and (7.1) gives $C_T(\tau) + \lambda_T(t) < a_d$. □

Remark 7.6 (how close to the method's limit). The certified coefficients are slightly larger than what the star bounds allow pointwise, because a cell bound loses a term of first order in the cell width. For orientation, floating-point quadrature at sampled thresholds (not certified) gives the following values. For $D = \{3, 4\}$ with $a_3 = F_3$, the smallest possible a_4 is about 0.25659; it is attained near $\tau \approx -0.089$ by a degree-4 vertex all of whose neighbors have degree 3. For $D = \{3, 4, 5\}$ with $a_3 = F_3$, the binding configuration is a degree-5 vertex with five neighbors of degree 3, near $\tau \approx -0.105$, and the smallest a_5 is about 0.36621. The triple S lies about $5 \cdot 10^{-4}$ inside the corresponding floating-point optimum.

The method does not appear able to bring the coefficient of n_4 down to F_4. On $(3, 4)$-biregular graphs of large girth, every edge joins a vertex of degree 3 to one of degree 4. A numerical evaluation of the frontier density of the sweep on such graphs, with $a_3 = F_3$, requires a coefficient of about 0.2533 at degree 4, above $F_4 \approx 0.2458$.

7.3 Expected frontier sizes

We now combine the star bounds, the certificate and the correlation estimate of Section 5. The explicit continuity estimate of Lemma 6.3, which is uniform in the threshold, turns the factor $1 - \varepsilon_R$ in (L1) into an additive error of $205\,\varepsilon_R$ per vertex; no slack of the certificate is needed for this step.

Lemma 7.7 (expected frontiers in the mixed case)**.** *Let (a_3, a_4, a_5) be one of $\mathsf{P}, \mathsf{R}, \mathsf{S}$, let H be a graph all of whose degrees lie in $\{3, 4, 5\}$, put $\beta(H) = a_3 n_3(H) + a_4 n_4(H) + a_5 n_5(H)$, and let R be such that $\varepsilon_R < 1$. Then for every $\tau \in [-3, 3]$,*

$$\mathbb{E}\,|\partial^+ S_\tau| \;\le\; \beta(H) + 205\,\varepsilon_R\,|V(H)|.$$

Proof. Let v be a vertex of type $T = (d; k_3, k_4, k_5)$, with neighbors $u_1, \dots, u_d$. The vector $(Y_v, Y_{u_1}, \dots, Y_{u_d})$ is centered Gaussian with unit variances, and by (L1) it is a d-star with lower

bounds $r'_T := (1-\varepsilon_R)r_T$, after ordering the neighbors as in r_T. The event $\{v \in \partial^+ S_\tau\}$ is its frontier event. By Lemma 6.2,

$$\mathbb{P}(v \in \partial^+ S_\tau) \le C_d(\tau; r'_T).$$

Every entry of r_T and of r'_T lies in $(0, r_{33}]$, and the entries differ by $\varepsilon_R(r_T)_i \le \varepsilon_R$. By Lemma 6.3(b), since $|\tau| \le 3$ and $d \le 5$,

$$\mathbb{P}(v \in \partial^+ S_\tau) \le C_T(\tau) + 205\,\varepsilon_R.$$

By Proposition 7.5 there is a transfer vector t with $C_T(\tau) + \lambda_T(t) \le a_d$ for every $\{3,4,5\}$-type T. Hence $\mathbb{P}(v \in \partial^+ S_\tau) + \lambda_{T(v)}(t) \le a_{d(v)} + 205\,\varepsilon_R$ for every vertex v, and Lemma 7.4 with $x_v = \mathbb{P}(v \in \partial^+ S_\tau)$ and $e = 205\,\varepsilon_R$ gives the claim, since $\mathbb{E}|\partial^+ S_\tau| = \sum_v \mathbb{P}(v \in \partial^+ S_\tau)$. □

7.4 Concentration

Three lemmas turn bounds on expectations at finitely many thresholds into a bound for one realization at all thresholds. The first (Lemma 7.8) is deterministic: it bounds a frontier at any threshold by the frontier at a neighboring grid point plus the number of labels between the two, and outside the grid by the number of labels beyond its ends. The plain sweep uses only its statements about outer frontiers; the median-switched sweep of Section 8 also uses those about inner frontiers.

Lemma 7.8 (grid)**.** *Let y be any labeling of a graph G, let $\tau_0 < \tau_1 < \cdots < \tau_K$, and for $0 \le j < K$ let $N_j = |\{v : \tau_j < y_v \le \tau_{j+1}\}|$.*

(i) *If $\tau_j \le \tau < \tau_{j+1}$, then $|\partial^+ S_\tau| \le |\partial^+ S_{\tau_{j+1}}| + N_j$ and $|\partial^- S_\tau| \le |\partial^- S_{\tau_j}| + N_j$.*

(ii) *If $\tau < \tau_0$, then $|\partial^+ S_\tau| \le \Delta(G) \cdot |\{v : y_v \le \tau_0\}|$ and $|\partial^- S_\tau| \le |\{v : y_v \le \tau_0\}|$.*

(iii) *If $\tau \ge \tau_K$, then $|\partial^+ S_\tau| \le |\{v : y_v > \tau_K\}|$.*

Proof. Let $v \in \partial^+ S_\tau$, so $y_v > \tau$ and some neighbor u (a *witness*) has $y_u \le \tau$. (i) If $y_v > \tau_{j+1}$, then $v \in \partial^+ S_{\tau_{j+1}}$ with the same witness, because $y_u \le \tau < \tau_{j+1}$. Otherwise $\tau_j \le \tau < y_v \le \tau_{j+1}$, and v is counted by N_j. For the inner frontier, let $u \in \partial^- S_\tau$, with a neighbor x such that $y_u \le \tau < y_x$. If $y_u \le \tau_j$, then $u \in S_{\tau_j}$ and $y_x > \tau \ge \tau_j$, so $u \in \partial^- S_{\tau_j}$. Otherwise $\tau_j < y_u \le \tau < \tau_{j+1}$, and u is counted by N_j. (ii) Map v to a witness u. Then $y_u \le \tau < \tau_0$, and each u is the image of at most $d(u) \le \Delta(G)$ vertices. For the inner frontier, $\partial^- S_\tau \subseteq S_\tau \subseteq S_{\tau_0}$. (iii) $y_v > \tau \ge \tau_K$. □

The second lemma is the concentration step shared by all sweeps of this paper. It combines the variance bound of Lemma 4.10, Chebyshev's inequality and a union bound over the finitely many counts. The counts of a sweep are frontier sizes, window counts $|\{v : \tau_j < Y_v \le \tau_{j+1}\}|$ and tail counts. Each is a sum over the vertices v of indicators that depend only on the labels on $N[v]$, hence on $(Z_w)_{w \in B_{\rho+1}(v)}$, where ρ is the locality radius of Lemma 5.6.

Lemma 7.9 (concentration of local counts)**.** *Let G be a graph with N vertices and $\Delta(G) \le \Delta$, where $\Delta \ge 2$, let (Y_v) be labels of G as in Definition 5.2 with underlying variables $(Z_w)_{w \in V(G)}$, let $\varrho \ge 0$ be an integer, and let $W_1, \dots, W_\mu$ be sums over $v \in V(G)$ of $[0,1]$-valued measurable functions of $(Z_w)_{w \in B_\varrho(v)}$. If $\lambda > 0$ and $\mu N \Delta^{2\varrho} < 2\lambda^2$, then some realization of (Z_w) has pairwise distinct labels and $|W_i - \mathbb{E}W_i| < \lambda$ for all i.*

Proof. By Lemma 4.10 with $a = 1$ and $A_v = B_\varrho(v)$, each W_i satisfies $\mathbb{P}(|W_i - \mathbb{E}W_i| \ge \lambda) \le N\Delta^{2\varrho}/(2\lambda^2)$. By the union bound, the probability that some W_i deviates from its mean by at least λ is at most $\mu N \Delta^{2\varrho}/(2\lambda^2) < 1$. By Lemma 5.6(ii) the labels are pairwise distinct with probability 1. So the event that the labels are pairwise distinct and every W_i lies within λ of its mean has positive probability, and in particular it is nonempty. □

The third lemma is the plain sweep itself: bounds on the expected outer frontiers at the thresholds in $[-3,3]$ give a path decomposition. In it, β is the target width, e is the error per vertex in the expected frontiers, and the grid of thresholds has K points per unit length. The hypothesis $\Delta\,\Phi(-3)\,N \le \beta$ covers the thresholds outside $[-3,3]$, and the last hypothesis is the condition of Lemma 7.9 for the $12K+2$ counts used in the proof, with $\lambda = N^{7/8}$.

Lemma 7.10 (plain sweep from expected frontiers)**.** *Let G be a graph with N vertices and $\Delta(G) \le \Delta$ for some $\Delta \ge 2$, let (Y_v) be labels of G as in Definition 5.2 with locality radius ρ (Lemma 5.6), let $K \ge 1$ be an integer, and let $\beta, e \ge 0$. Suppose that $\mathbb{E}|\partial^+ S_\tau| \le \beta + eN$ for every $\tau \in [-3,3]$, that $\Delta\,\Phi(-3)\,N \le \beta$, and that $(12K+2)\,\Delta^{2\rho+2} < 2N^{3/4}$. Then*

$$\mathrm{pw}(G) \;\le\; \beta + \Big(e + \frac{\varphi(0)}{K}\Big)N + \Delta\, N^{7/8}.$$

Proof. We fix a realization in which the counts at a grid of thresholds are all close to their means, and then bound the outer frontier at every threshold by Lemma 7.8. Let $\tau_j = -3 + j/K$ for $0 \le j \le 6K$, and consider the $12K+2$ counts

$$A_j = |\partial^+ S_{\tau_j}| \;\; (1 \le j \le 6K), \qquad N_j = |\{v : \tau_j < Y_v \le \tau_{j+1}\}| \;\; (0 \le j < 6K),$$
$$C_- = |\{v : Y_v \le -3\}|, \qquad C_+ = |\{v : Y_v > 3\}|.$$

Each is a sum over v of indicators that depend only on the labels on $N[v]$, hence on $(Z_w)_{w \in B_{\rho+1}(v)}$ (Lemma 5.6(iii),(iv)). The last hypothesis is the condition of Lemma 7.9 with $\varrho = \rho+1$, $\mu = 12K+2$ and $\lambda = N^{7/8}$, so some realization has pairwise distinct labels and every count within $N^{7/8}$ of its mean. Fix it. Since each Y_v is standard normal, $\mathbb{E}N_j \le N\varphi(0)/K$ and $\mathbb{E}C_\pm = N\Phi(-3)$, and $\mathbb{E}A_j \le \beta + eN$ by hypothesis. Let $\tau \in \mathbb{R}$.

- If $\tau \in [\tau_j, \tau_{j+1})$ for some $j < 6K$, then by Lemma 7.8(i), $|\partial^+ S_\tau| \le A_{j+1} + N_j \le \beta + (e + \varphi(0)/K)N + 2N^{7/8}$.
- If $\tau < -3$, then by Lemma 7.8(ii), $|\partial^+ S_\tau| \le \Delta\, C_- \le \Delta\,\Phi(-3)N + \Delta N^{7/8} \le \beta + \Delta N^{7/8}$.
- If $\tau \ge 3$, then by Lemma 7.8(iii), $|\partial^+ S_\tau| \le C_+ \le \Phi(-3)N + N^{7/8} \le \beta + N^{7/8}$.

Since $\Delta \ge 2$, every outer frontier has at most $\beta + (e + \varphi(0)/K)N + \Delta N^{7/8}$ vertices, and Lemma 4.1 for the injective labeling of the fixed realization gives the claim ($N \ge 1$ by the last hypothesis). □

7.5 Proofs of Theorems 7.1 and 7.2

We first prove the mixed bounds for graphs of minimum degree at least 3, then remove this assumption with the reductions of Definition 4.2, and finally prove Theorem 7.1(a).

Proposition 7.11 (mixed bounds for minimum degree 3)**.** *Let (a_3, a_4, a_5) be one of $\mathsf{P}, \mathsf{R}, \mathsf{S}$. There is a function $\rho^*(N) = o(N)$ such that every graph H on N vertices all of whose degrees lie in $\{3,4,5\}$ satisfies*

$$\mathrm{pw}(H) \;\le\; a_3 n_3(H) + a_4 n_4(H) + a_5 n_5(H) + \rho^*(N).$$

Proof. We let the label parameter R grow slowly with N and apply Lemma 7.10 to the expected frontiers of Lemma 7.7. Write $\beta(H) = a_3 n_3(H) + a_4 n_4(H) + a_5 n_5(H)$. Fix R_0 with $\varepsilon_R < 1$ for $R \ge R_0$, and for $N \ge 1$ let

$$R(N) := \max\Big\{R \in \{R_0, \dots, N\} :\; (12R+2)\,25^{\varrho_R+1} \le N^{1/2}\Big\},$$

with $R(N) := 0$ if the set is empty. For each fixed $R \geq R_0$ the condition holds for all large N, so $R(N) \to \infty$ as $N \to \infty$. Put

$$\rho^*(N) := \begin{cases} (205\,\varepsilon_{R(N)} + \varphi(0)/R(N))N + 5N^{7/8}, & R(N) \geq R_0, \\ N, & \text{otherwise.} \end{cases}$$

Since $\varepsilon_R \to 0$ and $R(N) \to \infty$, $\rho^*(N) = o(N)$. If $R(N) = 0$ the claim holds because $\mathrm{pw}(H) \leq N - 1$. So let $R := R(N) \geq R_0$ and use the labels with parameter R. By Lemma 7.7, $\mathbb{E}|\partial^+ S_\tau| \leq \beta(H) + 205\,\varepsilon_R N$ for every $\tau \in [-3, 3]$. Every vertex of H has degree at least 3, and $a_5 \geq a_4 \geq a_3 \geq F_3 > 0.14$ for all three triples (Proposition 6.9), so $\beta(H) \geq 0.14N > 5\Phi(-3)N$, as $5\Phi(-3) < 0.0068$. Finally $(12R + 2)\,25^{\varrho_R+1} \leq N^{1/2} < 2N^{3/4}$. Hence Lemma 7.10 applies with $\Delta = 5$, the locality radius ϱ_R of (L2), $K = R$, $\beta = \beta(H)$ and $e = 205\,\varepsilon_R$, and gives $\mathrm{pw}(H) \leq \beta(H) + (205\,\varepsilon_R + \varphi(0)/R)N + 5N^{7/8} = \beta(H) + \rho^*(N)$. □

Proof of Theorem 7.2 and of Theorem 7.1(b),(c). Let (a_3, a_4, a_5) be one of $\mathsf{P}, \mathsf{R}, \mathsf{S}$, and put $a_0 = a_1 = a_2 = 0$; then $0 = a_0 \leq a_1 \leq \cdots \leq a_5$ in all three cases. Let $\bar\rho$ be the running maximum (4.2) of the function ρ^* of Proposition 7.11, with $\rho^*(0) := 0$; it is nonnegative, nondecreasing and $o(n)$. By Proposition 7.11, every graph J with $\delta(J) \geq 3$ and $\Delta(J) \leq 5$ satisfies $\mathrm{pw}(J) \leq \sum_d a_d n_d(J) + \bar\rho(|V(J)|)$. Then Corollary 4.6 with $k = 5$ gives, for every nonempty graph G with $\Delta(G) \leq 5$,

$$\mathrm{pw}(G) \leq a_3 n_3 + a_4 n_4 + a_5 n_5 + \bar\rho(n) + 3\log_2 n + 1,$$

which is Theorem 7.2; the remainder depends only on n. (That corollary reduces G by a maximal sequence of the operations of Definition 4.2 and applies Lemma 4.5 without deletions.)

Theorem 7.1(b) and (c) are the cases P and S of Theorem 7.2 for graphs with $\Delta(G) \leq 4$, for which $n_5 = 0$. In such graphs no vertex type involves degree 5, so only the certified inequalities of the $\{3, 4\}$-types and the transfer t_{34} are used; in particular the case $d = 5$ of Lemma 6.5 is not needed for Theorem 7.1. □

For Theorem 7.1(a) we use the padded labels of degree 4 (Definition 5.2(a) with $p = 4$) with an integer radius $R \geq 2$. Every arc has length $\frac{1}{2}\ln 3$, so that $X_v = \sum_{w:\,\mathrm{dist}(v,w) \leq R} 3^{-\mathrm{dist}(v,w)/2} Z_w$ and $Y_v = X_v/\sqrt{\mathrm{Var}\,X_v}$. By Lemma 5.4, every edge uv of a graph with $\Delta \leq 4$ has $\mathrm{Corr}(Y_u, Y_v) \geq r_R := r_{44}(1 - 1/R)$, and by Lemma 5.6 the locality radius is $\rho = R$. The bound $F_4 n + o(n)$ needs no reductions: vertices of every degree from 0 to 4 are handled by padding the star (Corollary 6.8).

Proposition 7.12 (fixed radius)**.** *Let $R \geq 2$ be an integer and let G be a graph with $\Delta(G) \leq 4$ and $n \geq 3^{8R}$ vertices. Then*

$$\mathrm{pw}(G) \leq \Big(F_4 + \frac{179}{R}\Big)n + 4n^{7/8}.$$

Proof. We apply Lemma 7.10 with the ideal star bound. Use the padded labels with radius R. For a vertex v of degree $k \leq 4$, the vector $(Y_v; Y_u, u \in N(v))$ has unit variances and incident correlations at least r_R. By Corollary 6.8 with $d = 4$, (6.6) and Lemma 6.3(b) with $\bar r = r_{33} > r_{44}$ and $\varepsilon = r_{44} - r_R = r_{44}/R$, for every $\tau \in [-3, 3]$,

$$\mathbb{P}(v \in \partial^+ S_\tau) \leq f_{4,r_R}(\tau) \leq f_{4,r_{44}}(\tau) + \frac{205\,r_{44}}{R} \leq F_4 + \frac{178}{R},$$

so $\mathbb{E}|\partial^+ S_\tau| \leq (F_4 + 178/R)n$ for every $\tau \in [-3, 3]$. We check the remaining hypotheses of Lemma 7.10 with $\Delta = 4$, $\rho = R$, $K = R$, $\beta = F_4 n$ and $e = 178/R$. First, $4\Phi(-3)n < 0.0055\,n < F_4 n$ by Proposition 6.9. Second, $8(12R + 2) < 45^R$ for every $R \geq 2$ (at $R = 2$ this reads $208 < 2025$, and

the left side less than doubles when R grows by one), so $(12R+2)\,16^{R+1} < 2\cdot 45^R\cdot 16^R \le 2\cdot 729^R = 2\cdot 3^{6R} \le 2n^{3/4}$, since $n \ge 3^{8R}$. Hence Lemma 7.10 gives $\mathrm{pw}(G) \le F_4 n + (178+\varphi(0))n/R + 4n^{7/8} \le (F_4 + 179/R)n + 4n^{7/8}$, using $\varphi(0) < 0.399$. □

Proof of Theorem 7.1(a). For $n \ge 3^{16}$ put $R(n) = \lfloor \frac{1}{8}\log_3 n \rfloor \ge 2$, so that $n \ge 3^{8R(n)}$, and $\rho(n) = \frac{179}{R(n)}n + 4n^{7/8}$; for $n < 3^{16}$ put $\rho(n) = n$. Since $R(n) \to \infty$, $\rho(n) = o(n)$, and so is its running maximum $\bar\rho$ (4.2). By Proposition 7.12 with $R = R(n)$ for $n \ge 3^{16}$ and trivially otherwise, $\mathrm{pw}(G) \le F_4 n + \rho(n)$ for every graph G with $\Delta(G) \le 4$, which is the bound $F_4 n + o(n)$. For the bound with coefficients (F_4, F_4), let J be a graph with $\delta(J) \ge 3$ and $\Delta(J) \le 4$; then $|V(J)| = n_3(J) + n_4(J)$, so $\mathrm{pw}(J) \le F_4\big(n_3(J)+n_4(J)\big) + \bar\rho(|V(J)|)$. Corollary 4.6 with $k = 4$ and $(a_0, \dots, a_4) = (0, 0, 0, F_4, F_4)$ gives $\mathrm{pw}(G) \le F_4 n_3 + F_4 n_4 + \bar\rho(n) + 3\log_2 n + 1$ for every nonempty graph G with $\Delta(G) \le 4$. □

8 The median-switched sweep: maximum degree 3 and 4

This section proves the following two theorems, where $c_3 = \frac{3}{4\pi}\arccos\frac{5}{6} \approx 0.1398$. With the bootstrap of Section 9 they give the sequences C and M of Theorem 9.2.

Theorem 8.1 (maximum degree 3)**.** *Every graph G with $\Delta(G) \le 3$ satisfies* $\mathrm{pw}(G) \le c_3\, n + o(n)$.

Theorem 8.2 (maximum degree 4)**.** *Every graph G with $\Delta(G) \le 4$ satisfies* $\mathrm{pw}(G) \le c_3\, n_3 + 0.2627\, n_4 + o(n)$.

Theorem 8.2 contains Theorem 8.1, but its proof needs more: a switch lemma for stars with unequal correlation bounds (Lemma A.10), which we prove with a computer-assisted certificate (Appendix A.3.2), and a computer-assisted charging at nonzero thresholds (Proposition A.13). The proof of Theorem 8.1 uses no computation.

Compared with the plain sweep of Section 7, only the layout changes (Section 2.4): vertices still enter in increasing order of their geometric labels, but the layout keeps *inner* frontiers while the threshold is negative and *outer* frontiers while it is positive, and passes from one to the other at threshold zero by a short sequence of bags, the *connector*.

After the notation of Section 8.1, the proof of Theorem 8.1 has four steps. Section 8.2 constructs the layout (Lemma 8.4), with an ordering lemma for the connector (Lemma 8.3). Section 8.3 bounds the frontier probabilities away from threshold zero (Proposition 8.8). Section 8.4 bounds the expected values of L_0 and L_0', which control the connector bags at threshold zero (the switch lemma, Lemma 8.9). Section 8.5 adds the concentration argument and completes the proof. The proof does not use Slepian's inequality: at nonnegative thresholds a pointwise coupling suffices, and negative thresholds are reduced to nonnegative ones by the symmetry $Y \mapsto -Y$.

The proof of Theorem 8.2, in Appendix A.3.1, uses the same layout with a larger set at threshold zero, together with the mixed labels and the charging of Section 7. Finally, Section 10 turns the proof of Theorem 8.1 into an algorithm, and Appendix A.6 that of Theorem 8.2.

8.1 Setting

Throughout, G is a finite simple graph with $\Delta(G) \le 3$; it may be disconnected and may have vertices of degree 0, 1 or 2. Recall $r_* = r_{33} = \frac{2\sqrt{2}}{3}$ and $s_r = \sqrt{1-r^2}$, and put

$$b(r) = \frac{3r^2-1}{2}, \qquad \kappa(r) = \frac{3}{4\pi}\arccos b(r) \qquad (0 < r < 1). \tag{8.1}$$

For $r \in (0,1)$ we have $b(r) \in (-\frac{1}{2}, 1)$, so κ is well defined and continuous on $(0,1)$. Since $b(r_*) = \frac{3\cdot\frac{8}{9}-1}{2} = \frac{5}{6}$, we have $\kappa(r_*) = c_3$. By Lemma 8.6 below, $\kappa(r)$ is the value at threshold zero of the ideal star function $f_{3,r}$.

Labels. Fix an integer $R \geq 2$ and let $(Y_v)_{v\in V(G)}$ be the geometric labels at radius R, that is, the padded labels of degree 3 (Definition 5.2(a) with $p = 3$): for i.i.d. standard normal $(Z_w)_{w\in V(G)}$, $X_v = \sum_{w:\,\mathrm{dist}(v,w)\leq R} 2^{-\mathrm{dist}(v,w)/2} Z_w$ and $Y_v = X_v/\sqrt{\mathrm{Var}\, X_v}$. Put $r_R = r_*(1 - \frac{1}{R}) \in (0, r_*)$. We use three properties.

(L1) The vector $(Y_v)_v$ is centered Gaussian with unit variances, and every edge uv satisfies $\mathrm{Corr}(Y_u, Y_v) \geq r_R$ (Lemma 5.4, cubic case).

(L2) With probability 1, the labels are pairwise distinct, also for vertices in different components; see Lemma 5.6(ii).

(L3) Y_v is a function of $(Z_w)_{w\in B_R(v)}$ (Lemma 5.6(iii), with $\rho = R$); hence functions of Z restricted to disjoint vertex sets are independent.

(We take $R \geq 2$ only so that $r_R > 0$.) Since $-Y$ is again centered Gaussian with the same covariance, $-Y$ has the same distribution as Y; we call this *sign symmetry*. Every label has a continuous distribution, so for each fixed t the events $\{Y_v = t\}$ are null.

Threshold zero. For a labeling $y\colon V(G) \to \mathbb{R}$, the sublevel sets $S_t = \{v : y_v \leq t\}$ and their frontiers $\partial^+ S_t$, $\partial^- S_t$ are as in Section 4.2. Put $U = \partial^- S_0$ and $W = \partial^+ S_0$, and let D be the set of vertices of degree 3 all of whose neighbors lie on the other side of zero, that is, vertices v with $d(v) = 3$ and either $y_v \leq 0 < \min_{x\in N(v)} y_x$ or $y_v > 0 \geq \max_{x\in N(v)} y_x$. A vertex of D with $y_v \leq 0$ has a neighbor above zero and lies in U; one with $y_v > 0$ has a neighbor at or below zero and lies in W. Hence $D \subseteq U \cup W$. Put $U' = U \setminus D$, $W' = W \setminus D$ and

$$L_0 = |U| + |D \cap W| = |U'| + |D|, \qquad L_0' = |W| + |D \cap U| = |W'| + |D|. \tag{8.2}$$

The two expressions agree because U and W are disjoint and $D = (D \cap U) \cup (D \cap W)$. At threshold zero, the width of the median-switched layout of Section 8.2 is controlled by $\max(L_0, L_0')$ (Lemma 8.4).

8.2 The median-switched layout

This subsection proves the deterministic part of Theorem 8.1: a path decomposition whose width is controlled by inner frontiers below zero, outer frontiers above zero, and $\max(L_0, L_0')$ at zero (Lemma 8.4). The connector at zero needs an ordering lemma, which we state abstractly. Think of the elements of Q below as the vertices of W', to be added one at a time, and of the items as the vertices of U', each of which can be dropped once its at most two neighbors in W' are present.

Lemma 8.3 (connector order)**.** *Let Q be a finite set with k elements and let P be a finite set of items, each item p having a nonempty set $e(p) \subseteq Q$ of at most two endpoints. For an ordering $q_1, \ldots, q_k$ of Q and $0 \leq j \leq k$, let $c(j)$ be the number of items p with $e(p) \subseteq \{q_1, \ldots, q_j\}$. Then some ordering satisfies*

$$j - c(j) \leq \max(0,\, k - |P|) + 1 \qquad \textit{for all } 0 \leq j \leq k. \tag{8.3}$$

Proof. We use induction on k. If some nonempty proper subset of Q is closed (defined below), the instance splits into two smaller ones, which we order one after the other, the one with smaller excess first (Case 1). Otherwise each new element can be chosen to complete an item whose other endpoint is already present (Case 2).

If $k = 0$, then $P = \emptyset$ because endpoint sets are nonempty, and (8.3) reads $0 \le 1$. Let $k \ge 1$. Call a set $Y \subseteq Q$ *closed* if every item p satisfies $e(p) \subseteq Y$ or $e(p) \subseteq Q \setminus Y$. For a closed set A let $P_A = \{p : e(p) \subseteq A\}$ and define its *excess* $\operatorname{ex}(A) = |A| - |P_A|$; thus $\operatorname{ex}(Q) = k - |P|$.

Case 1: some nonempty proper subset Y of Q is closed. Then $Q \setminus Y$ is closed as well, and since endpoint sets are nonempty, every item lies in exactly one of P_Y and $P_{Q\setminus Y}$. Let A be the one of $Y, Q \setminus Y$ with the smaller excess (either one if they are equal) and B the other, so that

$$\operatorname{ex}(Q) = \operatorname{ex}(A) + \operatorname{ex}(B) \ge 2\operatorname{ex}(A).$$

The pairs (A, P_A) and (B, P_B) satisfy the hypotheses of the lemma with fewer than k elements, so by induction they have orderings $a_1, \dots, a_{|A|}$ and $b_1, \dots, b_{|B|}$ with counting functions c_A, c_B satisfying (8.3) with $\operatorname{ex}(A)$, respectively $\operatorname{ex}(B)$, in place of $k - |P|$. Order Q as $a_1, \dots, a_{|A|}, b_1, \dots, b_{|B|}$.

For $j \le |A|$, an item covered by $\{a_1, \dots, a_j\}$ has an endpoint in A and hence lies in P_A, so $c(j) = c_A(j)$ and $j - c(j) \le \max(0, \operatorname{ex}(A)) + 1$. If $\operatorname{ex}(A) \le 0$ this is at most 1; otherwise $\operatorname{ex}(Q) \ge 2\operatorname{ex}(A) > \operatorname{ex}(A)$. In both cases $j - c(j) \le \max(0, \operatorname{ex}(Q)) + 1$.

For $j = |A| + j'$ with $1 \le j' \le |B|$, all items of P_A are covered, and an item of P_B is covered exactly when $e(p) \subseteq \{b_1, \dots, b_{j'}\}$. Hence $c(j) = |P_A| + c_B(j')$ and

$$j - c(j) = \operatorname{ex}(A) + \bigl(j' - c_B(j')\bigr) \le \operatorname{ex}(A) + \max(0, \operatorname{ex}(B)) + 1.$$

If $\operatorname{ex}(B) \ge 0$, the right side is $\operatorname{ex}(Q) + 1 \le \max(0, \operatorname{ex}(Q)) + 1$. If $\operatorname{ex}(B) < 0$, then $\operatorname{ex}(A) \le \operatorname{ex}(B) < 0$ and the right side is less than 1. This proves (8.3) in Case 1.

Case 2: no nonempty proper subset of Q is closed. Choose $q_1 \in Q$ arbitrarily. Suppose $q_1, \dots, q_j$ have been chosen with $1 \le j < k$, and let $X = \{q_1, \dots, q_j\}$. Since X is nonempty and proper, it is not closed, so some item p has endpoints both in X and in $Q \setminus X$. As $|e(p)| \le 2$, we have $e(p) = \{x, y\}$ with $x \in X$ and $y \notin X$; put $q_{j+1} = y$. Then p is covered by $\{q_1, \dots, q_{j+1}\}$ but not by X, so $c(j+1) \ge c(j) + 1$. Since $c(1) \ge 0$, we get $c(j) \ge j - 1$ for every $1 \le j \le k$, and $c(0) = 0$. Hence $j - c(j) \le 1$ for all j, which implies (8.3). □

The layout itself has three phases. The inner phase is the vertex separation layout of Kinnersley [Kin92]: each bag consists of the new vertex and the inner frontier of the preceding prefix. The outer phase is the layout of Lemma 4.1. The connector between them is new.

Lemma 8.4 (median-switched layout)**.** *Let G be a graph with $\Delta(G) \le 3$, let $y\colon V(G) \to \mathbb{R}$ be injective, and let $w \ge 0$ be a real number such that*

(i) $|\partial^- S_t| \le w$ *for every* $t < 0$*;*

(ii) $|\partial^+ S_t| \le w$ *for every* $t > 0$*; and*

(iii) $\max(L_0, L_0') + 1 \le w$*, with* L_0, L_0' *as in* (8.2)*.*

Then $\operatorname{pw}(G) \le w$*.*

The proof builds the bags from any ordering of W' that satisfies (8.3), not only from that of Lemma 8.3, because the algorithm of Appendix A.5.8 uses another one.

The bags. Let G be a graph with $\Delta(G) \le 3$ and let $y\colon V(G) \to \mathbb{R}$ be injective. Let $v_1, \dots, v_n$ be the vertices in increasing order of y, let $\ell = |S_0|$, and write $T_i = \{v_1, \dots, v_i\}$ for $0 \le i \le n$. Let P be the set of vertices $u \in U'$ for which $e(u) := N(u) \cap W'$ is nonempty; these are the items, and $e(u)$ is the set of endpoints of u. Let $k = |W'|$, and let $q_1, \dots, q_k$ be an ordering of W'. The *inner bags* are $A_i = \{v_i\} \cup \partial^- T_{i-1}$ for $1 \le i \le \ell$, the *outer bags* are $A_i = \{v_i\} \cup \partial^+ T_i$ for $\ell < i \le n$, and the *connector bags* are

$$C_{-1} = U \cup (D \cap W), \qquad C_j = D \cup \{u \in P : e(u) \not\subseteq \{q_1, \dots, q_{j-1}\}\} \cup \{q_1, \dots, q_j\} \quad (0 \le j \le k).$$

In words, the connector starts from $C_{-1} = U' \cup D$ and drops the vertices of U' without neighbors in W'. It then adds the vertices of W' one at a time and drops each item one step after its last endpoint has been added. The vertices of D stay in every connector bag.

Lemma 8.5 (the bags of the layout)**.** *Let G, y, P and $q_1, \dots, q_k$ be as above. Then $|e(u)| \le 2$ for every $u \in P$. If moreover w satisfies* (i)–(iii) *of Lemma 8.4 and the ordering $q_1, \dots, q_k$ of W' satisfies* (8.3) *for $Q = W'$ and the items P, then $A_1, \dots, A_\ell, C_{-1}, C_0, \dots, C_k, A_{\ell+1}, \dots, A_n$ is a path decomposition of G of width at most w.*

Proof. A vertex $u \in U'$ is not in D, so either $d(u) \le 2$ or u has a neighbor at or below zero; in both cases u has at most two neighbors above zero, and $|e(u)| \le 2$.

For the second claim we check the size bound and the three axioms of a path decomposition (every vertex in a bag, every edge in a bag, the bags containing a vertex consecutive). Note that $T_i = S_t$ for every t with $y_{v_i} \le t < y_{v_{i+1}}$ (with $y_{v_0} = -\infty$ and $y_{v_{n+1}} = +\infty$), and $v_i \in S_0$ if and only if $i \le \ell$.

Inner bags. For $1 \le i \le \ell$ we have $y_{v_i} \le 0$, so $T_{i-1} = S_t$ for some $t < y_{v_i} \le 0$, and $|A_i| \le w + 1$ by (i).

Connector bags. A vertex of D lies in all connector bags, a vertex q_j of W' in $C_j, \dots, C_k$, an item $u \in P$ in $C_{-1}, \dots, C_J$, where $J \ge 1$ is the largest position of an element of $e(u)$ (so u is dropped one step after its last endpoint has been added), every other vertex of U in C_{-1} only, and no other vertex in any connector bag. So the connector bags containing a vertex of U form a prefix of $C_{-1}, C_0, \dots, C_k$, and those containing a vertex of W form a suffix.

We have $|C_{-1}| = L_0$, and $|C_0| = |D| + |P| \le |D| + |U'| = L_0$. For $1 \le j \le k$, the number of items still present in C_j is $|P| - c(j-1)$, where c is the counting function of Lemma 8.3 for the ordering $q_1, \dots, q_k$, so by (8.3) at $j - 1$,

$$|C_j| = |D| + |P| - c(j-1) + j \le |D| + |P| + \max(0, k - |P|) + 2 = |D| + \max(|P|, |W'|) + 2.$$

Since $|P| \le |U'|$, this is at most $\max(L_0, L_0') + 2 \le w + 1$ by (iii).

Outer bags. For $\ell < i \le n$ we have $y_{v_i} > 0$, so $T_i = S_t$ for some $t \ge y_{v_i} > 0$, and $|A_i| \le w + 1$ by (ii).

Every vertex v_i lies in A_i. If $\ell = 0$ or $\ell = n$, then $U = W = D = \emptyset$ and the connector bags are empty; the arguments below remain valid.

Intervals for lower vertices. Let $u = v_a$ with $a \le \ell$, and let b be the largest index of a neighbor of u ($b = 0$ if u is isolated). For $i \le \ell$ we have $u \in A_i$ if and only if $i = a$ or $a \le i - 1 < b$, so the inner bags containing u are $A_a, \dots, A_{\min(\max(a,b),\ell)}$. No outer bag contains u, because $u \in T_i$ for every $i > \ell$. If $u \notin U$, then $b \le \ell$ and u is in no connector bag, so its bags are consecutive. If $u \in U$, then $b > \ell$, so its bags are $A_a, \dots, A_\ell$ followed by a prefix of the connector bags.

Intervals for upper vertices. Let $x = v_c$ with $c > \ell$, and let a be the smallest index of a neighbor of x ($a = \infty$ if x is isolated). No inner bag contains x. For $i > \ell$ we have $x \in A_i$ if and only if $i = c$, or $i < c$ and x has a neighbor in T_i, that is, $a \le i < c$; so the outer bags containing x are

$A_{\max(a,\ell+1)}, \dots, A_c$ if $a < c$, and A_c alone otherwise. If $x \notin W$, then $a > \ell$ and x lies in no connector bag, so its bags are consecutive. If $x \in W$, then $a \le \ell$, so its bags are a suffix of the connector bags followed by $A_{\ell+1}, \dots, A_c$.

Edges. Let $v_a v_b$ be an edge with $a < b$. If $b \le \ell$, then $v_a \in \partial^- T_{b-1}$, because $v_a \in T_{b-1}$ and its neighbor v_b is not, so both ends lie in A_b. If $a > \ell$, then $v_b \in \partial^+ T_a$, because $v_b \notin T_a$ and its neighbor v_a is in T_a, so both ends lie in A_a. Otherwise $u = v_a \in U$ and $x = v_b \in W$. If $x \in D \cap W$, both ends lie in C_{-1}. If $u \in D \cap U$, both ends lie in C_k, which contains D and all of W'. Otherwise $u \in U'$ and $x \in W'$, so $u \in P$ and $x = q_j \in e(u)$ for some j; then $e(u) \not\subseteq \{q_1, \dots, q_{j-1}\}$, so $u \in C_j$, and $x \in C_j$.

Hence the sequence of bags is a path decomposition with all bags of size at most $w + 1$, that is, of width at most w. □

Proof of Lemma 8.4. By Lemma 8.5, every item $u \in P$ has one or two endpoints $e(u) \subseteq W'$, so Lemma 8.3 applies to $Q = W'$ and the items P and gives an ordering of W' that satisfies (8.3). For this ordering, Lemma 8.5 gives a path decomposition of G of width at most w, so $\mathrm{pw}(G) \le w$. □

8.3 The star comparison at nonnegative thresholds

This subsection bounds the probability that a vertex lies in an outer frontier at a threshold $t \ge 0$ by $\kappa(r)$ (Proposition 8.8). By sign symmetry the same bound applies to inner frontiers at thresholds $t \le 0$. The proof has three ingredients. A pointwise coupling lowers the incident correlations to r; the simplex theorem for three variables (Lemma 6.5) handles the neighbor residuals; and two facts about the ideal star function $f_{3,r}$ give the value $\kappa(r)$: its value at zero (Lemma 8.6) and its maximality at zero among nonnegative thresholds (Lemma 8.7).

The ideal 3-star. We use the ideal 3-star of Definition 6.7. Let $U^* = (U_1^*, U_2^*, U_3^*)$ be the regular simplex vector in three variables (Section 6): centered, with unit variances and all pairwise correlations $-\frac{1}{2}$, so that $S_3(a) = \mathbb{P}(U_i^* > a \ \forall i)$. It is realized as the projections of a standard planar Gaussian vector onto three unit vectors at mutual angles $\frac{2\pi}{3}$; in particular $U_1^* + U_2^* + U_3^* = 0$. For $r \in (0, 1)$ the ideal 3-star at correlation r is

$$Y_0^* \sim N(0,1) \text{ independent of } U^*, \qquad Y_i^* = r Y_0^* + s_r U_i^* \quad (i = 1, 2, 3).$$

Its incident correlations are r, its neighbor-pair correlations are $r^2 - \frac{1}{2} s_r^2 = b(r)$, and

$$Y_1^* + Y_2^* + Y_3^* = 3 r Y_0^*. \tag{8.4}$$

The difference of any two coordinates of the ideal star has positive variance, as $r < 1$ and $s_r > 0$, and each coordinate has a continuous distribution. So ties among the coordinates, or with a fixed threshold, have probability zero, and strict and non-strict inequalities may be exchanged in all events below. In particular the ideal star function $f_{3,r}$ of Definition 6.7, given by (6.5) with $d = 3$, satisfies $f_{3,r}(t) = \mathbb{P}(Y_0^* \ge t, \ \min_i Y_i^* \le t)$.

Both lemmas on $f_{3,r}$ below use Sheppard's formula (Lemma 4.9).

Lemma 8.6. *For every $r \in (0, 1)$, $f_{3,r}(0) = \kappa(r)$.*

Proof. If $Y_i^* > 0$ for all i, then $3 r Y_0^* > 0$ by (8.4), so $Y_0^* > 0$. Hence the event $\{\min_i Y_i^* > 0\}$ is contained in $\{Y_0^* \ge 0\}$, and

$$f_{3,r}(0) = \mathbb{P}(Y_0^* \ge 0) - \mathbb{P}(Y_0^* \ge 0, \ \min_i Y_i^* > 0) = \frac{1}{2} - \mathbb{P}(Y_1^*, Y_2^*, Y_3^* > 0).$$

The three neighbor coordinates have pairwise angles $\arccos b(r)$, so Lemma 4.9 gives $f_{3,r}(0) = \frac{1}{2} - \left(\frac{1}{2} - \frac{3 \arccos b(r)}{4\pi}\right) = \kappa(r)$. □

The next lemma shows that, among nonnegative thresholds, the ideal star function is largest at the median 0. Its proof needs no derivative and no numerical evaluation, and holds for every $r \in (0,1)$.

Lemma 8.7. *For every $r \in (0,1)$ and every $t \geq 0$, $f_{3,r}(t) \leq f_{3,r}(0)$.*

Proof. We write $f_{3,r}(0) - f_{3,r}(t)$ as a difference of two probabilities and bound the event that the smallest neighbor coordinate lies in $(0,t]$ by splitting according to which coordinate attains the minimum.

Let $M = \min_i Y_i^*$. If $M > t \geq 0$, then $3rY_0^* > 3t$ by (8.4), so $Y_0^* > t/r \geq t$. Hence $\{M > t\} \subseteq \{Y_0^* \geq t\}$, and since $f_{3,r}(t) = \mathbb{P}(Y_0^* \geq t,\ M \leq t)$, we get $f_{3,r}(t) = \mathbb{P}(Y_0^* \geq t) - \mathbb{P}(M > t)$ for all $t \geq 0$ and

$$f_{3,r}(0) - f_{3,r}(t) = \mathbb{P}(0 \leq Y_0^* < t) - \mathbb{P}(0 < M \leq t).$$

It remains to show $\mathbb{P}(0 < M \leq t) \leq \mathbb{P}(0 < Y_0^* \leq t)$; the two probabilities $\mathbb{P}(0 \leq Y_0^* < t)$ and $\mathbb{P}(0 < Y_0^* \leq t)$ are equal because Y_0^* has a continuous distribution.

If $0 < M \leq t$, then some coordinate attains the minimum, so the event is contained in the union of the three events $A_k = \{0 < Y_k^* \leq t,\ Y_k^* \leq Y_j^* \text{ for all } j \in \{1,2,3\}\}$. Fix k and write $b = b(r) \in (-\frac{1}{2}, 1)$. For $j \neq k$ put $R_j = Y_j^* - bY_k^*$. Then $\mathrm{Cov}(R_j, Y_k^*) = b - b = 0$, so the pair $(R_j)_{j \neq k}$ is independent of Y_k^*. On A_k we have $R_j \geq Y_k^* - bY_k^* = (1-b)Y_k^* > 0$ for both $j \neq k$. Hence

$$\mathbb{P}(A_k) \leq \mathbb{P}(0 < Y_k^* \leq t) \cdot \mathbb{P}(R_j > 0 \text{ for both } j \neq k).$$

The two residuals have variance $1 - b^2 > 0$ and covariance $b - 2b^2 + b^2 = b - b^2$, so their correlation is $\rho = \frac{b-b^2}{1-b^2} = \frac{b}{1+b}$. The function $b \mapsto \frac{b}{1+b} = 1 - \frac{1}{1+b}$ is increasing on $(-1, \infty)$ and equals $\frac{1}{2}$ at $b = 1$, so $\rho < \frac{1}{2}$ and $\arccos \rho > \frac{\pi}{3}$. By Lemma 4.9, applied to the residuals divided by their common standard deviation (which does not change their signs), $\mathbb{P}(R_j > 0 \text{ for both } j \neq k) < \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$. Since Y_k^* and Y_0^* are both standard normal, $\mathbb{P}(A_k) \leq \frac{1}{3}\mathbb{P}(0 < Y_0^* \leq t)$, and summing over k gives the claim. □

We now compare an arbitrary star with the ideal one. The case $d = 3$ of the simplex theorem (Lemma 6.5) states that every centered Gaussian vector (η_1, η_2, η_3) with unit variances, possibly with a singular correlation matrix, satisfies $\mathbb{P}(\eta_i > \theta\ \forall i) \geq S_3(\theta)$ for every $\theta \in \mathbb{R}$; equivalently, $\mathbb{P}(\min_i \eta_i \leq \theta) \leq \mathbb{P}(\min_i U_i^* \leq \theta)$. The proof below uses it for a vector with a repeated coordinate.

Proposition 8.8 (star bound at nonnegative thresholds)**.** *Let $r \in (0,1)$, let $0 \leq d \leq 3$, and let $(Y_0, Y_1, \ldots, Y_d)$ be a centered Gaussian vector with unit variances and $\mathrm{Corr}(Y_0, Y_i) \geq r$ for $1 \leq i \leq d$. Then for every $t \geq 0$,*

$$\mathbb{P}\big(Y_0 \geq t,\ \exists i \in \{1, \ldots, d\} :\ Y_i \leq t\big) \leq f_{3,r}(t) \leq \kappa(r).$$

Proof. The second inequality is Lemmas 8.6 and 8.7. For the first, we lower the incident correlations to r by a pointwise coupling, condition on Y_0, and apply the simplex theorem to the residuals. Put $a_i = \mathrm{Corr}(Y_0, Y_i) \geq r$. If $a_i = 1$ then $Y_i = Y_0$ with probability 1, and $\{Y_0 \geq t \geq Y_i\}$ is contained in the null event $\{Y_0 = t\}$; we discard such indices. If no index remains (in particular if $d = 0$), the probability is $0 \leq f_{3,r}(t)$. Otherwise, for each remaining i put $\eta_i = (Y_i - a_i Y_0)/s_{a_i}$. The variables η_i are standard normal, jointly Gaussian with Y_0, and uncorrelated with Y_0, so the vector of the η_i is independent of Y_0. If fewer than three indices remain, repeat one of the η_i until there are three coordinates (η_1, η_2, η_3); repetition does not change the event $\{\exists i : \eta_i \leq c\}$ for any c.

On the event $\{Y_0 \geq t\}$ consider $g(a) = (t - aY_0)/\sqrt{1-a^2}$ for $a \in [r, 1)$. Its derivative is

$$g'(a) = \frac{-Y_0(1-a^2) + a(t - aY_0)}{(1-a^2)^{3/2}} = \frac{at - Y_0}{(1-a^2)^{3/2}} \leq 0,$$

because $at \le t \le Y_0$ (here $t \ge 0$ is used). Hence $g(a_i) \le g(r)$, and

$$Y_i \le t \iff \eta_i \le g(a_i) \implies \eta_i \le g(r) = \frac{t - rY_0}{s_r}.$$

By the independence of η and Y_0 (Fubini) and Lemma 6.5 with $d = 3$,

$$\begin{aligned}\mathbb{P}(Y_0 \ge t,\ \exists i : Y_i \le t) &\le \int_t^\infty \varphi(z)\mathbb{P}\Big(\min_i \eta_i \le \frac{t - rz}{s_r}\Big)dz \\ &\le \int_t^\infty \varphi(z)\Big(1 - S_3\Big(\frac{t - rz}{s_r}\Big)\Big)dz = f_{3,r}(t),\end{aligned}$$

the last equality by (6.5). □

For $t < 0$ the monotonicity of g fails on part of the event $\{Y_0 \ge t\}$. This is why the plain sweep needs the general star comparison of Section 6. The median-switched layout uses outer frontiers only at $t > 0$ and inner frontiers only at $t < 0$, so it needs only the pointwise coupling above.

8.4 Level zero: the switch lemma

This subsection proves the switch lemma: if all incident correlations are at least r, each vertex contributes, in expectation, at most $\kappa(r)$ to the quantity L_0 of (8.2), which controls the connector bags at threshold zero. Applied to the negated labels, the lemma gives the same bound for L_0'. This is the same bound $\kappa(r)$ as at nonzero thresholds (Proposition 8.8). For a vertex v of degree d with neighbors $x_1, \dots, x_d$, the events $v \in U$ and $v \in D \cap W$ are $\{Y_v \le 0 < \max_i Y_{x_i}\}$ and $\{d = 3,\ Y_v > 0 \ge \max_i Y_{x_i}\}$, so the following lemma bounds the contribution of v to $\mathbb{E}L_0$.

Lemma 8.9 (switch lemma)**.** *Let $r \in (0,1)$, let $0 \le d \le 3$, and let $(Y_0, Y_1, \dots, Y_d)$ be a centered Gaussian vector with unit variances and $\mathrm{Corr}(Y_0, Y_i) \ge r$ for $1 \le i \le d$. Then*

$$\mathbb{P}\big(Y_0 \le 0,\ \exists i :\ Y_i > 0\big) + \mathbf{1}_{d=3}\mathbb{P}(Y_0 > 0,\ Y_i \le 0 \text{ for all } i\big) \le \kappa(r).$$

Idea of the proof. By sign symmetry the second term equals $\mathbb{P}(Y_0 \le 0 < \min_i Y_i)$, so the left side is $\mathbb{E}\big[\mathbf{1}\{Y_0 \le 0\}\lceil N/2\rceil\big]$, where N is the number of neighbors with positive label. After conditioning on $Y_0 = z < 0$ and lowering the incident correlations to r as in Proposition 8.8, it remains to show that, among all correlation matrices of three standard normal residuals, the probability that at least two of them exceed a level $h > 0$ is smallest for the equilateral matrix of U^* (Lemma 8.11). We prove this by a deformation. The smallest pairwise angles of the residual configuration are raised simultaneously ("water-filling") until the angles sum to 2π. Plackett's identity shows that the probability does not increase along the way. At the end the configuration is planar, and an exact formula and convexity finish the argument. The simplex theorem cannot be used here: the event "at least two exceed h" is not an orthant event, and the comonotone and countermonotone extremes pull in opposite directions.

We first record the tools: which angle triples are realizable, and Plackett's identity in the form needed here; the planar formula is part of the proof of Lemma 8.11.

Lemma 8.10 (angles of three unit vectors)**.** *Let $\alpha, \beta, \gamma \in [0, \pi]$ and let C be the symmetric 3×3 matrix with unit diagonal and off-diagonal entries $C_{23} = \cos\alpha$, $C_{13} = \cos\beta$, $C_{12} = \cos\gamma$. Then*

$$\det C = \big(\cos\alpha - \cos(\beta + \gamma)\big)\big(\cos(\beta - \gamma) - \cos\alpha\big). \tag{8.5}$$

Moreover, C is positive semidefinite if and only if $\alpha \le \beta + \gamma$, $\beta \le \alpha + \gamma$, $\gamma \le \alpha + \beta$ and $\alpha + \beta + \gamma \le 2\pi$, and positive definite if and only if these four inequalities are strict.

Proof. The identity is a direct expansion. That the four inequalities suffice follows from the signs of the two factors in (8.5), and that they are necessary from the triangle inequality for angles between unit vectors.

Expanding, $\det C = 1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma$. For the right side of (8.5), use $\cos(\beta+\gamma) + \cos(\beta-\gamma) = 2\cos\beta\cos\gamma$ and $\cos(\beta+\gamma)\cos(\beta-\gamma) = \cos^2\beta\cos^2\gamma - \sin^2\beta\sin^2\gamma = \cos^2\beta + \cos^2\gamma - 1$; then the product equals $\cos\alpha \cdot 2\cos\beta\cos\gamma - \cos^2\alpha - (\cos^2\beta + \cos^2\gamma - 1)$, which is the expansion.

Assume the four inequalities and put $m = \min(\beta+\gamma, 2\pi-\beta-\gamma) \in [0,\pi]$, so that $\cos(\beta+\gamma) = \cos m$ and $|\beta-\gamma| \le \alpha \le m$. Since cos is strictly decreasing on $[0,\pi]$, the first factor $\cos\alpha - \cos m$ of (8.5) is nonnegative, and positive exactly when $\alpha < m$, that is, when $\alpha < \beta+\gamma$ and $\alpha+\beta+\gamma < 2\pi$; the second factor $\cos|\beta-\gamma| - \cos\alpha$ is nonnegative, and positive exactly when $|\beta-\gamma| < \alpha$, that is, when $\beta < \alpha+\gamma$ and $\gamma < \alpha+\beta$. So $\det C \ge 0$, with equality exactly when one of the four inequalities is an equality. The 2×2 principal minors are $\sin^2 \ge 0$ and the diagonal entries are 1, so all principal minors are nonnegative and C is positive semidefinite.

Conversely, let C be positive semidefinite. Then C is the Gram matrix of unit vectors v_1, v_2, v_3 (for instance the columns of $C^{1/2}$), and α, β, γ are the angles between them. The angle between unit vectors is a metric on the sphere, which gives the three triangle inequalities, and applying the triangle inequality to v_1, $-v_3$, v_2 gives $\gamma \le (\pi-\beta) + (\pi-\alpha)$, the fourth inequality.

Finally, a positive semidefinite matrix is positive definite exactly when its determinant is positive. By the two preceding paragraphs, C is therefore positive definite exactly when the four inequalities hold and $\det C > 0$, that is, when they hold strictly. □

The second tool is Plackett's identity (Lemma 4.7), with $\varphi_2(\cdot,\cdot;c)$ the density of a pair of standard normals with correlation $c \in (-1,1)$. For positive definite 3×3 correlation matrices C, a level $h \in \mathbb{R}$ and $\eta \sim N(0,C)$, put $P_{ij}(C) = \mathbb{P}_C(\eta_i > h, \eta_j > h)$ and $P_{123}(C) = \mathbb{P}_C(\eta_1,\eta_2,\eta_3 > h)$, as functions of the three off-diagonal entries. Let $\{i,j,k\} = \{1,2,3\}$. The pair probabilities P_{ik} and P_{jk} depend only on the distributions of (η_i,η_k) and (η_j,η_k), hence not on C_{ij}. Lemma 4.7, applied to the two coordinates (η_i,η_j) and to all three coordinates, gives

$$\frac{\partial P_{ij}}{\partial C_{ij}} = \varphi_2(h,h;C_{ij}), \qquad \frac{\partial P_{123}}{\partial C_{ij}} = \varphi_2(h,h;C_{ij})\,\mathbb{P}_C\big(\eta_k > h \mid \eta_i = \eta_j = h\big), \tag{8.6}$$

and both derivatives are continuous in C. By (4.3), the conditional distribution of η_k given $\eta_i = \eta_j = h$ is normal with mean $h\,\frac{C_{ik}+C_{jk}}{1+C_{ij}}$ and positive variance.

For a correlation matrix C (positive semidefinite, unit diagonal) and $h > 0$ let

$$\Pi_C(h) = \mathbb{P}_C\big(\text{at least two of } \eta_1,\eta_2,\eta_3 \text{ exceed } h\big), \qquad \eta \sim N(0,C),$$

and let C_{eq} be the correlation matrix of U^* (all off-diagonal entries $-\frac{1}{2}$). The map $C \mapsto \Pi_C(h)$ is continuous on the set of all correlation matrices: if $C_n \to C$, then $N(0,C_n) \to N(0,C)$ weakly (by characteristic functions), and the boundary of the event is contained in the union of the hyperplanes $\{x_i = h\}$, which are null under $N(0,C)$ because every coordinate is standard normal.

Lemma 8.11 (two exceedances)**.** *For every 3×3 correlation matrix C and every $h > 0$, $\Pi_C(h) \ge \Pi_{C_{\mathrm{eq}}}(h)$.*

Proof. We reduce to positive definite C, deform C by raising its smallest angles until the angles sum to 2π, show that Π does not increase along this path, and compare the planar endpoint with C_{eq}.

Reduction to positive definite C. For $\xi \in (0,1]$, the matrix $(1-\xi)C + \xi I$ is a positive definite correlation matrix, and it tends to C as $\xi \downarrow 0$. By the continuity of Π, it suffices to prove the claim for positive definite C, which we now assume.

The water-filling path. Write $C_{ij} = \cos\theta_{ij}$ with $\theta_{ij} \in [0,\pi]$. By Lemma 8.10 and positive definiteness, the three triangle inequalities and the inequality $\theta_{12}+\theta_{13}+\theta_{23} \le 2\pi$ hold strictly. For $s \ge 0$ let $\theta_{ij}(s) = \max(\theta_{ij}, s)$ and $\Sigma(s) = \sum_{i<j}\theta_{ij}(s)$. The function Σ is continuous and nondecreasing, $\Sigma(0) < 2\pi$ and $\Sigma(\frac{2\pi}{3}) \ge 2\pi$; let $s^* \in (0, \frac{2\pi}{3}]$ be the smallest s with $\Sigma(s) = 2\pi$, and let $C(s)$ be the matrix with off-diagonal entries $\cos\theta_{ij}(s)$. On $[0, s^*]$ all angles lie in $[0,\pi]$.

For $0 \le s < s^*$, the matrix $C(s)$ is positive definite. Indeed $\Sigma(s) < 2\pi$. For a triangle inequality $\theta_{ij}(s) < \theta_{ik}(s) + \theta_{jk}(s)$: if $\theta_{ij}(s) = \theta_{ij}$, it follows from $\theta_{ij} < \theta_{ik} + \theta_{jk} \le \theta_{ik}(s) + \theta_{jk}(s)$; otherwise $\theta_{ij}(s) = s > 0$, and the right side is at least $2s > s$. Lemma 8.10 gives positive definiteness. Thus $C(s)$ moves within the positive definite matrices for $s < s^*$, and reaches a matrix $C(s^*)$ with angle sum exactly 2π, which is positive semidefinite by Lemma 8.10 (the non-strict inequalities persist by continuity).

Monotonicity along the path. The breakpoints $\{\theta_{12}, \theta_{13}, \theta_{23}\}$ divide $[0, s^*]$ into finitely many intervals. On the interior of each, the set of *moving* pairs, those with $\theta_{ij} < s$, is fixed; the moving pairs have $C_{ij}(s) = \cos s$, and the other entries are constant. By inclusion and exclusion,

$$\Pi_C(h) = P_{12}(C) + P_{13}(C) + P_{23}(C) - 2P_{123}(C),$$

because the indicator of "at least two" equals the number of exceeding pairs minus twice the indicator of "all three" (check $N = 2$: $1 - 0$; $N = 3$: $3 - 2$). By (8.6), for $s < s^*$ and each pair $\{i, j\}$ with third index k,

$$\frac{\partial \Pi_{C(s)}(h)}{\partial C_{ij}} = \varphi_2(h, h; C_{ij}(s))\Big(1 - 2\mathbb{P}_{C(s)}(\eta_k > h \mid \eta_i = \eta_j = h)\Big).$$

For a moving pair, $\theta_{ij}(s) = s \le \theta_{ik}(s), \theta_{jk}(s)$, so $C_{ik}(s), C_{jk}(s) \le C_{ij}(s)$, and $C_{ij}(s) = \cos s \ge -\frac{1}{2}$. Since $h > 0$, the conditional mean satisfies

$$h\,\frac{C_{ik}(s) + C_{jk}(s)}{1 + C_{ij}(s)} \le h\,\frac{2C_{ij}(s)}{1 + C_{ij}(s)} \le h,$$

the last step because $2c \le 1 + c$ for $c \le 1$ and $1 + c > 0$. A normal distribution with mean at most h gives probability at most $\frac{1}{2}$ to (h, ∞), so the partial derivative is nonnegative for every moving pair. By the chain rule, on the interior of each interval

$$\frac{d}{ds}\Pi_{C(s)}(h) = \sum_{\text{moving } \{i,j\}} \frac{\partial \Pi_{C(s)}(h)}{\partial C_{ij}} \cdot (-\sin s) \le 0,$$

since $\sin s \ge 0$ for $s \in [0,\pi]$. The function $s \mapsto \Pi_{C(s)}(h)$ is continuous on $[0, s^*]$ (by the continuity of Π), so by the mean value theorem on each interval it is nonincreasing on $[0, s^*]$, and $\Pi_C(h) \ge \Pi_{C(s^*)}(h)$.

The planar endpoint. Let v_1, v_2, v_3 be unit vectors with Gram matrix $C(s^*)$, and let $\theta_{ij} = \theta_{ij}(s^*)$ from now on. By (8.5) with angle sum 2π, $\cos(\beta+\gamma) = \cos(2\pi - \alpha) = \cos\alpha$, so $\det C(s^*) = 0$ and the vectors lie in a two-dimensional subspace; identify it with $\mathbb{R}^2$. If Γ is a standard Gaussian vector in $\mathbb{R}^2$, then $(\langle v_i, \Gamma\rangle)_i \sim N(0, C(s^*))$. Order the directions of the v_i around the circle and let $g_1, g_2, g_3 \ge 0$ be the three circular gaps between consecutive directions, $g_1 + g_2 + g_3 = 2\pi$, labeled so that g_k is the gap between the directions of v_i and v_j that does not contain the direction of v_k. The angle between v_i and v_j is $\theta_{ij} = \min(g_k, 2\pi - g_k) \le g_k$; since the angles sum to $2\pi = \sum_k g_k$, equality holds throughout, so $g_k = \theta_{ij} \le \pi$ for every k.

Write $\Gamma = \varrho(\cos\omega, \sin\omega)$ with $\varrho \ge 0$ and ω uniform on the circle, independent of ϱ. Given ϱ, the projection on v_i exceeds h if and only if $\varrho > h$ and ω lies within angular distance $\lambda = \arccos(h/\varrho) \in$

$(0, \frac{\pi}{2})$ of the direction of v_i. An arc of half-width λ around each direction gives three arcs of length $2\lambda < \pi$. No point lies in all three arcs: otherwise the three directions would lie in an arc of length $2\lambda < \pi$, and the complementary gap would exceed π. Two arcs around directions separated by the gap g_k overlap in an arc of length $(2\lambda - g_k)_+$ on that side, and not at all on the other side, whose length is $2\pi - g_k \geq \pi > 2\lambda$. Therefore

$$\Pi_{C(s^*)}(h) = \mathbb{E}\Big[\mathbf{1}\{\varrho > h\}\,\frac{1}{2\pi}\sum_{k=1}^{3}\big(2\lambda(\varrho) - g_k\big)_+\Big].$$

The function $g \mapsto (2\lambda - g)_+$ is convex, so by Jensen's inequality $\sum_k (2\lambda - g_k)_+ \geq 3(2\lambda - \frac{2\pi}{3})_+$, which is the value of the sum for the equal gaps $\frac{2\pi}{3}$ of the configuration of U^*. The same formula applies to C_{eq}, so $\Pi_{C(s^*)}(h) \geq \Pi_{C_{\mathrm{eq}}}(h)$. Together with the monotonicity, $\Pi_C(h) \geq \Pi_{C_{\mathrm{eq}}}(h)$. □

Proof of Lemma 8.9. Reduction to $d = 3$. If $d < 3$, the second term is absent. Add coordinates $Y_j = rY_0 + s_r\eta_j$ for $d < j \leq 3$, with fresh independent standard normal η_j. The extended vector satisfies the hypotheses with $d = 3$, its first event contains the original one, and its second term is nonnegative; so it suffices to treat $d = 3$.

The functional. Let $d = 3$ and $N = \#\{i : Y_i > 0\}$. By sign symmetry, $\mathbb{P}(Y_0 > 0,\ Y_i \leq 0\ \forall i) = \mathbb{P}(Y_0 < 0,\ Y_i \geq 0\ \forall i) = \mathbb{P}(Y_0 \leq 0,\ N = 3)$, the last step because the events differ by null events. Since $\lceil N/2 \rceil = \mathbf{1}\{N \geq 1\} + \mathbf{1}\{N = 3\}$ for $N \in \{0, 1, 2, 3\}$, the left side of the lemma equals

$$\Psi = \mathbb{E}\big[\mathbf{1}\{Y_0 \leq 0\}\lceil N/2 \rceil\big]. \tag{8.7}$$

We shall also use $\lceil N/2 \rceil = N - \mathbf{1}\{N \geq 2\}$, valid for the same values.

Coupling. Put $a_i = \mathrm{Corr}(Y_0, Y_i) \geq r$ and $s_i = \sqrt{1 - a_i^2}$. For $s_i > 0$ let $\eta_i = (Y_i - a_iY_0)/s_i$, and for $s_i = 0$ let η_i be an auxiliary standard normal variable independent of everything else. Then $\eta = (\eta_1, \eta_2, \eta_3)$ is a centered Gaussian vector with unit variances, independent of Y_0; let C be its correlation matrix. On $\{Y_0 < 0\}$ (the event $Y_0 = 0$ is null): if $s_i = 0$, then $Y_i = Y_0 < 0$ with probability 1; if $s_i > 0$, then $Y_i > 0$ if and only if $\eta_i > |Y_0|a_i/s_i$. The function $a \mapsto a/\sqrt{1 - a^2}$ is increasing on $(0, 1)$, so $|Y_0|a_i/s_i \geq |Y_0|r/s_r$. Hence, with probability 1 on $\{Y_0 < 0\}$,

$$N \leq N_h := \#\{i : \eta_i > h\}, \qquad h = h(Y_0) = \frac{r|Y_0|}{s_r} > 0. \tag{8.8}$$

Since $n \mapsto \lceil n/2 \rceil$ is nondecreasing and $\lceil n/2 \rceil = n - \mathbf{1}\{n \geq 2\}$, independence and Fubini give

$$\Psi \leq \int_{-\infty}^{0} \varphi(z)\,\mathbb{E}\lceil N_{h(z)}/2 \rceil\,dz = \int_{-\infty}^{0} \varphi(z)\Big(3\Phi\big(-h(z)\big) - \Pi_C\big(h(z)\big)\Big)dz, \qquad h(z) = \frac{r|z|}{s_r}. \tag{8.9}$$

Comparison and the ideal value. By Lemma 8.11, $\Pi_C(h) \geq \Pi_{C_{\mathrm{eq}}}(h)$ for every $h > 0$, so

$$\Psi \leq \int_{-\infty}^{0} \varphi(z)\Big(3\Phi\big(-h(z)\big) - \Pi_{C_{\mathrm{eq}}}\big(h(z)\big)\Big)dz. \tag{8.10}$$

For the ideal star $(Y_0^*, \ldots, Y_3^*)$ the coupling is an identity: its residuals are U^*, its incident correlations are r, and $Y_i^* > 0$ if and only if $U_i^* > h(Y_0^*)$ on $\{Y_0^* < 0\}$. So the right side equals Ψ for the ideal star, namely

$$\mathbb{P}(Y_0^* \leq 0 < \max_i Y_i^*) + \mathbb{P}(Y_0^* \leq 0 < \min_i Y_i^*).$$

The second probability is 0: by (8.4), $\min_i Y_i^* > 0$ forces $Y_0^* > 0$. By sign symmetry of the ideal star and the absence of ties, the first probability equals $\mathbb{P}(Y_0^* \geq 0,\ \min_i Y_i^* \leq 0) = f_{3,r}(0)$, which is $\kappa(r)$ by Lemma 8.6. Hence $\Psi \leq \kappa(r)$. □

8.5 Concentration and the proof of Theorem 8.1

We can now bound, at each fixed threshold, the expected sizes of the inner frontiers below zero and of the outer frontiers above zero, and the expected values of L_0 and L_0' at zero. To control all thresholds at once we use a finite grid of thresholds, the concentration of the finitely many grid counts (Lemma 7.9), and the deterministic interpolation of Lemma 7.8 between grid points, both for inner and for outer frontiers.

Proof of Theorem 8.1. We prove the following uniform statement: for every $\varepsilon > 0$ there is $N(\varepsilon)$ such that every graph G with $\Delta(G) \le 3$ and $n \ge N(\varepsilon)$ vertices satisfies $\mathrm{pw}(G) \le (c_3 + \varepsilon)n$. This implies Theorem 8.1: let $\rho(n)$ be the maximum of $\max(0, \mathrm{pw}(G) - c_3 n)$ over all graphs G with $\Delta(G) \le 3$ on n vertices (a finite set up to isomorphism); then $\rho(n) \le \varepsilon n$ for $n \ge N(\varepsilon)$, so $\rho(n) = o(n)$, and $\mathrm{pw}(G) \le c_3 n + \rho(n)$ for every such G.

For fixed ε, we choose the radius R and a grid of thresholds, bound the means of finitely many counts, fix a realization in which every count is close to its mean, and check the three conditions of Lemma 8.4.

Fix $\varepsilon > 0$. Since κ is continuous on $(0,1)$ and $r_R \to r_*$ as $R \to \infty$, choose an integer $R \ge 2$ with $\kappa(r_R) \le c_3 + \frac{\varepsilon}{3}$. Choose an integer $K_0 \ge 1$ with $\varphi(0)/K_0 \le \frac{\varepsilon}{3}$, and use the grid $\tau_j = -2 + j/K_0$ for $0 \le j \le K := 4K_0$. It contains $\tau_{2K_0} = 0$ and $\tau_K = 2$. Both R and K_0 depend only on ε. Let G have $\Delta(G) \le 3$ and n vertices, and let Y be the labels at radius R.

Counts. Consider the following $8K_0 + 4$ random variables, each a sum over $v \in V(G)$ of indicator variables:

- $|\partial^- S_{\tau_j}|$ for $0 \le j < 2K_0$ (thresholds $\tau_j < 0$);
- $|\partial^+ S_{\tau_j}|$ for $2K_0 < j \le 4K_0$ (thresholds $\tau_j > 0$);
- L_0 and L_0' from (8.2), with indicators $\mathbf{1}\{v \in U\} + \mathbf{1}\{v \in D \cap W\}$ and $\mathbf{1}\{v \in W\} + \mathbf{1}\{v \in D \cap U\}$ (each at most 1, as the two events are disjoint);
- the window counts $M_j := \#\{v : \tau_j < Y_v \le \tau_{j+1}\}$ for $0 \le j < 4K_0$ (the counts N_j of Lemma 7.8); and
- $\#\{v : Y_v \le -2\}$ and $\#\{v : Y_v > 2\}$.

The indicator for v in each count is determined by the degree of v and the labels on $N[v]$, hence by (L3) by $(Z_w)_{w \in B_{R+1}(v)}$.

Means. Fix v with neighbors $x_1, \dots, x_d$, $d \le 3$. By (L1), $(Y_v, Y_{x_1}, \dots, Y_{x_d})$ and its negative are centered Gaussian vectors with unit variances and incident correlations at least r_R. The events $v \in \partial^+ S_\tau$ for $\tau > 0$ and $v \in \partial^- S_\tau$ for $\tau < 0$ are contained in the events of Proposition 8.8 for this vector at τ and for its negative at $-\tau$. Up to the null events $\{Y_x = 0\}$ with $x \in N[v]$, the two events in the indicator of L_0 are the two events of Lemma 8.9 for this vector, and those of L_0' are the same events for its negative. So each of these indicators has mean at most $\kappa(r_R)$. Since moreover $\mathbb{P}(\tau_j < Y_v \le \tau_{j+1}) \le \varphi(0)/K_0$ and $\mathbb{P}(Y_v \le -2) = \mathbb{P}(Y_v > 2) = \Phi(-2) < 0.023$, summing over v shows that every inner or outer grid count and both L_0 and L_0' have mean at most $\kappa(r_R)n$, every M_j has mean at most $\varphi(0)n/K_0$, and both tail counts have mean at most $0.023n$.

Concentration. Choose $N(\varepsilon)$ so large that, for $n \ge N(\varepsilon)$,

$$(8K_0 + 4)\, 9^{R+1} < 2n^{1/2}, \qquad 2n^{3/4} \le \tfrac{\varepsilon}{3} n, \qquad n^{3/4} + 1 \le \tfrac{\varepsilon}{3} n, \qquad 0.023n + n^{3/4} \le c_3 n;$$

the last condition is possible because $0.023 < c_3$. All of these conditions depend only on ε. Let $n \ge N(\varepsilon)$. The first condition is the hypothesis of Lemma 7.9 for the $8K_0 + 4$ counts, with $\Delta = 3$,

$\varrho = R+1$ and $\lambda = n^{3/4}$, so some realization of Z has pairwise distinct labels and every count within $n^{3/4}$ of its mean. Fix it and let $y = Y$ for this realization.

Bounds at every threshold. Let $w = (c_3+\varepsilon)n$. For $t \in [-2,0)$, let j be such that $t \in [\tau_j, \tau_{j+1})$; then $j < 2K_0$, so $\tau_j < 0$, and Lemma 7.8(i) gives

$$|\partial^- S_t| \le |\partial^- S_{\tau_j}| + M_j \le \kappa(r_R)n + \frac{\varphi(0)}{K_0}n + 2n^{3/4} \le (c_3 + \tfrac{2\varepsilon}{3})n + 2n^{3/4} \le w.$$

For $t \in (0,2)$, let j be such that $t \in [\tau_j, \tau_{j+1})$; then $j \ge 2K_0$, so $\tau_{j+1} > 0$, and the same computation with the outer count at τ_{j+1} gives $|\partial^+ S_t| \le w$. For $t < -2$ and $t \ge 2$, Lemma 7.8(ii),(iii) give (inner, respectively outer) frontiers of size at most $0.023n + n^{3/4} \le c_3 n \le w$. Finally $\max(L_0, L_0') + 1 \le \kappa(r_R)n + n^{3/4} + 1 \le (c_3 + \frac{2\varepsilon}{3})n \le w$.

Conclusion. The labeling y is injective, and conditions (i)–(iii) of Lemma 8.4 hold with $w = (c_3+\varepsilon)n$. Hence $\mathrm{pw}(G) \le (c_3+\varepsilon)n$ for every graph G with $\Delta(G) \le 3$ and $n \ge N(\varepsilon)$. □

Every path decomposition is a tree decomposition, so the same bound holds for treewidth. The proof is existential; Section 10 turns it into an algorithm.

9 Pathwidth sequences and the bootstrap

The two previous sections bound the pathwidth of graphs of maximum degree at most $k \in \{3,4,5\}$ by $\sum_{d=3}^{k} a_d n_d + o(n)$, for several coefficient vectors $(a_3, \dots, a_k)$, none of which is best for all graphs (Theorems 7.1, 7.2, 8.1 and 8.2). This section states all of them in one format, the pathwidth sequence, and extends each of them to all graphs by the bootstrap lemma (Lemma 9.3). In this format, a vertex of degree at most 2 adds nothing to the bound, a vertex of degree $d \ge 3$ adds its coefficient a_d, and a vertex whose degree is beyond the given coefficients adds 1, which is what putting it into every bag adds to the width. Applied to the triple $\mathsf{S} = (\frac{5}{32}, \frac{1}{4}, \frac{11}{32})$ of Theorem 7.2, the bootstrap lemma gives the bound for all graphs (Theorem 1.3) and the bound in terms of edges (Corollary 1.4).

The sequences below use the constants

$$c_3 = \frac{3}{4\pi}\arccos\frac{5}{6} \approx 0.1398221, \qquad F_3 \approx 0.1401583, \qquad F_4 \approx 0.2457720,$$

where F_3 and F_4 are the maxima of the ideal 3-star and 4-star functions with edge correlations $\frac{2\sqrt{2}}{3}$ and $\frac{\sqrt{3}}{2}$, defined in Section 6; certified enclosures are given in Proposition 6.9.

Definition 9.1 (pathwidth sequence)**.** A sequence $a = (a_d)_{d\ge 3}$ of reals is a *pathwidth sequence* if there is a function $\rho(n) = o(n)$ such that every graph G satisfies $\mathrm{pw}(G) \le \sum_{d\ge 3} a_d\, n_d + \rho(n)$. A finite tuple $(a_3, \dots, a_k)$ stands for the sequence that continues it with $a_d = 1$ for all $d > k$.

So a finite tuple $(a_3, \dots, a_k)$ is a pathwidth sequence if and only if there is a function $\rho(n) = o(n)$ with $\mathrm{pw}(G) \le \sum_{d=3}^{k} a_d n_d + \sum_{d>k} n_d + \rho(n)$ for every graph G.

Two consequences of the definition are used below. A sequence that is entrywise at least a pathwidth sequence is again one. For a graph G of maximum degree at most k, a pathwidth sequence $(a_3, \dots, a_k)$ gives the bound $\mathrm{pw}(G) \le \sum_{d=3}^{k} a_d n_d + o(n)$.

Theorem 9.2 (pathwidth sequences)**.** *Each of the following six tuples is a pathwidth sequence:*

$$\mathsf{C} = (c_3), \qquad \mathsf{M} = (c_3,\, 0.2627), \qquad \mathsf{F} = (F_4,\, F_4),$$

$$\mathsf{S} = (\tfrac{5}{32},\, \tfrac{1}{4},\, \tfrac{11}{32}), \qquad \mathsf{P} = (F_3,\, 0.2575,\, 0.3667), \qquad \mathsf{R} = (0.1455,\, 0.2530,\, 0.3600).$$

For graphs of maximum degree at most k, the six bounds are those of Theorem 8.1 (C), Theorem 8.2 (M), Theorem 7.1(a) (F) and Theorem 7.2 (P, R, S), stated in Sections 7 and 8. In particular, every graph of maximum degree 3 satisfies $\mathrm{pw}(G) \le c_3 n_3 + o(n)$, and every graph of maximum degree 4 satisfies $\mathrm{pw}(G) \le c_3 n_3 + 0.2627\, n_4 + o(n)$ and $\mathrm{pw}(G) \le F_4(n_3 + n_4) + o(n)$.

The sequence M is entrywise below C, so M implies C. We list C nevertheless: its proof uses no computation, while that of M rests on computer-assisted certificates (Appendix A.3.1). Apart from this pair, none of the six sequences is entrywise below another, and each of M, F, P, R, S gives the smallest bound of the six for some graphs (Remark A.18).

9.1 The bootstrap lemma

Each of the six sequences has coefficients up to some degree $k \le 5$ only, so every vertex of degree above k adds 1 to its bound. The bootstrap lemma replaces these coefficients 1 by the entries of the following recurrence, which are at most 1, and smaller than 1 in every row of Table 4 except F.

Let $k \ge 3$ and let $0 \le b_3 \le \cdots \le b_k$ be reals, and put $b_0 = b_1 = b_2 := 0$. For every $d \ge 3$ for which b_d is defined, the *increment* h_d and the *running minimum* H_d are

$$h_d := b_d - b_{d-1}, \qquad H_d := \min_{3 \le i \le d} h_i.$$

Thus $h_3 = b_3$, the sequence H is nonincreasing, and $H_d \le h_d$ and $H_d \le h_3 = b_3$ for every $d \ge 3$. The *bootstrap* of $(b_3, \dots, b_k)$ is the infinite sequence $(b_d)_{d \ge 3}$ obtained by defining b_d for $d = k+1, k+2, \dots$ in turn by

$$b_d := \max\Big\{ \frac{1 + (d-1)\, b_{d-1} - H_{d-1}}{d},\ 1 - d\, H_{d-1} \Big\} \qquad (d > k). \tag{9.1}$$

This is well defined, since H_{d-1} depends only on $b_3, \dots, b_{d-1}$. For a graph G we write

$$\psi_b(G) := \sum_{v \in V(G)} b_{d_G(v)} = \sum_{d \ge 3} b_d\, n_d(G);$$

vertices of degree at most 2 contribute nothing. We drop the subscript b when the sequence is fixed.

Lemma 9.3 (bootstrap)**.** *Let $k \ge 3$ and let $0 \le b_3 \le \cdots \le b_k$ be reals with $b_k + H_k \le 1$. Suppose that there is a function $\rho(N) = o(N)$ such that every graph J with $\delta(J) \ge 3$ and $\Delta(J) \le k$ satisfies $\mathrm{pw}(J) \le \sum_{d=3}^{k} b_d\, n_d(J) + \rho(|V(J)|)$. Then the bootstrap of $(b_3, \dots, b_k)$ is a pathwidth sequence. It is nondecreasing, and all its entries are at most 1.*

The graphs J in Lemma 9.3 have no vertex of degree above k, so the bound on them holds whenever $(b_3, \dots, b_k)$ is a pathwidth sequence. Conversely, under the hypotheses of Lemma 9.3 the tuple $(b_3, \dots, b_k)$ is entrywise at least its bootstrap, which agrees with it up to degree k and has all entries at most 1; hence the tuple is a pathwidth sequence. This is how Theorem 9.2 follows from Theorems 7.1, 7.2, 8.1 and 8.2 (Appendix A.4.1). The bootstrap is used with $k = 3$ for the seed C, with $k = 4$ for M and F, and with $k = 5$ for P, R and S.

Reading the recurrence. By Lemma 9.6(ii) below, b_d is determined by

$$1 - b_d = (d-1)\, H_d + H_{d-1} \qquad (d > k). \tag{9.2}$$

Heuristically (Section 2.5), this identity says that deleting a vertex of degree d that has a neighbor of smaller degree lowers the potential ψ by at least 1, while putting the vertex back adds at most 1 to the width, since $\mathrm{pw}(G) \le \mathrm{pw}(G - v) + 1$. Indeed, the deletion lowers the potential by b_d for

Table 4: The pathwidth sequences of Corollary 9.5. The row named X is the bootstrap of the sequence X of Theorem 9.2, rounded up to four decimals; in the last column, the bootstrap is bounded by a closed form. The bootstrap of C, (0.1399, 0.4408, 0.5247, ...), is entrywise at least row M and is omitted. **Highlighted**: the smallest entry of its column. Grey: produced by the bootstrap (9.1). Black entries are those of Theorem 9.2, with c_3, F_3 and F_4 rounded up.

	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_d, $d \geq 11$
M	**0.1399**	0.2627	0.3857	0.4676	0.5319	0.5824	0.6232	0.6568	$1 - \frac{3.77}{d+1}$
F	0.2458	**0.2458**	1	1	1	1	1	1	1
P	0.1402	0.2575	0.3667	0.4541	0.5196	0.5715	0.6133	0.6478	$1 - \frac{3.87}{d+1}$
R	0.1455	0.2530	0.3600	0.4489	0.5149	0.5673	0.6096	0.6444	$1 - \frac{3.91}{d+1}$
S	$\frac{5}{32}$	$\frac{1}{4}$	$\frac{\mathbf{11}}{\mathbf{32}}$	**0.4375**	**0.5045**	**0.5581**	**0.6012**	**0.6368**	$1 - \frac{4}{d+1}$

the vertex itself, by at least H_{d-1} for that neighbor and by at least H_d for each of the other $d-1$ neighbors, whose degrees drop by one; by (9.2), these amounts add up to 1. The two terms of the maximum in (9.1) are the two cases of Lemma 9.6(iv). Regular graphs, in which no neighbor has smaller degree, are handled by the negative term of the potential ψ^- in Section 9.3.

Remark 9.4 (the recurrence of Edwards and McDermid). Take $k = 3$ and $b_3 = \alpha$ with $\frac{1}{6} \leq \alpha \leq \frac{3}{10}$. Then $H_3 = \alpha$ and $b_3 + H_3 = 2\alpha \leq 1$. By (9.1), $b_4 = \max\{\frac{1+2\alpha}{4}, 1 - 4\alpha\} = \frac{1+2\alpha}{4}$, because $1 + 2\alpha \geq 4 - 16\alpha$ is equivalent to $\alpha \geq \frac{1}{6}$. So $b_4 = g_\alpha(4)$ and $h_4 = \frac{1-2\alpha}{4} \leq \alpha$, whence $H_4 = h_4$. Moreover $H_3 \leq 3H_4$ is equivalent to $4\alpha \leq 3 - 6\alpha$, that is, $\alpha \leq \frac{3}{10}$. Hence Lemma 9.7 below applies with $j = 4$, and its last part gives $d\, b_d = (d-2)\, b_{d-1} + b_{d-2} + 1$ for all $d \geq 5$. So $b = g_\alpha$: Lemma 9.3 with $k = 3$ contains the recurrence of Edwards and McDermid on exactly their range of α. For $\alpha < \frac{1}{6}$ the second term wins and $b_4 = 1 - 4\alpha$; for $\alpha = c_3$ this is about 0.4407, worse than $g_{1/6}(4) = \frac{1}{3}$. This is why a smaller cubic coefficient alone does not improve the coefficients of higher degrees, and why Theorem 1.3 needs the base of maximum degree 5 from Theorem 7.2.

Prior work, and what is new. The induction is that of Fomin, Gaspers, Saurabh and Stepanov [FGSS09, Lemma 1], seeded with the cubic bound $\frac{n}{6} + o(n)$ of Fomin and Høie [FH06] and extended by Gaspers [Gas10, Table 6.1], in the form of the reduction theorem of Edwards and McDermid [EM15, Theorem 3.2], whose cost function g_α satisfies $d\, g_\alpha(d) = (d-2)\, g_\alpha(d-1) + g_\alpha(d-2) + 1$ and needs $\frac{1}{6} \leq \alpha \leq \frac{3}{10}$; the proofs of Lemma 9.9 and Proposition 9.10 follow theirs step by step. Three things are new. The seeds $b_3 \leq \cdots \leq b_k$ are arbitrary, with the increment of Edwards and McDermid replaced by the running minimum H_d, so that the cubic coefficient $c_3 < \frac{1}{6}$, which their theorem cannot use, and the seeds of Theorem 7.2 are admissible (Remark 9.4). The base is a per-degree bound for all graphs of minimum degree at least 3 and maximum degree at most k, not a bound for cubic graphs. And Lemma 4.5 replaces the two facts on reductions taken from Kneis, Mölle, Richter and Rossmanith [KMRR09].

The bootstraps of the six sequences. All six sequences of Theorem 9.2 satisfy the hypotheses of Lemma 9.3. Their bootstraps, rounded up, form Table 4; the bootstrap of C is entrywise at least that of M and is omitted.

Corollary 9.5 (the bootstraps)**.** *Each row of Table 4 is a pathwidth sequence.*

Among the rows of Table 4, the bootstrap of S has the smallest coefficient at every degree $d \geq 5$. Rounded up to closed forms, it gives the bound for all graphs of Theorem 1.3, and through it the bound in terms of edges of Corollary 1.4 (Sections 9.4 and 9.5). Which row of Table 4 is best for a given graph depends only on the proportions of its degrees (Remark A.18, Fig. 4).

The decompositions behind the bounds of Theorem 9.2 and Corollary 9.5 can be computed in polynomial time (Theorem 1.5). For the bootstrap of a sequence X of Theorem 9.2 with entries up to degree k, the algorithm runs the deletion process of the bootstrap lemma and decomposes the remaining components, of maximum degree at most k, by a sweep: the plain sweep for F, P, R and S, and the median-switched one for C and M (Appendices A.5.9 and A.6).

The proof of the bootstrap lemma. The proof involves no probability. A combinatorial induction deletes vertices of large degree one at a time; putting a deleted vertex back into every bag adds at most one to the width. The number of deletions is at most the potential $\psi(G)$ of the input graph G: each deletion, together with the reductions that follow it, lowers a modified potential by at least one. Vertices of degree at most 2 are removed by the reduction operations of Definition 4.2, and Lemma 4.5 bounds the effect of all reductions on the pathwidth at once, by an additive term $O(\log n)$. Section 9.2 proves the properties of the bootstrap that the induction needs, and Section 9.3 runs the deletion process and proves Lemma 9.3. Sections 9.4 and 9.5 derive Theorem 1.3 and Corollary 1.4. Appendix A.4 derives Theorem 9.2 and Corollary 9.5 from the lemma and compares the coefficients with the previous ones, and Appendix A.5.9 makes the deletion process algorithmic.

9.2 The extended sequence

We first show that the recurrence keeps the sequence nondecreasing and bounded by 1, and that it satisfies (9.2); this is all that the deletion process needs. We also record which term of the maximum is attained, because Lemma 9.7 and the proof of Proposition 9.13 use it. For $d > k$ put

$$\theta_d := \frac{1 - b_{d-1} - H_{d-1}}{d},$$

the increment h_d that the first term of the maximum in (9.1) would give.

Lemma 9.6 (the extended sequence)**.** *Let $k \geq 3$, let $0 \leq b_3 \leq \cdots \leq b_k$ be reals with $b_k + H_k \leq 1$, and extend them by* (9.1)*. Then $h_d \geq 0$ and $H_d \geq 0$ for every $d \geq 3$, and for every $d > k$:*

(i) $b_{d-1} \leq b_d \leq 1$;

(ii) $1 - b_d = (d-1)\, H_d + H_{d-1}$;

(iii) $b_d + H_d \leq 1$;

(iv) *if $\theta_d \leq H_{d-1}$, then b_d is the first term of the maximum in* (9.1) *and $h_d = H_d = \theta_d$; otherwise b_d is the second term and $H_d = H_{d-1} < h_d$.*

Proof. We first treat the seeds, and then argue by induction on $d > k$, comparing both terms of the maximum in (9.1) with b_{d-1}.

For $3 \leq d \leq k$ we have $h_d = b_d - b_{d-1} \geq 0$, because the seeds are nondecreasing and $b_2 = 0$; hence also $H_d \geq 0$.

Now let $d > k$. The induction hypothesis is that $H_{d-1} \geq 0$ and $b_{d-1} + H_{d-1} \leq 1$. For $d = k+1$ this holds by the first part and by the assumption $b_k + H_k \leq 1$. Put $\mu := H_{d-1} \geq 0$ and $\theta := \theta_d$,

so that $\theta = (1 - b_{d-1} - \mu)/d \geq 0$ by the induction hypothesis. Let p and q be the first and second terms of the maximum in (9.1). Then

$$p - b_{d-1} = \frac{1 - b_{d-1} - \mu}{d} = \theta,$$

and since $1 - b_{d-1} = d\theta + \mu$,

$$q - b_{d-1} = 1 - b_{d-1} - d\mu = d\theta - (d-1)\mu = \theta + (d-1)(\theta - \mu).$$

Case $\theta \leq \mu$. Then $q \leq p$, so $b_d = p$ and $h_d = \theta$. Hence $H_d = \min\{\mu, \theta\} = \theta$, and $1 - b_d = (1 - b_{d-1}) - \theta = (d-1)\theta + \mu = (d-1)H_d + H_{d-1}$.

Case $\theta > \mu$. Then $q > p$, so $b_d = q$ and $h_d = d\theta - (d-1)\mu > d\mu - (d-1)\mu = \mu$. Hence $H_d = \mu < h_d$, and $1 - b_d = d\mu = (d-1)H_d + H_{d-1}$.

This proves (ii) and (iv). In both cases $h_d \geq 0$ (as $\theta \geq 0$ and $\mu \geq 0$), so $b_d \geq b_{d-1}$ and $H_d \geq 0$. By (ii) and $d \geq 2$, $1 - b_d = (d-1)H_d + H_{d-1} \geq (d-1)H_d \geq H_d \geq 0$, which gives $b_d \leq 1$ and (iii). In particular the induction hypothesis holds for $d+1$. □

Once two consecutive running minima are within a factor 3 of each other, the first term of the maximum is attained at every later degree, and the running minima follow a linear recurrence. If, in addition, the increment h_j at the later degree j equals H_j, then from there on the extended sequence satisfies the recurrence of Edwards and McDermid. This is how the coefficients of Theorem 1.3 are computed.

Lemma 9.7 (the concave regime)**.** *Let b be as in Lemma 9.6, and suppose that $H_j \leq H_{j-1} \leq 3H_j$ for some $j > k$. Then for every $d > j$,*

$$d\,H_d = (d-3)\,H_{d-1} + H_{d-2}, \qquad h_d = H_d, \qquad H_d \leq H_{d-1} \leq 3H_d.$$

If moreover $h_j = H_j$, then $d\,b_d = (d-2)\,b_{d-1} + b_{d-2} + 1$ for every $d > j$.

Proof. We prove the three statements by induction on $d > j$, together with the hypothesis $H_{d-1} \leq H_{d-2} \leq 3H_{d-1}$, which holds for $d = j+1$ by assumption. Since $d - 1 \geq j > k$, Lemma 9.6(ii) at $d-1$ gives $1 - b_{d-1} = (d-2)H_{d-1} + H_{d-2}$, and therefore

$$\theta_d = \frac{1 - b_{d-1} - H_{d-1}}{d} = \frac{(d-3)\,H_{d-1} + H_{d-2}}{d}.$$

The inequality $\theta_d \leq H_{d-1}$ is equivalent to $H_{d-2} \leq 3H_{d-1}$, which holds. By Lemma 9.6(iv), $h_d = H_d = \theta_d$; this is the claimed recurrence for H_d, and $H_d \leq H_{d-1}$ because H is nonincreasing. Finally, $H_{d-1} \leq 3H_d = 3\theta_d$ is equivalent to $d\,H_{d-1} \leq 3(d-3)H_{d-1} + 3H_{d-2}$, that is, to $(9-2d)\,H_{d-1} \leq 3H_{d-2}$. This holds because $d \geq j+1 \geq k+2 \geq 5$ makes the left side nonpositive, while $H_{d-2} \geq 0$ by Lemma 9.6. So the hypothesis holds for $d+1$.

For the last statement, let $d > j$. Then $h_d = H_d$ by the first part, and $h_{d-1} = H_{d-1}$, by assumption if $d - 1 = j$ and by the first part if $d - 1 > j$. Substituting into Lemma 9.6(ii) gives $1 - b_d = (d-1)(b_d - b_{d-1}) + (b_{d-1} - b_{d-2})$, which rearranges to $d\,b_d = (d-2)\,b_{d-1} + b_{d-2} + 1$. □

9.3 The deletion process and the proof of the bootstrap lemma

Throughout this subsection, b is as in Lemma 9.6, and $\rho\colon \mathbb{N} \to [0, \infty)$ is a nondecreasing function such that every graph J with $\delta(J) \geq 3$ and $\Delta(J) \leq k$ satisfies

$$\mathrm{pw}(J) \leq \sum_{d=3}^{k} b_d\, n_d(J) + \rho(|V(J)|). \tag{9.3}$$

By Lemma 9.6, b is nondecreasing, $0 \le b_d \le 1$, and (9.2) holds for all $d > k$. We show that every graph G with $n \ge 1$ vertices satisfies

$$\mathrm{pw}(G) \le \sum_{d \ge 3} b_d \, n_d + b_3 + \rho(n) + 6 \log_2 n + 2, \tag{9.4}$$

and deduce Lemma 9.3.

The proof of (9.4) combines three bounds. The deletion process (Proposition 9.10) deletes vertices of large degree and applies reductions until every component is basic (Definition 9.8); the number of deleted vertices is at most the decrease of a modified potential ψ^-. Lemma 9.11 bounds the pathwidth of basic graphs, and Lemma 4.5 bounds the effect of the reductions. The three bounds are added up after Lemma 9.11.

Facts about reductions. We use the properties (RP1)–(RP3) of the reduction operations, stated after Definition 4.2. Since b is nondecreasing and nonnegative, (RP1) implies that an operation does not increase ψ. The effect of the reductions on the pathwidth is bounded by Lemma 4.5, applied to sequences of reduction operations and vertex deletions; recall that the empty graph has pathwidth -1 (Section 4.1).

The potential. For a graph G with $\delta(G) \ge 3$ and components $G_1, \dots, G_t$ put

$$\psi^-(G) := \psi(G) - \sum_{i=1}^{t} H_{\delta(G_i)},$$

and $\psi^-(G) := 0$ for the empty graph. Each component G_i contains a vertex of degree $\delta(G_i) \ge 3$, whose weight is $b_{\delta(G_i)} \ge h_{\delta(G_i)} \ge H_{\delta(G_i)}$, because $b_{\delta(G_i)-1} \ge 0$. Hence $\psi^-(G) \ge 0$, and also $\psi^-(G_i) \ge 0$ for each component, and ψ^- is additive over components. The negative term is the one of Edwards and McDermid [EM15]. It is needed when the deleted vertex lies in a regular component, where no neighbor has smaller degree (Case 1 in the proof of Lemma 9.9).

Basic graphs. The process stops at components of the following kind.

Definition 9.8 (basic graphs)**.** A connected graph J with $\delta(J) \ge 3$ is *basic* if $\Delta(J) \le k$, or if there is a vertex $v \in V(J)$ such that some maximal sequence of reduction operations applied to $J - v$ ends with the empty graph.

The deletion process of Proposition 9.10 below repeats one step: delete a vertex of maximum degree, then reduce. The next lemma shows that this step lowers ψ^- by at least one. The algorithm of Appendix A.5.9 uses it directly, without testing whether a component is basic.

Lemma 9.9 (one deletion)**.** *Let G be a connected graph with $\delta(G) \ge 3$ and $\Delta(G) \ge k+1$. Let w be a vertex of degree $\Delta(G)$ that has a neighbor of smaller degree, if such a vertex exists, and any vertex of degree $\Delta(G)$ otherwise. Let G^* be the final graph of a maximal sequence of reduction operations applied to $G - w$. If G^* is nonempty, then $1 + \psi^-(G^*) \le \psi^-(G)$.*

Proof. Write $\Delta := \Delta(G)$, so $\Delta \ge k+1 \ge 4$. Let $v_1, \dots, v_\Delta$ be the neighbors of w, ordered so that $d(v_1) \le d(v_2) \le \cdots \le d(v_\Delta)$; here and below, $d(\cdot)$ denotes the degree in G. Let $G' := G - w$. As G^* is nonempty, $\delta(G^*) \ge 3$ by (RP3). As G is connected, the claim reads

$$1 + \psi^-(G^*) \le \psi^-(G) = \psi(G) - H_{\delta(G)}. \tag{9.5}$$

We bound $1 + \psi^-(G^*)$ from above by $\psi(G)$ minus four terms $J_1, \dots, J_4$, and show by cases that $J_1 + \cdots + J_4 \geq H_{\delta(G)}$.

Rewriting the left side. Every neighbor v_i of w loses one degree, from $d(v_i) \geq 3$ to $d(v_i) - 1$, and its weight drops by $h_{d(v_i)}$. So

$$\psi(G') = \psi(G) - b_\Delta - \sum_{i=1}^{\Delta} h_{d(v_i)}.$$

Since $H \geq 0$ and some component of G^* has minimum degree $\delta(G^*)$, we have $\psi^-(G^*) \leq \psi(G^*) - H_{\delta(G^*)}$. Put

$$J_4 := \psi(G') - \psi(G^*) + H_{\delta(G^*)},$$

so that $\psi^-(G^*) \leq \psi(G') - J_4$. Since $\Delta > k$, identity (9.2) gives $1 - b_\Delta = (\Delta - 1)H_\Delta + H_{\Delta-1}$. Combining,

$$1 + \psi^-(G^*) \leq \psi(G) + (\Delta - 1)H_\Delta + H_{\Delta-1} - \sum_{i=1}^{\Delta} h_{d(v_i)} - J_4 = \psi(G) - (J_1 + J_2 + J_3 + J_4),$$

where

$$J_1 := h_{d(v_1)} - H_{\Delta-1}, \qquad J_2 := h_{d(v_2)} - H_\Delta, \qquad J_3 := \sum_{i=3}^{\Delta} (h_{d(v_i)} - H_\Delta).$$

Hence (9.5) follows once we show

$$J_1 + J_2 + J_3 + J_4 \geq H_{\delta(G)}. \tag{9.6}$$

Signs. For $3 \leq j \leq \Delta$ we have $h_j \geq H_j \geq H_\Delta$, because H is nonincreasing. As $3 \leq d(v_i) \leq \Delta$, this gives $J_2 \geq 0$ and $J_3 \geq 0$. Similarly $J_1 \geq 0$ whenever $d(v_1) \leq \Delta - 1$, since then $h_{d(v_1)} \geq H_{d(v_1)} \geq H_{\Delta-1}$. Finally $J_4 \geq 0$: by (RP1), $\psi(G^*) \leq \psi(G')$, and $H_{\delta(G^*)} \geq 0$.

Claim. If some $u \in V(G')$ has $d_{G'}(u) = j \geq 3$, then $J_4 \geq H_j$.

Indeed, suppose first that u is a vertex of G^*. By (RP1), $d_{G^*}(u) \leq j$, so $3 \leq \delta(G^*) \leq j$ and $H_{\delta(G^*)} \geq H_j$; together with $\psi(G') - \psi(G^*) \geq 0$ this gives $J_4 \geq H_j$. Otherwise u was removed by a reduction. Every vertex of G^* has weight in G^* at most its weight in G' by (RP1), and every removed vertex has nonnegative weight, so $\psi(G') - \psi(G^*) \geq b_j \geq h_j \geq H_j$; as $H_{\delta(G^*)} \geq 0$, again $J_4 \geq H_j$.

Case 1: G is Δ-regular. Then $\delta(G) = \Delta$ and $J_1 = h_\Delta - H_{\Delta-1}$. In G' the vertex v_1 has degree $\Delta - 1 \geq 3$, so $J_4 \geq H_{\Delta-1}$ by the claim. Hence $J_1 + J_4 \geq h_\Delta \geq H_\Delta = H_{\delta(G)}$, and (9.6) follows from $J_2, J_3 \geq 0$.

Case 2: G is not regular. Since G is connected, some edge joins a vertex of degree Δ to a vertex of smaller degree (on a path from a vertex of degree Δ to a vertex of smaller degree, take the first such edge). So w was chosen with a neighbor of smaller degree, $d(v_1) \leq \Delta - 1$, and $J_1 \geq 0$. All four terms are now nonnegative.

Case 2a: some $u \in V(G')$ has $3 \leq d_{G'}(u) \leq \delta(G)$. By the claim, $J_1 + J_2 + J_3 + J_4 \geq J_4 \geq H_{d_{G'}(u)} \geq H_{\delta(G)}$.

Case 2b: no such vertex exists. Since $\delta(G) < \Delta = d(w)$, there is a vertex $u \neq w$ of degree $\delta(G)$, and $d_{G'}(u) \in \{\delta(G) - 1, \delta(G)\}$. If $\delta(G) \geq 4$, then $3 \leq d_{G'}(u) \leq \delta(G)$, which is Case 2a. Hence $\delta(G) = 3$. Every vertex u of degree 3 is adjacent to w, since otherwise $d_{G'}(u) = 3$, which is Case 2a. In particular $d(v_1) = 3$, and $H_{\delta(G)} = H_3 = h_3$.

- *Case 2b-i: $d(v_\Delta) \geq 4$.* Then $3 \leq d_{G'}(v_\Delta) = d(v_\Delta) - 1 \leq \Delta - 1$, so by the claim $J_4 \geq H_{d_{G'}(v_\Delta)} \geq H_{\Delta-1}$. With $J_1 = h_3 - H_{\Delta-1}$ this gives $J_1 + J_4 \geq h_3 = H_{\delta(G)}$.

- *Case 2b-ii:* $d(v_\Delta) = 3$. Then all neighbors of w have degree 2 in G'. Let u be any vertex of the nonempty graph G^*. By (RP1), $d_{G'}(u) \geq d_{G^*}(u) \geq 3$, so u is not a neighbor of w; hence $d_{G'}(u) = d(u) \in [3, \Delta]$, and the claim gives $J_4 \geq H_{d(u)} \geq H_\Delta$. With $J_2 = h_3 - H_\Delta$ (note $d(v_2) = 3$) this gives $J_2 + J_4 \geq h_3 = H_{\delta(G)}$.

In every case (9.6) holds, which proves (9.5). □

The following proposition is the heart of the proof. It deletes vertices until every component is basic, and bounds the number of deleted vertices from above by the decrease of ψ^-.

Proposition 9.10 (deletion process)**.** *Every graph G with $\delta(G) \geq 3$ admits a finite sequence of vertex deletions and reduction operations that transforms G into a graph R all of whose components are basic, such that the set X of deleted vertices satisfies*

$$|X| + \psi^-(R) \leq \psi^-(G).$$

Proof. We argue by induction on $|V(G)|$. A disconnected graph is handled component by component. For a connected graph that is not basic, we delete one vertex w of maximum degree, reduce, and apply the induction hypothesis to the result; by Lemma 9.9, this deletion, with the reductions after it, lowers ψ^- by at least one.

If every component of G is basic, take the empty sequence.

G disconnected. Each component G_i has minimum degree at least 3 and fewer vertices than G, so by induction it admits a sequence with deleted set X_i and final graph R_i satisfying $|X_i| + \psi^-(R_i) \leq \psi^-(G_i)$. By (RP2), performing these sequences one after the other is a sequence of deletions and reductions on G. Its final graph is the disjoint union R of the R_i, whose components are all basic, and its deleted set is the disjoint union X of the X_i. Summing, and using additivity of ψ^-, gives $|X| + \psi^-(R) \leq \psi^-(G)$.

G connected and not basic. Then $\Delta := \Delta(G) \geq k + 1 \geq 4$, since otherwise G would be basic. Choose a vertex w of degree Δ that has a neighbor of smaller degree, if such a vertex exists, and any vertex of degree Δ otherwise, and let G^* be the final graph of a maximal sequence of reduction operations applied to $G - w$. Since G is not basic (take $v = w$ in Definition 9.8), G^* is not empty, so $\delta(G^*) \geq 3$ by (RP3). As $|V(G^*)| < |V(G)|$, the induction hypothesis gives a sequence from G^* to a graph R with basic components and deleted set X^*, such that $|X^*| + \psi^-(R) \leq \psi^-(G^*)$. The concatenated sequence (delete w, reduce maximally, then follow the sequence for G^*) transforms G into R and has deleted set $X = X^* \cup \{w\}$. By Lemma 9.9, $|X| + \psi^-(R) \leq 1 + \psi^-(G^*) \leq \psi^-(G)$. □

The pathwidth of a basic graph is bounded either by the hypothesis (9.3) or by Lemma 4.5.

Lemma 9.11 (basic graphs)**.** *Let J be a basic graph with N vertices. Then $\psi(J) \leq \psi^-(J) + b_3$. If $\Delta(J) \leq k$, then $\mathrm{pw}(J) \leq \psi^-(J) + b_3 + \rho(N)$. Otherwise $\mathrm{pw}(J) \leq 3 \log_2 N + 1$.*

Proof. As J is connected, $\psi(J) = \psi^-(J) + H_{\delta(J)} \leq \psi^-(J) + b_3$, because $H_d \leq h_3 = b_3$ for all $d \geq 3$. If $\Delta(J) \leq k$, then $\delta(J) \geq 3$ and (9.3) give $\mathrm{pw}(J) \leq \psi(J) + \rho(N)$, which with the first claim gives the bound.

Otherwise there is a vertex v such that a maximal sequence of reductions applied to $J - v$ ends with the empty graph. Apply Lemma 4.5 to $F = J$ and the sequence that deletes v and then performs these reductions: here $D = \{v\}$ and the final graph is empty, so $\mathrm{pw}(J) \leq 1 + (-1) + 3 \log_2 N + 1$. □

Proof of (9.4). Let G be a graph with $n \geq 1$ vertices. Apply a maximal sequence of reduction operations to G, and let R_0 be its final graph. If R_0 is not empty, then $\delta(R_0) \geq 3$ by (RP3); continue with the sequence of Proposition 9.10 for R_0. Together this is one sequence of deletions

and reductions from G to a graph R, with deleted set X; if R_0 is empty, then $R = R_0$ and $X = \emptyset$. We bound the three ingredients.

- *The potential.* By (RP1), $\psi(R_0) \le \psi(G)$, and $\psi^-(R_0) \le \psi(R_0)$ because $H \ge 0$. With Proposition 9.10, $|X| + \psi^-(R) \le \psi^-(R_0) \le \psi(G)$. (If R_0 is empty, both sides of the first inequality are 0.)

- *The final graph.* If R is empty, then $\mathrm{pw}(R) = -1$. Otherwise every component C of R is basic and has at most n vertices. By Lemma 9.11, monotonicity of ρ and $\log_2$, and nonnegativity of $\psi^-(C)$, b_3 and ρ,
$$\mathrm{pw}(C) \le \psi^-(C) + b_3 + \rho(n) + 3\log_2 n + 1.$$
The pathwidth of a graph is the maximum over its components (Section 4.1), and $\psi^-(C) \le \psi^-(R)$ by additivity and nonnegativity. Hence $\mathrm{pw}(R) \le \psi^-(R) + b_3 + \rho(n) + 3\log_2 n + 1$ in all cases.

- *The reductions.* By Lemma 4.5 applied to the whole sequence, with $D = X$, $\mathrm{pw}(G) \le |X| + \mathrm{pw}(R) + 3\log_2 n + 1$.

Adding up, $\mathrm{pw}(G) \le |X| + \psi^-(R) + b_3 + \rho(n) + 6\log_2 n + 2 \le \psi(G) + b_3 + \rho(n) + 6\log_2 n + 2$, and $\psi(G) = \sum_{d\ge 3} b_d n_d$ because $b_0 = b_1 = b_2 = 0$. □

Proof of Lemma 9.3. Let k and $b_3, \dots, b_k$ be as in Lemma 9.3, and let b be their bootstrap; it is as in Lemma 9.6. Let ρ_0 be the function ρ of Lemma 9.3, and let $\rho := \bar{\rho}_0$ be its running maximum (4.2), which is nonnegative, nondecreasing, at least ρ_0, and again $o(N)$. Then (9.3) holds, so (9.4) holds for every graph with $n \ge 1$ vertices, and its remainder $b_3 + \rho(n) + 6\log_2 n + 2$ is $o(n)$; the empty graph has pathwidth -1. So b is a pathwidth sequence. By Lemma 9.6(i), b is nondecreasing and $b_d \le 1$ for $d > k$; for $d \le k$, $b_d \le b_k \le 1 - H_k \le 1$, because $H_k \ge 0$ by Lemma 9.6. □

9.4 Proof of Theorem 1.3

We apply the explicit bound (9.4) with $k = 5$ to the triple S of Theorem 7.2. Throughout, $\sigma = (\sigma_d)_{d\ge 0}$ denotes the bootstrap of S: it has $\sigma_0 = \sigma_1 = \sigma_2 = 0$, $(\sigma_3, \sigma_4, \sigma_5) = \mathsf{S}$, and it is extended by (9.1) with $k = 5$. We write h_d and H_d for its increments and their running minima. The seeds are nondecreasing, with increments $h_3 = \frac{5}{32}$ and $h_4 = h_5 = \frac{3}{32}$; so $H_5 = \frac{3}{32}$ and $\sigma_5 + H_5 = \frac{14}{32} \le 1$, and σ is as in Lemma 9.6.

We compute σ exactly up to degree 8, and bound it from above by $1 - \frac{111}{28(d+1)}$ from degree 7 on and by $1 - \frac{4}{d+1}$ from degree 11 on (Proposition 9.13). These two bounds use the following lemma on the recurrence $d\,u_d = (d-2)\,u_{d-1} + u_{d-2}$. In the concave regime of Lemma 9.7, the deficits $u_d = 1 - b_d$ satisfy this recurrence. We use the lemma again in the proofs of Lemma 9.16 and of Corollary 9.5. Its first part says that the recurrence has a conserved quantity λ, which reduces it to the first-order rule $d\,u_d = \lambda - u_{d-1}$; the second part turns this rule into a two-sided bound on the weighted deficits $w_d = (d+1)\,u_d$.

Lemma 9.12 (weighted deficits)**.** *Let $d_0 \ge 3$, and let $(u_d)_{d\ge d_0-2}$ be reals with $d\,u_d = (d-2)\,u_{d-1} + u_{d-2}$ for all $d \ge d_0$. Put $w_d := (d+1)\,u_d$ and $\lambda := (d_0-1)\,u_{d_0-1} + u_{d_0-2}$. Then:*

(i) *$d\,u_d + u_{d-1} = \lambda$ for all $d \ge d_0 - 1$;*

(ii) *if $\lambda > 0$ and $\frac{d_1+1}{d_1+2}\,\lambda < w_{d_1} < \lambda$ for some $d_1 \ge d_0 - 1$, then*
$$\Big(1 - \frac{1}{d^2}\Big)\lambda < w_d < \lambda \qquad \text{for every } d > d_1.$$

Proof. (i) For $d \ge d_0$, $(d\,u_d + u_{d-1}) - \big((d-1)\,u_{d-1} + u_{d-2}\big) = d\,u_d - (d-2)\,u_{d-1} - u_{d-2} = 0$.

(ii) For $d \ge d_0$, the rule $d\,u_d = \lambda - u_{d-1}$ of (i) and $u_{d-1} = w_{d-1}/d$ give

$$w_d = \frac{d+1}{d}\Big(\lambda - \frac{w_{d-1}}{d}\Big), \qquad \text{so} \qquad w_d - \lambda = \frac{d\,\lambda - (d+1)\,w_{d-1}}{d^2}.$$

We argue by induction on $d > d_1$ under the hypothesis $\frac{d}{d+1}\lambda < w_{d-1} < \lambda$, which is the assumption for $d = d_1 + 1$; note that $d \ge d_0$ throughout. The lower bound on w_{d-1} makes the numerator negative, so $w_d < \lambda$. The upper bound gives $w_d - \lambda > \frac{d\lambda-(d+1)\lambda}{d^2} = -\frac{\lambda}{d^2}$. Finally $1 - \frac{1}{d^2} \ge \frac{d+1}{d+2}$ because $d^2 \ge d+2$ for $d \ge 2$; hence $\frac{d+1}{d+2}\lambda < w_d < \lambda$, which is the hypothesis for $d+1$. □

The next statement computes σ, records where the threshold $d \ge 11$ of Theorem 1.3 comes from and that it is sharp, and shows $\sigma_d \le a_d$ for every $d \ge 3$.

Proposition 9.13 (the bootstrap of S)**.** *The bootstrap σ of S satisfies $\sigma_d = \frac{3d-4}{32}$ for $3 \le d \le 6$,*

$$1 - \frac{129}{32\,(d+1)} < \sigma_d < 1 - \frac{129}{32\,(d+1)} + \frac{129}{32\,d^2(d+1)} \qquad (d \ge 5),$$

and, for $d \ge 3$, $\sigma_d \le 1 - \frac{4}{d+1}$ if and only if $d \ge 11$. Moreover, $\sigma_d \le a_d$ for every $d \ge 3$, where a_d are the coefficients of Theorem 1.3.

Proof. Steps 1 and 2 compute σ up to degree 6 and show that from degree 7 on it satisfies the recurrence of Edwards and McDermid. Steps 3 and 4 apply Lemma 9.12 to the deficits $u_d = 1 - \sigma_d$ and give the two closed-form bounds. The last paragraph collects the remaining claims.

Step 1: degrees up to 6. At $d = 6$, $\theta_6 = (1 - \sigma_5 - H_5)/6 = \frac{18/32}{6} = \frac{3}{32} = H_5$. By Lemma 9.6(iv), σ_6 is the first term of the maximum and $h_6 = H_6 = \frac{3}{32}$; so $\sigma_6 = \frac{11}{32} + \frac{3}{32} = \frac{7}{16}$. (The second term $1 - 6H_5 = \frac{7}{16}$ is equal.) Hence

$$\sigma_d = \frac{3d-4}{32} \qquad (3 \le d \le 6).$$

Step 2: the recurrence from degree 7. We have $j = 6 > k = 5$ and $H_6 = H_5 = \frac{3}{32}$, so $H_6 \le H_5 \le 3H_6$; also $h_6 = H_6$. By Lemma 9.7,

$$d\,\sigma_d = (d-2)\,\sigma_{d-1} + \sigma_{d-2} + 1 \qquad (d \ge 7).$$

For example, $\sigma_7 = \frac{5\cdot\frac{14}{32}+\frac{11}{32}+1}{7} = \frac{113}{224}$ and $\sigma_8 = \frac{6\cdot\frac{113}{224}+\frac{98}{224}+1}{8} = \frac{125}{224}$.

Step 3: the bound $1 - \frac{111}{28(d+1)}$. Put $u_d := 1 - \sigma_d$ and $w_d := (d+1)\,u_d$. Since $d = (d-2) + 1 + 1$, the recurrence of Step 2 is equivalent to $d\,u_d = (d-2)\,u_{d-1} + u_{d-2}$ for $d \ge 7$; it also holds at $d = 6$, because $6\sigma_6 = \frac{21}{8} = 4\sigma_5 + \sigma_4 + 1$. So Lemma 9.12(i) applies with $d_0 = 6$ and $\lambda = 5u_5 + u_4 = \frac{105}{32} + \frac{24}{32}$:

$$d\,u_d + u_{d-1} = \tfrac{129}{32} \qquad (d \ge 5).$$

Moreover $w_5 = w_6 = \frac{63}{16} = \frac{126}{32}$ and $\frac{7}{8}\cdot\frac{129}{32} = \frac{903}{256} < \frac{126}{32} < \frac{129}{32}$, so Lemma 9.12(ii) applies with $d_1 = 6$:

$$\Big(1 - \frac{1}{d^2}\Big)\frac{129}{32} < w_d < \frac{129}{32} \qquad (d \ge 7). \tag{9.7}$$

Now $w_7 = 8 \cdot \frac{111}{224} = \frac{111}{28}$, and for $d \ge 8$, $w_d > \frac{63}{64}\cdot\frac{129}{32} = \frac{8127}{2048} > \frac{111}{28}$. Hence $w_d \ge \frac{111}{28}$ for every $d \ge 7$, that is,

$$\sigma_d \le 1 - \frac{111}{28\,(d+1)} \qquad (d \ge 7),$$

with equality at $d = 7$.

Step 4: the bound $1 - \frac{4}{d+1}$. The inequality $\sigma_d \le 1 - \frac{4}{d+1}$ is $w_d \ge 4$. For $d \ge 12$, (9.7) gives $w_d > (1 - \frac{1}{144})\frac{129}{32} = \frac{18447}{4608} > 4$, as $18447 > 4 \cdot 4608 = 18432$. For $d = 11$, the rule $u_d = (\frac{129}{32} - u_{d-1})/d$ gives $u_7 = \frac{111}{224}$, $u_8 = \frac{99}{224}$, $u_9 = \frac{67}{168}$, $u_{10} = \frac{2441}{6720}$ and $u_{11} = \frac{24649}{73920}$, so $w_{11} = \frac{24649}{6160} > 4$ because $4 \cdot 6160 = 24640$. Hence

$$\sigma_d \le 1 - \frac{4}{d+1} \qquad (d \ge 11).$$

The remaining claims. The two-sided bound is (9.7) for $d \ge 7$, and for $d \in \{5, 6\}$ it reads $(1 - \frac{1}{d^2})\frac{129}{32} \le \frac{35}{36} \cdot \frac{129}{32} = \frac{4515}{1152} < \frac{126}{32} = w_d < \frac{129}{32}$. Step 4 gives $\sigma_d \le 1 - \frac{4}{d+1}$ for $d \ge 11$. Conversely, $w_3 = 4 \cdot \frac{27}{32} = \frac{27}{8}$, $w_4 = 5 \cdot \frac{3}{4} = \frac{15}{4}$, and the values of Step 4 give

$$(w_5, \dots, w_{10}) = \Big(\tfrac{63}{16}, \tfrac{63}{16}, \tfrac{111}{28}, \tfrac{891}{224}, \tfrac{335}{84}, \tfrac{26851}{6720}\Big),$$

all below 4 (for the last, $4 \cdot 6720 = 26880 > 26851$), so $\sigma_d > 1 - \frac{4}{d+1}$ for $3 \le d \le 10$. Finally, Steps 1, 3 and 4 give $\sigma_d \le a_d$ for every $d \ge 3$. □

The statement of Theorem 1.3 rounds the coefficients σ_d for $d \ge 7$ up to closed forms. The proof above gives $\sigma_7 = \frac{113}{224}$, $\sigma_8 = \frac{125}{224}$, $\sigma_9 = \frac{101}{168}$ and $\sigma_{10} = \frac{4279}{6720}$, which Table 7 compares with a_d.

Proof of Theorem 1.3. By Theorem 7.2 with the triple S, there is a function ρ_0 with $\rho_0(N) = o(N)$ such that every graph J with $\Delta(J) \le 5$, in particular every J with $\delta(J) \ge 3$ and $\Delta(J) \le 5$, satisfies $\mathrm{pw}(J) \le \sigma_3 n_3(J) + \sigma_4 n_4(J) + \sigma_5 n_5(J) + \rho_0(|V(J)|)$. Its running maximum $\rho := \bar{\rho}_0$ of (4.2) is nonnegative, nondecreasing, at least ρ_0 and again $o(N)$, so (9.3) holds with ρ. As σ is as in Lemma 9.6, (9.4) gives, for every graph G with $n \ge 1$ vertices,

$$\mathrm{pw}(G) \le \sum_{d \ge 3} \sigma_d \, n_d + \rho_4(n), \qquad \rho_4(n) := \tfrac{5}{32} + \rho(n) + 6 \log_2 n + 2. \tag{9.8}$$

The function ρ_4 is nondecreasing and $\rho_4(n) = o(n)$. By Proposition 9.13, $\sigma_d \le a_d$ for every $d \ge 3$, so $\mathrm{pw}(G) \le \sum_{d \ge 3} a_d n_d + \rho_4(n)$ for every graph G. □

9.5 Edges and average degree

This subsection derives the bound in terms of edges (Corollary 1.4) and a bound in terms of the average degree (Corollary 9.15) from Theorem 1.3, and compares the latter with the bound of Edwards and McDermid. Throughout, a_d ($d \ge 3$) are the coefficients of Theorem 1.3, and $a_0 = a_1 = a_2 := 0$. The edge bound follows from Theorem 1.3 because no coefficient a_d exceeds $\frac{7}{48} \cdot \frac{d}{2}$; the maximum ratio $2a_d/d$ is attained at $d = 6$.

Proof of Corollary 1.4. We first show that

$$a_d \le \frac{7}{48} \cdot \frac{d}{2} \qquad \text{for every } d \ge 3, \tag{9.9}$$

with equality only at $d = 6$, and that $2a_d/d \le \frac{113}{784}$ for $d \ne 6$. For $3 \le d \le 6$, $a_d = \frac{3d-4}{32} \le \frac{7d}{96}$ is equivalent to $9d - 12 \le 7d$, that is, to $d \le 6$, and $2a_d/d = \frac{3}{16} - \frac{1}{4d} \le \frac{11}{80} < \frac{113}{784}$ for $d \le 5$. For $d \ge 7$ we have $a_d \le 1 - \frac{111}{28(d+1)}$: with equality for $7 \le d \le 10$; for $d \ge 11$ because $4 > \frac{111}{28}$. Moreover

$$\frac{113}{784} \cdot \frac{d}{2} - \Big(1 - \frac{111}{28\,(d+1)}\Big) = \frac{(d-7)(113d - 664)}{1568\,(d+1)} \ge 0 \qquad (d \ge 7),$$

with equality at $d = 7$, and $\frac{113}{784} < \frac{7}{48}$. This proves both claims.

Now let G be a graph with m edges. Summing (9.9) over the vertices of G of degree at least 3, and using that vertices of degree at most 2 contribute nonnegative amounts to the right side,

$$\sum_{d\geq 3} a_d\, n_d \leq \frac{7}{96} \sum_{v\in V(G)} d_G(v) = \frac{7}{48}\, m.$$

So Theorem 1.3 gives $\mathrm{pw}(G) \leq \frac{7}{48}m + o(n)$. □

Remark 9.14 (remainder in m). The remainder of Corollary 1.4 can also be taken as a function of m: $\mathrm{pw}(G) \leq \frac{7}{48}m + \rho(m+1)$ for every graph G with m edges, where ρ is the remainder of Theorem 1.3, taken nondecreasing by (4.2). Indeed, the computation above applied to a component H of G gives $\mathrm{pw}(H) \leq \frac{7}{48}|E(H)| + \rho(|V(H)|) \leq \frac{7}{48}m + \rho(m+1)$, since $|V(H)| \leq |E(H)| + 1 \leq m+1$, and $\mathrm{pw}(G)$ is the maximum of $\mathrm{pw}(H)$ over the components. Taking the maximum over components is essential: a sum of remainders over many small components would not be $o(m)$. The same step shows that every bound of the form $\sum_d b_d n_d + o(n)$ with $b_d \leq \beta d/2$ for all d yields $\mathrm{pw}(G) \leq \beta m + o(m)$. The algorithms of Theorem 10.1 compute path decompositions of width at most $\frac{7}{48}m + \rho(2m)$, a remainder in m as well.

The same coefficients give a bound in terms of the average degree $\bar{d} = 2m/n$. Let $A\colon [0,\infty) \to \mathbb{R}$ be given by $A(x) := \frac{7}{96}x$ for $0 \leq x \leq 6$ and, on $[d, d+1]$ for each integer $d \geq 6$, by linear interpolation between $A(d) = a_d$ and $A(d+1) = a_{d+1}$; the two definitions agree at 6, since $a_6 = \frac{7}{16}$.

Corollary 9.15 (average degree)**.** *Every graph G with $n \geq 1$ vertices and average degree $\bar{d}$ satisfies $\sum_{d\geq 3} a_d n_d \leq A(\bar{d})\, n$, and hence* $\mathrm{pw}(G) \leq A(\bar{d})\, n + o(n)$*, where $A(\bar{d}) \leq 1 - \frac{111}{28(\bar{d}+1)}$ if $\bar{d} \geq 7$ and $A(\bar{d}) \leq 1 - \frac{4}{\bar{d}+1}$ if $\bar{d} \geq 11$. If $\delta(G) \geq 2$, then* $\mathrm{pw}(G) \leq \frac{3m-2n}{16} + o(n)$.

Proof. We first show that A is concave and lies above every coefficient a_d, and apply Jensen's inequality. The closed-form bounds on $A(\bar{d})$ then follow by comparing A with concave functions, and the bound for $\delta(G) \geq 2$ follows from $a_d \leq \frac{3d-4}{32}$.

The function A is continuous and piecewise linear. Its slopes are $\frac{7}{96}$ on $[0,6]$, $a_7 - a_6 = \frac{15}{224}$ on $[6,7]$, $a_{d+1} - a_d = \frac{111}{28(d+1)(d+2)}$ on $[d, d+1]$ for $7 \leq d \leq 9$, $a_{11} - a_{10} = \frac{25}{924}$ on $[10,11]$, and $a_{d+1} - a_d = \frac{4}{(d+1)(d+2)}$ on $[d, d+1]$ for $d \geq 11$. These slopes are nonincreasing, because $\frac{7}{96} > \frac{15}{224} > \frac{111}{28\cdot 72}$, the third expression decreases in d to $\frac{111}{3080} > \frac{25}{924}$ at $d = 9$, $\frac{25}{924} > \frac{4}{156}$, and the last expression decreases in d. Hence A is concave. Moreover $A(d) \geq a_d$ for every integer $d \geq 0$: for $d \leq 2$ because $a_d = 0$, for $3 \leq d \leq 5$ by (9.9), and for $d \geq 6$ with equality. Since the weights n_d/n sum to 1, Jensen's inequality gives

$$\sum_{d\geq 3} a_d n_d = \sum_{d\geq 0} a_d n_d \leq \sum_{d\geq 0} A(d)\, n_d \leq n\, A\Big(\sum_{d\geq 0} d\,\frac{n_d}{n}\Big) = n\, A(\bar{d}),$$

and Theorem 1.3 gives $\mathrm{pw}(G) \leq A(\bar{d})\, n + o(n)$. For $\bar{d} \geq 7$, let $q(x) := 1 - \frac{111}{28(x+1)}$; it is concave on $[0,\infty)$, and $a_d \leq q(d)$ for $d \geq 7$ (with equality for $d \leq 10$). On each interval $[d, d+1]$ with $d \geq 7$, A is the chord through (d, a_d) and $(d+1, a_{d+1})$, which lies below the chord of the concave function q and hence below q. So $A(\bar{d}) \leq q(\bar{d})$. The same argument with the concave function $1 - \frac{4}{x+1}$, which equals a_d for every $d \geq 11$, gives $A(\bar{d}) \leq 1 - \frac{4}{\bar{d}+1}$ for $\bar{d} \geq 11$.

For the last bound, note that $a_d \leq \frac{3d-4}{32}$ for every $d \geq 3$: with equality for $d \leq 6$, and for $d \geq 7$ because A is concave with slope $\frac{15}{224} < \frac{3}{32}$ on $[6,7]$, so that $a_d = A(d) \leq a_6 + \frac{15}{224}(d-6) \leq \frac{3d-4}{32}$.

So every graph satisfies $\mathrm{pw}(G) \le \sum_{d\ge3} \frac{3d-4}{32}\, n_d + o(n)$. Now let $\delta(G) \ge 2$, and write $n_{\ge3}$ for the number of vertices of degree at least 3. As $3d-4 = 2+3(d-2)$, $\sum_{d\ge2} n_d = n$ and $\sum_{d\ge2} d\, n_d = 2m$,

$$\sum_{d\ge3} a_d n_d \le \sum_{d\ge3} \frac{3d-4}{32}\, n_d = \frac{2n_{\ge3} + 3\sum_{d\ge2}(d-2)\, n_d}{32} = \frac{n_{\ge3} + 3(m-n)}{16} \le \frac{3m-2n}{16},$$

and Theorem 1.3 completes the proof. □

For $\bar d \le 6$ its main term is that of Corollary 1.4, since $A(\bar d)\, n = \frac{7}{96}\bar d\, n = \frac{7}{48}m$. For graphs with $\delta(G) \ge 2$ and $\bar d < 6$, the bound $\frac{3m-2n}{16}$ is smaller. The condition $\delta(G) \ge 2$ cannot be dropped from it, since adding isolated vertices makes $3m-2n$ negative without changing the pathwidth.

Edwards and McDermid prove $\mathrm{pw}(G) \le (g_{1/6}(\bar d) + o(1))n$ for connected graphs and for graphs with $\bar d \ge 5$, with $g_{1/6}$ interpolated linearly [EM15, Theorem 5.3]. For $\bar d \ge 6$, Corollary 9.15 is smaller, since both A and $g_{1/6}$ interpolate linearly between consecutive integers there and $a_d < g_{1/6}(d)$ for every $d \ge 3$ by the following lemma.

Lemma 9.16 (comparison with the previous coefficients)**.** *Let $g_{1/6}(2) = 0$, $g_{1/6}(3) = \frac16$ and $d\, g_{1/6}(d) = (d-2)\, g_{1/6}(d-1) + g_{1/6}(d-2) + 1$ for $d \ge 4$, the per-degree coefficients of Edwards and McDermid [EM15, §3.1.1 and Corollary 5.2]; for $3 \le d \le 6$ they are $\frac16, \frac13, \frac{13}{30}, \frac{23}{45}$, those of Fomin, Gaspers, Saurabh and Stepanov [FGSS09, Lemma 1]. Then the coefficients a_d of Theorem 1.3 satisfy $a_d < g_{1/6}(d)$ and $(d+1)(1-g_{1/6}(d)) < \frac72$ for every $d \ge 3$.*

Proof. Let $u_d := 1 - g_{1/6}(d)$ and $\omega_d := (d+1)u_d$. For $d \le 6$ the inequality $a_d < g_{1/6}(d)$ follows from the exact values: $\frac{5}{32} < \frac16$, $\frac14 < \frac13$, $\frac{11}{32} < \frac{13}{30}$ and $\frac{7}{16} < \frac{23}{45}$. The numbers u_d satisfy $d\, u_d = (d-2)\, u_{d-1} + u_{d-2}$ for $d \ge 5$, with $u_3 = \frac56$ and $u_4 = \frac23$. By Lemma 9.12(i) with $d_0 = 5$, $d\, u_d + u_{d-1} = 4u_4 + u_3 = \frac72$ for all $d \ge 4$, and since $\frac56 \cdot \frac72 = \frac{35}{12} < \omega_4 = \frac{10}{3} < \frac72$, Lemma 9.12(ii) with $d_1 = 4$ gives

$$\Big(1 - \frac{1}{d^2}\Big)\frac72 < \omega_d < \frac72 \qquad (d \ge 5),$$

which is the asymptotic form of Edwards and McDermid with an explicit error term. Hence $\omega_d < \frac72 < \frac{111}{28} \le (d+1)(1-a_d)$ for $d \ge 7$, where $(d+1)(1-a_d)$ is $\frac{111}{28}$ for $7 \le d \le 10$ and 4 for $d \ge 11$. So $1 - g_{1/6}(d) < 1 - a_d$, that is, $a_d < g_{1/6}(d)$, for all $d \ge 7$. Finally, $\omega_3 = \omega_4 = \frac{10}{3} < \frac72$, so $\omega_d < \frac72$ for every $d \ge 3$. □

10 Computing the decompositions

This section turns the width bounds of Theorems 1.3, 7.1, 7.2 and 8.1 and Corollary 1.4, and with Appendix A.6 also those of Theorems 8.2 and 9.2 and Corollary 9.5, into algorithms that compute the decompositions. The proofs of these bounds are probabilistic. The frontiers of a random Gaussian labeling have the right expected size at every threshold and concentrate around it, and Lemma 4.1, or the median-switched layout of Lemma 8.4, turns a good labeling into a path decomposition. We state Theorem 10.1, the precise form of Theorem 1.5, which gives a Las Vegas and a deterministic algorithm for each of these bounds, and outline its proof, which is in Appendix A.5. Part II uses this theorem, mostly its deterministic part, whenever an algorithm needs a decomposition of the width given by Part I.

10.1 Statement and outline

Model. The input graph G has vertices $1, \dots, n$ and is given by adjacency lists. Running times count bit operations on integers and rationals in binary; in the bootstrap and general cases below we also count operations on integers of $O(\log n)$ bits, which serve as vertex names and indices. The Las Vegas algorithms have access to independent uniform random bits. A path decomposition $(W_1, \dots, W_s)$ is output in interval form (Section 4.1), that is, as the number s and, for each vertex v, the indices $\mathrm{first}(v) \le \mathrm{last}(v)$ of the first and the last bag containing v. This takes $O(n \log n)$ bits when $s = O(n)$, as for all our decompositions. The bags themselves can have total size of order n^2, so the running times below refer to the interval form. The terms $o(1)$ in the exponents of running times are functions of n alone.

Cases. Each case of the following theorem consists of a class $\mathcal{C}$ of graphs, given by a condition on G, and a bound $\beta(G)$ for $G \in \mathcal{C}$:

- *cubic case*: $\Delta(G) \le 3$ and $\beta(G) = c_3 n$;
- *padded case*: $\Delta(G) \le 4$ and $\beta(G) = F_4 n$;
- *switched case*: $\Delta(G) \le 4$ and $\beta(G) = c_3 n_3 + 0.2627\, n_4$;
- *mixed cases*: $\Delta(G) \le 5$ and $\beta(G) = a_3 n_3 + a_4 n_4 + a_5 n_5$ with $(a_3, a_4, a_5) \in \{\mathsf{P}, \mathsf{R}, \mathsf{S}\}$;
- *bootstrap cases*: all graphs, and $\beta(G) = \sum_{d \ge 3} b_d n_d$, where $(b_d)_{d \ge 3}$ is the bootstrap of one of the six sequences $\mathsf{C}, \mathsf{M}, \mathsf{F}, \mathsf{P}, \mathsf{R}, \mathsf{S}$ of Theorem 9.2;
- *general case*: all graphs, and $\beta(G) = \sum_{d \ge 3} a_d n_d$ with the coefficients a_d of Theorem 1.3.

Theorem 10.1 (precise form of Theorem 1.5)**.** *In each case listed above there are a nondecreasing function $\rho\colon \mathbb{N} \to [0, \infty)$ with $\rho(n) = o(n)$ and two algorithms that, given $G \in \mathcal{C}$, output a path decomposition of G in interval form, of width at most $\beta(G) + \rho(n)$, and in the general case also at most $\frac{7}{48}m + \rho(2m)$:*

(a) *a Las Vegas algorithm, in expected $n^{1+o(1)}$ bit operations;*

(b) *a deterministic algorithm, in $n^{2+o(1)}$ bit operations.*

In the bootstrap and general cases, each algorithm uses in addition $O(n(n+m))$ operations on integers of $O(\log n)$ bits.

Here ρ is the remainder of the width bound. The bounds β are those of Theorem 8.1, Theorem 7.1(a), Theorem 8.2, Theorem 7.2, Lemma 9.3 and Theorem 1.3 without their remainders, and the second bound of the general case is that of Corollary 1.4. A graph with $\Delta(G) \le 4$ has $n_5 = 0$, so the mixed cases with P and S also give the bounds of Theorem 7.1(b) and (c). The bootstrap cases give the rows of Corollary 9.5 and the sequences of Theorem 9.2 on all graphs (Remark A.35); the general case is the bootstrap case of S in the rounded form of Theorem 1.3. Theorem 1.5 follows from Theorem 10.1.

Outline. The proof of each case follows the same pattern. The algorithm computes discrete labels $\tilde{Y}_v$, sorts the vertices by $(\tilde{Y}_v, v)$, and builds the layout of the case from the sorted order. The width of that layout is bounded by a finite family of *counts*, each the number of vertices satisfying a condition on the labels in their closed neighborhood (Lemma A.31 and Appendix A.5.8). Four ingredients, shared by all cases, bound the counts.

(1) *Exact local labels* (Appendices A.5.1 and A.5.2). We round the coefficients of Definition 5.2 to rationals, and we replace the Gaussian variables by a finite rational distribution coupled with the standard normal distribution (Lemma A.23).

(2) *Perturbation* (Appendix A.5.3). In this coupling every per-vertex probability used in Part I changes by $o(1)$, uniformly in the threshold (Lemma A.24 and Corollary A.25).

(3) *Uniform ideal means* (Appendix A.5.4). The grid of thresholds refines as n grows, so the bounds on expected frontier sizes are needed uniformly in the threshold (Lemma A.26). In the cubic, padded and mixed cases the bounds of Part I already have this form, with an error term $O(1/h)$; in the switched case, Lemma A.37 makes the error term at threshold zero explicit, as $O(\varepsilon_h^{1/3})$.

(4) *Derandomization* (Appendix A.5.5). The counts are close to their means with high probability. The method of conditional expectations (Erdős and Selfridge [ES73]; Raghavan [Rag88]), applied to the sum of the squared deviations of the counts, fixes the random variables one at a time so that every count is close to its mean (Lemma A.27). The locality radius grows only like $\frac{1}{8}\log_2\log_2 n$, so every conditional expectation is a sum of local terms, each computable exactly in $n^{o(1)}$ bit operations.

Derandomizing a local Gaussian rounding in this way is standard in outline; the work lies in the exact finite distribution and in the uniformity of the means.

Roadmap. The proof is in Appendix A.5. Appendices A.5.1 to A.5.5 develop the four ingredients in this order, and Appendix A.5.6 makes the construction of Lemma 4.5 algorithmic. Appendices A.5.7 to A.5.9 assemble the algorithms for the padded and mixed cases, the cubic case, and the bootstrap and general cases, Appendix A.6 treats the switched case and the base layouts of the bootstrap cases C and M, and Appendix A.5.10 collects the running times. The rest of this section computes orderings of small cutwidth for maximum degree 3, 4 and 5 and discusses the limitations.

10.2 Linear orderings of small cutwidth

This subsection computes, for graphs of maximum degree 3, linear orderings of cutwidth at most $c_3 n + o(n)$, and bisections with at most that many edges between the two sides (Corollary 10.2). Corollary 1.2 shows that such orderings and bisections exist, by bounding the cutwidth through a chain of known relations. Two links of that chain have constructive proofs, which turn an ordering of vertex separation w, such as the one read off a path decomposition of Theorem 10.1, into an ordering of cutwidth at most $w + 2$ (Lemma A.36, proved in Appendix A.5.11). For maximum degree 4 and 5, Theorem 10.3 derandomizes the cut sweep of Theorem 3.6 directly, with the machinery of Appendices A.5.1 to A.5.3 and A.5.5. The two routes differ because they start from different bounds. For maximum degree 3 the cut sweep gives only $\gamma_3 m \le c_* n$, while the route through vertex separation gives $c_3 n < c_* n$. For maximum degree 4 the same chain gives only $\mathrm{cw} \le 2\,\mathrm{pw} + O(1)$ [MS89, Theorem 3.2][EST94, Theorem 2.1], that is, $2F_4 n + o(n) > \frac{n}{3}$.

Corollary 10.2 (computing orderings and bisections)**.** *For graphs G with $\Delta(G) \le 3$, an ordering of $V(G)$ of cutwidth at most $c_3 n + o(n)$, and a bisection of G with at most $c_3 n + o(n)$ edges between its two sides, can be computed in deterministic time $n^{2+o(1)}$.*

The proof is in Appendix A.5.11.

Theorem 10.3 (deterministic cut sweep)**.** *Let $p \in \{4, 5\}$. For graphs G with $\Delta(G) \le p$ and m edges, an ordering of $V(G)$ of cutwidth at most $\gamma_p m + o(n)$, and a bisection of G with at most that many edges between its two sides, can be computed in deterministic time $n^{2+o(1)}$. For $p = 4$ the bound is $\frac{m}{6} + o(n) \le \frac{n}{3} + o(n)$.*

The proof is also in Appendix A.5.11.

10.3 Limitations

The guarantees of Theorem 10.1 are asymptotic: our analysis assumes $n \ge 2^{2^{40}}$, and some error terms are below 1 only when $\log_2 \log_2 n$ exceeds 10^4 or even 10^5, so the bounds say nothing for practical n. The terms $o(1)$ in the exponents also hide enumeration factors $e^{O(t^{1/4} \log t)}$ with $t = O(\log n)$, and practical efficiency is not claimed.

Part II
Algorithmic consequences

Part II derives faster exact algorithms from the width bounds of Part I. Before the proofs, which begin in Section 11, this opening gathers the new running-time bounds and says which of them the two recent preprints that claim to refute the ETH [Ope26a, Cla26] (Section 1.2) would supersede.

The tables. Table 2 lists the running-time bounds of this part in exponential space that, as far as we know, the preprints do not affect, and Table 5 lists those in exponential space that the preprints would supersede. Both tables follow the conventions stated after Theorem 1.6. "Previous" is the fastest bound that we are aware of.

Bounds in polynomial space. Neither table lists bounds in polynomial space; we state them here. Separate, Measure and Conquer with the separator constant c_3 (Theorem C.3) solves Max $(r, 2)$-CSP in polynomial space in time $r^{0.17587m+o(m)}$, in time $r^{0.16984m+o(m)}$ if $\Delta \leq 4$ and in time $r^{0.17935n+o(n)}$ on cubic instances. The previous bounds are $r^{9m/50}$, $r^{7m/40}$ and $r^{n/5}$ up to subexponential factors [GS17, Theorems 3.1–3.3]. As far as we know, the preprints do not affect these three bounds for Max $(r, 2)$-CSP. The same method counts the dominating sets of each size of a graph with $\Delta \leq 3$ in polynomial space and time $O(1.2178^n)$, against $O(1.2458^n)$ [GS17, Theorem 5.1].

By Theorem C.6, proved in Appendix C, the independent sets of a graph can be counted in polynomial space in time $O(1.1304^n)$ if $\Delta \leq 3$, $O(1.2037^n)$ if $\Delta \leq 4$ and $O(1.2341^n)$ in general, and its chromatic number can be computed in polynomial space in time $O(2.2341^n)$. The previous bounds, of Gaspers and Lee [GL23, Lemmas 5 and 6, Theorem 1], are $O(1.1389^n)$, $O(1.2070^n)$, $O(1.2356^n)$ and $O(2.2356^n)$. The previous analysis of the bound $O(1.1389^n)$ has open points (Appendix C); our bound $O(1.1304^n)$ includes the necessary corrections.

The reach of the preprints. The preprints reach other problems through the subexponential reduction families of Impagliazzo, Paturi and Zane [IPZ01], which make every problem in size-constrained SNP solvable in subexponential time if #3-SAT is, and through direct counting algorithms [Cla26, Corollary 7], [Ope26a, Corollary 1.3]. If correct, they supersede seventeen of our bounds: the twelve of Table 5 and the five in polynomial space above. For the chromatic number this uses inclusion–exclusion [BHK09]; for Max 2-SAT, Max k-Vertex Cover, Upper Domination, Edge Dominating Set and Odd Cycle Transversal on graphs of bounded degree, an encoding in size-constrained SNP or in k-SAT with $O(n)$ variables; for Independent Dominating Set, a result of Ganian, de Haan, Kanj and Szeider [GdHKS16, Corollary 6]. As far as we know, the problems of Table 2 lie outside this reach: they have real weights, a global connectivity or acyclicity constraint (whether Feedback Vertex Set is SERF-reducible to 3-SAT is open [GdHKS16]), or running times n^{ck} in a pattern size k, and we know of no SERF reduction from any of them to 3-SAT or #3-SAT.

Conventions for Part II. We write $O^*(f)$ for $f \cdot \mathrm{poly}(N)$, where N is the length of the input. All remainders $o(n)$, $o(m)$ and $o(k)$ are functions of the displayed parameter alone. The results of Part II are worst-case asymptotic statements, as are the records they improve. Numerical bases are rounded *up* in the last printed digit, so that $c^{n+o(n)} = O(b^n)$ holds for the printed b. Running times of the form $r^{\alpha n+o(n)}$ are stated with the exact exponent; their numerical values come from

Table 5: The running-time bounds of this part in exponential space that would be superseded if the preprints [Ope26a, Cla26] that claim to refute ETH are correct. Letters refer to the notes; the letters a–h have the same meaning as in Table 2, and the notes i and j are local to this table.

Problem	Previous	New	Where
Max 2-SAT, counting dominating sets and independent sets			
Max 2-SAT	$2^{m/6.321}$ [GS12]	$2^{m/6.857}$	Theorem 11.3
#Dominating Set, all graphs	1.5002^n [NvRvD14][a]	1.4815^n	Theorem 13.1
#Dominating Set (every size), $\Delta \leq 3$	$3^{n/6} < 1.2010^n$ [FH06]	$3^{c_3 n} < 1.1661^n$	Theorem 11.4
#Independent Sets, $\Delta \leq 3$	$2^{n/6} < 1.1225^n$ [FH06]	$2^{c_3 n} < 1.1018^n$	Theorem 11.4
#Independent Sets, all graphs	1.2330^n [GL23]	1.2287^n	Theorem C.6
Local problems on graphs of bounded degree (deterministic)			
Max k-Vertex Cover, $\Delta \leq 3$	1.123^n [CP12]	$2^{c_3 n} < 1.1018^n$	Theorem 11.4
Min. Independent Dominating Set, $\Delta \leq 3$	1.3351^n [BCEP13][i]	$3^{c_3 n} < 1.1661^n$	Theorem 11.4
Min. Independent Dominating Set, $\Delta \leq 4$	1.3351^n [BCEP13][i]	$3^{F_4 n} < 1.3100^n$	Theorem 11.4
Upper Domination, $\Delta \leq 3$	1.3481^n [BBC$^+$18]	$6^{c_3 n} < 1.2848^n$	Theorem 11.4
Edge Dominating Set[j], $\Delta \leq 3$	$3^{n/6} < 1.2010^n$ [FH06]	$3^{c_3 n} < 1.1661^n$	Theorem 11.4
Odd Cycle Transversal, $\Delta \leq 4$	$3^{n/3} < 1.4423^n$ [FGSS09]	$2^{n/3} < 1.2600^n$	Theorem 11.4
Odd Cycle Transversal, $\Delta \leq 5$	$3^{13n/30} < 1.6098^n$ [FGSS09]	$2^{0.5121n} < 1.4262^n$	Theorem 11.4

[a] A gap in the analysis of this algorithm allowed us only to reconstruct a running time of $O(1.5004^n)$ (Appendix D.5).
[i] A bound for all graphs. [j] Also Minimum Maximal Matching.

the accompanying code. By the *previous record* for a problem we mean the fastest algorithm we know of.

11 Dynamic programming over the decompositions

This section obtains exact algorithms by running dynamic programs over the path decompositions of Theorem 1.5. The running time of such a program is exponential in the width of the decomposition, so a smaller width bound gives a faster algorithm directly.

11.1 Problems, dynamic programs and the decompositions used

Max $(r, 2)$-CSP. An instance of Max $(r, 2)$-CSP consists of a simple graph G, the *constraint graph*, whose vertices are the *variables*; a score function $s_v \colon [r] \to \mathbb{R}$ for every variable v; and a score function $s_{uv} \colon [r]^2 \to \mathbb{R}$ for every edge $uv \in E(G)$. The task is to find an assignment $\varphi \colon V(G) \to [r]$ maximizing

$$\sum_{v \in V(G)} s_v(\varphi(v)) + \sum_{uv \in E(G)} s_{uv}(\varphi(u), \varphi(v)).$$

Following Scott and Sorkin we write r for the domain size. We write n for the number of variables and m for the number of constraints. Several constraints on the same pair of variables are added into one, so the constraint graph is simple and $|E(G)| \leq m$. Max 2-SAT ($r = 2$, with m the number of 2-clauses; unit clauses become vertex scores) and Max k-Colorable Subgraph ($r = k$) are special cases.

Semiring, ring and polynomial versions. Let $(\mathbb{S}, \oplus, \otimes)$ be a commutative semiring. The *semiring version* of the problem has score functions with values in $\mathbb{S}$ and asks for

$$Z = \bigoplus_{\varphi\colon V(G)\to[r]} \bigotimes_{v\in V(G)} s_v(\varphi(v)) \otimes \bigotimes_{uv\in E(G)} s_{uv}(\varphi(u),\varphi(v)). \tag{11.1}$$

More generally, each variable v may have its own finite domain D_v in place of $[r]$; the score functions are then defined on D_v and on $D_u \times D_v$, and (11.1) runs over the assignments with $\varphi(v) \in D_v$ for every v. Max $(r,2)$-CSP is the case of the semiring $(\mathbb{R}\cup\{-\infty\}, \max, +)$; counting the solutions of a binary constraint satisfaction problem is the case $(\mathbb{N}, +, \times)$ with scores in $\{0,1\}$. Scott and Sorkin [SS09] call the version over a commutative ring the *ring CSP* (RCSP), and the version over a polynomial ring the *polynomial CSP* (PCSP). These cover, among others, counting cuts of every size and the Ising partition function [SS09]. When we say that such a version is solved in time T, we mean T semiring operations; for the applications just named, each ring element has a representation of polynomial size and each operation takes polynomial time.

The following dynamic programs are standard; we state them for path decompositions in the form needed below. Part (a) is the dynamic program behind [BKM21, Lemmas 2 and 4] and [SS09, §9]; the 3^w dynamic program for dominating sets is due, for tree decompositions, to van Rooij, Bodlaender and Rossmanith [vRBR09], and part (c) is standard and folklore. In this section, vertex weights are real numbers.

Lemma 11.1 (dynamic programs over a path decomposition)**.** *Let G be a graph with n vertices, given with a path decomposition of width w in interval form.*

(a) *For a semiring version with constraint graph G and domains of size at most r, the value Z of* (11.1) *can be computed with $O(n\,(w+2)\,r^{w+1})$ semiring operations, keeping $O(r^{w+1})$ semiring elements at a time; for Max $(r,2)$-CSP with constraint graph G, an optimal assignment can be found in time and space $r^{w+1}\,\mathrm{poly}(n)$.*

(b) *One can count the dominating sets of each size, find a dominating set of minimum weight and count the dominating sets of minimum weight, in time and space $3^{w+1}\,\mathrm{poly}(n)$.*

(c) *One can count the independent sets, find a maximum weight independent set and count the maximum weight independent sets in time and space $2^{w+1}\,\mathrm{poly}(n)$.*

Each part keeps a table indexed by the partial solutions on the vertices of the current bag along a nice sequence of introduce and forget steps.

The decompositions. Most results of this section run one of the dynamic programs of Lemma 11.1, the others algorithms from the literature, over a path decomposition. The decompositions are in interval form and are computed in polynomial time by the deterministic algorithm of Theorem 10.1(b); when we cite Theorem 1.5 below, we mean this algorithm. For a graph G the width is at most $W(G)+\rho(n)$, with $\rho(n)=o(n)$, where W is one of the following bounds (Theorems 1.3, 7.1, 7.2 and 8.1 and Corollary 1.4):

- $c_3 n$ if $\Delta(G)\le 3$, with $c_3 = \frac{3}{4\pi}\arccos\frac{5}{6} < 0.1398222$;
- $F_4 n$, $F_3 n_3 + 0.2575\,n_4$ or $\frac{5}{32}n_3 + \frac{1}{4}n_4$ if $\Delta(G)\le 4$, with $F_4 < 0.24577198$ (Proposition 6.9); we write $\beta_4(G) := \min\{F_4 n,\ F_3 n_3 + 0.2575\,n_4,\ \frac{5}{32}n_3 + \frac{1}{4}n_4\}$;
- $a_3 n_3 + a_4 n_4 + a_5 n_5$ for each of the triples $\mathsf{P}, \mathsf{R}, \mathsf{S}$ if $\Delta(G)\le 5$;

- $\beta(G) := \sum_{d\geq 3} a_d n_d$ for every G, with the coefficients a_d of Theorem 1.3;
- $\frac{7}{48}m$ for every G.

Running Theorem 1.5 once for each of several bounds and keeping the narrowest decomposition realizes their minimum, for instance $\beta_4(G)$. Since the width is at most $W(G)$ plus the remainder, a dynamic program with $c^{w+1}\operatorname{poly}(n)$ steps runs in time $c^{W(G)+o(n)}\operatorname{poly}(n)$ (with $o(m)$ for the bound $\frac{7}{48}m$), for every constant c and also for $c = r$ growing with the input.

Two parts of Theorem 11.3 evaluate a bound on the final graph R of a maximal sequence of reduction operations (Definition 4.2) applied to G; then R is empty or has minimum degree at least 3 by (RP3). Reduction operations increase no degree by (RP1), so they do not increase the number of edges either. Lemma A.28, with no vertex deletions, turns a path decomposition of R of width w in interval form into one of G of width at most $w + 3\log_2 n + 1$, in polynomial time.

11.2 Pathwidth in terms of the vertices of degree three

Theorem 8.1 bounds the pathwidth of a graph of maximum degree 3 in terms of the number of all its vertices. Some algorithms, such as the algorithm for counting dominating sets of Section 13, need a bound in the number of vertices of degree 3 alone; the polynomial-space analyses of Section 12 use the separator version of such a bound in Lemma C.2. The next lemma converts Theorem 8.1 into such a bound: the vertices of degree at most 2 add only a term of order $\log n_3$ to the width.

Lemma 11.2 (pathwidth by the number of degree-3 vertices)**.** *There is a nondecreasing function $\rho\colon \mathbb{N} \to \mathbb{R}$ with $\rho(k) = o(k)$ such that every graph H with $\Delta(H) \leq 3$ satisfies*

$$\operatorname{pw}(H) \leq c_3\, n_3(H) + \rho\big(n_3(H)\big).$$

A path decomposition of this width can be computed in polynomial time.

The proof, in Appendix A.7.1, removes the trees outside the 2-core, suppresses the vertices of degree 2 of the core, and applies Theorem 8.1 to the resulting cubic multigraph; putting back the suppressed paths costs 2 in the width, and the trees $O(\log n_3)$.

11.3 Max $(r, 2)$-CSP and its relatives

The first theorem is the most direct consequence of the width bounds: Max $(r, 2)$-CSP, and every semiring version of it, is solved by the dynamic program of Lemma 11.1(a) over the narrowest available decomposition of the constraint graph. In terms of the number m of constraints, this lowers the exponent of the previous record $r^{(13/75)m+o(m)}$ [SS07, Cor. 21] to $\frac{7}{48}m + o(m)$. Part (1) of the theorem is the most general statement; the other parts are stated in the parameters of the previous records. Part (4) uses the average degree $\bar{d} := 2|E(G)|/n$ of the constraint graph G and the concave function $A\colon [0,\infty) \to \mathbb{R}$ defined before Corollary 9.15: $A(x) = \frac{7}{96}x$ for $0 \leq x \leq 6$, and A interpolates the coefficients a_d linearly between consecutive integers $d \geq 6$.

Theorem 11.3 (Max $(r, 2)$-CSP)**.** *Let G be the constraint graph of an instance of Max $(r, 2)$-CSP, or of a semiring version, with n variables and m constraints. The instance can be solved in time and space*

(1) $O^*\big(r^{\beta(R)+o(n)}\big)$ *for the final graph R of some maximal sequence of reduction operations applied to G, and hence $O^*\big(r^{\beta(G)+o(n)}\big)$;*

(2) $O^*\big(r^{(7/48)m+o(m)}\big)$*;*

(3) $O^*\big(r^{c_3 n_3 + o(n)}\big)$ *if* $\Delta(G) \le 3$, *and* $O^*\big(r^{\beta_4(G)+o(n)}\big)$ *and* $O^*\big(r^{3F_4 m/(2+16F_4)+o(m)}\big)$ *if* $\Delta(G) \le 4$;

(4) $O^*\big(r^{n\,A(\bar d)+o(n)}\big)$, *and hence*

$$O^*\Big(r^{(1-\frac{111}{28(\bar d+1)})n+o(n)}\Big) \text{ if } \bar d \ge 7, \qquad O^*\Big(r^{(1-\frac{4}{\bar d+1})n+o(n)}\Big) \text{ if } \bar d \ge 11.$$

Numerically, $\frac{48}{7} = 6.857\ldots$, so part (2) reads $r^{m/6.857+o(m)}$, and $\frac{3F_4}{2+16F_4} < 0.12429$. For cubic constraint graphs ($m = \frac{3}{2}n$), part (3) reads $r^{(2c_3/3)m+o(m)}$ with $\frac{2c_3}{3} < 0.09322$.

Proof. In each part we compute a path decomposition of the constraint graph and run Lemma 11.1(a) with $D_v = [r]$; its running time and space are $r^{w+1}\operatorname{poly}(n)$ for a decomposition of width w, so it suffices to bound w. Isolated vertices of G are handled separately in time $O(rn)$ (each contributes a factor $\bigoplus_a s_v(a)$), so we may assume that G has none; then $n \le 2|E(G)| \le 2m$.

(1) Apply reduction operations to G until none applies, and let R be the final graph. By Theorem 1.3 and Theorem 1.5 applied to R, we compute a decomposition of R of width at most $\beta(R)+\rho(|V(R)|)$, and by Lemma A.28 with no deletions one of G of width at most $\beta(R)+\rho(|V(R)|)+3\log_2 n + 1 = \beta(R) + o(n)$, since ρ is nondecreasing (Theorem 10.1). Reduction operations only remove vertices and do not increase degrees by (RP1), and $0 \le a_d \le a_{d+1}$ for all $d \ge 0$ (with $a_0 = a_1 = a_2 = 0$); hence $\beta(R) \le \beta(G)$.

(2) By Theorem 10.1 (general case), $w \le \frac{7}{48}|E(G)| + \rho(2|E(G)|) \le \frac{7}{48}m + o(m)$.

(3) If $\Delta(G) \le 3$, Lemma 11.2 gives $w \le c_3 n_3(G) + o(n)$. If $\Delta(G) \le 4$, Theorem 7.1 and Theorem 1.5 give $w \le \beta_4(G)+o(n)$, by running the algorithm for each of the three forms and keeping the narrowest decomposition. For the bound in m, let R be the final graph of a maximal sequence of reduction operations applied to G; it is empty or has $\delta(R) \ge 3$, and $\Delta(R) \le 4$. If R is empty, $w = O(\log n) = o(m)$ by Lemma A.28 with no deletions. Otherwise put $n' := |V(R)|$, $m' := |E(R)|$ and $x := n_4(R)/n' \in [0,1]$. Then $n_3(R) = (1-x)n'$ and $2m' = 3n_3(R) + 4n_4(R) = (3+x)n'$, so

$$\frac{F_4 n'}{m'} = \frac{2F_4}{3+x}, \qquad \frac{\frac{5}{32}n_3(R) + \frac{1}{4}n_4(R)}{m'} = \frac{5+3x}{16(3+x)}.$$

The first ratio decreases in x, and the second increases, since $\frac{d}{dx}\frac{5+3x}{3+x} = \frac{4}{(3+x)^2} > 0$. Hence the minimum of the two is at most their common value at the crossing point $x^* = \frac{32F_4-5}{3} \in (0,1)$, which is $\frac{2F_4}{3+x^*} = \frac{3F_4}{2+16F_4}$. This expression increases in F_4, and $F_4 < 0.24577198$ gives $\frac{3F_4}{2+16F_4} < 0.12429$ (at the crossing, $x^* \approx 0.9549$). So Theorem 7.1(a),(c) and Theorem 1.5 applied to R give a decomposition of R of width at most $\frac{3F_4}{2+16F_4}m' + \rho(n')$, where $n' \le \frac{2}{3}m'$ and $m' \le |E(G)| \le m$, because reduction operations do not increase the number of edges. By Lemma A.28 with no deletions, $w \le \frac{3F_4}{2+16F_4}m + o(m) + 3\log_2(2m) + 1$, which is $\frac{3F_4}{2+16F_4}m + o(m)$.

(4) Part (1) gives the running time $O^*(r^{\beta(G)+o(n)})$ for the given instance. By Corollary 9.15, applied to the constraint graph G with all its n vertices (isolated ones included), $\beta(G) = \sum_{d\ge 3} a_d n_d \le n\,A(\bar d)$, where $A(\bar d) \le 1 - \frac{111}{28(\bar d+1)}$ if $\bar d \ge 7$ and $A(\bar d) \le 1 - \frac{4}{\bar d+1}$ if $\bar d \ge 11$. □

This improves the previous records in sparse instances. Other methods solve dense instances faster. For integer scores of polynomial size, Williams's algorithm [Wil05] runs in time $O^*(r^{\omega n/3})$, where ω is the matrix multiplication exponent (domain r: [SS09, Thm. 18]). With $\omega \le \frac{9}{4}$ [Ope26b], which Pass proves with integer arithmetic [Pas26, Theorem 1.1], as Williams's algorithm needs, this algorithm is faster than ours when $\bar d > 15$. For very large $\bar d$, an algorithm of Golovnev and Kutzkov has a smaller exponent [EM15, p. 942]: for Boolean domains it runs in polynomial space and time $O^*\big(2^{(1-\alpha\ln\bar d/\bar d)n}\big)$ for every constant $\alpha \in (0,1)$ and every sufficiently large $\bar d = o(n)$ [GK14, Thm. 4].

Their threshold on $\bar{d}$ is not explicit, so we do not determine that crossover. Hence part (4) is a record only for moderate $\bar{d}$, and for integer scores of polynomial size only for $\bar{d} \leq 15$.

11.4 Domination, independent sets and vertex covers in graphs of bounded degree

This subsection runs dynamic programs for domination, independent sets and five further local problems over the decompositions of Lemma 11.2 (for $\Delta \leq 3$) and Theorem 7.1 (for $\Delta \leq 4$), and an algorithm for Odd Cycle Transversal over the orderings of Theorem 3.6. We use the dynamic programs of Lemma 11.1(b),(c) and five previous algorithms. Most previous bounds come from the same dynamic programs over decompositions of width $\frac{n}{6}$ for $\Delta \leq 3$ [FH06] and $\frac{n}{3}$ for $\Delta \leq 4$ [FGSS09], which the new widths $c_3 n_3 + o(n) \leq c_3 n + o(n)$ and $\beta_4(G) + o(n) \leq F_4 n + o(n)$ replace. For instance, the dominating sets of each size of a graph with $\Delta \leq 3$ are counted in time $O(1.1661^n)$ instead of $3^{n/6} < 1.2010^n$.

Theorem 11.4 (bounded-degree graph problems)**.** *Let G be a graph with n vertices.*

(a) *If $\Delta(G) \leq 3$, then in time and space $O^*(3^{c_3 n_3 + o(n)}) = O(1.1661^n)$ one can count the dominating sets of G of each size, find a minimum weight dominating set and count the dominating sets of minimum weight.*

(b) *A maximum weight independent set of G can be found, and the maximum weight independent sets counted, in time and space $O^*(2^{c_3 n_3 + o(n)}) = O(1.1018^n)$ if $\Delta(G) \leq 3$, and $O^*(2^{\beta_4(G) + o(n)}) = O(1.1858^n)$ if $\Delta(G) \leq 4$.*

(c) *If $\Delta(G) \leq 3$, then, in time and space $O^*(2^{c_3 n_3 + o(n)}) = O(1.1018^n)$, one can find k vertices that cover the most edges (Max k-Vertex Cover); in time and space $O^*(3^{c_3 n_3 + o(n)}) = O(1.1661^n)$, a minimum independent dominating set; and in time and space $O^*(6^{c_3 n_3 + o(n)}) = O(1.2848^n)$, a minimal dominating set of maximum size (Upper Domination). If $\Delta(G) \leq 4$, a minimum independent dominating set can be found in time and space $O^*(3^{\beta_4(G) + o(n)}) = O(1.3100^n)$.*

(d) *If $\Delta(G) \leq 3$, a minimum edge dominating set and a minimum maximal matching of G can be found in time and space $O^*(3^{c_3 n_3 + o(n)}) = O(1.1661^n)$.*

(e) *A minimum odd cycle transversal of G can be found by a deterministic algorithm in time $O^*(2^{m/6 + o(n)}) = O(1.2600^n)$ if $\Delta(G) \leq 4$ and in time $O^*(2^{\gamma_5 m + o(n)}) = O(1.4262^n)$ if $\Delta(G) \leq 5$, where $m = |E(G)|$ and $\gamma_5 < 0.20484$ is the constant of Theorem 3.6.*

Proof. If $\Delta(G) \leq 3$, Lemma 11.2 gives a path decomposition of width at most $c_3 n_3 + o(n)$ in polynomial time; if $\Delta(G) \leq 4$, Theorem 7.1 and Theorem 1.5 give one of width at most $\beta_4(G) + o(n)$. Run Lemma 11.1(b) for (a) and Lemma 11.1(c) for (b). For (c) we run, on these decompositions, the dynamic program for Max k-Vertex Cover in time $O^*(2^{\mathrm{pw}})$ and exponential space that Della Croce and Paschos combine with the bound of Fomin and Høie [CP12, Corollary 2], the dynamic program for Independent Dominating Set in time $O(n^3 3^{\mathrm{tw}})$ of van Rooij [vR11, Proposition 11.11], and that for Upper Domination in time $O(6^{\mathrm{pw}}\,\mathrm{pw})$ of Dublois, Lampis and Paschos [DLP22, Theorem 10]. For (d) we run the dynamic program of van Rooij and Bodlaender, which finds a minimum weight maximal matching in time $O^*(3^{\mathrm{pw}})$ on a path decomposition of width pw [vRB12, Lemma 3]. A minimum maximal matching is an edge dominating set, and a minimum edge dominating set can be turned in polynomial time into a maximal matching of the same size [YG80]. For (e), Theorem 10.3 computes, in deterministic polynomial time, an ordering of cutwidth at most $\frac{m}{6} + o(n) \leq \frac{n}{3} + o(n)$ or $\gamma_5 m + o(n) \leq \frac{5}{2}\gamma_5 n + o(n)$, and Bojikian, Chekan, Hegerfeld and Kratsch find a minimum odd

cycle transversal deterministically in time $2^{\mathrm{ctw}} n^{O(1)}$ for a given ordering of cutwidth ctw [BCHK22, Theorem 36], with $2^{1/3} < 1.25993$ and $2^{5\gamma_5/2} < 1.42611$, where $\gamma_5 = \frac{1}{\pi}\arccos\frac{4}{5}$. The numerical bases follow from $n_3 \le n$, $F_4 n$ being one of the three forms, $3^{c_3} < 1.16604$, $2^{c_3} < 1.10177$, $6^{c_3} < 1.28471$, $2^{F_4} < 1.18573$ and $3^{F_4} < 1.30998$. □

11.5 Counting subgraphs

Curticapean, Dell and Marx reduce counting the subgraphs of a host graph that are isomorphic to a pattern to counting homomorphisms from the graphs of the *spasm* of the pattern (defined below), and the treewidth of these graphs governs the running time. For a pattern with k edges, their bound $k^{O(k)} n^{0.174k+o(k)}$ comes from the Scott–Sorkin bound $\frac{13}{75}k + o(k)$ on the treewidth of graphs with k edges; Remark 9.14 lowers the exponent to $\frac{7}{48}k + o(k)$ (Theorem 11.5).

For a graph H, the spasm of H is the set of graphs H/π obtained from H by choosing a partition π of $V(H)$ into independent sets and identifying the vertices of each part, keeping one copy of parallel edges [CDM17]. Since every part is independent, H/π is simple without loops, and every edge of H/π is the image of an edge of H, so $|E(H/\pi)| \le |E(H)|$.

Theorem 11.5 (counting subgraphs)**.** *There is a function $\rho(k) = o(k)$ such that, given a pattern graph H with k edges and a host graph G with n vertices, the number of subgraphs of G isomorphic to H can be computed in time $k^{O(k)}\, n^{\frac{7}{48}k+\rho(k)}$.*

In particular this holds for counting paths, cycles and matchings with k edges. Numerically, $\frac{7}{48} < 0.14584$.

Proof. By [CDM17, Thm. 1.1], the number of subgraphs of G isomorphic to H can be computed in time $k^{O(k)} n^{t+1}$, where t is the largest treewidth of a graph in the spasm of H. Every graph F in the spasm is simple and has at most k edges. By Remark 9.14, $\mathrm{tw}(F) \le \mathrm{pw}(F) \le \frac{7}{48}|E(F)| + \rho(|E(F)|+1)$, where ρ is the nondecreasing remainder of Remark 9.14; so $t \le \frac{7}{48}k + \rho(k+1)$, and the theorem holds with the function $k \mapsto \rho(k+1)+1$ in place of ρ. (The algorithm needs tree decompositions of the graphs of the spasm; they can be computed by brute force in time depending on k only, or by Theorem 1.5.) □

11.6 Bucketing: permutation patterns and k-opt

This subsection lowers two exponents of the form $\frac{k}{4}$ to $\frac{253}{1215}k + o(k)$ (note that $\frac{253}{1215} < 0.2083$): for Permutation Pattern Matching with a pattern of length k, from the bound $n^{k/4}$ of Berendsohn, Kozma and Marx [BKM21], and for k-opt Optimization for TSP, from the bound $n^{(1/4+\varepsilon_k)k}$, with $\varepsilon_k \to 0$, of Cygan, Kowalik and Socała [CKS19]. On graphs of bounded degree, Remark 9.14 also lowers the exponent $(\frac{23}{135} + \varepsilon_k)k$ of k-opt Optimization [BIJK19b] to $(\frac{7}{48} + \varepsilon_k)k$ (Theorem 11.8). We first define the two problems.

In *Permutation Pattern Matching* the input is a pattern $\pi\colon [k] \to [k]$ and a text $\tau\colon [n] \to [n]$, both permutations, with $k \le n$. An *occurrence* of π in τ is a strictly increasing map $f\colon [k] \to [n]$ such that $\tau(f(i)) < \tau(f(j))$ if and only if $\pi(i) < \pi(j)$; the decision version asks whether an occurrence exists, and the counting version for their number. In *k-opt Optimization* for TSP the input is an edge-weighted graph with a Hamiltonian cycle, and the task is to find a k-opt move (removing k edges of the cycle and adding k edges so as to obtain a Hamiltonian cycle) of maximum gain [CKS19, BIJK19b].

Cygan, Kowalik and Socała [CKS19] introduced a *bucketing* argument for k-opt, and Berendsohn, Kozma and Marx [BKM21] transferred it to Permutation Pattern Matching. It trades the number

of guesses against the pathwidth of a graph made of a cycle cover (or a path) and some edges of the path $1, 2, \ldots, k$: coarser buckets mean fewer guesses but larger domains. Both previous bounds $n^{k/4}$ come from a coefficient $\frac{1}{3}$ for graphs of maximum degree 4. A single smaller coefficient for maximum degree 4, such as F_4, improves the bound obtained without bucketing to $n^{F_4 k}$, but it does not combine with bucketing to anything better: inside the trade-off, the pathwidth bound has to grow with the number of path edges. The per-degree bounds of Theorem 7.1(b),(c) and Theorem 7.2 have exactly this property. We use the triple R of Theorem 7.2, with $n_5 = 0$; within our argument, adding the bounds of Theorems 7.1 and 8.2 does not lower the exponent further (see Appendix A.7.2). We first bound the pathwidth of the graphs that arise (Lemma 11.6), then solve the trade-off (Lemma 11.7), and finally apply both lemmas to the two problems (Theorem 11.8).

For an integer $k \geq 1$ and a real $s \in [0, k]$ put

$$p(s) := \min\{0.291\, s,\ 0.038\, k + 0.215\, s\};$$

k is fixed wherever p is used. We write $P_k := \{\{i, i+1\} : i \in [k-1]\}$ for the edge set of the path $1, 2, \ldots, k$.

Lemma 11.6 (cycle cover plus path edges)**.** *There is a function $\rho(k) = o(k)$ such that the following holds. Let C be a multigraph on $[k]$, possibly with loops, of maximum degree at most 2, and let $A \subseteq P_k$ with $|A| = s$. Then the simple graph H underlying $C \cup A$ satisfies $\mathrm{pw}(H) \leq p(s) + \rho(k)$, and a path decomposition of H of this width can be computed in polynomial time.*

The bucketing exponent is a maximum over s of a weighted sum of the number $k - s$ of guessed buckets and the width bound $p(s)$; the weight α is the exponent of the bucket width. The next lemma chooses α.

Lemma 11.7 (the trade-off)**.** *Let $\alpha = \frac{200}{243}$. For every $0 \leq s \leq k$,*

$$(1-\alpha)(k-s) + \alpha\, p(s) \leq \tfrac{253}{1215} k.$$

Moreover, no $\alpha \in [0, 1]$ gives a smaller bound than $\frac{253}{1215} k$ on $\max_{0 \leq s \leq k}[(1-\alpha)(k-s) + \alpha p(s)]$.

The proofs of the two lemmas, and a check that the other bounds of Part I do not lower the exponent within this argument, are in Appendix A.7.2.

The two lemmas give the following bounds for the problems defined at the start of this subsection.

Theorem 11.8 (permutation patterns and k-opt)**.** *There are a function $\rho(k) = o(k)$ and numbers $\varepsilon_k \to 0$ such that:*

(a) *the decision and counting versions of Permutation Pattern Matching, for a pattern of length k and a text of length n, can be solved in time $2^{O(k)}\, n^{\frac{253}{1215} k + \rho(k)}$;*

(b) *for every fixed k, k-opt Optimization can be solved in time $O(n^{(253/1215 + \varepsilon_k)k})$ on n-vertex edge-weighted graphs, and in time $O(n^{(7/48 + \varepsilon_k)k})$ on those of bounded maximum degree.*

The proof, in Appendix A.7.2, follows Berendsohn, Kozma and Marx for (a) and Cygan, Kowalik and Socała for (b), with Lemmas 11.6 and 11.7 in place of their pathwidth bounds, and with Remark 9.14 for bounded degree.

12 Separate, Measure and Conquer in polynomial space

The dynamic programs of Section 11 keep a table for every bag of a path decomposition and therefore use exponential space. In polynomial space the width bounds enter in a different way. A path decomposition of small width contains a small balanced separator, and a branching algorithm can prefer to branch on the vertices of that separator: once the separator is exhausted, the instance falls apart into two independent halves. Gaspers and Sorkin [GS17] built a measure-and-conquer analysis around this idea, which they call *Separate, Measure and Conquer*, and Gaspers and Lee [GL23] used it to count independent sets. In both analyses the separator constant—the coefficient $\frac{1}{6}$ of Fomin and Høie [FH06]—enters a mathematical program in one or two places. We replace it by c_3 and solve the programs again. This gives the following results, all in polynomial space:

- Max $(r, 2)$-CSP in time $r^{0.17935n}$ on cubic instances, $r^{0.16984m}$ on instances of maximum degree 4 and $r^{0.17587m}$ on all instances, up to $o(\cdot)$ terms in the exponent, and counting dominating sets of subcubic graphs in time $O(1.2178^n)$ (Theorem C.3), against the bounds $r^{n/5}$, $r^{7m/40}$, $r^{9m/50}$ and $O(1.2458^n)$ of Gaspers and Sorkin;
- counting independent sets in time $O(1.1304^n)$ for maximum degree 3, in time $O(1.2037^n)$ for maximum degree 4 and $O(1.2341^n)$ in general, and computing the chromatic number in time $O(2.2341^n)$ (Theorem C.6), against the bounds $O(1.1389^n)$, $O(1.2070^n)$, $O(1.2356^n)$ and $O(2.2356^n)$ of Gaspers and Lee.

Appendix C explains the technique, shows where the separator constant enters, and proves both theorems.

13 Counting dominating sets

This section shows how the smaller pathwidth bounds speed up an algorithm that counts dominating sets by branching first and then switching to dynamic programming: with the new bounds, the switch can come one step earlier. We start from the exponential-space algorithm of Nederlof, van Rooij and van Dijk [NvRvD14] for counting dominating sets in general graphs. It counts set covers by branching, and switches to dynamic programming over a path decomposition of the incidence graph once the instance is sparse. Its stated running time is $O(1.5002^n)$. We prove the following theorem.

Theorem 13.1 (counting dominating sets)**.** *The numbers of dominating sets of each size of an n-vertex graph can be computed deterministically in time and space $O(1.4815^n)$. If the vertices carry real weights with $n^{O(1)}$ distinct subset sums, the same holds for the numbers of dominating sets of each size and each weight.*

Our algorithm is the one of Nederlof, van Rooij and van Dijk with two changes. First, it switches to dynamic programming one step earlier, when all non-exceptional sets have size at most 4 instead of 3. Second, it uses a longer list of exceptional configurations.

While rerunning the original analysis we found a gap in it: one family of branching cases is not covered by the previous recurrences, and with the previous weights some of these cases violate the claimed branching factor. The gap is in the proof, not in the algorithm: re-optimizing the weights with all cases included shows that the original algorithm runs in time $O(1.5004^n)$ (Proposition D.10). Appendix D.5 describes the gap; our analysis uses the corrected set of recurrences.

Table 6: Types $(r_1, \ldots, r_5)$ of exceptional elements of frequency 5. The previous list $\mathcal{X}_{14}$ [NvRvD14, Overview 1] and the sixteen types added in $\mathcal{X}_{30}$.

$\mathcal{X}_{14}$	(1,4,0,0,0)	(1,3,1,0,0)	(1,2,2,0,0)	(1,1,3,0,0)	(1,0,4,0,0)
	(1,3,0,1,0)	(1,2,1,1,0)	(0,5,0,0,0)	(0,4,1,0,0)	(0,3,2,0,0)
	(0,2,3,0,0)	(0,1,4,0,0)	(0,4,0,1,0)	(0,3,1,1,0)	
added in $\mathcal{X}_{30}$	(1,0,0,4,0)	(1,0,1,3,0)	(1,0,2,1,1)	(1,0,2,2,0)	
	(1,0,3,0,1)	(1,0,3,1,0)	(1,1,0,2,1)	(1,1,0,3,0)	
	(1,1,1,1,1)	(1,1,1,2,0)	(1,1,2,0,1)	(1,1,2,1,0)	
	(1,2,0,1,1)	(1,2,0,2,0)	(1,2,1,0,1)	(1,3,0,0,1)	

13.1 The algorithm

Set cover formulation. A *set cover instance* consists of a finite universe $\mathcal{U}$ of *elements*, a finite family $\mathcal{S}$ of subsets of $\mathcal{U}$ called *sets*, and a *multiplicity* $\text{mult}(s) \geq 1$ for every $s \in \mathcal{S}$; a set s stands for $\text{mult}(s)$ interchangeable copies. A *cover of size* ℓ is a choice of ℓ copies, of possibly several sets, whose union is $\mathcal{U}$. The task is to compute the numbers c_ℓ of covers of size ℓ for all ℓ. The *incidence graph* I of the instance is the bipartite graph on $\mathcal{S} \cup \mathcal{U}$ in which $s \in \mathcal{S}$ and $e \in \mathcal{U}$ are adjacent when $e \in s$. We write $d(s) = d_I(s) = |s|$ for the *size* of a set and $d(e) = d_I(e)$ for the *frequency* of an element (the number of sets containing it), and $N_I(e)$ for the sets containing e. In the thesis and in the journal version, sets and elements are the red and blue vertices of a red-blue graph; this is the same object.

A graph G gives the instance with $\mathcal{U} = V(G)$ and one set $N[v]$ of multiplicity one for each $v \in V(G)$ (true twins u, v, with $N[u] = N[v]$, give two identical sets, which the first reduction rule below merges). A vertex set D is dominating if and only if the copies $\{N[v] : v \in D\}$ form a cover, so c_ℓ is the number of dominating sets of size ℓ.

Exceptional configurations. The algorithm leaves certain elements of frequency 5 to the dynamic program. The *type* of an element e of frequency 5 all of whose sets have size at most 5 is the tuple $(r_1, \ldots, r_5)$, where r_i is the number of sets of size i containing e; thus $r_1 + \cdots + r_5 = 5$. (Here and below, r_i counts neighbors of a vertex of I by degree; it is unrelated to the domain size r of Section 11.) Let $\mathcal{X}$ be a list of types and $\sigma \in \{3, 4\}$. An element is *exceptional* if it has frequency 5 and its type is in $\mathcal{X}$. A set is *exceptional* if its size lies in $\{\sigma+1, \ldots, 5\}$ and it contains an exceptional element. We always assume

$$r_i = 0 \ \text{ for } i > \sigma + 1 \text{ and } r_1 \leq 1 \quad \text{for every type } (r_1, \ldots, r_5) \in \mathcal{X}, \tag{13.1}$$

so that every set that contains an exceptional element has size at most $\sigma + 1$; in particular, every exceptional set has size exactly $\sigma + 1$.

The previous algorithm has $\sigma = 3$ and the list $\mathcal{X}_{14}$ of fourteen types of Table 6 [vR11, Overview 8.1, p. 150], [NvRvD14, Overview 1, p. 727]; with it, exceptional sets have size 4. Our algorithm has $\sigma = 4$ and the list $\mathcal{X}_{30}$, which adds sixteen types with $r_1 = 1$. Both lists satisfy (13.1): the types of $\mathcal{X}_{14}$ have $r_5 = 0$, and those of $\mathcal{X}_{30}$ have no set of size above 5.

The algorithm $\mathcal{A}_{\sigma,\mathcal{X}}$. On an instance $(\mathcal{S},\mathcal{U},\mathrm{mult})$ the algorithm applies the first applicable rule of the following list and recurses. We write $C(\cdot)$ for the vector $(c_\ell)_\ell$ returned by a recursive call, and

$$\mathrm{take}_s(C)_\ell = \sum_{i=1}^{\mathrm{mult}(s)} \binom{\mathrm{mult}(s)}{i} c_{\ell-i} \tag{13.2}$$

for the correction that accounts for choosing at least one of the copies of a set s (with $c_j = 0$ for $j < 0$) [NvRvD14, Proposition 26].

(R1) *Identical sets.* If two sets $s_1 \neq s_2$ have the same elements, delete s_2 and set $\mathrm{mult}(s_1) := \mathrm{mult}(s_1) + \mathrm{mult}(s_2)$.

(R2) *Forced set.* If an element has frequency 1, let s be its set; return $\mathrm{take}_s(C(\mathcal{S} \setminus \{s\},\mathcal{U} \setminus s))$, where every remaining set loses its elements in s.

(R3) *Subsumed element.* If two elements $e_1 \neq e_2$ satisfy $N_I(e_1) \subseteq N_I(e_2)$ (every set containing e_1 contains e_2), delete e_2 from $\mathcal{U}$ and from all sets.

If no reduction rule applies, let s be a non-exceptional set of maximum size and e a non-exceptional element of maximum frequency (with $d(s) = 0$ or $d(e) = 0$ if there is none).

(A) *Switch.* If $d(s) \leq \sigma$ and $d(e) \leq 4$, compute a path decomposition of I and count the covers by dynamic programming over it.

(B) *Branch on a set.* Otherwise, if $d(s) > d(e)$ or $d(e) \leq 4$, return $\mathrm{take}_s(C(\mathcal{S} \setminus \{s\},\mathcal{U} \setminus s)) + C(\mathcal{S} \setminus \{s\},\mathcal{U})$ (take s, or discard it).

(C) *Branch on an element.* Otherwise return $C(\mathcal{S},\mathcal{U}\setminus\{e\}) - C(\mathcal{S}\setminus N_I(e),\mathcal{U}\setminus\{e\})$ (make e optional, minus forbid e).

This is [vR11, Algorithm 8.2, p. 148] and [NvRvD14, Algorithm 4, p. 725] when $\sigma = 3$ and $\mathcal{X} = \mathcal{X}_{14}$; there, the switch is line 11 and line 12, respectively. The order of the reduction rules plays no role in the analysis.

How the switch works. The algorithm monitors two quantities: the largest size of a non-exceptional set and the largest frequency of a non-exceptional element. While either is large, branching on that set or element lowers the measure (13.3) defined below by a large amount in both branches. Once both are small, each further branching would lower the measure only a little, but the incidence graph is then sparse (maximum degree 5, with the vertices of degree 5 confined to the exceptional configurations), and its pathwidth is small. The threshold σ decides whether sets of size $\sigma + 1$ are branched on or handed to the dynamic program. The previous threshold $\sigma = 3$ branches on sets of size 4; we take $\sigma = 4$ and hand them to the dynamic program. The exceptional configurations make the same choice for elements of frequency 5: branching on an element whose sets are mostly very small lowers the measure only a little, and the dynamic program handles such an element at little extra running time.

Lemma 13.2 (correctness)**.** *For every $\sigma \in \{3, 4\}$ and every list $\mathcal{X}$, the algorithm $\mathcal{A}_{\sigma,\mathcal{X}}$ terminates and returns $(c_\ell)_\ell$, the numbers of covers of each size.*

Proof. Every rule produces instances with fewer vertices in the incidence graph, so the recursion terminates. Correctness follows by induction on $|V(I)|$; we check each rule in turn. (R1): a choice of copies of s_1 and s_2 is a choice of copies of the merged set, with the same union. (R2): every

cover contains a copy of s; covers containing $i \geq 1$ copies of s correspond to a choice of i of the $\mathrm{mult}(s)$ copies together with a cover of $\mathcal{U} \setminus s$ by the other sets, which gives (13.2). (R3): every cover contains a copy of a set containing e_1, which contains e_2; so a choice of copies covers $\mathcal{U}$ if and only if it covers $\mathcal{U} \setminus \{e_2\}$. (B): covers either contain a copy of s (counted as in (R2)) or none. (C): a choice of copies that covers $\mathcal{U} \setminus \{e\}$ either covers e or uses no set containing e, and the latter choices are exactly the covers of $\mathcal{U} \setminus \{e\}$ by $\mathcal{S} \setminus N_I(e)$; so the number of covers of $\mathcal{U}$ of each size is the difference. (A): the dynamic program over a path decomposition of width p counts the covers of each size, with multiplicities, for every instance, in time $O^*(2^p)$ [vR11, Proposition 8.10], [NvRvD14, Proposition 27]; vertices of degree 0 of I (empty sets and elements of frequency 0) are handled by a factor $(1+x)^{\mathrm{mult}(s)}$ in the size-generating function, or by returning 0. □

The measure. The analysis measures an instance by

$$k = \sum_{e \in \mathcal{U}} v(d(e)) + \sum_{s \in \mathcal{S}} w(d(s)), \tag{13.3}$$

where merged sets count once. The weights satisfy the following conditions, which are those of the thesis [vR11, p. 151]; write $\Delta v(i) = v(i) - v(i-1)$ and $\Delta w(i) = w(i) - w(i-1)$.

(W1) $v(0) = v(1) = w(0) = 0$; $v(i) = v(7)$ and $w(i) = w(7)$ for $i \geq 7$; $v(2) > 0$ and $w(1) > 0$.

(W2) $\Delta v(i) \geq 0$ for $i \geq 2$ and $\Delta w(i) \geq 0$ for $i \geq 1$.

(W3) $\Delta v(i) \geq \Delta v(i+1)$ for $i \geq 2$ and $\Delta w(i) \geq \Delta w(i+1)$ for $i \geq 1$.

(W4) $2\,\Delta v(5) \leq v(2)$.

By (W1)–(W2) the weights are nonnegative and nondecreasing, so none of (R1)–(R3) increases the measure: each deletes vertices of I and lowers sizes or frequencies of others. A graph with n vertices gives an instance of measure at most $(v(7) + w(7))\,n$.

13.2 Outline of the analysis

The running-time analysis follows the measure-and-conquer scheme of the thesis [vR11, §8.5] in three steps, which Appendix D carries out for both thresholds at once. First, branching recurrences bound the number of instances of each measure on which the dynamic program is called [vR11, Lemma 8.11]; the branching vectors of degree above 9 reduce to those of degree at most 9. Second, a linear program in the numbers of sets and elements of each degree bounds the pathwidth of these instances [vR11, Lemma 8.12]; its objectives are the pathwidth bounds of Part I, applied to the incidence graph. Third, the two bounds combine into the running time $\alpha^{(v(7)+w(7))n} n^{O(1)}$ whenever $2^L < \alpha$ for the value L of the linear program [vR11, Theorem 8.13].

For $\mathcal{A}_{4,\mathcal{X}_{30}}$ we take $\alpha = \frac{13}{10}$ and rational weights with $v(7) + w(7) = 1.4980028$ (Table 12). All 30,972 branching vectors of degree at most 9 hold at α, the linear program has value $L \approx 0.37851$ with $2^L < \alpha$, and $\alpha^{1.4980028} < 1.4815$. Each of these steps is checked in exact rational or interval arithmetic. The decisive input is the coefficient $\frac{1}{4}$ of Theorem 1.3 at degree 4, in place of the $\frac{1}{3}$ of Fomin, Gaspers, Saurabh and Stepanov [FGSS09, Lemma 1]: the worst instances of the linear program are close to incidence graphs of cubic graphs. For the weighted counts, the multiplicity of a set becomes a generating polynomial in size and weight, and the measure and running time are those of the unweighted case (Appendix D.3).

14 Connectivity problems

Many problems with a global connectivity or acyclicity constraint can be solved in time single exponential in the width of a given decomposition, and several of them are NP-hard on graphs of maximum degree 3 or 4. This section runs such algorithms on the decompositions of Theorem 1.5 and on the orderings of Corollary 10.2 and Theorem 3.6. This gives new running-time bounds for these problems on graphs of maximum degree 3 to 6, on graphs with a given number of edges, and for three parameterized problems (Theorem 14.1 and the paragraph on Almost Forest Deletion at the end of this section).

The algorithms that we use are the following. The Cut&Count technique of Cygan, Nederlof, Pilipczuk, Pilipczuk, van Rooij and Wojtaszczyk gives Monte Carlo algorithms in time $c^{\mathrm{tw}}n^{O(1)}$ for a given tree decomposition of width tw [CNP+11, Theorem 1.2]. The base is $c = 3$ for Feedback Vertex Set, Connected Vertex Cover and Steiner Tree, $c = 4$ for Connected Dominating Set and eight further problems on undirected graphs, and $c = 6$ for the directed versions of Min Cycle Cover, Longest Path and Longest Cycle and for Exact k-Leaf Outbranching. Cygan, Kratsch and Nederlof decide Hamiltonicity by a Monte Carlo algorithm in time $(2+\sqrt{2})^{\mathrm{pw}}n^{O(1)}$ for a given path decomposition of width pw [CKN12, Theorem 4.3], and Longest Path in the same time [CKN12, Theorem 4.6, stated with a proof sketch]. The rank-based approach of Bodlaender, Cygan, Kratsch and Nederlof gives deterministic algorithms, for instance in time $(2+2^{\omega/2})^{\mathrm{pw}}n^{O(1)}$ for the traveling salesperson problem with nonnegative integer edge weights, where ω is the matrix multiplication exponent [BCKN15, Theorem 3.18]. Bojikian, Chekan, Hegerfeld and Kratsch [BCHK22] and Bojikian, Chekan and Kratsch [BCK25] give algorithms of the same kind for a given linear ordering of cutwidth ctw, with smaller bases.

Combined with the pathwidth bound $\frac{n}{6}$ of Fomin and Høie [FH06] and the per-degree bounds of Fomin, Gaspers, Saurabh and Stepanov [FGSS09] ($\frac{n}{3}$ for maximum degree 4 and $\frac{13n}{30}$ for maximum degree 5), these algorithms gave the fastest algorithms known to us for several of these problems on graphs of maximum degree 3, 4 and 5. Cygan et al. list the Cut&Count bounds [CNP+11, Corollary 1.9]. Cygan, Kratsch and Nederlof [CKN12, Corollary 4.5] state theirs for maximum degree 3, and Bodlaender et al. [BCKN15, Corollary 3.19] theirs for cubic graphs. Kowalik and Majewski tabulate those for Hamiltonicity and the traveling salesperson problem [KM20, Table 1]. The pathwidth bound of Fomin, Gaspers, Saurabh and Stepanov in terms of the numbers of vertices of each degree also bounds the width of the instances that remain after branching in parameterized algorithms for Co-Path Packing [LX26] and k-Co-Path Set [SvdP16]. With the bounds of Part I we obtain the following.

Theorem 14.1 (connectivity problems)**.** *Polynomial factors are omitted, all algorithms use exponential space, and the Monte Carlo algorithms have one-sided error at most* $\frac{1}{2}$*.*

(a) *Feedback Vertex Set and Connected Vertex Cover can be solved by Monte Carlo algorithms in time* $2^{m/6+o(n)} \leq 2^{n/3+o(n)} = O(1.2600^n)$ *on graphs with* m *edges and maximum degree* 4*, and in time* $2^{\gamma_5 m+o(n)} = O(1.4262^n)$ *on graphs of maximum degree* 5*, where* $\gamma_5 < 0.20484$ *is the constant of Theorem 3.6. Steiner Tree with unit weights can be solved by Monte Carlo algorithms in time* $3^{c_3 n+o(n)} = O(1.1661^n)$ *and* $3^{F_4 n+o(n)} = O(1.3100^n)$ *on graphs of maximum degree* 3 *and* 4*. Feedback Vertex Set can be solved on graphs of maximum degree* 6 *by a Monte Carlo algorithm in time* $3^{7n/16+o(n)} = O(1.6172^n)$*. Feedback Vertex Set can be solved on graphs of maximum degree* 4 *deterministically in time* $5^{F_4 n+o(n)} = O(1.4853^n)$*.*

(b) *Connected Dominating Set, Connected Feedback Vertex Set, Connected Odd Cycle Transversal, Exact* k*-Leaf Spanning Tree (in particular Minimum and Maximum Leaf Spanning Tree),*

Maximum Full Degree Spanning Tree, Graph Metric Traveling Salesperson, and the undirected versions of Min Cycle Cover and Longest Cycle can be solved by Monte Carlo algorithms in time $4^{c_3 n+o(n)} = O(1.2139^n)$, $4^{F_4 n+o(n)} = O(1.4060^n)$ *and* $4^{11n/32+o(n)} = O(1.6105^n)$ *on graphs of maximum degree* 3, 4 *and* 5. *Minimum Leaf Spanning Tree (equivalently, Maximum Internal Spanning Tree) can be solved on graphs of maximum degree* 6 *by a Monte Carlo algorithm in time* $4^{7n/16+o(n)} = O(1.8341^n)$, *and Connected Dominating Set on graphs of maximum degree* 3 *by a Monte Carlo algorithm in time* $3^{c_3 n+o(n)} = O(1.1661^n)$. *Maximum Leaf Spanning Tree has the bounds of Connected Dominating Set.*

(c) *Hamiltonian Cycle can be solved by Monte Carlo algorithms in time* $(1+\sqrt{2})^{c_3 n+o(n)} = O(1.1312^n)$ *on graphs of maximum degree* 3, *in time* $(1+\sqrt{2})^{m/6+o(n)} \leq (1+\sqrt{2})^{n/3+o(n)} = O(1.3416^n)$ *on graphs with* m *edges and maximum degree* 4, *and in time* $(2+\sqrt{2})^{11n/32+o(n)} = O(1.5252^n)$ *on graphs of maximum degree* 5. *Undirected Longest Path can be solved by Monte Carlo algorithms in time* $(2+\sqrt{2})^{c_3 n+o(n)} = O(1.1874^n)$, $(2+\sqrt{2})^{F_4 n+o(n)} = O(1.3523^n)$ *and* $(2+\sqrt{2})^{11n/32+o(n)} = O(1.5252^n)$ *on graphs of maximum degree* 3, 4 *and* 5.

(d) *With* $\omega \leq \frac{9}{4}$ *[Ope26b, AvdBZ26], the traveling salesperson problem with nonnegative integer edge weights can be solved deterministically in time* $(1+2^{\omega/2})^{c_3 n+o(n)} = O(1.1757^n)$ *on graphs of maximum degree* 3, *and in time* $(2+2^{\omega/2})^{F_4 n+o(n)} = O(1.4214^n)$, $(2+2^{\omega/2})^{11n/32+o(n)} = O(1.6352^n)$ *and* $(2+2^{\omega/2})^{7n/16+o(n)} = O(1.8699^n)$ *on graphs of maximum degree* 4, 5 *and* 6.

(e) *The directed versions of Min Cycle Cover, Longest Path and Longest Cycle, and Exact* k*-Leaf Outbranching, can be solved by Monte Carlo algorithms in time* $6^{c_3 n+o(n)} = O(1.2848^n)$ *and* $6^{F_4 n+o(n)} = O(1.5533^n)$ *on directed graphs whose underlying undirected graph has maximum degree* 3 *and* 4.

(f) *Co-Path Packing (delete at most* k *vertices so that every component of the remaining graph is a path) can be solved by a Monte Carlo algorithm in time* $O^*(2.8192^k)$, *and* k*-Co-Path Set (delete at most* k *edges so that every component of the remaining graph is a path) by a Monte Carlo algorithm in time* $O^*(1.5423^k)$.

(g) *On graphs with* m *edges, Hamiltonian Cycle can be solved by a Monte Carlo algorithm in time* $(2+\sqrt{2})^{7m/48+o(m)} = O(1.1962^m)$, *and the traveling salesperson problem with nonnegative integer edge weights deterministically in time* $(2+2^{\omega/2})^{7m/48+o(m)} = O(1.2320^m)$, *with* ω *as in* (d). *For average degree at most* d, *that is* $m \leq \frac{d}{2}n$, *these are* $O(1.0937^{dn})$ *and* $O(1.1100^{dn})$.

The proof, in Appendix A.7.3, runs the algorithms listed above on the path decompositions of Theorem 10.1 and on the orderings of Corollary 10.2 and Theorem 10.3. The path-decomposition bounds for maximum degree 5 and 6 and those of (f) and (g) use Theorem 1.3, and hence Theorem 7.2. The cutwidth bounds for maximum degree 4 and 5 use the orderings of Theorem 3.6, which Theorem 10.3 computes in deterministic polynomial time.

Almost Forest Deletion. Gowda, Lonkar, Panolan, Patel and Saurabh ask for at most k vertices whose deletion leaves a graph that becomes a forest after the deletion of at most ℓ further edges. One of their Monte Carlo algorithms runs in time $O^*(3^k \cdot 1.78^\ell)$ and exponential space [GLP+24, Theorem 1.2(4)]. It applies the Cut&Count algorithm to a tree decomposition of width $k + \frac{3}{5.769}\ell + O(\log \ell)$, obtained from the bound of Kneis et al. on a graph with at most 3ℓ edges [GLP+24, Lemma 3.34]. With Remark 9.14 in place of that bound, the width is $k + \frac{7}{16}\ell + o(\ell)$, and since $3^{7/16} < 1.6172$ the running time becomes $O^*(3^k \cdot 1.62^\ell)$.

15 Conclusion and open problems

For twenty years, the bound $\frac{n}{6}+o(n)$ of Fomin and Høie [FH06] was the best bound on the pathwidth of n-vertex graphs of maximum degree 3, and many exact algorithms stated their running times through it. We replaced the bisection behind it by a threshold sweep through correlated Gaussian labels. Comparing each star with an ideal star, and switching from inner to outer frontiers at the median, brought the bound down to $c_3 n + o(n)$ with $c_3 = \frac{3}{4\pi}\arccos\frac{5}{6} \approx 0.1398$. A charging scheme certified by computer and a bootstrap to large degrees extended the method to all graphs, with the coefficient $\frac{1}{4}$ in place of $\frac{1}{3}$ at degree 4 and the bound $\frac{7}{48}m$ per edge in place of $\frac{13}{75}m$. The decompositions can be computed in polynomial time. This solves the open problem of van Rooij [vR11, Open Problem 14.2], which asks for tree decompositions of cubic graphs of width $c_t n + o(n)$, for some constant $c_t < \frac{1}{6}$, computable in polynomial time: we obtain $c_t = c_3$, even for path decompositions. With the new decompositions, many exact algorithms whose running times depend on these widths have smaller running-time bounds (Part II). We close with the questions that the Gaussian sweep leaves open.

The constant for maximum degree 3. What is the smallest constant c such that every graph of maximum degree 3 has pathwidth at most $cn + o(n)$? It lies in $[0.1032, c_3]$, and the lower end comes from random cubic graphs [LM23]. Since c_3 is exactly what the median-switched sweep gives on cubic graphs of large girth, an improvement must change the rule at threshold zero, for instance by letting a vertex above the threshold carry the crossing edges of all lower neighbors whose only upper neighbor it is. On the *ideal wave*, the centered Gaussian vector on the rooted tree T_2 (root o with three children u_1, u_2, u_3, each with two children) with covariance $2^{-d/2}(1+\frac{d}{3})$ at distance d, which is the limit of our labels on cubic graphs of large girth, this rule saves $3p_{\mathrm{ch}}$ per vertex at threshold zero, where $p_{\mathrm{ch}} = 0.00252477819777\ldots$ (certified in interval arithmetic) is the probability that u_1 and u_2 both have o as their only neighbor above zero.[2]

Conjecture 15.1 (ideal-wave bound)**.** *Every cubic graph G satisfies* $\mathrm{pw}(G) \le (c_3 - 3p_{\mathrm{ch}})n + o(n)$*, where* $c_3 - 3p_{\mathrm{ch}} = 0.13224778941\ldots$.

Proving it, or any improvement of c_3, seems to require a quantitative stability version of the star comparison.

Maximum degree 4. For maximum degree 3, the median switch lowers the coefficient of the plain sweep, the maximum $F_3 \approx 0.1402$ of the ideal 3-star function, to the value c_3 of that function at threshold zero. For maximum degree 4 we prove $F_4 n + o(n)$, where $F_4 \approx 0.2458$ is the maximum of the ideal 4-star function; its value at threshold zero is 0.2423. Is there a rule at threshold zero whose width bound on 4-regular graphs approaches $0.2423n$? Our median switch, analyzed star by star, cannot approach 0.2423: the switch quantity Ψ of a vertex with four neighbors of degree 4 can equal $\frac{1}{4} > F_4$ (see the discussion after Lemma A.10). It does give the bound $c_3 n_3 + 0.2627\, n_4$ (Theorem 8.2), which is smaller than the bounds of Theorem 7.1 when $n_4 < \frac{F_3 - c_3}{0.0052}\, n_3 \approx 0.0646\, n_3$. This form of the method cannot lower 0.2627 much below 0.2618 (Remark A.15).

Degree 6 and the bound per edge. Can the coefficient $a_6 = \frac{7}{16}$ of Theorem 1.3 be lowered? For degrees 6 and more, all our coefficients come from the bootstrap. The constant $\frac{7}{48}$ of Corollary 1.4

[2] This is the ideal-tree value of the "isolated helper" event of Lichev and Mitsche [LM23]; the number 0.002818666 printed there is not this value, and its authors have confirmed that it was not accurate (L. Lichev and D. Mitsche, personal communication, 2026).

is $2a_6/6$, and every other ratio $2a_d/d$ is at most $\frac{113}{784}$, attained at $d = 7$ (proof of Corollary 1.4); so a better coefficient at degree 6 alone would lower the bound per edge.

Treewidth. For any of the classes considered here, is the worst-case coefficient of treewidth smaller than that of pathwidth? The Gaussian sweep produces path decompositions, and it has given us no gain specific to treewidth at leading order.

Acknowledgments

This research was supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP210103849). This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney [PVC10]. We thank the UNSW AI Institute for supporting the co-algorithmist framework, a framework for human–AI research collaboration in algorithms and theoretical computer science that is under development. We thank Yuexi He for interesting discussions around the treewidth and pathwidth of structured (but not inherently sparse) graphs occurring in the context of a parallel research project.

AI methodology statement

This research was carried out with assistance from large language model systems: Codex (OpenAI) in the early stages, and Claude (Anthropic) thereafter, in the versions ChatGPT 6 Astra, Sol 5.6, Claude Opus 5.5, Sonnet 5.5, Haiku 5.5 and Fable 5.1. GitHub Copilot assisted in project maintenance. These tools materially affected every section of the paper.

- They proposed and developed ideas in discussion with the author, searched the literature, carried out mathematical derivations and proofs, wrote the code for the experiments, the computer certificates and the Lean development, drafted, reviewed, and revised the text.
- OpenAI Codex was also used for developing the research-workflow tooling of the project.
- GitHub Copilot was used for code completion, reviews of pull requests and commit messages. Pull requests were also reviewed by Claude Haiku 5.5.

The author directed the research, contributed ideas, checked the proofs and the Lean statements, edited the manuscript and extended the literature search. The author verified the correctness and originality of all content, including references.

A Deferred proofs

This appendix collects the proofs deferred from the main text. The first five subsections restate the deferred statements, with their numbers in the main text, and prove them: Lemma 4.5 in Appendix A.1; Proposition 6.9 and the tools of the certificate behind Proposition 7.5 in Appendix A.2; Theorem 8.2 in Appendix A.3; Theorem 9.2 and Corollary 9.5 in Appendix A.4; and Theorems 10.1 and 10.3 and Corollary 10.2 in Appendix A.5. The next, Appendix A.6, restates nothing: it completes the proof of Theorem 10.1 in the cases that use the median-switched layout. The last, Appendix A.7, proves Lemmas 11.2, 11.6 and 11.7 and Theorems 11.8 and 14.1 of Part II; the analyses behind Theorems C.3, C.6 and 13.1 have appendices of their own (Appendices C and D).

A.1 Proof of the reduction lemma

This subsection proves Lemma 4.5, with the reduction operations of Definition 4.2, their properties (RP1)–(RP3) and the two facts Lemmas 4.3 and 4.4 of Section 4.3.

Lemma 4.5 (undoing reductions)**.** *Let F be a nonempty graph. Apply to F any finite sequence of steps, each of which is a reduction operation or the deletion of a vertex of the current graph. Let D be the set of vertices removed by deletions and R the final graph. Then:*

(a) $\mathrm{pw}(F) \le |D| + \mathrm{pw}(R) + 3\log_2 |V(F)| + 1$*;*

(b) *every edge of $F - D$ with both ends in $V(R)$ is an edge of R, and for every component C of $(F - D) - V(R)$, the set $N_{F-D}(C)$ is empty, a single vertex, or the two ends of an edge of R, and* $\mathrm{tw}(F[C]) \le 2$*.*

Proof. The proof follows the sequence step by step. The vertices removed so far by reductions induce components in F minus the vertices deleted so far. We maintain as an invariant that each of these components has treewidth at most 2 and attaches to the current graph at no vertex, at one vertex, or at the two ends of an edge. At the end, a path decomposition of the final graph R is extended to one of $F - D$ by inserting path decompositions of these components, each of width at most $3\log_2 |V(F)|$ by Lemma 4.4, next to bags that contain their attachments. Adding D to every bag then gives the claim.

The invariants. Write $n = |V(F)|$. Let H_t be the current graph after t steps, D_t the set of vertices deleted in the first t steps, and $F_t := F - D_t$. Then $V(H_t) \subseteq V(F_t)$, but H_t need not be a subgraph of F_t, since (R3) adds edges. For a component C of $F_t - V(H_t)$ put $S_C := N_{F_t}(C)$; these are vertices of H_t. We maintain two invariants:

(I1) every edge of F_t with both ends in $V(H_t)$ is an edge of H_t;

(I2) for every component C of $F_t - V(H_t)$: $|S_C| \le 2$; if $|S_C| = 2$ then S_C spans an edge of H_t; and the graph Q_C^+ obtained from $Q_C := F_t[C \cup S_C]$ by adding the edge on S_C when $|S_C| = 2$ has treewidth at most 2.

Both hold for $t = 0$, when $H_0 = F_0 = F$ and there are no components.

A deletion of a vertex $w \in V(H_t)$ removes w from both graphs. Edges of F_{t+1} inside $V(H_{t+1})$ are edges of F_t inside $V(H_t)$ avoiding w, hence edges of $H_t - w = H_{t+1}$; so (I1) holds. The graph $F_t - V(H_t)$ does not change, each S_C loses at most the vertex w, an edge on S_C that avoids w survives in H_{t+1}, and each new Q_C^+ is a subgraph of the old one. So (I2) holds.

A reduction operation at a vertex v. Let $N := N_{H_t}(v)$, so $|N| \le 2$. Here $F_{t+1} = F_t$ and $V(H_{t+1}) = V(H_t) \setminus \{v\}$. Edges of F_t inside $V(H_{t+1})$ are edges of H_t not incident with v, and these

survive in H_{t+1}; so (I1) holds. The components C of $F_t - V(H_t)$ with $v \notin S_C$ remain components, with the same S_C, the same Q_C^+, and an edge on S_C that avoids v and therefore survives. The components C with $v \in S_C$ merge with v into one new component C'. We check (I2) for C'.

- $S_{C'} \subseteq N$. A neighbor of v in F_t outside C' lies in $V(H_t)$, so by (I1) it lies in N. A neighbor of a merged C outside C' lies in $S_C \setminus \{v\}$; if this set is nonempty, $S_C = \{v, u\}$ spans an edge of H_t by (I2), so $u \in N$.

- If $|S_{C'}| = 2$, then $N = S_{C'} = \{x, y\}$, so v had degree 2 and the operation was (R3) or (R4). In both cases $xy \in E(H_{t+1})$.

- Let K be the graph on $\{v\} \cup N$ with the edges from v to N, and with the edge xy if $N = \{x, y\}$; it has at most three vertices, so $\mathrm{tw}(K) \leq 2$. Every edge of $Q_{C'} = F_t[C' \cup S_{C'}]$ lies in K or in some Q_C with C merged: edges at v go to N or into some merged C; edges at a vertex of a merged C stay in $C \cup S_C$; and an edge inside $S_{C'}$ exists only if $S_{C'} = N = \{x, y\}$, in which case it is the edge xy of K. The added edge on $S_{C'}$ (if $|S_{C'}| = 2$) is also an edge of K. Hence $Q_{C'}^+$ is a subgraph of the union of K and the graphs Q_C^+ for the merged C. Each Q_C^+ meets K exactly in S_C (the set C is disjoint from $V(H_t) \supseteq \{v\} \cup N$, and $S_C \subseteq \{v\} \cup N$ by the first item), and different Q_C^+ meet only inside K. Each S_C is the vertex v or an edge vu with $u \in N$, which is an edge of both K and Q_C^+. Gluing the graphs Q_C^+ to K one at a time, Lemma 4.3 gives $\mathrm{tw}(Q_{C'}^+) \leq 2$.

The final decomposition. At the end, $H_T = R$ and $F_T = F - D$. Every component C of $(F - D) - V(R)$ induces a graph $F[C] \subseteq Q_C^+$ of treewidth at most 2 on at most n vertices, so $\mathrm{pw}(F[C]) \leq 3 \log_2 n$ by Lemma 4.4. Take a path decomposition $W_1, \ldots, W_s$ of R of width $\mathrm{pw}(R)$ (no bags if R is empty). For each C with $S_C \neq \emptyset$ choose i with $S_C \subseteq W_i$, which is possible because S_C is a vertex or spans an edge of R by (I2). Immediately after W_i insert the bags $W_i \cup P$ for the bags P of a path decomposition of $F[C]$ of minimum width, the components assigned to the same i one after another. Append the path decompositions of the components with $S_C = \emptyset$ at the end. The result is a path decomposition of $F - D$:

- every edge inside $V(R)$ is an edge of R by (I1) and lies in some W_i; every other edge has an end in some C and its other end in $C \cup S_C$, so it lies in an inserted bag $W_i \cup P$ or in a bag of $F[C]$;

- a vertex of C occurs only in the consecutive bags inserted for C;

- a vertex $x \in V(R)$ occurs in a consecutive run of the bags W_i, and every bag inserted after W_i contains W_i; so x occurs in the bags inserted after W_i exactly when $x \in W_i$, and its bags remain consecutive.

Every bag has at most $(\mathrm{pw}(R) + 1) + (3 \log_2 n + 1)$ vertices, so $\mathrm{pw}(F - D) \leq \mathrm{pw}(R) + 3 \log_2 n + 1$; this also holds when R is empty, since then $\mathrm{pw}(R) = -1$. Adding D to every bag gives $\mathrm{pw}(F) \leq |D| + \mathrm{pw}(R) + 3 \log_2 n + 1$, which implies (a).

Part (b) is (I1) and (I2) for $t = T$, since $H_T = R$, $F_T = F - D$, $S_C = N_{F-D}(C)$ and $F[C] \subseteq Q_C^+$. □

A.2 The star-bound certificate

This subsection describes the computer-assisted proof of Proposition 7.5: the triples $\mathsf{P}, \mathsf{R}, \mathsf{S}$ are admissible for $D = \{3, 4, 5\}$ (Definition 7.3). Admissibility asks that for *every* threshold $\tau \in [-3, 3]$

there is a transfer vector t such that, for every type T of degree d, the reference star bound $C_T(\tau)$ plus the transfer term $\lambda_T(t)$ is at most the coefficient a_d. The certificate replaces the continuum of thresholds by finitely many cells. It bounds each reference star bound from above through a lower Riemann sum of explicit lower bounds for the orthant infimum $P_{\min}$ of Section 6. The code also certifies the pair $(F_3, 0.2575)$ for $D = \{3, 4\}$ directly, although its admissibility follows from that of P by restriction to $D = \{3, 4\}$.

Appendices A.2.1 and A.2.2 collect the tools of the star bounds that the certificate uses and prove the enclosures of F_3 and F_4. Appendix A.2.3 proves that the replacement by cells and sums is valid. Appendix A.2.4 describes the rigorous arithmetic, Appendix A.2.5 the search for cells and the results, Appendix A.2.6 how the runs are reproduced, and Appendix A.2.7 what is trusted.

A.2.1 Tools from the star bounds

This subsection collects the lower bounds on $P_{\min}$ for unequal thresholds and the recursion for the simplex orthant that the certificate and Appendix A.2.2 use; the notation is that of Section 6.

Unequal thresholds. When the lower bounds r_i of a star differ, the thresholds $a_i(z) = (\tau - r_i z)/s_i$ of the center bound differ, and the simplex theorem (Lemma 6.5) does not apply directly. The certificate of Appendix A.2 uses the following two lower bounds on the infimum $P_{\min}$. The first holds for every $d \in \{3, 4, 5\}$ and is exact for equal thresholds; the second is specific to $d = 3$ and is much sharper for unequal thresholds. For $d \in \{3, 4, 5\}$ and $a \in \mathbb{R}^d$ put

$$\Lambda_d(a) := \max\Big\{0,\ 1 - \sum_i \Phi(a_i),\ \max_{1 \le j \le d}\Big(S_d(a_j) - \sum_{i:\, a_i > a_j} \big(\Phi(a_i) - \Phi(a_j)\big)\Big)\Big\}.$$

Lemma A.1 (simplex truncation)**.** *Let $d \in \{3, 4, 5\}$ and $a \in \mathbb{R}^d$. Then $P^{(d)}_{\min}(a) \ge \Lambda_d(a)$.*

Proof. Let U be a centered Gaussian vector in $\mathbb{R}^d$ with unit variances. The probability $\mathbb{P}(U_i > a_i\ \forall i)$ is nonnegative, and it is at least $1 - \sum_i \Phi(a_i)$ by the union bound $\mathbb{P}(\exists i :\ U_i \le a_i) \le \sum_i \Phi(a_i)$. Let $\lambda \in \mathbb{R}$. If $U_i > \lambda$ for all i but $U_k \le a_k$ for some k, then $\lambda < U_k \le a_k$, so in particular $a_k > \lambda$. Hence

$$\{U_i > \lambda\ \forall i\} \ \subseteq\ \{U_i > a_i\ \forall i\} \ \cup \bigcup_{k:\, a_k > \lambda} \{\lambda < U_k \le a_k\},$$

and $\mathbb{P}(\lambda < U_k \le a_k) = \Phi(a_k) - \Phi(\lambda)$. With $\mathbb{P}(U_i > \lambda\ \forall i) \ge S_d(\lambda)$ from Lemma 6.5, this gives

$$\mathbb{P}(U_i > a_i\ \forall i) \ \ge\ S_d(\lambda) - \sum_{i:\, a_i > \lambda} \big(\Phi(a_i) - \Phi(\lambda)\big).$$

Taking $\lambda = a_j$ for each j and then the infimum over U gives $P^{(d)}_{\min}(a) \ge \Lambda_d(a)$. □

For equal thresholds the bound is exact: every sum over $\{i : a_i > a_j\}$ is empty, and $S_d(a) \ge \max\{0, 1 - d\,\Phi(a)\}$ by the union bound applied to U^*, so $\Lambda_d(a, \dots, a) = S_d(a) = P^{(d)}_{\min}(a, \dots, a)$ by Lemma 6.5.

For three variables, the rank reduction of Lemma 6.4, which allows unequal thresholds, leads to a planar picture that can be counted exactly. The following bound is used for the stars of a degree-3 vertex whose neighbors do not all have the same degree.

Lemma A.2 (rank reduction and arc count, $d = 3$)**.** *Let $a \in \mathbb{R}^3$. For $\rho > 0$ put $\gamma_i(\rho) = \arccos \operatorname{clip}(-a_i/\rho, -1, 1) \in [0, \pi]$, where $\operatorname{clip}(x, -1, 1) = \max\{-1, \min\{1, x\}\}$. Then for all $0 \le \rho_a \le \rho_b$,*

$$P^{(3)}_{\min}(a) \;\ge\; \int_{\rho_a}^{\rho_b} \Big(1 - \sum_{i=1}^{3} \frac{\gamma_i(\rho)}{\pi}\Big) \rho\, e^{-\rho^2/2}\, d\rho.$$

Proof. Let U be a centered Gaussian vector in $\mathbb{R}^3$ with unit variances and correlation matrix C.

Rank reduction. By Lemma 6.4 it suffices to bound $\mathbb{P}(U_i > a_i\ \forall i)$ from below when C is singular, that is, $\operatorname{rank} C \le 2$. Then U has the distribution of $(\langle v_i, G\rangle)_i$ for unit vectors $v_1, v_2, v_3 \in \mathbb{R}^2$ and a standard Gaussian vector G in $\mathbb{R}^2$.

Arc count. Write $G = \rho(\cos\vartheta, \sin\vartheta)$, so ρ has density $\rho e^{-\rho^2/2}$ on $(0, \infty)$ and ϑ is uniform and independent of ρ. Given $\rho > 0$, the event $\{U_i \le a_i\}$ says $\langle -v_i, (\cos\vartheta, \sin\vartheta)\rangle \ge -a_i/\rho$. If $-a_i/\rho > 1$ this set of directions is empty, if $-a_i/\rho < -1$ it is the whole circle, and otherwise it is a closed arc of half-angle $\arccos(-a_i/\rho)$ around the direction of $-v_i$. In all three cases its normalized length is $\gamma_i(\rho)/\pi$ (up to a null set). By the union bound, the conditional probability of $\{U_i > a_i\ \forall i\}$ given ρ is at least $\max\{0, 1 - \sum_i \gamma_i(\rho)/\pi\}$. Integrating against the density of ρ, and dropping the nonnegative contribution of $\rho \notin [\rho_a, \rho_b]$,

$$\mathbb{P}(U_i > a_i\ \forall i) \;\ge\; \int_{\rho_a}^{\rho_b} \Big(1 - \sum_{i} \frac{\gamma_i(\rho)}{\pi}\Big) \rho e^{-\rho^2/2}\, d\rho.$$

The right side does not depend on U, so it also bounds the infimum $P_{\min}(a)$. □

The interval $[\rho_a, \rho_b]$ is a free parameter; any choice gives a valid bound. The certificate takes $\rho_a = \max\{0, a_1, a_2, a_3\}$, below which some $\gamma_i = \pi$ and the integrand is negative, and ρ_b close to the first zero of the integrand beyond ρ_a. The integral has a closed form in terms of Owen's T function (Lemma A.8).

Evaluating the simplex orthant. The functions S_d are evaluated through the reflection $T_d(h) = S_d(-h)$, the distribution function of the maximum of the regular simplex vector. It satisfies a one-dimensional recursion in d, which is the basis of all numerical evaluations of S_3, S_4, S_5 in this paper.

Lemma A.3 (simplex recursion)**.** *For $h \ge 0$, $T_2(h) = 2\Phi(h) - 1$, and for $d \ge 3$,*

$$T_d(h) = \int_0^h d\,\varphi(x)\, T_{d-1}\Big(\sqrt{\tfrac{d}{d-2}}\, x\Big)\, dx;$$

for $h \le 0$, $T_d(h) = 0$. In particular each T_d is continuous and nondecreasing, and $S_d(a) = T_d(-a)$ is continuous and nonincreasing.

Proof. Let $X = U^*$ be the regular simplex vector in d variables. For $d = 2$ we have $X_2 = -X_1$ (the correlation is -1), so $\max_i X_i = |X_1|$ and $T_2(h) = \mathbb{P}(|X_1| < h) = 2\Phi(h) - 1$ for $h \ge 0$. The case $h \le 0$ is (6.4).

Let $d \ge 3$ and $h \ge 0$, and put $\rho = -\frac{1}{d-1}$. For $j \ne k$, $\operatorname{Var}(X_j - X_k) = 2 - 2\rho > 0$, so the maximum is attained at a unique index almost surely. By symmetry,

$$T_d(h) = \sum_{i=1}^{d} \mathbb{P}(X_i < h,\ X_j \le X_i\ \forall j \ne i) = d\,\mathbb{P}(X_1 < h,\ X_j \le X_1\ \forall j \ge 2).$$

Write $X_j = \rho X_1 + R_j$ for $j \ge 2$. The residual vector $(R_2, \dots, R_d)$ is centered Gaussian and independent of X_1, with $\operatorname{Var} R_j = 1-\rho^2$ and $\operatorname{Cov}(R_j, R_k) = \rho - \rho^2$ for $j \ne k$; its correlations are $(\rho-\rho^2)/(1-\rho^2) = \rho/(1+\rho) = -\frac{1}{d-2}$. So $\tilde R = R/\sqrt{1-\rho^2}$ is the regular simplex vector in $d-1$ variables. The condition $X_j \le X_1$ reads $R_j \le (1-\rho)X_1$, that is, $\tilde R_j \le X_1\sqrt{(1-\rho)/(1+\rho)} = X_1\sqrt{d/(d-2)}$. By the independence of X_1 and $\tilde R$ and Fubini's theorem,

$$T_d(h) = d\int_{-\infty}^{h} \varphi(x)\,\mathbb{P}\Big(\max_j \tilde R_j \le \sqrt{\tfrac{d}{d-2}}\,x\Big)dx = \int_{-\infty}^{h} d\,\varphi(x)\,T_{d-1}\Big(\sqrt{\tfrac{d}{d-2}}\,x\Big)dx,$$

using that the maximum of $\tilde R$ has a continuous distribution. The integrand vanishes for $x \le 0$ by (6.4), which gives the formula. Continuity and monotonicity follow by induction on d. □

A.2.2 The constants F_3 and F_4

The constants F_3 and F_4 are evaluated with explicit one-dimensional integral formulas, which we now record. For F_3 we use the geometry of the planar regular simplex; for F_4 we use the recursion of Lemma A.3.

Lemma A.4 (formula for the ideal 3-star)**.** *Let $r \in (0,1)$, $s = s_r$ and $\tau \in \mathbb{R}$. Put $z_\ell = \max\{\tau, \tau/r\}$ and, for $\vartheta \in [0, \pi/3]$, $\sigma = s\cos\vartheta$ and $q = \sigma^2 + r^2$. Then $P_{3,r}(\tau) := \mathbb{P}(Y_k^* > \tau \text{ for } 0 \le k \le 3)$ equals*

$$P_{3,r}(\tau) = \Phi(-z_\ell) - \frac{3}{\pi}\int_0^{\pi/3} \frac{\sigma}{\sqrt{q}}\, e^{-\tau^2/(2q)}\,\Phi\Big(-\frac{\sqrt{q}}{\sigma}\Big(z_\ell - \frac{r\tau}{q}\Big)\Big)d\vartheta,$$

and $f_{3,r}(\tau) = \Phi(-\tau) - P_{3,r}(\tau)$. Moreover $P_{3,r}$ is nonincreasing.

Proof. Given $Z = z$, all Y_i^* exceed τ iff $\langle -w_i, G\rangle < h := (rz-\tau)/s$ for $i = 1,2,3$. Since $\sum_i \langle -w_i, G\rangle = 0$, this is impossible if $h \le 0$. For $h > 0$ it says that G lies in the open equilateral triangle T_h with inradius h and outer unit normals $-w_1, -w_2, -w_3$. In polar coordinates, the ray in a direction at angle $\vartheta \in [-\pi/3, \pi/3]$ from the nearest outer normal leaves T_h at distance $h/\cos\vartheta$, and $\mathbb{P}(|G| < x) = 1 - e^{-x^2/2}$; by the six-fold symmetry,

$$\mu(h) := \mathbb{P}(G \in T_h) = \frac{3}{\pi}\int_0^{\pi/3}\Big(1 - e^{-h^2/(2\cos^2\vartheta)}\Big)d\vartheta.$$

The conditions $z > \tau$ (for Y_0^*) and $h > 0$ together say $z > z_\ell$. By Tonelli's theorem,

$$P_{3,r}(\tau) = \int_{z_\ell}^{\infty} \varphi(z)\,\mu\Big(\frac{rz-\tau}{s}\Big)dz = \Phi(-z_\ell) - \frac{3}{\pi}\int_0^{\pi/3}\int_{z_\ell}^{\infty} \varphi(z)\, e^{-(rz-\tau)^2/(2\sigma^2)}\,dz\,d\vartheta.$$

Completing the square, $z^2 + (rz-\tau)^2/\sigma^2 = (q/\sigma^2)(z - r\tau/q)^2 + \tau^2/q$, because $q - r^2 = \sigma^2$. Hence the inner integral equals $(\sigma/\sqrt{q})\, e^{-\tau^2/(2q)}\,\Phi(-(\sqrt{q}/\sigma)(z_\ell - r\tau/q))$, which is the formula. The identity for $f_{3,r}$ is (6.1), and $P_{3,r}$ is nonincreasing because the event shrinks as τ grows. □

Lemma A.5 (formula for the ideal 4-star)**.** *Let $r = r_{44} = \frac{\sqrt{3}}{2}$ and $s = s_r = \frac{1}{2}$. Put $g(x) = 4\varphi(x)\,T_3(\sqrt{2}\,x)$ for $x \ge 0$ and $x_0(\tau) = \max\{0, (r-1)\tau/s\}$. Then $P_4(\tau) := \mathbb{P}(Y_k^* > \tau \text{ for } 0 \le k \le 4)$ satisfies*

$$P_4(\tau) = \Phi(-\tau)\,T_4(x_0(\tau)) + \int_{x_0(\tau)}^{\infty} g(x)\,\Phi\Big(-\frac{\tau + sx}{r}\Big)dx,$$

$f_{4,r} = \Phi(-\tau) - P_4(\tau)$ is continuously differentiable, and

$$f_{4,r}'(\tau) = -\varphi(\tau) + \varphi(\tau)\,T_4(x_0(\tau)) + \frac{1}{r}\int_{x_0(\tau)}^{\infty} g(x)\,\varphi\Big(\frac{\tau + sx}{r}\Big)dx.$$

Proof. By Lemma A.3, $T_4(h) = \int_0^h g(x)\,dx$ for $h \ge 0$ and $T_4 = 0$ on $(-\infty, 0]$; recall $T_4(h) = S_4(-h)$. Conditioning on Z as in (6.5),

$$P_4(\tau) = \int_\tau^\infty \varphi(z)\, T_4\Big(\frac{rz-\tau}{s}\Big)dz = \int_0^\infty g(x) \int \varphi(z)\, \mathbf{1}\Big\{z > \tau,\ z > \frac{\tau + sx}{r}\Big\}\, dz\, dx$$

by Tonelli's theorem, since $T_4((rz-\tau)/s) = \int_0^\infty g(x)\mathbf{1}\{x < (rz-\tau)/s\}dx$. The inner integral is $\Phi(-\max\{\tau, (\tau+sx)/r\})$, and $(\tau+sx)/r \ge \tau$ exactly when $x \ge (r-1)\tau/s$. Splitting the outer integral at $x_0(\tau)$ gives the formula. For the derivative, differentiate under the integral sign (the integrand and its τ-derivative are dominated by g and $g\,\varphi(0)/r$). When $x_0(\tau) > 0$, the two terms coming from the moving limit x_0 cancel, because $(\tau + sx_0)/r = \tau$ there; when $x_0 = 0$ they are absent, and $T_4(0) = 0$. This gives $P_4'(\tau) = -\varphi(\tau)T_4(x_0) - \frac{1}{r}\int_{x_0}^\infty g(x)\varphi((\tau+sx)/r)\,dx$, which is continuous in τ, and $f_{4,r}' = -\varphi - P_4'$. □

Proposition 6.9 (enclosures of F_3 and F_4)**.** *The suprema F_3 and F_4 of Definition 6.7 satisfy*

$$0.14015825233145216369394051401 72 \ \le\ F_3\ \le\ 0.1401582524,$$
$$0.2457719746988\ \le\ F_4\ \le\ 0.24577197475.$$

The proof is computer-assisted. All four bounds are obtained in ball arithmetic with the Arb library [Joh17] (through its Python interface `python-flint` 0.9.0), which encloses every real number computed in an interval with rigorously controlled rounding, and integrates analytic functions with rigorous error bounds. Nothing else is trusted beyond the formulas of Lemmas A.4 and A.5 and the correctness of Arb.

Proof. Each lower bound comes from an enclosure of the ideal star function at one threshold. Each upper bound combines a tail estimate outside a bounded interval of thresholds with a bound on each cell of a finite cover of that interval.

Lower bounds. A supremum is at least any value. Evaluating $f_{3,r_{33}}$ at a point $\hat\tau \approx -0.0681371311$ with the formula of Lemma A.4 (for $\tau \le 0$ it simplifies to $f_{3,r}(\tau) = \frac{3}{\pi}\int_0^{\pi/3} \frac{\sigma}{\sqrt{q}} e^{-\tau^2/(2q)} \Phi(-\tau\frac{q-r}{\sigma\sqrt{q}})\,d\vartheta$), in 200-bit Arb arithmetic, gives an enclosure of radius below 10^{-59} around

$$0.14015825233145216369394051401725\ldots.$$

Evaluating $f_{4,r_{44}}$ at -0.16367 with the formula of Lemma A.5 (with T_3, T_4 computed by nested Arb integrals from Lemma A.3) gives $0.2457719746989 \pm 7.9 \cdot 10^{-14}$.

Upper bound for F_3. Write $f = f_{3,r_{33}}$ and $P = P_{3,r_{33}}$. For $\tau \le -2$, $f(\tau) \le \mathbb{P}(\min_i Y_i^* \le \tau) \le 3\Phi(-2) < 0.069$; for $\tau \ge 2$, $f(\tau) \le \Phi(-2) < 0.023$. On an interval $[\alpha, \beta] \subseteq [-2, 2]$,

$$f(\tau) = \Phi(-\tau) - P(\tau) \le \Phi(-\alpha) - P(\beta) \qquad (\tau \in [\alpha, \beta]),$$

because $\Phi(-\cdot)$ and P are nonincreasing (Lemma A.4). The computation covers $[-2, 0]$ and $[0, 2]$ by 830,610 intervals, obtained by bisection, on each of which the Arb enclosure of $\Phi(-\alpha) - P(\beta)$ lies below 0.1401582524. This proves $F_3 \le 0.1401582524$ without any information on the location of the maximizer.

Upper bound for F_4. Write $f = f_{4,r_{44}}$. For $\tau \ge 0.7$, $f(\tau) \le \Phi(-\tau) \le \Phi(-0.7) < 0.24197$; for $\tau \le -2.5$, $f(\tau) \le \mathbb{P}(\min_i Y_i^* \le \tau) \le 4\Phi(-2.5) < 0.0249$. The interval $[-2.5, 0.7]$ is split at 0 and covered by 97 cells, obtained by bisection from cells of width $\frac{1}{16}$ (the smallest has width 2^{-17}). On a cell $[t_0, t_1]$ with midpoint m, the mean value theorem gives $f(\tau) \le f(m) + \max\{D^+(t_1 - m), D^-(t_0 - m)\}$, where $[D^-, D^+]$ is an Arb enclosure of f' on the cell, computed from Lemma A.5. The integrals over x are cut at $x = 10$; the omitted part is at most $\int_{10}^\infty g \le 4\Phi(-10)$ (respectively $4\varphi(0)\Phi(-10)/r$ for f'), and is added to the enclosure as an interval. Every cell bound lies below 0.24577197475; the largest, $0.2457719747237\ldots$, occurs on the cell of width 2^{-16} around -0.16364. □

A.2.3 Cells and discretization

Two lemmas reduce the check to finitely many computations: one replaces a cell of thresholds by its two endpoints, and the other replaces the integral in the reference star bound by a finite sum. Fix a type T of degree d with reference bounds $r = r_T \in (0,1)^d$ (Definition 7.3), and write $a(z;t) = a(z;t,r)$ for the vector of thresholds $a_i(z;t) = (t - r_i z)/s_i$, $s_i = \sqrt{1-r_i^2}$, of Lemma 6.2 at threshold t. Put

$$J_T(t) := \int_t^\infty \varphi(z)\, P_{\min}(a(z;t))\, dz, \qquad \text{so that} \qquad C_T(t) = \Phi(-t) - J_T(t).$$

The first lemma shows that a bound for a whole cell of thresholds needs only one evaluation of J_T, at the right end of the cell.

Lemma A.6 (cells)**.** *For $t_0 \le \tau \le t_1$, $C_T(\tau) \le \Phi(-t_0) - J_T(t_1)$.*

Proof. First, $\Phi(-\tau) \le \Phi(-t_0)$. Second, J_T is nonincreasing: each threshold $a_i(z;t) = (t - r_i z)/s_i$ is increasing in t, $P_{\min}$ is nonincreasing in each threshold (Lemma 6.3(a)), the integrand is nonnegative, and the domain of integration shrinks as t grows. Hence $J_T(\tau) \ge J_T(t_1)$. □

The bound of Lemma A.6 can exceed $C_T(\tau)$ by a term of first order in the cell width, roughly $(d+2)\varphi(t_0)(t_1 - t_0)$. This is why the certified coefficients are slightly larger than the floating-point optima of the same star bounds (Remark 7.6).

Lemma A.7 (discretization)**.** *For every $t \in \mathbb{R}$ and all points $t = z_0 < z_1 < \cdots < z_N$,*

$$J_T(t) \;\ge\; \sum_{k=0}^{N-1} P_{\min}(a(z_k;t))\big(\Phi(z_{k+1}) - \Phi(z_k)\big).$$

Proof. Since $r_i > 0$, each $a_i(z;t)$ is decreasing in z, so $z \mapsto P_{\min}(a(z;t))$ is nondecreasing (Lemma 6.3(a)). For every $z \in [z_k, z_{k+1}]$, the integrand of $J_T(t)$ is therefore at least $\varphi(z) P_{\min}(a(z_k;t))$. Integrate over each interval $[z_k, z_{k+1}]$ and drop the nonnegative contribution of $z > z_N$. □

In the sum of Lemma A.7, each value $P_{\min}(a(z_k;t))$ is replaced by a computable lower bound $L_T(a(z_k;t))$:

- if all reference bounds of T are equal, that is, all neighbors have the same degree (9 types), the thresholds are equal and $L_T(a,\dots,a) = \max\{0,\ 1 - d\,\Phi(a),\ S_d(a)\} = S_d(a)$, which equals $P_{\min}$ (Lemma 6.5);
- if $d \in \{4,5\}$ and the reference bounds are not all equal, $L_T = \Lambda_d$, the simplex truncation of Lemma A.1;
- if $d = 3$ and the reference bounds are not all equal (seven types), $L_T = \max\{\Lambda_3, A\}$, where A is the arc bound of Lemma A.2 with radii $\rho_a = \max\{0, a_1, a_2, a_3\}$ and $\rho_b \in (\rho_a, 12]$. The radius ρ_b is located by floating-point bisection near the first zero of the integrand beyond ρ_a; any ρ_b gives a valid bound. To reduce the running time, A is evaluated only at the points where a floating-point estimate of it exceeds Λ_3 by more than 10^{-10}; at the other points $L_T = \Lambda_3$. Either way $L_T \le P_{\min}$, by Lemmas A.1 and A.2; the estimate only decides which valid bound is computed.

For the arc bound, the integral over ρ is evaluated in closed form with Owen's function [Owe56]

$$T_{\mathrm{Owen}}(h,x)=\frac{1}{2\pi}\int_0^x \frac{e^{-h^2(1+y^2)/2}}{1+y^2}\,dy \qquad (h,x\in\mathbb{R}).$$

Lemma A.8 (closed form of the arc integral)**.** *Let $|c|\le\rho_1\le\rho_2$. If $c\neq 0$, then*

$$\int_{\rho_1}^{\rho_2}\arccos\frac{c}{\rho}\;\rho e^{-\rho^2/2}\,d\rho=\Big[-e^{-\rho^2/2}\arccos\frac{c}{\rho}\Big]_{\rho_1}^{\rho_2}+\operatorname{sign}(c)\,2\pi\big(T_{\mathrm{Owen}}(|c|,y_2)-T_{\mathrm{Owen}}(|c|,y_1)\big),$$

where $y_j=\sqrt{\rho_j^2-c^2}/|c|$; if $c=0$, the integral equals $\frac{\pi}{2}(e^{-\rho_1^2/2}-e^{-\rho_2^2/2})$. Consequently, for $a\in\mathbb{R}^3$, $c_i=-a_i$ and $\rho_a\ge\max\{0,a_1,a_2,a_3\}$, the bound of Lemma A.2 equals

$$e^{-\rho_a^2/2}-e^{-\rho_b^2/2}-\frac{1}{\pi}\sum_{i=1}^{3}\int_{\max\{\rho_a,|c_i|\}}^{\rho_b}\arccos\frac{c_i}{\rho}\;\rho e^{-\rho^2/2}\,d\rho,$$

where an integral with lower limit above ρ_b is read as 0.

Proof. For $c=0$ the integrand is $\frac{\pi}{2}\rho e^{-\rho^2/2}$. For $c\neq 0$, integrate by parts with $v=-e^{-\rho^2/2}$; since $\frac{d}{d\rho}\arccos\frac{c}{\rho}=\frac{c}{\rho\sqrt{\rho^2-c^2}}$ for $\rho>|c|$ (integrable at $\rho=|c|$), the remaining integral is $\int_{\rho_1}^{\rho_2}e^{-\rho^2/2}\frac{c}{\rho\sqrt{\rho^2-c^2}}\,d\rho$. The substitution $\rho^2=c^2(1+y^2)$, $y\ge 0$, gives $\rho\,d\rho=c^2y\,dy$ and $\sqrt{\rho^2-c^2}=|c|y$, so the integrand becomes $\operatorname{sign}(c)\,e^{-c^2(1+y^2)/2}(1+y^2)^{-1}dy$, whose integral from y_1 to y_2 is $\operatorname{sign}(c)\,2\pi(T_{\mathrm{Owen}}(|c|,y_2)-T_{\mathrm{Owen}}(|c|,y_1))$.

For the second claim, $\int_{\rho_a}^{\rho_b}\rho e^{-\rho^2/2}d\rho=e^{-\rho_a^2/2}-e^{-\rho_b^2/2}$, and $\gamma_i(\rho)=\arccos\operatorname{clip}(c_i/\rho,-1,1)$. If $c_i>0$ and $\rho<c_i$, then $\gamma_i(\rho)=\arccos 1=0$, so the integral over γ_i starts at $\max\{\rho_a,c_i\}$. If $c_i\le 0$, then $|c_i|=\max\{0,a_i\}\le\rho_a$, so on $[\rho_a,\rho_b]$ the clip is inactive. □

A.2.4 Rigorous arithmetic

For each type T and each right end t_1 of a cell, the accompanying code computes a rational number $I_T^{\mathrm{lo}}(t_1)$ such that

$$I_T^{\mathrm{lo}}(t_1)\;\le\;\sum_{k=0}^{N-1}P_{\min}\big(a(z_k;t_1)\big)\big(\Phi(z_{k+1})-\Phi(z_k)\big)\;\le\;J_T(t_1), \tag{A.1}$$

the second inequality being Lemma A.7. The first inequality rests on the following facts about the computation.

(a) *Grids.* All cell endpoints are multiples of 2^{-17} in $[-3,3]$. The points of Lemma A.7 are $z_k=t_1+k\delta_z$, $0\le k\le N$, with $N\delta_z=10$ and $\delta_z=2^{-14}$ ($\delta_z=2^{-13}$ for the seven types that use the arc bound). So every z_k is a multiple of 2^{-17} in $[-3,13]$.

(b) *Masses.* Enclosures of Φ and φ at all multiples of 2^{-17} in $[-8,13]$ are computed once in Arb [Joh17] (80-bit precision) and rounded outward to double precision. The mass $\Phi(z_{k+1})-\Phi(z_k)$ is bounded below by the lower enclosure at z_{k+1} minus the upper enclosure at z_k, rounded down and clamped at 0.

(c) *Values of Φ elsewhere.* At a point a off the grid, with x the grid point below a and $\delta=a-x\in[0,2^{-17})$, Taylor's theorem and $|\varphi'|\le\varphi(1)<0.242$ give $|\Phi(a)-\Phi(x)-\varphi(x)\delta|\le 0.121\,\delta^2$. Upper and lower bounds for $\Phi(a)$ are formed from the grid enclosures accordingly, with outward rounding. Below -8 and above 13 the trivial monotone bounds are used. These bounds were also compared with Arb at 800 random points.

(d) *Thresholds.* The reference bounds r_{ij} and the numbers $s_{ij} = \sqrt{1 - r_{ij}^2}$ are double-precision numbers within $4 \cdot 10^{-16}$ of their Arb values (checked at start-up). Each computed threshold $(t_1 - r_i z_k)/s_i$ is increased by $3 \cdot 10^{-13}$, which exceeds its rounding error for $|z| \le 13$ and $s_i \ge \frac{1}{3}$. So the code evaluates its lower bound at thresholds $a' \ge a(z_k; t_1)$ coordinatewise, and $P_{\min}(a') \le P_{\min}(a(z_k; t_1))$ by monotonicity.

(e) *Simplex tables.* The simplex functions enter through the functions T_d of Lemma A.3, with $S_d(a) = T_d(-a)$. Lower tables $T_d^{\mathrm{lo}} \le T_d$ for $d = 2, 3, 4, 5$ are built on the grid $x_k = k\,2^{-19}$, $0 \le x_k \le 12$. For $d = 2$, T_2^{lo} is the lower Arb enclosure of $2\Phi(x_k) - 1$. For $d \ge 3$, Lemma A.3 writes T_d as an integral of $d\,\varphi(x) T_{d-1}(\sqrt{d/(d-2)}\,x)$. On $[x_j, x_{j+1}]$ this integrand is at least $d\,\varphi(x_{j+1})\,T_{d-1}^{\mathrm{lo}}(g_j)$, where g_j is the largest grid point not exceeding $\sqrt{d/(d-2)}\,x_j$, because φ is decreasing on $[0, \infty)$ and T_{d-1} is nondecreasing. These terms, with $\varphi(x_{j+1})$ replaced by the lower end of its Arb enclosure and the product rounded down, are floored to multiples of 2^{-61} and summed exactly in integers. A lookup of $S_d(a) = T_d(-a)$ rounds $-a$ down to the grid (and uses 0 for $-a \le 0$ and $T_d^{\mathrm{lo}}(12)$ for $-a > 12$), which is valid because T_d is nondecreasing.

(f) *Lower bounds for $P_{\min}$.* The functions Λ_d and S_d are evaluated from the tables and the bounds of (c), each term rounded in the direction that lowers the result. The arc bound is evaluated from the closed form of Lemma A.8 as an Arb ball at 64-bit precision; its inputs a_i, ρ_a and ρ_b are double-precision numbers and hence exact. The exponentials, arccosines and square roots are Arb functions. The Owen term $\int_{y_1}^{y_2} e^{-c^2(1+y^2)/2}(1+y^2)^{-1}\,dy = 2\pi(T_{\mathrm{Owen}}(|c|, y_2) - T_{\mathrm{Owen}}(|c|, y_1))$ is enclosed by Arb's rigorous numerical integration of this integrand, which is analytic off the poles $\pm i$, between the exact midpoints of the balls for y_1 and y_2. The result is then widened by the radii of these balls, which is valid because the integrand is at most 1. For $|c| < 2^{-30}$ the enclosure $[\arctan y_2 - \arctan y_1 - \frac{c^2}{2}(y_2 - y_1),\ \arctan y_2 - \arctan y_1]$, which follows from $1 - u \le e^{-u} \le 1$, is used instead. The lower end of the resulting ball, rounded down, is the value of the arc bound. In the run reported below, the arc bound was enclosed at about $1.2 \cdot 10^8$ points, all with radius below 10^{-13}.

(g) *The sum.* Each product of a lower bound and a mass is rounded down, floored to a multiple of 2^{-62}, and the products are summed exactly in integers.

A.2.5 Charging on cells and results

The main run treats four *problems*, one for each coefficient vector: the pair $(F_3, 0.2575)$ for $D = \{3, 4\}$ and the triples P, R, S for $D = \{3, 4, 5\}$. Each problem is solved by a search for cells with transfer vectors, and its output is checked by a second program.

Cell search. For each problem the interval $[-3, 3]$ starts as 192 cells of width $\frac{1}{32}$. For a cell $[t_0, t_1]$ the code forms, for every type T other than the pure cubic type T_0, the number $\widehat{C}_T = \widehat{\Phi}(-t_0) - I_T^{\mathrm{lo}}(t_1)$, where $\widehat{\Phi}(-t_0)$ is the upper Arb enclosure of $\Phi(-t_0)$ in double precision. By Lemma A.6 and (A.1), $C_T(\tau) \le \widehat{C}_T$ for all $\tau \in [t_0, t_1]$. The type T_0 needs no check, since $C_{T_0} \le F_3 \le a_3$ and $\lambda_{T_0} = 0$. The code then solves, in floating point, the linear program that minimizes the largest violation of the inequalities $\widehat{C}_T + \lambda_T(t) \le a_d$ over transfer vectors $t \in [-1, 1]^3$; we used the HiGHS solver through SciPy [VGO+20]. The solver's output is used only as a candidate: the cell is accepted if, in exact rational arithmetic, every inequality holds for this t with slack at least 10^{-12}. Otherwise the cell is bisected.

Exact recheck. A second, short program reads only the stored cells, transfers and numbers $I_T^{\mathrm{lo}}(t_1)$. It checks in exact rational arithmetic that the cells tile $[-3, 3]$, that $|t_{ij}| \le 1$, and that every inequality $\widehat{\Phi}(-t_0) - I_T^{\mathrm{lo}}(t_1) + \lambda_T(t) < a_d$ holds. Here $\widehat{\Phi}(-t_0)$ is a fresh Arb upper bound

for $\Phi(-t_0)$, computed at 100-bit precision and rounded up to a multiple of 10^{-15}. For the pair and for P, whose coefficient at degree 3 is F_3 itself, the degree-3 inequalities are checked with $a_3 = 0.140158252331452163693940514 0172$, a lower bound for F_3 (Proposition 6.9), so that they prove $\widehat{C}_T + \lambda_T(t) \le F_3$. For R and S, the omission of T_0 uses $F_3 \le 0.1401582524 < a_3$.

Results. All four problems are certified: in each, the accepted cells tile $[-3, 3]$, no cell failed, and the recheck passes. The smallest slack is the minimum over all cells and all checked types of $a_d - \widehat{C}_T - \lambda_T(t)$, as computed by the recheck. The column "types" gives the number of types checked (all D-types except T_0), and "min. width" the smallest cell width.

Problem	coefficients	types	cells	min. width	min. slack
pair	(F_3, 0.2575)	8	890	2^{-13}	$2.76 \cdot 10^{-7}$
P	(F_3, 0.2575, 0.3667)	45	977	2^{-13}	$1.04 \cdot 10^{-7}$
R	(0.1455, 0.2530, 0.3600)	45	375	2^{-10}	$7.09 \cdot 10^{-6}$
S	($\frac{5}{32}$, $\frac{1}{4}$, $\frac{11}{32}$)	45	1012	2^{-14}	$1.24 \cdot 10^{-7}$

The slack is a property of the cells after bisection stopped, not the distance of a coefficient vector from the boundary of the admissible set. Smaller coefficient vectors were also tried. For example, $(0.1444, 0.2514, 0.3583)$ in place of R could not be certified: cells still failed at width 2^{-17}, so this vector is at or beyond the reach of this certificate.

A fifth problem, used for Theorem 8.2, is run separately by the same program. It is the pair $(c_3, 0.2627)$ for $D = \{3, 4\}$ on $[0, \frac{35}{32}]$ only, with c_3 replaced by its rational lower bound 0.139822124008. The type T_0 is omitted, because at nonnegative thresholds Proposition 8.8 bounds it instead. The search starts from 35 cells of width $\frac{1}{32}$ and certifies the problem with 62 cells, the smallest of width 2^{-8}, and smallest slack $1.06 \cdot 10^{-4}$ (Proposition A.13). The exact recheck covers this problem as well.

A.2.6 Reproducibility

A single driver script in the accompanying code reproduces every computation used in the proofs of Theorems 7.1 and 7.2, and hence of Theorem 1.3 and Corollary 1.4, as well as the charging certificate of Theorem 8.2. The switch-lemma certificates of Appendix A.3.2 have their own programs and are not part of this script. The driver has two modes for these computations:

- *quick* (about one minute): the exact recheck of the stored cells for all four problems, and exact rational checks of the coefficient arithmetic used elsewhere in the paper;
- *full* (about 75 minutes on 16 cores): in addition, the Arb computations of the lower and upper bounds of F_3 and of the enclosure of F_4 (Proposition 6.9), and the certificate computation itself, including the fifth problem of Proposition A.13 (about 4 minutes). The certificate computation rewrites the stored cell files, and a second exact recheck follows.

The runs reported here used CPython 3.14.7, `python-flint` 0.9.0, NumPy 2.5.3, SciPy 1.18.1 and mpmath 1.4.1 on a 16-core Apple M3 Max computer. The certificate computation took 3661 s with 14 worker processes, the enclosure of F_4 475 s, the global upper bound of F_3 189 s, and each exact recheck about one second. Repeated runs reproduced the stored certificate data (the cells, transfers and I^{lo} values of the four problems, and the cells of the F_4 enclosure) bit for bit, as verified by SHA-256 checksums.

A.2.7 What is trusted

The certificate relies on the following software, data, outputs and mathematical inputs.

- *Arb and exact arithmetic.* The computation rests on the correctness of Arb/FLINT, including its rigorous numerical integration, and of Python's exact rational arithmetic, and on the derivations of Section 6 and of this subsection. Floating point is used only for choices. The linear program for the transfers is solved in floating point, but its output enters only through the exact check. The radius ρ_b and the points at which the arc bound is evaluated are chosen in floating point, and every choice gives a valid lower bound. No floating-point value of a special function enters an inequality.
- *Stored data.* The exact recheck uses the stored numbers $I_T^{\mathrm{lo}}(t_1)$, which only the main certificate program computes.
- *Reporting.* The Arb programs for F_3 and F_4 report their enclosures in their output; the driver checks only exit codes, so their outputs must be read.
- *Mathematical inputs outside the computation.* The certificate uses the simplex theorem (Lemma 6.5) through S_d and Λ_d. For $d = 4$ and $d = 5$ this rests on Mulgund's theorem [Mul26], whose status Remark 6.6 describes.

A.3 Maximum degree 4: proof of Theorem 8.2 and its certificates

This subsection proves Theorem 8.2, which we restate below (Appendix A.3.1). The charging at nonzero thresholds that the proof uses (Proposition A.13) is computed by the programs of Appendix A.2. The remaining parts, Appendices A.3.2 to A.3.5, describe the computer-assisted part of the proof of Lemma A.10.

Theorem 8.2 (maximum degree 4)**.** *Every graph G with $\Delta(G) \le 4$ satisfies* $\mathrm{pw}(G) \le c_3\, n_3 + 0.2627\, n_4 + o(n)$.

A.3.1 Proof of Theorem 8.2

The proof of Theorem 8.1 has three ingredients: the layout of Lemma 8.4, the bound $\kappa(r)$ at nonzero thresholds (Proposition 8.8), and the switch lemma at threshold zero (Lemma 8.9). For maximum degree 4 we keep the layout, with a larger set at threshold zero, and use the mixed labels of Definition 5.2(b), whose edge correlations depend on the degrees of the two ends. A vertex of degree 3 with a neighbor of degree 4 then has weaker correlation bounds than one in a cubic region, and the bound on its expected contribution to the frontiers exceeds c_3. As in Section 7, the excess is sent to the neighbors of degree 4, and the coefficient 0.2627 must absorb it.

- At threshold zero, the expected contribution of a vertex is a switch-lemma quantity for its star, whose correlation bounds may be unequal. For a vertex of degree 3 whose neighbors all have the same degree, the bounds are equal and Lemma 8.9 applies. For the other types, two of degree 3 and five of degree 4, we prove explicit bounds with a computer-assisted certificate (Lemma A.10). Each vertex of degree 3 sends the same amount $s \approx 0.01308$ to each neighbor of degree 4.
- At nonzero thresholds, the frontier probabilities are bounded by the star bounds of Section 7, except for vertices of degree 3 all of whose neighbors have degree 3, which are bounded by c_3 through Proposition 8.8. The charging is certified as in Proposition 7.5, now only on $[0, \frac{35}{32}]$, with c_3 in place of F_3 (Proposition A.13).

Throughout, $c_3 = \frac{3}{4\pi}\arccos\frac{5}{6}$, the numbers r_{ij} are those of (5.3), and s is the excess of the bound of Lemma A.10(a) over c_3:

$$s := 0.1529003 - c_3 \approx 0.0130782.$$

The layout. For a graph G with $\Delta(G) \le 4$ and a labeling y, let $U = \partial^- S_0$ and $W = \partial^+ S_0$ as in Section 8.1, and let $\hat{D}$ be the set of vertices with at least three neighbors on the other side of zero, that is, vertices v with $y_v \le 0$ and at least three neighbors above zero, or $y_v > 0$ and at least three neighbors at or below zero. For $\Delta(G) \le 3$ this is the set D of Section 8.1. A vertex of $\hat{D}$ with $y_v \le 0$ lies in U, and one with $y_v > 0$ lies in W, so $\hat{D} \subseteq U \cup W$. Put

$$\hat{L}_0 = |U| + |\hat{D} \cap W|, \qquad \hat{L}'_0 = |W| + |\hat{D} \cap U|.$$

Lemma A.9 (median-switched layout, maximum degree 4)**.** *Lemma 8.4 holds for graphs G with $\Delta(G) \le 4$, with $\hat{L}_0$ and $\hat{L}'_0$ in place of L_0 and L'_0.*

Proof. The proofs of Lemmas 8.4 and 8.5 use the degree bound at one point only: in the first claim of Lemma 8.5, that every vertex $u \in U \setminus D$ has at most two neighbors above zero, so that Lemma 8.3 applies to the items $u \in P$ with $e(u) = N(u) \cap W'$. For $u \in U \setminus \hat{D}$ this holds by the definition of $\hat{D}$. The rest of the two proofs uses only that $\hat{D} \subseteq U \cup W$, the identities $\hat{L}_0 = |U \setminus \hat{D}| + |\hat{D}|$ and $\hat{L}'_0 = |W \setminus \hat{D}| + |\hat{D}|$ (which follow as in (8.2)), and the placement of $\hat{D}$ in every connector bag $C_0, \dots, C_k$ and of $\hat{D} \cap W$ in C_{-1}. In particular, a crossing edge uw with $y_u \le 0 < y_w$ lies in C_{-1} if $w \in \hat{D}$, in C_k if $u \in \hat{D}$, and otherwise in the bag C_j with $w = q_j \in e(u)$; none of these steps uses the degree of u or w. □

Threshold zero. For a centered Gaussian vector $(Y_0, Y_1, \dots, Y_d)$ with unit variances let $N = \#\{i \ge 1 : Y_i > 0\}$ and

$$\Psi(Y_0; Y_1, \dots, Y_d) = \mathbb{E}\big[\mathbf{1}\{Y_0 \le 0\}\,\lceil N/2 \rceil\big].$$

For $d \le 4$ we have $\lceil N/2 \rceil = \mathbf{1}\{N \ge 1\} + \mathbf{1}\{N \ge 3\}$. If Y is a labeling with the sign symmetry of Section 8.1 and continuous marginals, then for a vertex v with neighbors $x_1, \dots, x_d$, $d \le 4$,

$$\begin{aligned} \mathbb{P}(v \in U) + \mathbb{P}(v \in \hat{D} \cap W) &= \mathbb{P}(Y_v \le 0 < \max_i Y_{x_i}) + \mathbb{P}(Y_v > 0,\ \#\{i : Y_{x_i} \le 0\} \ge 3) \\ &= \Psi(Y_v; Y_{x_1}, \dots, Y_{x_d}), \end{aligned} \tag{A.2}$$

because the second probability equals $\mathbb{P}(Y_v < 0,\ N \ge 3)$ by sign symmetry. For $d \le 3$ the left side of Lemma 8.9 is the same quantity, so Lemma 8.9 states $\Psi \le \kappa(r)$ when all incident correlations are at least r.

Lemma A.10 (switch lemma, unequal bounds)**.** *Let $(Y_0, Y_1, \dots, Y_d)$ be a centered Gaussian vector with unit variances and $\mathrm{Corr}(Y_0, Y_i) \ge r_i$ for $1 \le i \le d$.*

(a) *If $d = 3$ and (r_1, r_2, r_3) is a permutation of (r_{33}, r_{33}, r_{34}), then $\Psi \le 0.1529003$.*

(b) *If $d = 3$ and (r_1, r_2, r_3) is a permutation of (r_{33}, r_{34}, r_{34}), then $\Psi \le 0.165616$.*

(c) *If $d = 4$, $0 \le j \le 4$, and j of the bounds r_i equal r_{34} and the other $4 - j$ equal r_{44}, then $\Psi \le 0.2627 - j\,s$.*

The numerical maxima of Ψ in the five patterns of (c) are 0.250000, 0.245466, 0.235687, 0.221375 and 0.208185 for $j = 0, \dots, 4$, and those in (a) and (b) are 0.1528953 and 0.1656115. So the bound in (a) is within $5 \cdot 10^{-6}$ of the numerical maximum, and (c) is tight to within 10^{-3} at $j = 2$. Unlike in Lemma 8.9, the maximizers are not the regular simplex in general; for $j = 0$, for instance, three equal residuals and a fourth opposite give $\Psi = \frac{1}{4}$.

Proof. Reduction to equality in the incident correlations. Write $Z = Y_0$ and $a_i = \mathrm{Corr}(Z, Y_i) \geq r_i$. For $a_i < 1$ put $W_i = (Y_i - a_i Z)/s_{a_i}$, which is standard normal and independent of Z; for $a_i = 1$ let W_i be an auxiliary standard normal variable independent of everything else. Then $W = (W_1, \ldots, W_d)$ is a centered Gaussian vector with unit variances, independent of Z. On $\{Z < 0\}$ and for $a_i < 1$, $Y_i > 0$ holds exactly when $W_i > a_i|Z|/s_{a_i}$; since $a \mapsto a/s_a$ is increasing, this implies $Y_i' := r_i Z + s_{r_i} W_i > 0$. For $a_i = 1$, $Y_i = Z < 0$ is never positive on $\{Z < 0\}$. Hence on $\{Z < 0\}$ the number N for Y is at most the number N' for $Y' = (Z, Y_1', \ldots, Y_d')$, and since $\lceil \cdot/2 \rceil$ is nondecreasing and $\{Z = 0\}$ is null, $\Psi(Y) \leq \Psi(Y')$. The vector Y' has incident correlations exactly r_i, and its residuals have the correlation matrix C of W. So it suffices to bound, for each pattern, the function

$$G_r(C) = \Psi\big(Z;\, r_1 Z + s_{r_1} W_1, \ldots, r_d Z + s_{r_d} W_d\big), \qquad W \sim N(0, C) \text{ independent of } Z \sim N(0,1), \tag{A.3}$$

over the compact convex set E_d of $d \times d$ correlation matrices C.

The bounds on E_d. These are proved by computer-assisted certificates in Arb ball arithmetic, described in Appendices A.3.2 to A.3.5. Each certificate covers a parametrization of E_d by boxes. On each box, either G_r is bounded through exact formulas for its value and gradient (Plackett's identity), or the box is shown to miss E_d. Every bound is a strict inequality in Arb. □

The labels satisfy the bounds of Lemmas A.10 and 8.9 only up to a factor $1 - \varepsilon_R$. For $d \leq 4$ and $r \in (0,1)^d$ let $M_d(r)$ be the supremum of Ψ over all centered Gaussian vectors $(Y_0, \ldots, Y_d)$ with unit variances and $\mathrm{Corr}(Y_0, Y_i) \geq r_i$ for all i. The next lemma shows that lowering the bounds by a factor $1 - \gamma$ raises M_d only slightly.

Lemma A.11 (lowering the bounds)**.** *For $d \leq 4$, $r \in (0,1)^d$ and $0 < \gamma < 1$, $M_d((1-\gamma)r) \leq M_d(r) + 3d\,\gamma^{1/3}$.*

Proof. Let $(Y_0, \ldots, Y_d)$ be centered Gaussian with unit variances and $a_i = \mathrm{Corr}(Y_0, Y_i) \geq (1-\gamma) r_i$. Write $Y_i = \langle u_i, X \rangle$ with unit vectors $u_0, \ldots, u_d \in \mathbb{R}^{d+1}$ and X standard Gaussian. If $a_i \geq r_i$ put $u_i' = u_i$. Otherwise $0 < a_i < r_i < 1$, the vector $w_i = u_i - a_i u_0$ has $|w_i| = s_{a_i} > 0$, and $u_i' = r_i u_0 + s_{r_i} w_i/|w_i|$ is a unit vector with $\langle u_0, u_i' \rangle = r_i$ and

$$|u_i - u_i'|^2 = 2 - 2(a_i r_i + s_{a_i} s_{r_i}) \leq 2 - 2(a_i r_i + s_{r_i}^2) = 2 r_i (r_i - a_i) \leq 2\gamma.$$

Put $Y_i' = \langle u_i', X \rangle$; the vector $(Y_0, Y_1', \ldots, Y_d')$ is one of the vectors in the supremum that defines $M_d(r)$. Let $t = \gamma^{1/3}$. The events $\{Y_i > 0\}$ and $\{Y_i' > 0\}$ differ only if $|Y_i| \leq t$ or $|Y_i - Y_i'| > t$. Since Y_i is standard normal and $Y_i - Y_i'$ is centered normal with variance at most 2γ, this has probability at most $2t\varphi(0) + 2\gamma/t^2 < 3\gamma^{1/3}$. Changing the signs of k neighbors changes $\lceil N/2 \rceil$ by at most k, so $\Psi(Y_0; Y_1, \ldots, Y_d) \leq \Psi(Y_0; Y_1', \ldots, Y_d') + 3d\gamma^{1/3} \leq M_d(r) + 3d\gamma^{1/3}$. □

Setting. For the rest of this subsection H is a graph with $\delta(H) \geq 3$ and $\Delta(H) \leq 4$, with N vertices, and $\beta(H) := c_3 n_3(H) + 0.2627\, n_4(H)$. The labels Y are the mixed labels of Definition 5.2(b) with parameter R. Properties (L1) and (L2) of Section 7 hold, since $\Delta(H) \leq 5$. We use the $\{3,4\}$-types, reference bounds r_T and transfer terms λ_T of Definition 7.3, with the transfer vector written $t = (t_{34}, t_{35}, t_{45})$; only t_{34} matters here. By Lemma 5.6(i), $-Y$ has the same distribution as Y, so the sign symmetry of Section 8.1 holds.

Lemma A.12 (threshold zero, maximum degree 4)**.** *For every $\eta > 0$ there is R_0 such that for every $R \geq R_0$ and every graph H with $\delta(H) \geq 3$ and $\Delta(H) \leq 4$ on N vertices,*

$$\mathbb{E}\hat{L}_0 \leq \beta(H) + \eta N \qquad \text{and} \qquad \mathbb{E}\hat{L}_0' \leq \beta(H) + \eta N.$$

Proof. We bound the expected contribution of a vertex to $\hat{L}_0$ for each type, sum these bounds with Lemma 7.4, and obtain $\hat{L}_0'$ by sign symmetry. Let v have type T, with k neighbors of degree 4 if $d(v) = 3$ and j neighbors of degree 3 if $d(v) = 4$. By (A.2), (L1) and the definition of M_d, $\mathbb{P}(v \in U) + \mathbb{P}(v \in \hat{D} \cap W) \le M_{d(v)}((1-\varepsilon_R) r_T)$. By Lemma 8.9 (applied with $r = r_{33}$ when $k = 0$, and with $r = r_{34}$ when $k = 3$) and by Lemma A.10, $M_{d(v)}(r_T) \le \beta_T$, where

type	$d=3,\, k=0$	$k=1$	$k=2$	$k=3$	$d=4,\, j$
β_T	$\kappa(r_{33}) = c_3$	$c_3 + s$	0.165616	$\kappa(r_{34}) < 0.1778364$	$0.2627 - j\,s$
$\beta_T + \lambda_T(s,0,0)$	c_3	c_3	< 0.1394597	< 0.1386019	0.2627

with $\lambda_T(s,0,0) = -k\,s$ for $d = 3$ and $j\,s$ for $d = 4$. The second row is at most $a_{d(v)}$ for $(a_3, a_4) = (c_3, 0.2627)$, because $c_3 > 0.1398221$. There are nine types, and by Lemma A.11 and $\varepsilon_R \to 0$ there is R_0 such that $M_d((1-\varepsilon_R) r_T) \le \beta_T + 12\,\varepsilon_R^{1/3} \le \beta_T + \eta$ for every type T and every $R \ge R_0$. Hence $\mathbb{P}(v \in U) + \mathbb{P}(v \in \hat{D} \cap W) + \lambda_{T(v)}(s,0,0) \le a_{d(v)} + \eta$ for every vertex v, and Lemma 7.4 with $D = \{3,4\}$ and $e = \eta$ gives the bound on $\mathbb{E}\hat{L}_0 = \sum_v (\mathbb{P}(v \in U) + \mathbb{P}(v \in \hat{D} \cap W))$. The events $\{v \in W\}$ and $\{v \in \hat{D} \cap U\}$ are, up to null events, the events $\{v \in U\}$ and $\{v \in \hat{D} \cap W\}$ for the labeling $-Y$, which has the same distribution; so the same bound holds for $\mathbb{E}\hat{L}_0'$. □

Nonzero thresholds. Away from zero the frontier events are those of Section 7. Only thresholds $\tau \ge 0$ are needed, since negative ones follow by sign symmetry. At these thresholds the pure cubic type $T_0 = (3; 3, 0, 0)$ is bounded by c_3 through Proposition 8.8 instead of by F_3, and above $\frac{35}{32}$ no charging is needed at all.

Proposition A.13 (certified charging at nonnegative thresholds)**.** *Let $(a_3, a_4) = (c_3, 0.2627)$. For every $\tau \in [0, \frac{35}{32}]$ there is a transfer vector t with $C_T(\tau) + \lambda_T(t) \le a_d$ for every $\{3,4\}$-type $T = (d; k_3, k_4, 0)$ other than T_0.*

Proof. This is a further problem of the certificate of Proposition 7.5, computed and rechecked by the same programs (Appendix A.2), with $[0, \frac{35}{32}]$ in place of $[-3, 3]$ and the eight $\{3,4\}$-types other than T_0; Appendix A.2.5 reports it as the fifth problem. The degree-3 inequalities are checked against the rational number $0.139822124008 < c_3$, a lower bound for c_3. The 62 cells tile $[0, \frac{35}{32}]$, their smallest width is 2^{-8}, and the smallest slack in (7.1) is $1.06 \cdot 10^{-4}$. Let τ lie in a cell $[t_0, t_1]$ with stored transfer vector t, and let $T \ne T_0$ be a type of degree d. Then Lemmas A.6 and A.7 and the construction of I_T^{lo} give $C_T(\tau) \le \Phi(-t_0) - I_T^{\mathrm{lo}}(t_1)$, and (7.1) gives $C_T(\tau) + \lambda_T(t) < a_d$. □

The factor $1 - \varepsilon_R$ in the correlation bounds adds an error term at nonzero thresholds. For $\varepsilon_R < 1$ put $e_R := 205\,\varepsilon_R + \kappa(r_{33}(1-\varepsilon_R)) - c_3$, where the constant 205 comes from Lemma 6.3(b). Since κ is continuous and decreasing on $(0,1)$ (as $b(r)$ in (8.1) is increasing), $e_R \ge 0$ and $e_R \to 0$ as $R \to \infty$.

Lemma A.14 (expected frontiers, maximum degree 4)**.** *Let $\varepsilon_R < 1$. Then $\mathbb{E}|\partial^+ S_\tau| \le \beta(H) + e_R N$ for every $\tau > 0$, and $\mathbb{E}|\partial^- S_\tau| \le \beta(H) + e_R N$ for every $\tau < 0$.*

Proof. First let $0 < \tau \le \frac{35}{32}$, and let t be the transfer vector of Proposition A.13. Let v have type $T \ne T_0$. By (L1), its star is a $d(v)$-star with lower bounds $(1-\varepsilon_R) r_T$, so Lemma 6.2 and Lemma 6.3(b) (as $\tau < 3$, $d(v) \le 4 \le 5$, and the entries of r_T lie in $(0, r_{33}]$) give $\mathbb{P}(v \in \partial^+ S_\tau) \le C_T(\tau) + 205\,\varepsilon_R$, and hence $\mathbb{P}(v \in \partial^+ S_\tau) + \lambda_T(t) \le a_{d(v)} + e_R$. A vertex of type T_0 has three neighbors of degree 3, so by (L1) the vector $(Y_v, Y_{x_1}, Y_{x_2}, Y_{x_3})$ satisfies the hypothesis of Proposition 8.8 with $r = r_{33}(1-\varepsilon_R)$, and $\mathbb{P}(v \in \partial^+ S_\tau) \le \kappa(r_{33}(1-\varepsilon_R)) \le c_3 + e_R$, with $\lambda_{T_0} = 0$. Lemma 7.4 with $e = e_R$ gives the claim. If $\tau > \frac{35}{32}$, then $\mathbb{P}(v \in \partial^+ S_\tau) \le \mathbb{P}(Y_v > \tau) \le \Phi(-\frac{35}{32}) < 0.13704 < c_3$ for every v, so

$\mathbb{E}|\partial^+ S_\tau| \le c_3 N \le \beta(H)$. Finally, for $\tau < 0$ the event $\{v \in \partial^- S_\tau\}$ is, up to null events, the event $\{v \in \partial^+ S_{-\tau}\}$ for the labeling $-Y$, which has the same distribution. □

Proof of Theorem 8.2. The proof has two parts. The first proves the bound with remainder ηN for graphs of minimum degree at least 3, by a concentration argument as for Theorem 8.1; the second extends it to all graphs of maximum degree 4 with Corollary 4.6.

Minimum degree 3. We first show: for every $\eta > 0$ there is $N(\eta)$ such that every graph H with $\delta(H) \ge 3$, $\Delta(H) \le 4$ and $N \ge N(\eta)$ vertices satisfies $\mathrm{pw}(H) \le \beta(H) + \eta N$. Fix $\eta > 0$. We begin by fixing the parameters and the counts to be controlled. Choose R with $\varepsilon_R < 1$, $e_R \le \frac{\eta}{4}$, and $R \ge R_0$ for $\frac{\eta}{4}$ in Lemma A.12; choose an integer $K_0 \ge 1$ with $\varphi(0)/K_0 \le \frac{\eta}{4}$; and let $\tau_j = -2 + j/K_0$ for $0 \le j \le 4K_0$. Consider the $8K_0 + 4$ counts $|\partial^- S_{\tau_j}|$ for $0 \le j < 2K_0$, $|\partial^+ S_{\tau_j}|$ for $2K_0 < j \le 4K_0$, $\hat{L}_0$ and $\hat{L}_0'$, the window counts $M_j = \#\{v : \tau_j < Y_v \le \tau_{j+1}\}$ for $0 \le j < 4K_0$, and the tail counts $\#\{v : Y_v \le -2\}$ and $\#\{v : Y_v > 2\}$. Each is a sum over v of $\{0,1\}$-valued terms (for $\hat{L}_0$ and $\hat{L}_0'$ because $U \cap W = \emptyset$) that depend only on the labels on $N[v]$ and the degrees, hence, by (L2), only on $(Z_w)_{w \in B_{\varrho_R+1}(v)}$. Choose $N(\eta)$ such that for $N \ge N(\eta)$

$$(8K_0 + 4)\, 16^{\varrho_R+1} < 2N^{1/2}, \qquad 2N^{3/4} \le \tfrac{\eta}{4} N, \qquad 0.023N + N^{3/4} \le c_3 N.$$

Next, by Lemma 7.9 with $\Delta = 4$, $\varrho = \varrho_R + 1$ and $\lambda = N^{3/4}$, some realization has pairwise distinct labels and every count within $N^{3/4}$ of its mean; fix it. By Lemmas A.12 and A.14, every inner count at $\tau_j < 0$ and every outer count at $\tau_j > 0$ is at most $\beta(H) + \frac{\eta}{4}N + N^{3/4}$, both $\hat{L}_0$ and $\hat{L}_0'$ are at most $\beta(H) + \frac{\eta}{4}N + N^{3/4}$, every window count is at most $\frac{\eta}{4}N + N^{3/4}$, and the two tail counts are at most $0.023N + N^{3/4}$.

Finally, we bound every frontier of this realization. Let $t \in [-2, 2) \setminus \{0\}$, and let j be such that $t \in [\tau_j, \tau_{j+1})$. If $t < 0$, then $j < 2K_0$, and Lemma 7.8(i) bounds $|\partial^- S_t|$ by the inner count at τ_j plus M_j. If $t > 0$, then $\tau_{j+1} > 0$, and Lemma 7.8(i) bounds $|\partial^+ S_t|$ by the outer count at τ_{j+1} plus M_j. In both cases the bound is at most $\beta(H) + \frac{\eta}{2}N + 2N^{3/4} \le \beta(H) + \eta N$. For $t < -2$ and $t \ge 2$, Lemma 7.8(ii),(iii) bound the inner, respectively outer, frontier by a tail count, which is at most $0.023N + N^{3/4} \le c_3 N \le \beta(H)$ because every vertex of H has degree 3 or 4. Also $\max(\hat{L}_0, \hat{L}_0') + 1 \le \beta(H) + \frac{\eta}{4}N + N^{3/4} + 1 \le \beta(H) + \eta N$. Lemma A.9 with $w = \beta(H) + \eta N$ gives $\mathrm{pw}(H) \le \beta(H) + \eta N$.

All graphs. Let $\rho_0(N)$ be the largest value of $\max\{0, \mathrm{pw}(H) - \beta(H)\}$ over the graphs H on N vertices with $\delta(H) \ge 3$ and $\Delta(H) \le 4$ (zero if there are none). By the first part, $\rho_0(N) \le \eta N$ for $N \ge N(\eta)$, so $\rho_0(N) = o(N)$, and so is its running maximum $\bar{\rho}$ (4.2), which is nondecreasing. Corollary 4.6 with $k = 4$ and $(a_0, \dots, a_4) = (0, 0, 0, c_3, 0.2627)$ gives $\mathrm{pw}(G) \le c_3 n_3 + 0.2627\, n_4 + \bar{\rho}(n) + 3\log_2 n + 1$ for every nonempty graph G with $\Delta(G) \le 4$. □

Remark A.15 (the coefficient 0.2627). The coefficient 0.2627 is a target, not the limit of the method. The numerical maxima after Lemma A.10 show that, in this form, the bounds of Lemma A.10(c) cannot be lowered below about $0.2618 - j\,s$, the binding pattern being $j = 2$. The charging at nonzero thresholds has slack about 10^{-4} (Proposition A.13). The target lies just below $(1+6c_3)/7 \approx 0.2627047$, the value at which the two branches of the bootstrap recurrence (9.1) meet at $d = 5$. Compared with Theorem 7.1, Theorem 8.2 gives the smaller bound exactly when $n_4 < \frac{F_3 - c_3}{0.0052}\, n_3 \approx 0.0646\, n_3$ (against (b); Remark A.18). Appendix A.6 turns the proof into an algorithm (Theorem 10.1, switched case).

A.3.2 The certificates of the switch lemma

The rest of this subsection describes the computer-assisted part of the proof of Lemma A.10. After the reduction to equal incident correlations in that proof, it remains to bound the function G_r

of (A.3) on the set E_d of $d \times d$ correlation matrices. This is done for $d = 3$ and the two patterns of Lemma A.10(a),(b), and for $d = 4$ and the five patterns $j = 0, \dots, 4$ of Lemma A.10(c). Write V for the bound to be proved and $s_i = s_{r_i} = \sqrt{1 - r_i^2}$. The two cases use different parametrizations and formulas (Appendix A.3.3) but the same kind of certificate (Appendix A.3.4). Appendix A.3.5 gives the results and lists what is trusted.

A.3.3 Parametrization and formulas

Let $W \sim N(0, C)$ be independent of $Z \sim N(0, 1)$, let $Y_i = r_i Z + s_i W_i$, and let Σ be the correlation matrix of $(Y_1, \dots, Y_d)$, so that $\Sigma_{ij} = r_i r_j + s_i s_j C_{ij}$ for $i \neq j$.

Degree 3: angles. Write $C_{ij} = \cos\theta_{ij}$ with $\theta = (\theta_{12}, \theta_{13}, \theta_{23}) \in [0, \pi]^3$. By Lemma 8.10, C is a correlation matrix exactly when θ lies in the convex polytope Θ cut out by the three triangle inequalities and $\theta_{12} + \theta_{13} + \theta_{23} \leq 2\pi$, and C is positive definite exactly in the interior of Θ. Let $\Sigma^{(0)}$ be the correlation matrix of (Z, Y_1, Y_2, Y_3). It is positive definite whenever C is, since its Schur complement in Z is $\operatorname{diag}(s_i)\, C \operatorname{diag}(s_i)$. Since $\lceil N/2 \rceil = \mathbf{1}\{N \geq 1\} + \mathbf{1}\{N = 3\}$ for $N \leq 3$,

$$G_r = \mathbb{P}(Z \leq 0,\ N \geq 1) + \mathbb{P}(Z \leq 0,\ N = 3) = \tfrac{1}{2} - P_4(\Sigma^{(0)}) + P_4(\Sigma^{(0)\prime}),$$

where P_4 is the probability of the positive orthant and $\Sigma^{(0)\prime}$ is $\Sigma^{(0)}$ with the signs of the row and column of Z reversed. Along the path $R_t = (1 - t)I + t\Sigma^{(0)}$, Lemma 4.7 gives the value as an integral,

$$P_4(\Sigma^{(0)}) = \frac{1}{16} + \int_0^1 \sum_{i<j} \frac{\Sigma^{(0)}_{ij}}{2\pi\sqrt{1 - t^2(\Sigma^{(0)}_{ij})^2}} \Big(\frac{1}{4} + \frac{\arcsin \pi_{kl}(t)}{2\pi} \Big) dt,$$

where $\{k, l\}$ is the complementary pair of indices and $\pi_{kl}(t)$ is the partial correlation of k and l given i and j under R_t. Here the conditional orthant probability of two coordinates is given by Lemma 4.9. In the interior of Θ, Lemma 4.7 and the chain rule $\partial\Sigma_{ij}/\partial\theta_{ij} = -s_i s_j \sin\theta_{ij}$ give

$$\frac{\partial G_r}{\partial \theta_{ij}} = -s_i s_j \sin\theta_{ij}\, \varphi_2(0, 0; \Sigma_{ij})\, (p'_{ij} - p_{ij}),$$

where p_{ij} and p'_{ij} are the conditional probabilities, given $Y_i = Y_j = 0$, that the two remaining coordinates of the vectors with correlation matrices $\Sigma^{(0)}$ and $\Sigma^{(0)\prime}$ are positive; both are closed forms by Lemma 4.9. When $r_i = r_j$, the factor $s_i s_j \sin\theta_{ij}\, \varphi_2(0, 0; \Sigma_{ij})$ is rewritten as $s_i\sqrt{1 + \cos\theta_{ij}}/\big(2\pi\sqrt{1 + \Sigma_{ij}}\big)$ (using $1 - \Sigma_{ij} = s_i^2(1 - \cos\theta_{ij})$), which stays bounded as $\theta_{ij} \to 0$.

Degree 4: the elliptope. Use the six entries $c = (c_{12}, c_{13}, c_{14}, c_{23}, c_{24}, c_{34})$ of C as coordinates. The parameter set is the elliptope $E_4 \subset [-1, 1]^6$, which is compact and convex, with the positive definite matrices as its interior. For every $v \in \mathbb{R}^4$, $v^{\mathsf{T}} C v \geq 0$ is a linear inequality in c valid on E_4; with v a vector of binary fractions it has rational coefficients, and these *cuts* are used to describe E_4 from outside. Since G_r is invariant under the permutations of the neighbors that preserve the bounds, we may also impose linear symmetry constraints: for $j \in \{0, 4\}$, c_{12} is the largest entry and $c_{13} \geq c_{14}, c_{23}, c_{24}$; for $j \in \{1, 3\}$, the three neighbors with equal bounds are $2, 3, 4$ and $c_{12} \geq c_{13} \geq c_{14}$; for $j = 2$, the neighbors with bound r_{34} are $1, 2$ and $c_{13} \geq c_{14}, c_{23}, c_{24}$. The next lemma reduces the value of Ψ, and hence of G_r, to a sign moment.

Lemma A.16 (sign moment)**.** *For a centered Gaussian vector $(Z, Y_1, \dots, Y_4)$ with unit variances and $\operatorname{Corr}(Z, Y_i) = a_i$,*

$$\Psi(Z; Y_1, \dots, Y_4) = \sum_{i=1}^{4} \frac{\arccos a_i}{4\pi} + \frac{1}{8} - \frac{K}{8}, \qquad K = \mathbb{E}\big[\operatorname{sgn} Y_1 \operatorname{sgn} Y_2 \operatorname{sgn} Y_3 \operatorname{sgn} Y_4\big].$$

Proof. With probability 1, all coordinates are nonzero. Write $\lceil N/2 \rceil = \frac{N}{2} + \frac{1}{2}\mathbf{1}\{N \text{ odd}\}$. By Lemma 4.9, $\mathbb{E}[\mathbf{1}\{Z \le 0\}N] = \sum_i \mathbb{P}(Z \le 0 < Y_i) = \sum_i \frac{\arccos a_i}{2\pi}$. Next, $\mathbf{1}\{N \text{ odd}\} = \frac{1}{2}(1 - \Pi)$ with $\Pi = \prod_i \operatorname{sgn} Y_i$, because $\Pi = (-1)^{4-N} = (-1)^N$; and $\mathbf{1}\{Z \le 0\} = \frac{1}{2}(1 - \operatorname{sgn} Z)$. The means of $\operatorname{sgn} Z$ and of $\operatorname{sgn} Z \cdot \Pi$ vanish, as both are odd functions of a centered Gaussian vector. Hence $\mathbb{E}[\mathbf{1}\{Z \le 0\}\mathbf{1}\{N \text{ odd}\}] = \frac{1}{4}(1 - K)$, and the identity follows. □

After the reduction to equal incident correlations, $a_i = r_i$, so the first term in Lemma A.16 is a constant, and K is a function of Σ. Along $R_t = (1-t)I + t\Sigma$, Lemma 4.7 gives

$$K(\Sigma) = \int_0^1 \sum_{p<q} \frac{4}{\pi^2} \frac{\Sigma_{pq} \arcsin \pi_{kl}(t)}{\sqrt{1 - t^2\Sigma_{pq}^2}}\, dt,$$

since $\partial K/\partial \Sigma_{pq}$ is $4\varphi_2(0,0;\Sigma_{pq})$ times the conditional mean of $\operatorname{sgn} Y_k \operatorname{sgn} Y_l$ given $Y_p = Y_q = 0$, which is $\frac{2}{\pi}\arcsin$ of the conditional correlation. In the interior of E_4, therefore,

$$\frac{\partial G_r}{\partial c_{pq}} = -\frac{s_p s_q}{2\pi^2} \frac{\arcsin \pi_{kl\cdot pq}}{\sqrt{1 - \Sigma_{pq}^2}},$$

where $\pi_{kl\cdot pq}$ is the partial correlation in Σ. If $r_p = r_q$, then $1 - \Sigma_{pq} = s_p^2(1 - c_{pq})$, and the derivative is at most $K_{pq}/\sqrt{1 - c_{pq}}$ in absolute value, with $K_{pq} = s_p/(4\pi\sqrt{1 + \Sigma_{pq}^{\min}})$ for the smallest value $\Sigma_{pq}^{\min}$ of Σ_{pq} on the region considered. Integrated along a segment from x' to x, this coordinate contributes at most $2K_{pq}|\sqrt{1 - x_{pq}} - \sqrt{1 - x'_{pq}}|$. This square-root bound handles the corners where two residuals with equal bounds coincide. Near such corners the gradient has no limit, and the following bound is used instead.

Lemma A.17 (pair bound)**.** *In the setting of Lemma A.16, for every pair $\{p,q\}$ with complementary pair $\{k,l\}$,*

$$\Psi \le \sum_{i=1}^{4} \frac{\arccos a_i}{4\pi} + \frac{1}{8} - \frac{\arcsin \Sigma_{kl}}{4\pi} + \frac{\arccos \Sigma_{pq}}{4\pi}.$$

Proof. With probability 1, $\operatorname{sgn} Y_p \operatorname{sgn} Y_q = 1 - 2 \cdot \mathbf{1}\{Y_p Y_q < 0\}$, so

$$K = \mathbb{E}[\operatorname{sgn} Y_k \operatorname{sgn} Y_l] - 2\,\mathbb{E}[\operatorname{sgn} Y_k \operatorname{sgn} Y_l;\ Y_p Y_q < 0].$$

By Lemma 4.9 the first term is $\frac{2}{\pi}\arcsin \Sigma_{kl}$, and the second has absolute value at most $2\mathbb{P}(Y_p Y_q < 0) = \frac{2}{\pi}\arccos \Sigma_{pq}$. Substitute into Lemma A.16. □

On a box, the right side is largest at the lower ends of c_{kl} and c_{pq}, so the pair bound needs only one evaluation per box.

A.3.4 Box certificates

In this subsection the *domain* is Θ for $d = 3$, and for $d = 4$ the set of points of E_4 that satisfy the symmetry constraints.

The root box ($[0,\pi]^3$ for $d = 3$, $[-1,1]^6$ for $d = 4$) is covered by finitely many closed boxes forming a guillotine partition, found by a floating-point branch-and-bound search. A separate program checks that the stored boxes tile the root box exactly. It rebuilds the partition by finding, at each step, a coordinate value that separates the boxes, so the certificate does not rely on the bookkeeping of the search.

Some points of the domain need no boxes, because G_r is continuous. Continuity holds because the distribution of the vector in (A.3) depends weakly continuously on C, and the function $y \mapsto \mathbf{1}\{y_0 \le 0\}\lceil N(y)/2 \rceil$ is bounded and continuous outside the hyperplanes $\{y_i = 0\}$, a null set for every distribution in the family, so the continuous mapping theorem applies. For $d = 3$, the floating-point root box misses the points of Θ with a coordinate in $(\mathrm{fl}(\pi), \pi]$, where $\mathrm{fl}(\pi) < \pi$ is the floating-point value of π. These points are limits of points of Θ inside the root box, since Θ is convex with nonempty interior, so they need no boxes. For $d = 4$, the half-open boxes $\prod_m [\mathrm{lo}_m, \mathrm{hi}_m)$ partition $[-1, 1)^6$, which contains the interior of the domain. Since the domain is convex with nonempty interior and G_r is continuous, it suffices to bound G_r on the interior.

Every box B is then verified in Arb ball arithmetic, regardless of how the search classified it, in one of the following ways; a box that fails is subdivided and the parts are verified again.

- *Empty.* The box misses the domain. For $d = 3$ one constraint of Θ is violated on all of B, checked exactly. For $d = 4$ a Farkas certificate, a nonnegative combination of the box bounds, the symmetry constraints and the stored cuts, is checked in exact rational arithmetic; if the stored cuts do not suffice, cuts at eigenvectors of non-positive-definite points are added. For $d = 4$ a box can also be discarded because an exact weak-duality bound shows $x_m \ge \mathrm{hi}_m$ on the domain for some coordinate m; then the domain meets the half-open box in no point. This occurs where symmetry and the box bounds force an equality.

- *Pair bound* ($d = 4$). Lemma A.17 for some pair, evaluated at the lower box ends, is $< V$.

- *Bounded.* Fix a point x_1 in the interior of the domain, in or near B (for $d = 4$, positive definiteness is checked by leading principal minors in Arb, and x_1 satisfies the symmetry constraints exactly). Let H be the smallest box containing B and x_1. Let $[g_{c,m} - g_{r,m},\ g_{c,m} + g_{r,m}]$ enclose the partial derivative of G_r in the coordinate m over the intersection of H with the interior of the domain. Let M be an upper bound for $\max g_c \cdot x$ over the points x of B that satisfy the linear constraints; it is certified by a dual feasible solution of the linear program, evaluated exactly or in ball arithmetic. For x in B and in the domain, the half-open segment $[x_1, x)$ lies in the interior of the domain (by convexity), where G_r is differentiable, so integrating the gradient along it gives

$$G_r(x) \le G_r(x_1) + (M - g_c \cdot x_1) + \sum_m g_{r,m} \max_{y \in B} |y_m - x_{1,m}|,$$

  plus, for $d = 4$, the term $\sum 2K_{pq} \max_{y \in B} |\sqrt{1 - y_{pq}} - \sqrt{1 - x_{1,pq}}|$ for the coordinates treated by the square-root bound instead of the gradient enclosure. The box is accepted if the right side is $< V$ in Arb.

- *Pushed to a face* ($d = 3$). Near the face $F = \{\theta_{12} + \theta_{13} + \theta_{23} = 2\pi\}$ the conditional variances degenerate and the gradient formula fails. Moving along $(1, 1, 1)$ preserves the triangle inequalities, so every $x \in B \cap \Theta$ is joined to a point $x^* \in F$ by a segment in Θ. The verifier certifies, by affine-arithmetic gradient enclosures on sub-boxes, that $\sum_m \partial G_r / \partial \theta_m > 0$ on the prism swept by B towards F, so $G_r(x) \le G_r(x^*)$. It then bounds G_r on the face patch by covering it with parallelogram cells, each handled like a bounded box with a linear term maximized at a vertex.

Enclosures. Values $G_r(x_1)$ are enclosed by integrating the Plackett representations above with Arb's rigorous integration (`acb.integral`, whose `analytic` flags certify analyticity of the integrand) on $[0, 1 - 10^{-8}]$. The rest of the path is bounded by the elementary integral $\int dt / \sqrt{1 - t^2\rho^2}$, an

arcsine difference, since every conditional term is bounded ($\frac{1}{4} + \frac{\arcsin}{2\pi} \in [0, \frac{1}{2}]$, $|\arcsin| \le \frac{\pi}{2}$). For $d = 4$ the integrand is evaluated without matrix inverses: along R_t each conditional covariance given (p, q) is $A_{ab}(t)/(1 - t^2\Sigma_{pq}^2)$ for a cubic polynomial A_{ab}, and the substitution $t = 1 - e^{-u}$ moves poles near $t = 1$ away from the path. Gradients are enclosed in ball arithmetic and, for $d = 4$, also in affine arithmetic over Arb (with rigorous remainders for products, reciprocals and square roots), on up to 16 sub-boxes; the results are intersected. An undetermined partial correlation is replaced by the whole interval $[-1, 1]$ (equivalently, a conditional probability by $[0, 1]$), and if every enclosure fails for $d = 4$, the analytic bound $s_p s_q/(4\pi\sqrt{1 - \max|\Sigma_{pq}|^2})$ on the derivative is used.

A.3.5 Results and what is trusted

All seven bounds are certified. Every comparison is a strict inequality in Arb, and every box list passes the coverage check.

pattern	V	boxes	largest certified bound	numerical maximum
(r_{33}, r_{33}, r_{34})	0.1529003	8 498	0.1529002993	0.1528953
(r_{33}, r_{34}, r_{34})	0.165616	21 082	0.1656159972	0.1656115
$d = 4$, $j = 0$	0.2627	368 779	0.2626999832	0.250000
$d = 4$, $j = 1$	$0.2627 - s$	2 366 198	0.2496218225	0.245466
$d = 4$, $j = 2$	$0.2627 - 2s$	5 450 439	0.2365436480	0.235687
$d = 4$, $j = 3$	$0.2627 - 3s$	5 415 446	0.2234654716	0.221375
$d = 4$, $j = 4$	$0.2627 - 4s$	2 738 755	0.2103872958	0.208185

The column "boxes" counts the boxes of the search; the verification subdivides some of them further. The largest certified bounds lie just below V because the search accepts a box as soon as its bound is below V; they are not estimates of the maximum. For $d = 3$, 27 boxes of the first pattern touch or approach the face F, where the numerical maximum 0.1528953 lies $5 \cdot 10^{-6}$ below V; they were closed by the push to the face, with 331 face cells. For $d = 4$, the first verification pass left 14 boxes open, and recursive subdivision with the same certificates closed them. Twelve of these had been labeled empty by the search although they contain feasible points, because its cutting-plane loop had stopped early. The verifier for $d = 4$ compared its bounds with a value of V computed in double precision, with c_3 rounded to 0.139822124008. This value differs from the exact V by less than $2 \cdot 10^{-14}$, far below the smallest margin, $2.7 \cdot 10^{-11}$ at $j = 2$.

What is trusted. As in Appendix A.2.7, the certificates rest on the correctness of Arb/FLINT (`python-flint` 0.9.0), including its rigorous integration, of exact rational arithmetic, and of the derivations above. The HiGHS linear-program solver is used only to find dual and Farkas certificates, which are rechecked exactly or in ball arithmetic. Each verifier has a single implementation. An earlier verification of the case $d = 3$ had a defect: an undetermined partial correlation (a not-a-number enclosure) was turned into an exact value, which collapsed the gradient enclosures near the face F. It was found by an audit and corrected, and the results above are those of the corrected verifier, which treats an undetermined enclosure as the whole interval. The box lists for $d = 3$ and per-part summaries for $d = 4$ are part of the accompanying code. The full box lists for $d = 4$, about 11 GB, are not; the search programs regenerate them.

A.4 Pathwidth sequences: proofs, and comparison with previous bounds

This subsection proves the two results on pathwidth sequences stated in Section 9: Theorem 9.2, which lists six pathwidth sequences, and Corollary 9.5, which states that every row of Table 4 is a pathwidth sequence; the rows are the bootstraps of five of the six sequences, rounded up, with the

last column bounded by closed forms. Appendix A.4.2 then compares the coefficients of Theorem 1.3 with the previous ones.

We use the notation of Section 9: pathwidth sequences are as in Definition 9.1, the bootstrap is the extension (9.1), h_d and H_d are the increments of a bootstrap and their running minima (Section 9.1), and σ is the bootstrap of S (Section 9.4). The constants c_3, F_3 and F_4 are those of Section 9.

A.4.1 Pathwidth sequences: proofs of Theorem 9.2 and Corollary 9.5

Both statements follow from Lemma 9.3, applied with the six sequences of Theorem 9.2 as seeds. The first proof checks the hypotheses of the lemma; the second computes the bootstraps.

Theorem 9.2 (pathwidth sequences)**.** *Each of the following six tuples is a pathwidth sequence:*

$$\mathsf{C} = (c_3), \qquad \mathsf{M} = (c_3,\, 0.2627), \qquad \mathsf{F} = (F_4,\, F_4),$$

$$\mathsf{S} = (\tfrac{5}{32},\, \tfrac{1}{4},\, \tfrac{11}{32}), \qquad \mathsf{P} = (F_3,\, 0.2575,\, 0.3667), \qquad \mathsf{R} = (0.1455,\, 0.2530,\, 0.3600).$$

Proof of Theorem 9.2. For each of the six sequences $X = (X_3, \dots, X_k)$ we check the hypotheses of Lemma 9.3 with $(b_3, \dots, b_k) = X$.

First, all six are nondecreasing with $X_3 \geq 0$.

Next, $X_k + H_k \leq 1$ in all six cases. For C ($k = 3$), $X_3 + H_3 = 2c_3$. For M ($k = 4$), $H_4 = h_4 = 0.2627 - c_3 < h_3$, so $X_4 + H_4 = 0.2627 + (0.2627 - c_3)$. For F ($k = 4$), $H_4 = h_4 = 0$, so $X_4 + H_4 = F_4$. For P, R and S ($k = 5$), $X_5 + H_5$ equals $0.3667 + 0.1092$, $0.3600 + 0.1070$ and $\frac{11}{32} + \frac{3}{32}$, respectively; the increments of P and R are listed in Remark A.20(a), and those of S are $\frac{5}{32}, \frac{3}{32}, \frac{3}{32}$. All these values are at most 1.

Finally, we check the bound on the graphs J in Lemma 9.3. Let J be a graph with $\delta(J) \geq 3$ and $\Delta(J) \leq k$. If $X = \mathsf{C}$, then J is cubic, $|V(J)| = n_3(J)$, and Theorem 8.1 gives $\mathrm{pw}(J) \leq c_3 n_3(J) + o(|V(J)|)$. If $X = \mathsf{M}$, Theorem 8.2 gives $\mathrm{pw}(J) \leq c_3 n_3(J) + 0.2627\, n_4(J) + o(|V(J)|)$. If $X = \mathsf{F}$, then $|V(J)| = n_3(J) + n_4(J)$, and Theorem 7.1(a) gives $\mathrm{pw}(J) \leq F_4(n_3(J) + n_4(J)) + o(|V(J)|)$. If $X \in \{\mathsf{P}, \mathsf{R}, \mathsf{S}\}$, Theorem 7.2 gives $\mathrm{pw}(J) \leq \sum_{d=3}^{5} X_d n_d(J) + o(|V(J)|)$.

So Lemma 9.3 shows that the bootstrap b of X is a pathwidth sequence with all entries at most 1. Since $b_d = X_d$ for $d \leq k$, the sequence X, which continues with $X_d = 1$ for $d > k$ (Definition 9.1), is entrywise at least b, hence a pathwidth sequence. □

Corollary 9.5 (the bootstraps)**.** *Each row of Table 4 is a pathwidth sequence.*

Proof of Corollary 9.5. By Lemma 9.3, whose hypotheses were checked in the previous proof, the bootstraps of the six sequences are pathwidth sequences. A sequence that is entrywise at least a pathwidth sequence is one, so it suffices to show that every entry of Table 4 is at least the corresponding entry of the bootstrap. For the black entries this holds by $c_3 < 0.1399$, $F_3 < 0.1402$ and $F_4 < 0.2458$ (Proposition 6.9). The grey entries for $4 \leq d \leq 10$ are the exact values of (9.1) rounded up to four decimals. The accompanying script `bootstrap_table.py` computes them in exact rational arithmetic and checks the branch of the maximum at every step; it also computes the bootstrap of C, which the table omits. We record what the computation consists of, row by row, and then bound the entries of the last column.

Row F. Here $H_4 = 0$, so (9.1) gives $b_5 = \max\{\frac{1+3F_4}{5}, 1\} = 1$, then $H_5 = 0$ again, and $b_d = 1$ for all $d \geq 5$.

Row C. Here $k = 3$ and $H_3 = c_3$. Since $\frac{1}{8} < c_3 < \frac{1}{7}$, the recurrence takes the second term of the maximum at $d = 4$ and the first term at $d = 5, 6$:

$$b_4 = 1 - 4c_3, \qquad b_5 = 1 - \tfrac{17}{5}c_3, \qquad b_6 = 1 - \tfrac{44}{15}c_3,$$

with $H_4 = c_3$ (as $h_4 = 1 - 5c_3 > c_3$), $H_5 = h_5 = \frac{3}{5}c_3$ and $H_6 = h_6 = \frac{7}{15}c_3$. So $H_6 \le H_5 \le 3H_6$ and $h_6 = H_6$, and Lemma 9.7 with $j = 6$ gives $d\,b_d = (d-2)\,b_{d-1} + b_{d-2} + 1$ for $d \ge 7$. Hence $u_d := 1 - b_d = \gamma_d c_3$ for $d \ge 4$, with rationals γ_d that satisfy $d\gamma_d = (d-2)\gamma_{d-1} + \gamma_{d-2}$ for $d \ge 7$, and b_d is rounded up using a rigorous lower bound for c_3.

Row M. Here $k = 4$ and $H_4 = 0.2627 - c_3$. Since $0.2627 < \frac{1+6c_3}{7}$, the recurrence takes the second term of the maximum at $d = 5$, and then the first term:

$$b_5 = 1 - 5\,(0.2627 - c_3), \qquad b_6 = \tfrac{13}{3}c_3 - \tfrac{4151}{30000},$$

with $H_5 = H_4$ (as $h_5 = 5c_3 - 0.5762 > H_4$) and $H_6 = h_6 = \frac{5254}{30000} - \frac{2}{3}c_3$. So $H_6 \le H_5 \le 3H_6$, and Lemma 9.7 with $j = 6$ gives $d\,b_d = (d-2)\,b_{d-1} + b_{d-2} + 1$ for $d \ge 7$. Every entry is of the form $\alpha_d + \beta_d c_3$ with rationals α_d, β_d. The script decides each branch of the maximum and each running minimum at both ends of a rigorous enclosure of c_3, and rounds up with the same enclosure.

Rows P, R *and* S. For P and R the computation of Remark A.20(a) applies: the first term of the maximum is attained from $d = 6$ on, and $d\,b_d = (d-2)\,b_{d-1} + b_{d-2} + 1$ for $d \ge 7$. For P the entries for $d \ge 6$ do not depend on the exact value of F_3, since $H_5 = h_5$ whenever $0.1092 \le F_3 \le 0.1483$. Row S is the bootstrap σ of Section 9.4.

The last column. For $X \in \{\mathsf{C}, \mathsf{M}, \mathsf{P}, \mathsf{R}, \mathsf{S}\}$ put $u_d := 1 - b_d$ and $w_d := (d+1)u_d$. By the above, $d\,u_d = (d-2)\,u_{d-1} + u_{d-2}$ for $d \ge 7$, so Lemma 9.12 applies with $d_0 = 7$ and

$$\lambda = 6u_6 + u_5 = 21c_3, \qquad \tfrac{81437}{10000} - 31c_3, \qquad \tfrac{3909}{1000}, \qquad \tfrac{3947}{1000}, \qquad \tfrac{129}{32}$$

for C, M, P, R and S respectively; for M, $\lambda \approx 3.80921$. The hypothesis $\frac{11}{12}\lambda < w_{10} < \lambda$ of Lemma 9.12(ii) with $d_1 = 10$ holds in all five cases (checked by the script). Hence $w_d > (1 - \frac{1}{d^2})\lambda \ge \frac{120}{121}\lambda$ for every $d \ge 11$, that is, $b_d < 1 - \frac{120\lambda/121}{d+1}$. Since $\frac{120}{121} \cdot 21c_3 > 2.91$, $\frac{120}{121} \cdot 3.8092 > 3.77$, $\frac{120}{121} \cdot 3.909 > 3.87$ and $\frac{120}{121} \cdot 3.947 > 3.91$, the entries of the last column for M, P and R, and the closed form $1 - \frac{2.91}{d+1}$ for C, bound the bootstrap from above. For S, $b_d = \sigma_d \le 1 - \frac{4}{d+1}$ for $d \ge 11$ by Proposition 9.13, and for F the entry is 1. □

We also justify the highlighting in Table 4: the highlighted entries are the column minima. At $d = 3$ and $d = 4$ this can be read off the table. For $d \ge 5$, row S is below rows P and R (Remark A.20(a)). It is also below row M: for $5 \le d \le 10$ by the table, and for $d \ge 11$ because the bootstrap of M has $w_d < \lambda < 3.81$ by Lemma 9.12, while that of S has $w_d \ge 4$. The omitted bootstrap of C is entrywise at least row M: the script checks this exactly for $d \le 60$, and for $d > 60$ its weights satisfy $w_d < 21c_3 < 2.94$, while those of M satisfy $w_d > \frac{120}{121} \cdot 3.8092$. In the last column, the closed forms all have the shape $1 - \frac{c}{d+1}$, so the minimum there is the one with the largest c.

Remark A.18 (which row gives the smallest bound). The rows of Table 4 are pairwise incomparable, and the best of them for a given graph depends only on the proportions of its degrees. Two rows x, y compare as $\sum_d x_d n_d \le \sum_d y_d n_d$ if and only if $\sum_d (y_d - x_d) n_d \ge 0$. The numbers below are computed in floating point from the exact sequences and rounded; they are illustrations, not claims used elsewhere. With the abbreviation $\delta_d^X := b_d^X - b_d^{\mathsf{S}}$, where b^X is the bootstrap of X,

$$\begin{aligned}
\mathsf{P} \le \mathsf{S} &\iff 0.01609\,n_3 \ge 0.0075\,n_4 + 0.02295\,n_5 + \textstyle\sum_{d\ge6} \delta_d^{\mathsf{P}}\, n_d,\\
\mathsf{R} \le \mathsf{S} &\iff 0.01075\,n_3 \ge 0.003\,n_4 + 0.01625\,n_5 + \textstyle\sum_{d\ge6} \delta_d^{\mathsf{R}}\, n_d,\\
\mathsf{P} \le \mathsf{R} &\iff 0.00534\,n_3 \ge 0.0045\,n_4 + 0.0067\,n_5 + \textstyle\sum_{d\ge6} (\delta_d^{\mathsf{P}} - \delta_d^{\mathsf{R}})\, n_d.
\end{aligned}$$

Maximum degree 3. Row M is best; it coincides with C there.

Maximum degree 4. Each row is best on an interval of the ratio n_4/n_3:

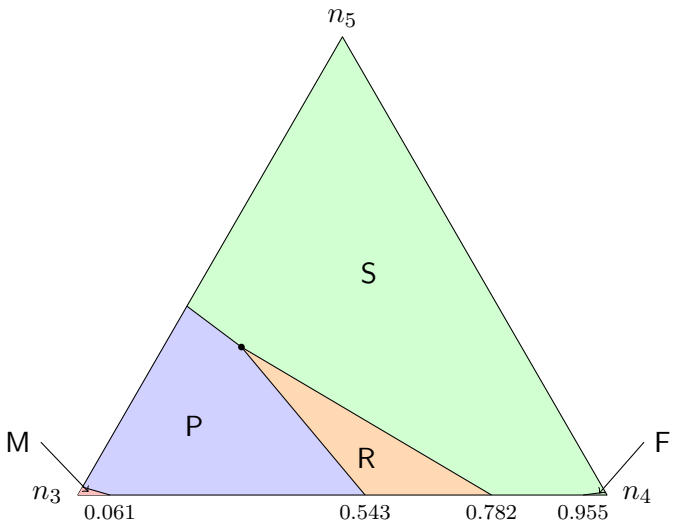


Figure 4: The row of Table 4 with the smallest bound for graphs of maximum degree at most 5, as a function of $(n_3, n_4, n_5)/(n_3 + n_4 + n_5)$ (Remark A.18; computed in floating point). A corner stands for the graphs all of whose vertices of degree at least 3 have that degree; the labels on the bottom edge give the proportion of degree-4 vertices among those of degree 3 or 4. The region of F is too thin to be visible.

best row	n_4/n_3	bound per vertex of degree 3 or 4
M	$[0,\ 0.0646]$	0.1398 to 0.1473
P	$[0.0646,\ 1.187]$	0.1473 to 0.2038
R	$[1.187,\ 3.583]$	0.2038 to 0.2295
S	$[3.583,\ 21.17]$	0.2295 to 0.2458
F	$[21.17,\ \infty]$	0.2458

So M is best only if fewer than about 6.1% of the vertices of degree 3 or 4 have degree 4, and F only if fewer than about 4.5% of the vertices of degree at least 3 have degree 3. The bound per vertex of degree 3 or 4 never exceeds F_4.

Maximum degree 5. Figure 4 shows the best row as a function of the proportions of n_3, n_4, n_5. Rows P, R and S meet at $(n_3, n_4, n_5) \approx (0.529, 0.148, 0.323)$. On the edge $n_4 = 0$ of the triangle, row M is best for $n_5/n_3 \le 0.0178$, row P for $0.0178 \le n_5/n_3 \le 0.70$, and row S beyond.

Larger degrees. Without the bootstrap, every seed would have coefficient 1 at these degrees. With it, the differences between the rows are small: δ_d^{P} is 0.0166, 0.0151, 0.0134, 0.0121, 0.0111 and δ_d^{R} is 0.0114, 0.0105, 0.0093, 0.0084, 0.0076 for $d = 6, \dots, 10$, and both behave like $(\frac{129}{32} - \lambda)/(d+1)$, where the conserved quantity $\lambda = 6(1 - b_6^X) + (1 - b_5^X)$ of Lemma 9.12 is 3.909 for $X = \mathsf{P}$ and 3.947 for $X = \mathsf{R}$. So row P remains best for graphs with many vertices of degree 3 and a fair number of vertices of large degree. For example, if every vertex has degree 3 or 10, row P is best for $0.0374\, n_3 \le n_{10} \le 1.46\, n_3$; row M is best below this range and row S above it.

A.4.2 Comparison with previous bounds, and refinements

This subsection compares the coefficients a_d of Theorem 1.3 with the previous ones (Remark A.19) and records refinements that follow from the same argument (Remark A.20).

Table 7: Per-degree coefficients, rounded to five decimals: the previous $g_{1/6}(d)$ of [EM15], which for $d \le 6$ are those of [FGSS09] (where every degree $d \ge 7$ has coefficient 1), the bootstrap σ_d of S, and the coefficients a_d of Theorem 1.3. The last row gives, for each sequence x, the coefficient $\max_d 2x_d/d$ of m (their ratio in the last column).

d	$g_{1/6}(d)$	σ_d	a_d	$a_d/g_{1/6}(d)$
3	0.16667	0.15625	0.15625	0.938
4	0.33333	0.25000	0.25000	0.750
5	0.43333	0.34375	0.34375	0.793
6	0.51111	0.43750	0.43750	0.856
7	0.56984	0.50446	0.50446	0.885
8	0.61627	0.55804	0.55952	0.908
9	0.65375	0.60119	0.60357	0.923
10	0.68463	0.63676	0.63961	0.934
$d \ge 11$	$1 - \frac{7/2}{d+1} + O(d^{-3})$	$1 - \frac{129/32}{d+1} + O(d^{-3})$	$1 - \frac{4}{d+1}$	$\to 1$
per edge	$\frac{13}{75} \approx 0.17333$	$\frac{7}{48} \approx 0.14583$	$\frac{7}{48}$	0.841

Remark A.19 (comparison with previous coefficients). The previous per-degree coefficients are those of Fomin, Gaspers, Saurabh and Stepanov [FGSS09, Lemma 1] for $d \le 6$, namely $\frac{1}{6}, \frac{1}{3}, \frac{13}{30}, \frac{23}{45}$ at degrees 3 to 6, and the sequence $g_{1/6}(d)$ of Edwards and McDermid [EM15, Corollary 5.2] (recalled in Lemma 9.16), which agrees with them for $d \le 6$ and continues with $g_{1/6}(7) = \frac{359}{630}$, $g_{1/6}(8) = \frac{1553}{2520}$, $g_{1/6}(9) = \frac{14827}{22680}$, $g_{1/6}(10) = \frac{155273}{226800}$. Edwards and McDermid show $g_{1/6}(d) = 1 - \frac{7/2}{d+1} + O(d^{-3})$ [EM15, Section 3.1.2]. The previous bound in terms of edges is $\frac{13}{75}m + o(m)$: for treewidth by Scott and Sorkin [SS07, Corollary 21], and for pathwidth by Kneis, Mölle, Richter and Rossmanith [KMRR09]; it also follows from the bound of Edwards and McDermid, since $\max_d 2g_{1/6}(d)/d = \frac{13}{75}$, attained at $d = 5$. Table 7 compares these values with the bootstrap σ of S and the coefficients a_d of Theorem 1.3.

By Lemma 9.16, the coefficients of Theorem 1.3 are below $g_{1/6}(d)$ at every degree $d \ge 3$, and the previous sequence never reaches the coefficient $1 - \frac{4}{d+1}$ that Theorem 1.3 has from degree 11 on.

The relative improvement over $g_{1/6}$ (one minus the last column of Table 7) is largest at degree 4 and shrinks for large degrees. The deficit $1 - a_d$ is $\frac{111/28}{d+1}$ for $7 \le d \le 10$ and $\frac{4}{d+1}$ for $d \ge 11$, instead of the deficit $\frac{7/2}{d+1} + O(d^{-3})$ of $g_{1/6}$. The bootstrap σ itself has deficit $\frac{w_d}{d+1}$ with $\frac{111}{28} \le w_d < \frac{129}{32}$ for $d \ge 7$ and $w_d \to \frac{129}{32}$ by Proposition 9.13, so the ratio of the deficit of σ to that of $g_{1/6}$ tends to $\frac{129/32}{7/2} = \frac{129}{112} \approx 1.152$. At degree 3 the coefficient $\frac{5}{32}$ of Theorem 1.3 is above the cubic constant $c_3 \approx 0.13982$ of Theorem 8.1; graphs in which vertices of degree 3 dominate get smaller bounds from the refinements of Remark A.20.

Remark A.20 (refinements). The following bounds follow from Theorem 7.2 and Lemmas 9.3 and 9.7, each with a remainder $o(n)$. Part (a) is the computation behind rows P and R of Corollary 9.5.

(a) The bootstraps of P *and* R. Apply Lemma 9.3 with $k = 5$ to the triples $\mathsf{P} = (F_3, 0.2575, 0.3667)$ and $\mathsf{R} = (0.1455, 0.2530, 0.3600)$ of Theorem 7.2. Both triples are nondecreasing. Their increments are $(h_3, h_4, h_5) \approx (0.14016, 0.11734, 0.10920)$ for P (using $F_3 < 0.1401583$) and (0.1455, 0.1075, 0.1070) for R, so $H_5 = h_5$ equals 0.1092 and 0.1070, and $b_5 + H_5 \le 1$. At $d = 6$, $\theta_6 = (1 - b_5 - H_5)/6$ equals 0.08735 for P and 0.08883… for R; in both cases $\theta_6 \le H_5 \le 3\theta_6$. By Lemma 9.6(iv), $h_6 = H_6 = \theta_6$, so $b_6 = b_5 + \theta_6$, and Lemma 9.7 with $j = 6$ gives

$d\,b_d = (d-2)\,b_{d-1} + b_{d-2} + 1$ for $d \geq 7$. Rounded up, the resulting sequences σ^{P} and σ^{R} are rows P and R of Table 4; in particular $\sigma_6^{\mathsf{P}} = 0.45405$ and $\sigma_6^{\mathsf{R}} = 0.44883\ldots$.

Every graph satisfies $\mathrm{pw}(G) \leq \sum_{d\geq 3} \sigma_d^X n_d + o(n)$ for each $X \in \{\mathsf{P}, \mathsf{R}, \mathsf{S}\}$, where $\sigma^{\mathsf{S}} := \sigma$, hence $\mathrm{pw}(G) \leq \min_X \sum_{d\geq 3} \sigma_d^X n_d + o(n)$. The sequences from P and R improve on σ only at degree 3: for $d \geq 4$, $\sigma_d < \sigma_d^{\mathsf{R}}$ and $\sigma_d < \sigma_d^{\mathsf{P}}$. (For $4 \leq d \leq 6$ this follows from $\sigma_4 = \frac{1}{4} < 0.2530$, $\sigma_5 = \frac{11}{32} < 0.3600$ and $\sigma_6 = \frac{7}{16} < 0.4488$. For $d \geq 7$ the differences satisfy the homogeneous recurrence $d\,e_d = (d-2)\,e_{d-1} + e_{d-2}$, whose coefficients are positive, and they are positive at $d = 5, 6$.) They do not improve Corollary 1.4 either, since σ_6^{P} and σ_6^{R} exceed $\frac{7}{16}$, so $\max_d 2\sigma_d^X/d \geq \sigma_6^X/3 > \frac{7}{48}$.

(b) Average degree. The increments of σ^{P}, σ^{R} and σ are nonincreasing from $d = 3$ on: for $d \leq 6$ by the increments computed in (a) and by $\sigma_d = \frac{3d-4}{32}$ (Proposition 9.13), and for $d \geq 7$ because $h_d = H_d \leq H_{d-1} = h_{d-1}$ by Lemma 9.7. So for each X the linear interpolation $\bar{\sigma}^X \colon [2, \infty) \to \mathbb{R}$ of the points (d, σ_d^X), $d \geq 2$, where $\sigma_2^X = 0$, is concave. Recall the average degree $\bar{d} = 2m/n$ (Section 9.5). Let $n_{\geq 3}$ be the number of vertices of degree at least 3 and, if $n_{\geq 3} \geq 1$, let $\hat{d} := \frac{1}{n_{\geq 3}} \sum_{d\geq 3} d\,n_d$ be their average degree. Jensen's inequality for $\bar{\sigma}^X$ bounds $\sum_{d\geq 3} \sigma_d^X n_d$ from above in two ways: by $\bar{\sigma}^X(\bar{d})\,n$, with the weights n_d/n for $d \geq 2$, if $\delta(G) \geq 2$; and by $\bar{\sigma}^X(\hat{d})\,n_{\geq 3}$, with the weights $n_d/n_{\geq 3}$ for $d \geq 3$. With (a) this gives

$$\mathrm{pw}(G) \leq \min_X \bar{\sigma}^X(\bar{d})\,n + o(n) \quad \text{if } \delta(G) \geq 2, \qquad \mathrm{pw}(G) \leq \min_X \bar{\sigma}^X(\hat{d})\,n_{\geq 3} + o(n) \quad \text{if } n_{\geq 3} \geq 1.$$

The second bound needs no condition on the vertices of degree at most 2. The condition $\delta(G) \geq 2$ in the first cannot be dropped, since adding isolated vertices can lower $\bar{d}$ to 2, where the bound is $o(n)$, without changing the pathwidth.

The minimum in the first bound is attained by P for $2 \leq \bar{d} \leq 3.5428$, by R for $3.5428 \leq \bar{d} \leq 3.7818$, and by S from 3.7818 on; the breakpoints are numerical. Indeed, on $[2, 3]$ all three interpolations pass through $(2, 0)$, and σ_3^{P} is the smallest value at 3; on $[3, 4]$ the crossings are those of the three chords; and from 4 on, σ_d is the smallest value at every integer, by (a). For example, the minimum is $F_3 < 0.1402$ at $\bar{d} = 3$ and about 0.1988 at $\bar{d} = 3.5$, compared with $\frac{7}{96}\bar{d}$, that is, 0.21875 and $0.2552\ldots$, from Corollary 1.4.

A.5 Proof of the algorithmic theorem

This subsection proves Theorem 10.1, restated below, following the outline and the roadmap of Section 10.1, and then Corollary 10.2 and Theorem 10.3.

Theorem 10.1 (precise form of Theorem 1.5)**.** *In each case listed above there are a nondecreasing function $\rho \colon \mathbb{N} \to [0, \infty)$ with $\rho(n) = o(n)$ and two algorithms that, given $G \in \mathcal{C}$, output a path decomposition of G in interval form, of width at most $\beta(G) + \rho(n)$, and in the general case also at most $\frac{7}{48}m + \rho(2m)$:*

(a) *a Las Vegas algorithm, in expected $n^{1+o(1)}$ bit operations;*

(b) *a deterministic algorithm, in $n^{2+o(1)}$ bit operations.*

In the bootstrap and general cases, each algorithm uses in addition $O(n(n+m))$ operations on integers of $O(\log n)$ bits.

Besides the Las Vegas and deterministic algorithms of the theorem, the case proofs also analyze *the randomized algorithm*, which takes one sample of the labels within a budget of random draws and outputs its layout without checking the counts. The Las Vegas algorithm draws samples in the same way, without the budget, and checks the counts of each sample until one passes. The analysis also gives the following statement, which Theorem 10.1 does not need.

Remark A.21 (a randomized algorithm). In each case of Theorem 10.1, the randomized algorithm outputs, in $n(\log n)^{O(1)}$ bit operations (in the bootstrap and general cases, with the $O(n(n+m))$ further operations of the theorem), a path decomposition of width at most $\beta(G)+\rho(n)$, and in the general case also at most $\frac{7}{48}m+\rho(2m)$, except with probability $\pi(n)$, or $\pi(n_{\geq 1})$ in the bootstrap and general cases, where $\pi(k)\to 0$ as $k\to\infty$ and $n_{\geq 1}$ is the number of non-isolated vertices of G. This follows from the case proofs below and Appendix A.5.10.

A.5.1 Model, parameters and exact local labels

This subsection fixes the parameters as functions of n, recalls the labels of Definition 5.2 in the multiplicative form used by the algorithms, and shows how their coefficients are computed. The algorithms use these coefficients, rounded to rationals, together with the finite distribution of Appendix A.5.2 in place of the Gaussian variables.

Parameters. For an integer $n\geq 3$ put

$$t:=\lceil\log_2(n+2)\rceil,\qquad h:=\max\{1,\lfloor\tfrac{1}{8}\log_2 t\rfloor\},\qquad K:=h,\qquad \gamma:=t^{-2},\qquad M:=2t^3+1,$$

and

$$\xi:=12\,t^{-7/4},\qquad p_t:=12\,t^{-11/4},\qquad \omega_d:=2(d+1)\,\varphi(0)\,\xi+p_t\quad(d\geq 0).$$

Here h is the locality radius of the labels, $1/K$ is the step of the grid of thresholds, γ is the step of the finite distribution and M its number of points, and ξ, p_t, ω_d are the error terms of Lemma A.24 and Corollary A.25. The parameters are always computed from the order of the *input* graph (in the bootstrap and general cases from $n_{\geq 1}$), also when they are used for a subgraph.

The threshold n_2. Put $n_2:=2^{2^{40}}$. For $n\geq n_2$ we have $t\geq 2^{40}$, so $h=\lfloor\frac{1}{8}\log_2 t\rfloor\geq 5$ and

$$4^h=2^{2h}\leq 2^{\frac{1}{4}\log_2 t}=t^{1/4}. \tag{A.4}$$

Moreover the following inequalities hold for all $n\geq n_2$:

(N1) $\xi<1/K$;

(N2) $\xi\varphi(0)+p_t\leq 5\cdot 10^{-5}$;

(N3) $8\mu\,t^{1/2}n^{-1/2}\leq\frac{1}{2}$ for $\mu=12K+3$, and hence for every $\mu\leq 12K+3$.

Indeed, $K=h\leq\log_2 t\leq t$, so $\xi K\leq 12t^{-3/4}<1$, and $\xi\varphi(0)+p_t\leq 24\,t^{-7/4}\leq 24\cdot 2^{-70}$. For (N3), put $L=\log_2 n\geq 2^{40}$; then $t\leq L+1\leq 2L$ and $\mu\leq 12t+3\leq 13t$, so $8\mu t^{1/2}\leq 104\,t^{3/2}\leq 104\,(2L)^{3/2}\leq 2^{L/2-1}=\frac{1}{2}n^{1/2}$.

For $n<n_2$ we shall put $\rho(n):=n$ and $\pi(n):=0$; then every width claim is trivial, because every path decomposition of a graph with n vertices has width at most $n-1$. All probabilistic and quantitative estimates below therefore assume $n\geq n_2$.

Weights and path products. Each case fixes an integer weight $q(x)\in\{2,3,4\}$ for every vertex x of the graph to be labeled:

- cubic case: $q\equiv 2$; padded case: $q\equiv 3$; padded case of degree 5, used only for the cut sweep of Theorem 10.3: $q\equiv 4$;

- mixed cases, applied to graphs with $\delta \geq 3$ and $\Delta \leq 5$, and switched case, applied to graphs with $\delta \geq 3$ and $\Delta \leq 4$: $q(x) = d(x) - 1$.

For vertices v, w let $P(v, w)$ be the minimum of the products $\prod_{i=0}^{k-1} q(x_i)$ over all v–w paths $v = x_0, x_1, \ldots, x_k = w$, with $P(v, v) = 1$. The *support* $B(v)$ of v is the ball $B_h(v)$ in the cubic and padded cases (of degree 4 and 5), and $\{w : P(v, w) \leq 2^h\}$ in the mixed and switched cases. Since $q \geq 2$, a path with k edges has product at least 2^k; hence every $w \in B(v)$ satisfies $\operatorname{dist}(v, w) \leq h$, so $B(v) \subseteq B_h(v)$ in all cases.

Ideal labels. Let $(Z_w)_w$ be independent standard normal variables and put

$$X_v := \sum_{w \in B(v)} P(v, w)^{-1/2} Z_w, \qquad \sigma_v^2 := \operatorname{Var} X_v = \sum_{w \in B(v)} \frac{1}{P(v, w)} \ \ (\geq 1), \qquad Y_v := X_v / \sigma_v.$$

These are the labels of Definition 5.2 with truncation radius h.

- In the cubic case $P(v, w) = 2^{\operatorname{dist}(v,w)}$, so the coefficients are the geometric coefficients $2^{-\operatorname{dist}(v,w)/2}$, truncated at distance h.
- In the padded case $P(v, w) = 3^{\operatorname{dist}(v,w)}$, so the coefficients are $3^{-\operatorname{dist}(v,w)/2}$, truncated at distance h; in the padded case of degree 5, $P(v, w) = 4^{\operatorname{dist}(v,w)}$.
- In the mixed and switched cases, the arc length $\ell(x \to y) = \frac{1}{2} \ln(d(x) - 1) = \frac{1}{2} \ln q(x)$ gives the directed distance $D(v, w) = \frac{1}{2} \ln P(v, w)$, so $e^{-D(v,w)} = P(v, w)^{-1/2}$, and $P(v, w) \leq 2^h$ is the truncation $D(v, w) \leq R$ with $R := \frac{h}{2} \ln 2$.

This is the path-product form of Remark 5.3. Thus the integer h plays the role of the radius parameter R of Section 5. The cubic and padded cases use the padded labels of Definition 5.2(a) (with $p = 3$ and $p = 4$, and $p = 5$ in the padded case of degree 5) with $R = h$, and the mixed cases use the mixed labels of Definition 5.2(b) with $R = \frac{h}{2} \ln 2$. In this subsection we write h throughout.

The next lemma restates the edge correlation bounds of Section 5 for these labels. For $h \geq 5$ put $\varepsilon_h := 1/h$ in the cubic and padded cases and $\varepsilon_h := 8/(h - 4)$ in the mixed and switched cases, and for an edge uv put $r_{uv} := r_* = r_{33}$ in the cubic case, $r_{uv} := r_{44}$ in the padded case, $r_{uv} := r_{55}$ in the padded case of degree 5, and $r_{uv} := r_{d(u)d(v)}$ in the mixed and switched cases.

Lemma A.22 (edge correlations)**.** *Let $h \geq 5$. Every edge uv satisfies* $\operatorname{Corr}(Y_u, Y_v) \geq r_{uv}(1 - \varepsilon_h)$.

Proof. In the cubic and padded cases this is Lemma 5.4 with radius h (with $p = 3$, 4 and 5). In the mixed and switched cases it is Lemma 5.5 with $\Delta_0 = 5$ and radius $R = \frac{h}{2} \ln 2$: the arc lengths lie between $\ell_{\min} = \frac{1}{2} \ln 2$ and $\ell_{\max} = \frac{1}{2} \ln 4 = \ln 2$, and its error term (5.4) is

$$\varepsilon_R = \frac{2\ell_{\max}^2}{\ell_{\min}(R - 2\ell_{\max})} = \frac{2(\ln 2)^2}{\frac{1}{2} \ln 2 \cdot (\frac{h}{2} - 2) \ln 2} = \frac{8}{h - 4},$$

which is positive and finite for $h > 4$. □

Ball sizes. For a graph with $\Delta \leq 5$, the first bound of (4.1) gives $|B_k(v)| \leq 1 + 5 \sum_{i=0}^{k-1} 4^i \leq 2 \cdot 4^k$ for every $k \geq 0$. With (A.4), for $n \geq n_2$,

$$|B(v)| \leq 2t^{1/4}, \qquad |B_{h+1}(v)| \leq 8t^{1/4}, \qquad |B_{2h+2}(v)| \leq 32\, t^{1/2}. \tag{A.5}$$

Computing the coefficients. For each v, a breadth-first search to depth h finds $B_h(v)$. In the cubic and padded cases $P(v,w) = q^{\mathrm{dist}(v,w)}$. In the mixed and switched cases we run Dijkstra's algorithm from v inside $B_h(v)$, with path products of integers in place of path sums (all factors are at least 1, so Dijkstra's algorithm applies). This computes $P(v,w)$ correctly for every w with $P(v,w) \le 2^h$: a minimizing path has product at most 2^h, hence at most h edges, so all its vertices lie in $B_h(v)$. For the other vertices of $B_h(v)$ the computed value is at least the true value and so also exceeds 2^h.

The algorithms use rational roundings of the coefficients and of the standard deviations. Put

$$\tilde{c}_{vw} := \frac{\lfloor t^3 P(v,w)^{-1/2} \rfloor}{t^3} \quad (w \in B(v)), \qquad \tilde{\sigma}_v := \frac{\lfloor t^3 \sigma_v \rfloor}{t^3}.$$

Both are computed with integer square roots, using $\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor$ for $x \ge 0$: $\lfloor t^3 P^{-1/2} \rfloor = \lfloor \sqrt{\lfloor t^6/P \rfloor} \rfloor$ and $\lfloor t^3 \sigma_v \rfloor = \lfloor \sqrt{\lfloor t^6 \sigma_v^2 \rfloor} \rfloor$, where $\sigma_v^2 = \sum_{w \in B(v)} 1/P(v,w)$ is an exact rational. Then

$$0 \le \tilde{c}_{vw} \le 1, \qquad |\tilde{c}_{vw} - P(v,w)^{-1/2}| \le t^{-3}, \qquad \tilde{\sigma}_v \ge 1, \qquad |\tilde{\sigma}_v - \sigma_v| \le t^{-3}, \tag{A.6}$$

because $P \ge 1$ and $\sigma_v \ge 1$. The numbers $P(v,w) \le 4^h \le t^{1/4}$ and the numerators of $\tilde{c}_{vw}, \tilde{\sigma}_v$ have $O(\log t)$ bits. For $n \ge n_2$ the work per vertex is $t^{O(1)}$ bit operations by (A.5); for $n < n_2$ it is bounded by a constant. In total, computing the coefficients takes $n(\log n)^{O(1)}$ bit operations.

A.5.2 A finite distribution for the Gaussian variables

The algorithms replace each Z_w by a random variable with finitely many rational values and rational probabilities, so that labels can be sampled and compared, and conditional expectations computed, in exact arithmetic. This subsection defines that distribution, couples it with the standard normal distribution, and defines the discrete labels that the algorithms sort.

The distribution. The finite distribution puts on each point of a grid of step γ in $[-t,t]$ a weight proportional to a rational approximation of the standard normal density. For $|j| \le t^3$ put $x_j := j\gamma$, so that $|x_j| \le t$; there are $M = 2t^3 + 1$ such points. Let $g_j := 1/T_j$, where $T_j = \sum_{k<4t^2} (x_j^2/2)^k/k!$ is the Taylor polynomial of $e^{x_j^2/2}$ with $4t^2$ terms. Since $y := x_j^2/2 \le t^2/2$ and $k! \ge (k/e)^k$, the omitted terms satisfy

$$0 \le e^y - T_j \le \sum_{k \ge 4t^2} \Big(\frac{ey}{k}\Big)^k \le \sum_{k \ge 4t^2} \Big(\frac{e}{8}\Big)^k \le 2 \cdot 0.34^{4t^2},$$

so $T_j = e^y(1-\delta_j)$ with $0 \le \delta_j \le 2 \cdot 0.34^{4t^2} \le \frac{1}{2}$, and $|g_j - e^{-y}| = e^{-y}\delta_j/(1-\delta_j) \le 4 \cdot 0.34^{4t^2} \le t^{-8}$ for $t \ge 2$. Put

$$A_j := \lfloor t^8 g_j \rfloor, \qquad A := \sum_{|i| \le t^3} A_i,$$

and let ζ be the random variable with $\mathbb{P}(\zeta = x_j) = A_j/A$. Since $g_0 = 1$, we have $A_0 = t^8 > 0$. Every A_j is an integer in $[0, t^8]$ and $A \le 3t^{11}$, so the probabilities have $O(\log t)$ bits. Computing T_j exactly takes $t^{O(1)}$ bit operations (it is a sum of $4t^2$ rationals whose numerators and denominators have $O(t^2 \log t)$ bits), so computing the whole distribution takes $t^{O(1)} = (\log n)^{O(1)}$ bit operations.

Lemma A.23 (coupling)**.** *Let $t \ge 5$. There is a coupling of ζ with a standard normal variable Z such that $\mathbb{P}(|Z - \zeta| > \gamma) \le t^{-3}$.*

Proof. We round Z to the grid, show that the distribution of the rounding is within total variation distance $t^{-3}/10$ of the distribution of ζ, and then glue a maximal coupling of the two distributions to Z.

Let $\hat{Z}$ be the rounding of Z to the grid: $\hat{Z} = x_j$ if $Z \in [x_j - \frac{\gamma}{2}, x_j + \frac{\gamma}{2})$ for some $|j| \le t^3$, and $\hat{Z} = \star$ (an extra point) otherwise. Write ν for the distribution of $\hat{Z}$, so $\nu(x_j) = p_j := \mathbb{P}(Z \in [x_j - \frac{\gamma}{2}, x_j + \frac{\gamma}{2}))$, and $p_{\text{out}} := \nu(\star) \le \mathbb{P}(|Z| \ge t) \le e^{-t^2/2}$.

Midpoint rule. Put $\bar{q}_j := \gamma\varphi(x_j)$. Since $\varphi''(z) = (z^2 - 1)\varphi(z)$ satisfies $\sup|\varphi''| = \varphi(0) < 0.4$, the midpoint rule gives $|p_j - \bar{q}_j| \le \frac{\gamma^3}{24} \cdot 0.4 = \frac{t^{-6}}{60}$. Summed over the $M \le 3t^3$ points this is at most $t^{-3}/20$.

Rational weights. Put $\tilde{a}_j := \gamma A_j/(t^8\sqrt{2\pi})$. Since $|A_j/t^8 - g_j| \le t^{-8}$ and $|g_j - e^{-x_j^2/2}| \le t^{-8}$, we have $|A_j/t^8 - e^{-x_j^2/2}| \le 2t^{-8}$, so $|\tilde{a}_j - \bar{q}_j| \le \gamma \cdot 2t^{-8}/\sqrt{2\pi} \le t^{-10}$; summed over j this is at most $3t^{-7}$.

Total variation. Hence

$$\sum_j |\tilde{a}_j - p_j| + p_{\text{out}} \le e_t := \frac{t^{-3}}{20} + 3t^{-7} + e^{-t^2/2} \le \frac{t^{-3}}{10} \qquad (t \ge 5),$$

because $3t^{-7} \le t^{-3}/40$ and $e^{-t^2/2} \le t^{-3}/40$ for $t \ge 5$. The distribution of ζ is $(\tilde{a}_j/\sum_i \tilde{a}_i)_j$, since $\tilde{a}_j$ is proportional to A_j. Normalizing changes the ℓ_1 distance to ν by at most $|1 - \sum_i \tilde{a}_i| = |\sum_i(\tilde{a}_i - p_i) - p_{\text{out}}| \le e_t$. So the ℓ_1 distance between the distribution of ζ and ν (on the points x_j and $\star$) is at most $2e_t$, and the total variation distance is at most $e_t \le t^{-3}/10$.

Coupling. Take a maximal coupling $(\hat{Z}', \zeta)$ of ν and the distribution of ζ, so $\mathbb{P}(\hat{Z}' \ne \zeta) \le t^{-3}/10$, and glue it to $(Z, \hat{Z})$ along $\hat{Z} = \hat{Z}'$ (draw Z from its conditional distribution given $\hat{Z} = \hat{Z}'$). On the event $\zeta = \hat{Z}$, the common value is a grid point x_j, because ζ never equals $\star$, and then $|Z - \zeta| \le \gamma/2$. Hence $\mathbb{P}(|Z - \zeta| > \gamma) \le \mathbb{P}(\zeta \ne \hat{Z}) \le t^{-3}$. □

Discrete labels. Let $(\tilde{Z}_w)_w$ be independent copies of ζ and

$$\tilde{X}_v := \sum_{w \in B(v)} \tilde{c}_{vw}\tilde{Z}_w, \qquad \tilde{Y}_v := \tilde{X}_v/\tilde{\sigma}_v.$$

Each $\tilde{Y}_v$ is a rational of the form $a/(b\,t^2)$ with integers a, b of $O(\log t)$ bits, so two labels, or a label and a rational threshold with $O(\log t)$ bits, are compared in $(\log t)^{O(1)}$ bit operations. The algorithms order the vertices by $(\tilde{Y}_v, v)$ lexicographically; ties between labels are broken by the vertex index. For a rational τ write

$$\tilde{S}_\tau := \{v : \ \tilde{Y}_v < \tau\},$$

a prefix of this order. Note the strict inequality.

Sampling. To sample ζ, draw $\lceil \log_2 A \rceil$ random bits, forming an integer U; if $U \ge A$, draw again; otherwise return the x_j with $\sum_{i<j} A_i \le U < \sum_{i \le j} A_i$, found by binary search in the prefix sums. A draw succeeds with probability $A/2^{\lceil \log_2 A \rceil} > \frac{1}{2}$, so a sample needs fewer than two draws in expectation, each of $O(\log t)$ bits and $O(\log M)$ comparisons. To keep the running time of the randomized algorithms bounded in the worst case, they draw at most $2N\lceil \log_2 n \rceil$ times in total for N samples; if this budget is exhausted they output the trivial decomposition with one bag $V(G)$. By Markov's inequality the budget is exhausted with probability less than $1/\lceil \log_2 n \rceil$. The Las Vegas algorithms have no budget.

A.5.3 Perturbation

We couple the discrete labels with the ideal labels so that, in the closed neighborhood of any fixed vertex, all labels move by at most ξ except with probability at most p_t (Lemma A.24). As a consequence, the per-vertex probabilities that bound the counts transfer from the ideal to the discrete labels, up to a shift of the threshold by ξ and an additive error that tends to zero as $n \to \infty$ (Corollary A.25). So the analysis of the algorithms can use bounds on the ideal labels, which Appendix A.5.4 derives from Part I.

Lemma A.24 (perturbation)**.** *Let $n \geq n_2$, and let H be a graph with $\Delta(H) \leq 5$ whose labels are built with the parameters of n. There is a coupling of $(\tilde{Y}_v)_v$ with the ideal labels $(Y_v)_v$ such that every vertex v satisfies*

$$\mathbb{P}\big(|\tilde{Y}_u - Y_u| > \xi \text{ for some } u \in N[v]\big) \leq p_t.$$

Proof. Couple each pair $(\tilde{Z}_w, Z_w)$ by Lemma A.23 ($t \geq 2^{40} \geq 5$), independently over w. The labels in $N[v]$ use at most $|N[v]| \cdot 2t^{1/4} \leq 12t^{1/4}$ variables by (A.5), so with probability at least $1 - 12t^{1/4}t^{-3} = 1 - p_t$ all of them satisfy $|\tilde{Z}_w - Z_w| \leq \gamma = t^{-2}$. On this event, for $u \in N[v]$, using (A.6), $|\tilde{Z}_w| \leq t$ and $|B(u)| \leq 2t^{1/4}$,

$$\begin{aligned} |\tilde{X}_u - X_u| &\leq \sum_{w \in B(u)} \Big(|\tilde{c}_{uw} - P(u,w)^{-1/2}| \, |\tilde{Z}_w| + P(u,w)^{-1/2} |\tilde{Z}_w - Z_w| \Big) \\ &\leq 2t^{1/4}(t^{-2} + t^{-2}) = 4t^{-7/4}. \end{aligned}$$

Also $|\tilde{X}_u| \leq 2t^{1/4} \cdot t$, and $|1/\tilde{\sigma}_u - 1/\sigma_u| = |\sigma_u - \tilde{\sigma}_u|/(\sigma_u \tilde{\sigma}_u) \leq t^{-3}$ since $\sigma_u, \tilde{\sigma}_u \geq 1$. Hence

$$|\tilde{Y}_u - Y_u| \leq \frac{|\tilde{X}_u - X_u|}{\sigma_u} + |\tilde{X}_u| \Big| \frac{1}{\tilde{\sigma}_u} - \frac{1}{\sigma_u} \Big| \leq 4t^{-7/4} + 2t^{5/4}t^{-3} = 6t^{-7/4} \leq \xi. \qquad \square$$

The coupling transfers per-vertex probabilities from ideal to discrete labels. Recall that $S_\tau = \{v : Y_v \leq \tau\}$ for the ideal labels, and $\tilde{S}_\tau = \{v : \tilde{Y}_v < \tau\}$ for the discrete ones.

Corollary A.25 (transfer)**.** *In the setting of Lemma A.24, let v be a vertex of degree d and $\tau \in \mathbb{R}$. Then:*

(a) $\mathbb{P}(v \in \partial^+ \tilde{S}_\tau) \leq \mathbb{P}(v \in \partial^+ S_{\tau - \xi}) + \omega_d$;

(b) $\mathbb{P}(v \in \partial^- \tilde{S}_\tau) \leq \mathbb{P}(v \in \partial^- S_{\tau + \xi}) + \omega_d$;

(c) *for every family $\mathcal{A}$ of subsets of $N[v]$, $\mathbb{P}(\{u \in N[v] : \tilde{Y}_u < 0\} \in \mathcal{A}) \leq \mathbb{P}(\{u \in N[v] : Y_u < 0\} \in \mathcal{A}) + \omega_d$;*

(d) *for $\tau \leq \tau'$: $\mathbb{P}(\tau \leq \tilde{Y}_v < \tau') \leq (\tau' - \tau + 2\xi)\varphi(0) + p_t$, and $\mathbb{P}(\tilde{Y}_v < \tau) \leq \Phi(\tau) + \xi\varphi(0) + p_t$, $\mathbb{P}(\tilde{Y}_v \geq \tau) \leq \Phi(-\tau) + \xi\varphi(0) + p_t$.*

Proof. Let $\mathcal{E}$ be the event that $|\tilde{Y}_u - Y_u| \leq \xi$ for all $u \in N[v]$, and $\mathcal{W}$ the event that no Y_u with $u \in N[v]$ lies in $[\tau - \xi, \tau + \xi]$. Each Y_u is standard normal, so $\mathbb{P}(\mathcal{W}^c) \leq (d+1) \cdot 2\xi\varphi(0)$, and $\mathbb{P}(\mathcal{E}^c) \leq p_t$ by Lemma A.24. Hence $\mathbb{P}(\mathcal{E}^c \cup \mathcal{W}^c) \leq \omega_d$, and it suffices to prove the inclusions (a)–(c) on $\mathcal{E} \cap \mathcal{W}$.

(a) Let $v \in \partial^+ \tilde{S}_\tau$: $\tilde{Y}_v \geq \tau$ and some neighbor u has $\tilde{Y}_u < \tau$. Then $Y_v \geq \tau - \xi$, so $Y_v > \tau + \xi > \tau - \xi$ on $\mathcal{W}$; and $Y_u < \tau + \xi$, so $Y_u < \tau - \xi$ on $\mathcal{W}$. Thus $v \in \partial^+ S_{\tau - \xi}$.

(b) Let $v \in \partial^- \tilde{S}_\tau$: $\tilde{Y}_v < \tau$ and some neighbor x has $\tilde{Y}_x \geq \tau$. Then $Y_v < \tau + \xi$, and $Y_x \geq \tau - \xi$, so $Y_x > \tau + \xi$ on $\mathcal{W}$. Thus $v \in \partial^- S_{\tau + \xi}$.

(c) Take $\tau = 0$. On $\mathcal{E} \cap \mathcal{W}$, every $u \in N[v]$ has $|Y_u| > \xi \geq |\tilde{Y}_u - Y_u|$, so $\tilde{Y}_u < 0$ if and only if $Y_u < 0$. Hence the two sets in (c) coincide on $\mathcal{E} \cap \mathcal{W}$.

(d) uses only the event $|\tilde{Y}_v - Y_v| \leq \xi$, which fails with probability at most p_t: on it, $\tau \leq \tilde{Y}_v < \tau'$ implies $\tau - \xi \leq Y_v < \tau' + \xi$, an event of probability at most $(\tau' - \tau + 2\xi)\varphi(0)$; $\tilde{Y}_v < \tau$ implies $Y_v < \tau + \xi$; and $\tilde{Y}_v \geq \tau$ implies $Y_v \geq \tau - \xi$. □

A.5.4 Uniform ideal means

This subsection bounds the expected frontier sizes of the ideal labels uniformly in the threshold (Lemma A.26). The counts of the layouts are taken at the thresholds of a grid whose step $1/K$ tends to zero with n, and Corollary A.25 shifts these thresholds by ξ. We therefore need the bounds of Part I on expected frontier sizes for all thresholds in an interval at once, with an error that tends to zero as $h \to \infty$. In the cubic, padded and mixed cases the bounds of Part I have this form. In the cubic and padded cases the per-vertex bounds hold for every threshold. In the mixed cases Lemma 7.7 holds for every threshold in $[-3, 3]$, because the continuity of the center bound (Lemma 6.3) turns the factor $1 - \varepsilon_R$ in the correlation bounds into an additive error $205\,\varepsilon_R$ per vertex. The lemma below restates these bounds for the ideal labels of radius h.

In the lemma, $r_h := r_*(1 - 1/h)$, where $r_* = r_{33}$ is the edge correlation bound of the cubic case, and $\varepsilon_{\mathrm{mix}}(h) := 205\,\varepsilon_h = 1640/(h-4)$, where $\varepsilon_h = 8/(h-4)$ is the error term of Lemma A.22 in the mixed cases.

Lemma A.26 (uniform ideal means)**.** *Let $h \geq 5$ and let (Y_v) be the ideal labels of radius h of a graph H with N vertices.*

(a) Cubic case, $\Delta(H) \leq 3$. *Every vertex v satisfies $\mathbb{P}(v \in \partial^+ S_\tau) \leq \kappa(r_h)$ for every $\tau \geq 0$, $\mathbb{P}(v \in \partial^- S_\tau) \leq \kappa(r_h)$ for every $\tau \leq 0$, and*

$$\mathbb{P}\big(Y_v < 0,\ \exists u \in N(v) : Y_u \geq 0\big) + \mathbf{1}_{d(v)=3}\mathbb{P}\big(Y_v \geq 0,\ Y_u < 0\ \forall u \in N(v)\big) \leq \kappa(r_h),$$
$$\mathbb{P}\big(Y_v \geq 0,\ \exists u \in N(v) : Y_u < 0\big) + \mathbf{1}_{d(v)=3}\mathbb{P}\big(Y_v < 0,\ Y_u \geq 0\ \forall u \in N(v)\big) \leq \kappa(r_h).$$

(b) Padded case, $\Delta(H) \leq 4$. *For every $\tau \in \mathbb{R}$, $\mathbb{E}|\partial^+ S_\tau| \leq (F_4 + 178/h)\,N$.*

(c) Mixed cases, $\delta(H) \geq 3$ *and* $\Delta(H) \leq 5$. *For every $\tau \in [-3, 3]$, $\mathbb{E}|\partial^+ S_\tau| \leq \beta(H) + \varepsilon_{\mathrm{mix}}(h)\,N$.*

Proof. (a) By Lemma A.22, the labels of v and its $d(v) \leq 3$ neighbors form a centered Gaussian vector with unit variances and incident correlations at least r_h. For $\tau \geq 0$, $\{v \in \partial^+ S_\tau\} \subseteq \{Y_v \geq \tau,\ \exists u \in N(v) : Y_u \leq \tau\}$, whose probability is at most $\kappa(r_h)$ by Proposition 8.8. For $\tau \leq 0$, apply Proposition 8.8 to the labels $-Y$, which have the same incident correlations, at the threshold $-\tau \geq 0$: the event $v \in \partial^- S_\tau$ is contained in $\{-Y_v \geq -\tau,\ \exists x \in N(v) : -Y_x \leq -\tau\}$. The first of the two displayed inequalities is Lemma 8.9, because its events differ from those of Lemma 8.9 only by events on which some label equals 0, which have probability zero. The second is Lemma 8.9 applied to $-Y$, in the same way.

(b) By Lemma A.22, the labels of v and its $d(v) \leq 4$ neighbors form a vector as in Corollary 6.8 with $d = 4$ and $r = r_{44}(1 - 1/h)$. So $\mathbb{P}(v \in \partial^+ S_\tau) \leq f_{4,r_{44}(1-1/h)}(\tau)$. For $|\tau| \leq 3$ this is at most $f_{4,r_{44}}(\tau) + 205\,r_{44}/h \leq F_4 + 178/h$ by (6.6), Lemma 6.3(b) with $\bar{r} = r_{33}$ and Definition 6.7; for $|\tau| > 3$ it is at most $4\Phi(-3) < F_4$ (Proposition 6.9). Sum over v.

(c) If $h \leq 12$, then $\varepsilon_{\mathrm{mix}}(h) \geq 205$ and the claim is trivial, since $\mathbb{E}|\partial^+ S_\tau| \leq N$. Let $h \geq 13$, so that $\varepsilon_h < 1$. In the mixed cases the ideal labels are the mixed labels of Definition 5.2(b) with $R = \frac{h}{2}\ln 2$, and $\varepsilon_R = \varepsilon_h$ for this R (proof of Lemma A.22). The triple of the case is one of $\mathsf{P}, \mathsf{R}, \mathsf{S}$, and all degrees of H lie in $\{3, 4, 5\}$. So Lemma 7.7 gives $\mathbb{E}|\partial^+ S_\tau| \leq \beta(H) + 205\,\varepsilon_h N = \beta(H) + \varepsilon_{\mathrm{mix}}(h)\,N$ for every $\tau \in [-3, 3]$. □

A.5.5 Counts and derandomization

This subsection shows that the counts are close to their means with high probability, and that the method of conditional expectations finds discrete labels for which they are close to their means for certain. The randomized and Las Vegas algorithms use the first statement, and the deterministic algorithm uses the second.

Counts. Let H be a graph with N vertices and $\Delta(H) \le 5$, labeled with the parameters of some $n \ge \max\{N, n_2\}$. A *count* is a random variable $X_c = \sum_{v \in V(H)} I_{c,v}$, where each $I_{c,v} \in \{0,1\}$ is a function of $(\tilde{Z}_w)_{w \in B_{h+1}(v)}$ that can be evaluated from these values in $t^{O(1)}$ bit operations. For example, $\mathbf{1}\{v \in \partial^+ \tilde{S}_\tau\}$ is such an indicator for every rational τ with $O(\log t)$ bits, because it depends on the labels in $N[v]$, and $\tilde{Y}_u$ depends on $\tilde{Z}_w$ for $w \in B(u) \subseteq B_h(u) \subseteq B_{h+1}(v)$.

Greedy fixing. Let $X_1, \dots, X_\mu$ be counts, $m_c := \mathbb{E}X_c$, and

$$\Psi := \sum_{c=1}^{\mu} (X_c - m_c)^2.$$

Greedy fixing is the following deterministic procedure: fix $\tilde{Z}_w$ for $w \in V(H)$ one at a time, each time choosing a value x_j with $A_j > 0$ that minimizes the conditional expectation of Ψ given the values fixed so far.

Lemma A.27 (concentration and derandomization)**.** *In this setting the following hold.*

(a) $\mathbb{E}\Psi = \sum_c \operatorname{Var} X_c \le 8\mu N t^{1/2}$.

(b) *For every $\lambda > 0$, the probability that $X_c > m_c + \lambda$ for some c is at most $8\mu N t^{1/2}/\lambda^2$.*

(c) *Greedy fixing ends with values of all $\tilde{Z}_w$, each of positive probability, for which $\Psi \le 8\mu N t^{1/2}$, and hence $|X_c - m_c| \le \sqrt{8\mu N}\, t^{1/4}$ for every c.*

(d) *The means m_c can be computed exactly with $\mu N\, n^{o(1)}$ bit operations, and greedy fixing uses $\mu N^2 n^{o(1)}$ bit operations.*

Proof. (a) Apply Lemma 4.10 with $a = 1$ and $A_v = B_{h+1}(v)$: two indicators $I_{c,u}$ and $I_{c,v}$ can be dependent only if $B_{h+1}(u) \cap B_{h+1}(v) \ne \emptyset$, that is, if $v \in B_{2h+2}(u)$. By (A.5) there are at most $32t^{1/2}N$ such ordered pairs, so $\operatorname{Var} X_c \le 8Nt^{1/2}$, and $\mathbb{E}\Psi$ is the sum of these variances.

(b) Chebyshev's inequality and a union bound over the μ counts.

(c) Let $\mathcal{F}$ be the values fixed before the step that fixes $\tilde{Z}_w$. Then

$$\mathbb{E}[\Psi \mid \mathcal{F}] = \sum_j \frac{A_j}{A}\, \mathbb{E}[\Psi \mid \mathcal{F},\ \tilde{Z}_w = x_j],$$

an average over the values of positive probability, so the minimizing value chosen by greedy fixing does not increase the conditional expectation. Initially it is $\mathbb{E}\Psi \le 8\mu N t^{1/2}$ by (a). At the end all variables are fixed and Ψ is a number at most $8\mu N t^{1/2}$; each of its terms $(X_c - m_c)^2$ is at most Ψ.

(d) Given fixed values $\mathcal{F}$,

$$\mathbb{E}\big[(X_c - m_c)^2 \mid \mathcal{F}\big] = \Big(\sum_v \mathbb{E}[I_{c,v} \mid \mathcal{F}] - m_c\Big)^2 + \sum_u \sum_{v \in B_{2h+2}(u)} \operatorname{Cov}(I_{c,u}, I_{c,v} \mid \mathcal{F}),$$

because the unfixed variables are independent, so the conditional covariance vanishes when $B_{h+1}(u) \cap B_{h+1}(v) = \emptyset$. Each term involves at most $|B_{h+1}(u) \cup B_{h+1}(v)| \le 16t^{1/4}$ variables by (A.5), and is computed exactly by enumerating the at most $M^{16t^{1/4}}$ assignments of the unfixed ones among them, with probabilities that are products of the rationals A_j/A. Since $M \le 3t^3$ and $t \le 2\log_2 n$,

$$M^{16t^{1/4}} = \exp\big(O(t^{1/4}\log t)\big) = \exp\big(O((\log n)^{1/4}\log\log n)\big) = n^{o(1)}.$$

All numbers involved are rationals whose denominators are powers of A with exponent $O(t^{1/4})$, and whose numerators have $n^{o(1)}$ bits; each indicator is evaluated in $t^{O(1)}$ bit operations. So each term takes $n^{o(1)}$ bit operations, and one evaluation of $\mathbb{E}[\Psi \mid \mathcal{F}]$, with its $O(\mu N t^{1/2})$ terms, takes $\mu N\, n^{o(1)}$. The means m_c are computed in the same way with nothing fixed. Greedy fixing makes N steps with at most M candidates each, which gives $\mu N^2 n^{o(1)}$ bit operations. □

In all our layouts $\mu = O(K) = O(\log\log n)$, so the bounds in (d) become $N\, n^{o(1)}$ and $N^2 n^{o(1)}$ bit operations. Below we use only upper bounds of the form $X_c \le m_c + \lambda$, in the following form.

The three algorithms. Let $n \ge n_2$, $N \le n$ and $\mu \le 12K+3$, and put $\theta_{\mathrm{D}}(n) := \sqrt{8\mu n}\, t^{1/4}$ and $\theta_{\mathrm{R}}(n) := n^{3/4}$. By Lemma A.27(c), greedy fixing ends with every count at most its mean plus $\theta_{\mathrm{D}}(n)$. By Lemma A.27(b) with $\lambda = n^{3/4}$, one sample has every count at most its mean plus $\theta_{\mathrm{R}}(n)$ except with probability at most $8\mu t^{1/2} n^{-1/2}$, which is at most $\frac{1}{2}$ by (N3). So the randomized algorithm, with its sampling budget, fails with probability at most $\pi(n) := 8(12K+3)t^{1/2}n^{-1/2} + 1/\lceil \log_2 n \rceil$ (as $\mu \le 12K+3$), and the Las Vegas algorithm, which resamples until every count is at most its mean plus $n^{3/4}$, needs at most two samples in expectation. In the padded, mixed and cubic cases below, it therefore suffices to bound the width when every count is at most its mean plus $\theta \in \{\theta_{\mathrm{D}}(n), \theta_{\mathrm{R}}(n)\}$.

A.5.6 Undoing reductions

In the switched, mixed, bootstrap and general cases the algorithm first applies the reduction operations of Definition 4.2, possibly interleaved with vertex deletions, and decomposes the final graph. The proof of Lemma 4.5 constructs from this a decomposition of the original graph. This subsection shows that the construction can be carried out in near-linear time (Lemma A.28). If the final graph is empty, its path decomposition has no bags and width -1.

Lemma A.28 (undoing reductions)**.** *Let F, D and R be as in Lemma 4.5, where F has n vertices and m edges. Given F, D, $V(R)$ and a path decomposition of R of width w with s bags, a path decomposition of F of width at most $|D| + w + 3\log_2 n + 1$ with at most $s+n$ bags can be computed with $O(m + s + n\log n)$ operations on integers of $O(\log n)$ bits, where both decompositions are in interval form.*

The algorithm does not need the sequence of operations itself; the sequence only guarantees the structure of Lemma 4.5(b). We first record two standard facts, in the form in which the algorithm uses them. Reduction operations take a graph of treewidth at most 2 to the empty graph and yield a tree decomposition of width at most 2 (Claim A.29). Every tree decomposition of a graph with N' vertices has a bag whose removal leaves components of at most $N'/2$ vertices, and such a bag can be found in time linear in the size of the decomposition (Claim A.30).

Elimination bags. Suppose that reduction operations are applied to a graph Q until none applies, and that the final graph is empty. Let $u_1, u_2, \dots$ be the vertices of Q in the order of their removal, let Γ_i be the set of neighbors of u_i in the current graph when u_i is removed, and put $X_i := \{u_i\} \cup \Gamma_i$.

The parent of node i is the node j of smallest index with $u_j \in \Gamma_i$ (then $j > i$), and node i is a root if $\Gamma_i = \emptyset$. We call the sets X_i, with this forest, the *elimination bags* of the sequence.

Claim A.29. *Let Q be a graph with $\mathrm{tw}(Q) \le 2$. Applying reduction operations to Q in any order until none applies ends with the empty graph, and the elimination bags, with the roots of their forest joined into a tree, form a tree decomposition of Q of width at most 2.*

Proof. First, the process ends with the empty graph. Each operation produces a minor of the current graph, since (R3) contracts the edge vx and the others delete a vertex; so every current graph has treewidth at most 2. A nonempty graph of treewidth at most k has a vertex of degree at most k: in a tree decomposition of width at most k in which no bag is contained in a neighboring bag (contract tree edges violating this), either there is one bag, or a leaf bag contains a vertex outside its neighbor's bag, and all neighbors of that vertex lie in the leaf bag. By (RP3), every vertex of degree at most 2 admits an operation, so the process continues until the graph is empty.

Next, we show that the elimination bags form a tree decomposition. In every operation, Γ_i is a clique of the graph after the operation: it has at most one vertex, or it is $\{x, y\}$ and xy is present in (R4) and added in (R3). Every edge of Q joins u_i to Γ_i, where u_i is its end removed first, because edges disappear only with their ends; so it lies in X_i. Let $u_a \in \Gamma_i$ with $a > i$, and let j be the parent of i. If $j \ne a$, then $j < a$, and $u_a, u_j \in \Gamma_i$ are adjacent after step i; this edge survives until step j, so $u_a \in \Gamma_j$. Repeating, the nodes containing u_a include a path from i to a in the forest, so the nodes containing any vertex form a connected subtree. Joining the roots by arbitrary edges keeps this property. □

Claim A.30. *Let Q be a graph with $N' \ge 1$ vertices and let $(\mathcal{T}, (X_\nu)_\nu)$ be a tree decomposition of Q. One can find, in time linear in $|V(\mathcal{T})| + \sum_\nu |X_\nu|$, a node ν such that every component of $Q - X_\nu$ has at most $N'/2$ vertices.*

Proof. Root $\mathcal{T}$, and assign each vertex x to the node $\mathrm{top}(x)$ closest to the root among the nodes whose bags contain x. For a node ν let $\omega(\nu)$ be the number of vertices assigned to the subtree $\mathcal{T}_\nu$ rooted at ν. Both are computed by one traversal. For a child ν' of ν, the vertices assigned to $\mathcal{T}_{\nu'}$ are exactly the vertices that occur in $\mathcal{T}_{\nu'}$ but not in X_ν: the nodes containing a vertex form a subtree, so a vertex occurring both in $\mathcal{T}_{\nu'}$ and outside it occurs in X_ν. Similarly, the vertices outside X_ν not assigned to $\mathcal{T}_\nu$ occur only outside $\mathcal{T}_\nu$. A component of $Q - X_\nu$ is connected and avoids X_ν, so all its vertices occur in the same component of $\mathcal{T} - \nu$; hence it has at most $\omega(\nu')$ vertices for some child ν', or at most $N' - \omega(\nu)$ vertices.

To find ν, start at the root, and while the current node has a child ν' with $\omega(\nu') > N'/2$, move to that child. At the final node ν, all children have $\omega(\nu') \le N'/2$, and $N' - \omega(\nu) < N'/2$ (if ν is not the root, we moved to ν because $\omega(\nu) > N'/2$; if ν is the root, $\omega(\nu) = N'$). □

Proof of Lemma A.28. For each component C of $(F - D) - V(R)$ put $S_C := N_{F-D}(C) \subseteq V(R)$. By Lemma 4.5(b), S_C is empty, a single vertex, or the two ends of an edge of R, and $\mathrm{tw}(F[C]) \le 2$. The algorithm decomposes each $F[C]$ separately, with width $O(\log n)$, inserts this decomposition, with that bag added to each of its bags, after a bag of the decomposition of R that contains S_C (at the end if $S_C = \emptyset$), and adds D to every bag. In detail, it has four steps.

(1) *Components.* Compute the components C of $(F - D) - V(R)$ and their sets S_C by breadth-first search, in $O(n + m)$ time.

(2) *Decomposing a component.* Fix C, with $N_C = |C|$. Apply reduction operations to $F[C]$ in any order, keeping the vertices of degree at most 2 in a queue and testing adjacency for (R3) and

(R4) with a balanced search tree of edges. By Claim A.29 this ends with the empty graph, and its elimination bags form a tree decomposition of $F[C]$ of width at most 2 with N_C nodes. A graph of treewidth at most 2 has fewer than $2N_C$ edges, and every operation adds at most one edge, so this step takes $O(N_C \log n)$ operations.

(3) *Balanced recursion.* Given a connected graph $Q = F[C']$, $C' \subseteq C$, with a tree decomposition of width at most 2, find by Claim A.30 a bag X such that every component of $Q - X$ has at most $|C'|/2$ vertices. For each component C'' of $Q - X$, restrict the tree decomposition to C'': intersect every bag with C'' and keep the nodes whose bags meet C''. These nodes form a subtree, because $Q[C'']$ is connected and the nodes containing the two ends of an edge share a node, so this is a tree decomposition of $Q[C'']$ of width at most 2. Recurse on each component, concatenate the resulting path decompositions, and add X to every bag (if $Q - X$ is empty, output the single bag X). By induction on $|C'|$, the result has width at most $3\lfloor \log_2 |C'| \rfloor$: if $Q - X$ is empty, the width is $|X| - 1$, which is 0 if $|C'| = 1$ and at most 2 otherwise; if not, each C'' has $\lfloor \log_2 |C''| \rfloor \leq \lfloor \log_2 |C'| \rfloor - 1$ because $|C''| \leq |C'|/2$, and adding the at most three vertices of X to bags of width at most $3(\lfloor \log_2 |C'| \rfloor - 1)$ gives the bound. It has at most $|C'|$ bags, since the bags X are nonempty. At one level of the recursion the components are disjoint, and a bag of at most three vertices meets at most three of them, so the work per level is linear in N_C; there are at most $\lfloor \log_2 N_C \rfloor + 1$ levels. In interval form, the vertices of X receive the interval of the whole block produced by the call. This step takes $O(N_C \log N_C)$ operations.

(4) *Insertion.* Let $W_1, \ldots, W_s$ be the decomposition of R. For each C with $S_C \neq \emptyset$ choose the index $i_C := \max_{x \in S_C} \mathrm{first}(x)$; then $S_C \subseteq W_{i_C}$, because S_C is a vertex, or the two ends of an edge of R, whose intervals intersect. Immediately after W_{i_C} insert the bags $W_{i_C} \cup P$ for the bags P of the decomposition of $F[C]$ from step 3, the components with the same index one after another. Append the decompositions of the components with $S_C = \emptyset$ at the end, and finally add D to every bag.

Correctness. The result is a path decomposition of F. Every vertex lies in a bag and D in all bags, which covers the edges at D. An edge of $F - D$ inside $V(R)$ is an edge of R by Lemma 4.5(b), so it lies in some W_i. Any other edge of $F - D$ joins a vertex of some C to a vertex of $C \cup S_C$, so it lies in a bag $W_{i_C} \cup P$ or P. A vertex of C occurs only in bags added for C, and there in consecutive bags, since step 3 gives a path decomposition of $F[C]$. A vertex $x \in V(R)$ occurs in a run of consecutive bags W_i and in the bags inserted after them, since a bag inserted after W_i contains x exactly when $x \in W_i$. As step 3 gives width at most $3 \log_2 n$, the width is at most $|D| + w + 3 \log_2 n + 1$, with at most $s + \sum_C N_C \leq s + n$ bags. (If no bag results, which happens only when $F - D$ is empty, output the single bag D.)

Interval form and running time. Let ℓ_i be the total number of bags inserted after W_i, and $\mathrm{new}(i) := i + \sum_{i' < i} \ell_{i'}$ the new index of W_i. A vertex x of R with interval $[\mathrm{first}(x), \mathrm{last}(x)]$ gets $[\mathrm{new}(\mathrm{first}(x)), \mathrm{new}(\mathrm{last}(x)) + \ell_{\mathrm{last}(x)}]$, because the bags inserted after W_i contain x exactly when $x \in W_i$. A vertex of a component C gets its interval within the block of C, shifted by the position of the block, and every vertex of D gets the interval of all bags. Computing the ℓ_i, the prefix sums and the intervals takes $O(n + s)$ time. Summing the running times of the four steps, using $\sum_C N_C \leq n$, gives $O(m + s + n \log n)$. □

A.5.7 The plain sweep: padded and mixed cases

This subsection proves Theorem 10.1 in the padded and mixed cases. The layout is the one of Lemma 4.1 for the order of the discrete labels, and its width is controlled by frontier sizes on the

grid

$$\tau_j := -3 + \frac{j}{K} \qquad (0 \le j \le 6K),$$

together with the numbers of labels in the windows between grid points and in the two tails. Lemma A.31 bounds the width by these counts, and Proposition A.32 bounds the width when every count is at most its mean plus some $\theta \ge 0$, using the bounds on the means from Appendices A.5.3 and A.5.4. The algorithms ensure this condition, with high probability or for certain, by Lemma A.27; the proof of Theorem 10.1 in these cases ends the subsection.

The layout. Let H be a graph with N vertices and $\Delta(H) \le 5$, with $\delta(H) \ge 3$ in the mixed cases, labeled with the parameters of some $n \ge N$. Let $v_1, \dots, v_N$ be the vertices in increasing order of $(\tilde{Y}_v, v)$ and $T_i := \{v_1, \dots, v_i\}$. The *plain-sweep layout* consists of the bags $W_i := \{v_i\} \cup \partial^+ T_i$, $1 \le i \le N$. By Lemma 4.1, applied to any injective labeling that induces this order, it is a path decomposition of H of width $\max_i |\partial^+ T_i|$. By the definition of the bags, $v_j \in W_i$ if and only if $i = j$, or $i < j$ and v_j has a neighbor in T_i. So v_j occupies the bags $W_{f_j}, \dots, W_j$, where f_j is the smallest index k with $k = j$ or $v_k \in N(v_j)$, and the interval form is computed in $O(N)$ time from the order.

Counts. The plain sweep uses the $\mu = 12K + 3$ counts

$$\begin{aligned} X_j^{\mathrm{fr}} &:= |\partial^+ \tilde{S}_{\tau_j}| = \big|\{v : \ \tilde{Y}_v \ge \tau_j, \ \exists u \in N(v) : \tilde{Y}_u < \tau_j\}\big| && (0 \le j \le 6K), \\ X_j^{\mathrm{win}} &:= \big|\{v : \ \tau_j \le \tilde{Y}_v < \tau_{j+1}\}\big| && (0 \le j < 6K), \\ X^{\mathrm{lo}} &:= \big|\{v : \ \tilde{Y}_v < -3\}\big|, \qquad X^{\mathrm{hi}} := \big|\{v : \ \tilde{Y}_v \ge 3\}\big|. \end{aligned}$$

Each is a count in the sense of Appendix A.5.5.

The following analogue of Lemma 7.8 works directly with the discrete labels, whose sublevel sets $\tilde{S}_\tau$ are defined by a strict inequality.

Lemma A.31 (grid)**.** *Every i satisfies $|\partial^+ T_i| \le \max\{5X^{\mathrm{lo}}, \ X^{\mathrm{hi}}, \ \max_{0 \le j < 6K}(X_{j+1}^{\mathrm{fr}} + X_j^{\mathrm{win}})\}$.*

Proof. Let $y := \tilde{Y}_{v_i}$. Every vertex of T_i has discrete label at most y, and every vertex outside T_i has discrete label at least y. Let $u \in \partial^+ T_i$, with a neighbor $z \in T_i$; so $\tilde{Y}_u \ge y \ge \tilde{Y}_z$.

- If $y < -3$, then z is counted by X^{lo}. Each such z is a neighbor of at most 5 vertices u, so $|\partial^+ T_i| \le 5X^{\mathrm{lo}}$.
- If $y \ge 3$, then u is counted by X^{hi}.
- Otherwise $\tau_j \le y < \tau_{j+1}$ for some $0 \le j < 6K$. If $\tilde{Y}_u \ge \tau_{j+1}$, then $u \in \partial^+ \tilde{S}_{\tau_{j+1}}$, because $\tilde{Y}_z \le y < \tau_{j+1}$; so u is counted by X_{j+1}^{fr}. Otherwise $\tau_j \le y \le \tilde{Y}_u < \tau_{j+1}$, and u is counted by X_j^{win}. □

Let $\varepsilon(h)$ be the error term of Lemma A.26: $\varepsilon(h) := 178/h$ in the padded case and $\varepsilon(h) := \varepsilon_{\mathrm{mix}}(h)$ in the mixed cases. Put

$$\rho_0(n) := \varepsilon(h) + 14\,\xi\varphi(0) + 2p_t + \frac{\varphi(0)}{K}.$$

Proposition A.32 (plain sweep)**.** *Let $n \ge n_2$ and $\theta \ge 0$. If every count of the plain sweep is at most its mean plus θ, then the plain-sweep layout of H has width at most $\beta(H) + \rho_0(n)N + 5\theta$.*

Proof. Frontier counts. Let $1 \le j \le 6K$. By Corollary A.25(a), $\mathbb{E}X_j^{\mathrm{fr}} \le \mathbb{E}|\partial^+ S_{\tau_j-\xi}| + \omega_5 N$. Since $\tau_j \in [-3+1/K, 3]$ and $\xi < 1/K$ by (N1), the threshold $\tau_j - \xi$ lies in $[-3, 3]$, and Lemma A.26 gives $\mathbb{E}X_j^{\mathrm{fr}} \le \beta(H) + (\varepsilon(h) + 12\xi\varphi(0) + p_t)N$ (in the padded case $\beta(H) = F_4 N$).

Windows. By Corollary A.25(d), $\mathbb{E}X_j^{\mathrm{win}} \le (\varphi(0)/K + 2\xi\varphi(0) + p_t)N$.

Hence $X_{j+1}^{\mathrm{fr}} + X_j^{\mathrm{win}} \le \beta(H) + \rho_0(n)N + 2\theta$ for every $0 \le j < 6K$.

Tails. By Corollary A.25(d) and (N2), $\mathbb{E}X^{\mathrm{lo}}$ and $\mathbb{E}X^{\mathrm{hi}}$ are at most $(\Phi(-3) + \xi\varphi(0) + p_t)N \le 0.0014\,N$. So $5X^{\mathrm{lo}}$ and X^{hi} are at most $0.007N + 5\theta$. In every case $\beta(H) \ge 0.14N$: in the padded case $F_4 > 0.14$, and in the mixed cases every vertex has degree at least 3 and every coefficient is at least 0.14. So $0.007N + 5\theta \le \beta(H) + 5\theta$.

By Lemma A.31, every $|\partial^+ T_i|$, and hence the width, is at most $\beta(H) + \rho_0(n)N + 5\theta$. □

The algorithms. Let G be a graph of the padded or a mixed case, with $n \ge 3$ vertices (graphs with $n \le 2$ are decomposed by one bag). The randomized, Las Vegas and deterministic algorithms consist of the following four steps and differ only in step 3.

(1) In the padded case put $H := G$. In the mixed cases apply reduction operations to G until none applies and let H be the final graph; it is empty or has $\delta(H) \ge 3$ by (RP3), and $\Delta(H) \le \Delta(G)$ by (RP1). If H is empty, go to step 4.

(2) Compute the parameters from n, the coefficients $\tilde{c}_{vw}$ and $\tilde{\sigma}_v$ of H, and the finite distribution.

(3) *Randomized:* sample all $\tilde{Z}_w$, compute the labels, sort, and build the plain-sweep layout. *Las Vegas:* also compute the means m_c of the counts; if $n \ge n_2$ and some count exceeds $m_c + n^{3/4}$, sample again. *Deterministic:* fix the $\tilde{Z}_w$ by greedy fixing (Lemma A.27(c)), then build the layout.

(4) In the mixed cases apply Lemma A.28 to $F = G$, the performed reductions ($D = \emptyset$), and the layout of H (no bags if H is empty).

Proof of Theorem 10.1 in the padded and mixed cases. Let $n \ge n_2$ (for $n < n_2$ the claims are trivial, as explained in Appendix A.5.1). Since operations remove vertices and do not increase degrees by (RP1), and the coefficients are nonnegative and nondecreasing in the degree with $a_0 = a_1 = a_2 = 0$, we have $\beta(H) \le \beta(G)$. Let

$$\theta_{\mathrm{D}}(n) := \sqrt{8(12K+3)n}\, t^{1/4}, \qquad \theta_{\mathrm{R}}(n) := n^{3/4},$$

$$\rho(n) := \rho_0(n)\, n + 5\max\{\theta_{\mathrm{D}}(n), \theta_{\mathrm{R}}(n)\} + 3\log_2 n + 1$$

for $n \ge n_2$, and $\rho(n) := n$ for $n < n_2$; we replace ρ by its running maximum (4.2). Then $\rho(n) = o(n)$: we have $\rho_0(n) \to 0$, because $h \to \infty$, $\varepsilon(h) \to 0$, $\xi, p_t \to 0$ and $K = h$; and $\theta_{\mathrm{D}}(n) = O(\sqrt{n \log\log n}\,(\log n)^{1/4})$.

When every count is at most its mean plus $\theta \in \{\theta_{\mathrm{D}}(n), \theta_{\mathrm{R}}(n)\}$, with $\mu = 12K + 3$ in the paragraph "The three algorithms" of Appendix A.5.5, Proposition A.32 bounds the width of the layout of H by $\beta(G) + \rho_0(n)n + 5\theta$, and Lemma A.28 adds at most $3\log_2 n + 1$; so the width is at most $\beta(G) + \rho(n)$. That paragraph gives the three algorithms, with failure probability $\pi(n) \to 0$ for the randomized one.

The running times are derived in Appendix A.5.10. □

A.5.8 The median-switched layout: cubic case

This subsection proves Theorem 10.1 in the cubic case. The layout is the median-switched layout of Lemma 8.4: inner frontiers below threshold zero, a connector at zero, and outer frontiers above zero. Two points need care. First, the layout must be built for the discrete order, which is a lexicographic order with ties broken by vertex index rather than an order of distinct real labels; an auxiliary injective labeling y handles this. Second, the connector order of Lemma 8.3 must be computed efficiently (Lemma A.33). The counts and their means then bound the width as in Appendix A.5.7.

Let G be a graph with $\Delta(G) \le 3$ and $n \ge 3$ vertices; vertices of degree 0, 1 and 2 are allowed and no reductions are applied. The three algorithms are those of Appendix A.5.7 with $H := G$ and without steps 1 and 4, and with the layout below in place of the plain-sweep layout and the counts below in place of its counts. The labels use the weight $q \equiv 2$, and the counts use the grid

$$\tau_j := -2 + \frac{j}{K} \qquad (0 \le j \le 4K), \qquad \text{so that } \tau_{2K} = 0.$$

An injective labeling for the discrete order. Let Σ be the finite set $\{\tilde{Y}_v : v \in V(G)\} \cup \{0\} \cup \{\tau_j : 0 \le j \le 4K\}$, let $\vartheta > 0$ be so small that ϑn is less than every positive difference of two elements of Σ, and put

$$y_v := \tilde{Y}_v + \vartheta v.$$

The algorithms never compute y; it is used only in the analysis.

- *y is injective and induces the order of $(\tilde{Y}_v, v)$.* If $\tilde{Y}_u < \tilde{Y}_v$, then $\tilde{Y}_v - \tilde{Y}_u > \vartheta n > \vartheta(u - v)$, so $y_u < y_v$; if $\tilde{Y}_u = \tilde{Y}_v$ and $u < v$, then $y_u < y_v$.
- *Sublevel sets at the grid.* For every $\tau \in \Sigma$,

$$S_\tau = \{v : \ y_v \le \tau\} = \tilde{S}_\tau = \{v : \ \tilde{Y}_v < \tau\}.$$

  If $\tilde{Y}_v < \tau$, then $\tau - \tilde{Y}_v > \vartheta n \ge \vartheta v$, so $y_v < \tau$; if $\tilde{Y}_v \ge \tau$, then $y_v > \tilde{Y}_v \ge \tau$, since $v \ge 1$.

Here and below, S_τ, $\partial^- S_\tau$, $\partial^+ S_\tau$, U, W, D, U', W', L_0 and L_0' are as in Section 8.1, for the labeling y. We write τ for the threshold called t there, because t is a parameter of this section. By the second item,

$$U = \partial^- S_0 = \{u : \ \tilde{Y}_u < 0, \ \exists x \in N(u) : \tilde{Y}_x \ge 0\},$$
$$W = \partial^+ S_0 = \{x : \ \tilde{Y}_x \ge 0, \ \exists u \in N(x) : \tilde{Y}_u < 0\},$$

and D is the set of vertices of degree 3 all of whose neighbors lie on the other side of 0, where the two sides are $\{\tilde{Y} < 0\}$ and $\{\tilde{Y} \ge 0\}$.

The connector in linear time. The connector of Lemma 8.4 needs an order of W' satisfying (8.3). The proof of Lemma 8.3 is a recursion on closed parts; the following order is computed directly. As in Lemma 8.3, let Q be a finite set and P a finite set of items, each item p having a nonempty set $e(p) \subseteq Q$ of at most two endpoints. Let $Q_1, \dots, Q_r$ be the vertex sets of the connected components of the hypergraph on Q whose edges are the sets $e(p)$, $p \in P$, indexed so that their *excesses* $\mathrm{ex}_l := |Q_l| - |\{p \in P : e(p) \subseteq Q_l\}|$ are nondecreasing in l. A *greedy order* of Q_l starts with any element of Q_l and, as long as some item has one endpoint among the elements chosen so far and one outside them, adds the latter. A *component order* of Q is a concatenation of greedy orders of $Q_1, \dots, Q_r$, in this order.

Lemma A.33 (connector in linear time)**.** *Every component order is an ordering of Q that satisfies* (8.3)*, and a component order can be computed in $O(|Q| + |P|)$ time.*

Proof. Every endpoint set is nonempty, connected in the hypergraph, and hence contained in exactly one Q_l; so $\sum_l \mathrm{ex}_l = |Q| - |P|$. Fix l and let X be the set of chosen elements of Q_l at some point, with $\emptyset \neq X \subsetneq Q_l$. Since Q_l is connected, some edge $e(p)$ meets both X and $Q_l \setminus X$; as $|e(p)| \leq 2$, p has one endpoint in X and one outside. So the greedy order continues until it exhausts Q_l, and each step after the first covers an item that was not covered before; in particular, a component order is an ordering of Q. Let $c_l(j')$ be the number of items with $e(p)$ contained in the first j' chosen elements of Q_l; then $c_l(j') \geq j' - 1$ for $j' \geq 1$, so $j' - c_l(j') \leq 1$ for all $0 \leq j' \leq |Q_l|$. In particular $\mathrm{ex}_l = |Q_l| - c_l(|Q_l|) \leq 1$.

Now let $j = \sum_{l<i} |Q_l| + j'$ with $0 \leq j' \leq |Q_i|$. The first j elements cover all items of $Q_1, \ldots, Q_{i-1}$, and the items of Q_i counted by $c_i(j')$; no other item, since each item lies in one component. Hence

$$j - c(j) = \sum_{l<i} \mathrm{ex}_l + (j' - c_i(j')) \leq \sum_{l<i} \mathrm{ex}_l + 1.$$

If $\sum_{l<i} \mathrm{ex}_l > 0$, then $\mathrm{ex}_l > 0$ for some $l < i$, so $\mathrm{ex}_{l'} \geq \mathrm{ex}_l > 0$ for all $l' \geq l$ by the ordering, and $\sum_{l<i} \mathrm{ex}_l \leq \sum_l \mathrm{ex}_l = |Q| - |P|$. Otherwise the sum is at most 0. In both cases $j - c(j) \leq \max(0, |Q| - |P|) + 1$.

The components are found by breadth-first search in $O(|Q| + |P|)$ time; the greedy orders are breadth-first searches as well. The excesses are integers in $[-|P|, 1]$, so counting sort orders the components in $O(|Q| + |P|)$ time. □

The layout. Let $v_1, \ldots, v_n$ be the vertices in increasing order of $(\tilde{Y}_v, v)$, which is the increasing order of y, and $\ell := |\tilde{S}_0|$. The algorithms build the bags of Lemma 8.5: the inner bags $A_1, \ldots, A_\ell$, the connector bags $C_{-1}, C_0, \ldots, C_k$ with $k = |W'|$, and the outer bags $A_{\ell+1}, \ldots, A_n$. The connector bags use a component order of W', computed by Lemma A.33 with $Q = W'$ and the items $P = \{u \in U' : N(u) \cap W' \neq \emptyset\}$, where $e(u) = N(u) \cap W'$; each item has at most two endpoints by Lemma 8.5. By Lemmas A.33 and 8.5, under the hypotheses of Lemma 8.4 these bags form a path decomposition of width at most w. The sets U, W, D, U', W' and P are computed in $O(n)$ time from the signs of the labels. By the definitions of the bags, those containing a vertex are determined by its index, the smallest and largest indices of its neighbors, its position in the connector order, and the largest position of an element of $e(u)$. So the interval form is computed in $O(n)$ time after sorting.

Counts. The cubic case uses the $\mu = 8K + 4$ counts

$$\begin{aligned}
X_j^{\mathrm{in}} &:= |\partial^- \tilde{S}_{\tau_j}| = |\{u : \ \tilde{Y}_u < \tau_j, \ \exists x \in N(u) : \tilde{Y}_x \geq \tau_j\}| && (0 \leq j < 2K),\\
X_j^{\mathrm{out}} &:= |\partial^+ \tilde{S}_{\tau_j}| && (2K < j \leq 4K),\\
X_j^{\mathrm{win}} &:= |\{v : \ \tau_j \leq \tilde{Y}_v < \tau_{j+1}\}| && (0 \leq j < 4K),\\
X^{\mathrm{lo}} &:= |\{v : \ \tilde{Y}_v < -2\}|, \qquad X^{\mathrm{hi}} := |\{v : \ \tilde{Y}_v \geq 2\}|,
\end{aligned}$$

and L_0, L_0'. The last two are counts in the sense of Appendix A.5.5: $L_0 = \sum_v \mathbf{1}\{v \in U \cup (D \cap W)\}$, a disjoint union since $U \cap W = \emptyset$, and each indicator depends only on the set $\{u \in N[v] : \tilde{Y}_u < 0\}$; similarly for $L_0' = \sum_v \mathbf{1}\{v \in W \cup (D \cap U)\}$.

Means. Let $n \geq n_2$ and $r_h = r_*(1 - 1/h)$, as in Appendix A.5.4. Then:

- for $j < 2K$ we have $\tau_j + \xi \leq -1/K + \xi < 0$ by (N1), so Corollary A.25(b) and Lemma A.26(a) give $\mathbb{E}X_j^{\text{in}} \leq (\kappa(r_h) + \omega_3)n$;
- for $j > 2K$ we have $\tau_j - \xi \geq 1/K - \xi > 0$, so Corollary A.25(a) and Lemma A.26(a) give $\mathbb{E}X_j^{\text{out}} \leq (\kappa(r_h) + \omega_3)n$;
- Corollary A.25(c) and the two switch inequalities of Lemma A.26(a) give $\mathbb{E}L_0, \mathbb{E}L_0' \leq (\kappa(r_h) + \omega_3)n$;
- Corollary A.25(d) gives $\mathbb{E}X_j^{\text{win}} \leq (\varphi(0)/K + 2\xi\varphi(0) + p_t)n$, and with (N2), $\mathbb{E}X^{\text{lo}}, \mathbb{E}X^{\text{hi}} \leq (\Phi(-2) + \xi\varphi(0) + p_t)n < 0.023\,n$.

Proof of Theorem 10.1 in the cubic case. Let $n \geq n_2$ and suppose that every count is at most its mean plus θ. We check the hypotheses of Lemma 8.4 for y with

$$w := (\kappa(r_h) + \omega_3 + \varphi(0)/K + 2\xi\varphi(0) + p_t)n + 2\theta + 1.$$

Inner frontiers. Let $\tau < 0$. If $\tau < -2$, then $\partial^- S_\tau \subseteq S_\tau \subseteq S_{-2} = \tilde{S}_{-2}$, so $|\partial^- S_\tau| \leq X^{\text{lo}} \leq 0.023n + \theta$. Otherwise $\tau \in [\tau_j, \tau_{j+1})$ for some $0 \leq j < 2K$. Let $u \in \partial^- S_\tau$, so $y_u \leq \tau < y_x$ for some neighbor x. If $y_u \leq \tau_j$, then $u \in \partial^- S_{\tau_j}$; otherwise $\tau_j < y_u \leq \tau_{j+1}$. Since $S_{\tau_j} = \tilde{S}_{\tau_j}$ and $S_{\tau_{j+1}} \setminus S_{\tau_j} = \{v : \tau_j \leq \tilde{Y}_v < \tau_{j+1}\}$, this gives $|\partial^- S_\tau| \leq X_j^{\text{in}} + X_j^{\text{win}} \leq w - 1$.

Outer frontiers. Let $\tau > 0$. If $\tau \geq 2$, then $\partial^+ S_\tau \subseteq V(G) \setminus S_2$, so $|\partial^+ S_\tau| \leq X^{\text{hi}} \leq 0.023n + \theta$. Otherwise $\tau \in [\tau_j, \tau_{j+1})$ for some $2K \leq j < 4K$. Let $x \in \partial^+ S_\tau$, so $y_x > \tau \geq y_u$ for some neighbor u. If $y_x > \tau_{j+1}$, then $x \in \partial^+ S_{\tau_{j+1}}$, and $j + 1 > 2K$; otherwise $\tau_j \leq \tau < y_x \leq \tau_{j+1}$. So $|\partial^+ S_\tau| \leq X_{j+1}^{\text{out}} + X_j^{\text{win}} \leq w - 1$.

Threshold zero. $\max(L_0, L_0') + 1 \leq (\kappa(r_h) + \omega_3)n + \theta + 1 \leq w$.

Tails. $0.023n + \theta \leq w$, because $\kappa(r_h) > c_3 > 0.023$; indeed κ is decreasing, since $b(r)$ is increasing and arccos is decreasing.

So Lemma 8.5 applies, and the decomposition built by the algorithms has width at most w. Put

$$\rho(n) := (\kappa(r_h) - c_3 + \omega_3 + \varphi(0)/K + 2\xi\varphi(0) + p_t)n + 2\max\{\theta_{\text{D}}(n), \theta_{\text{R}}(n)\} + 1$$

for $n \geq n_2$, with $\theta_{\text{D}}(n) = \sqrt{8(8K+4)n}\, t^{1/4}$ and $\theta_{\text{R}}(n) = n^{3/4}$, and $\rho(n) := n$ for $n < n_2$; we replace ρ by its running maximum (4.2). Since κ is continuous and $r_h \to r_*$, we have $\kappa(r_h) \to c_3$, and $\rho(n) = o(n)$.

The paragraph "The three algorithms" of Appendix A.5.5, with $\mu = 8K + 4$, shows that the algorithms ensure this with $\theta \in \{\theta_{\text{D}}(n), \theta_{\text{R}}(n)\}$, the randomized one except with probability $\pi(n)$; then the width is at most $c_3 n + \rho(n)$. □

A.5.9 The deletion process: bootstrap and general cases

This subsection proves Theorem 10.1 in the bootstrap cases and in the general case, by making the deletion process of Section 9 algorithmic. The bootstrap lemma is proved through the explicit bound (9.4), and the proof of that bound, through Proposition 9.10, is already a procedure: delete a vertex of maximum degree, reduce, and repeat until every component is basic. The induction of Edwards and McDermid [EM15], which Proposition 9.10 generalizes, is of the same kind. The algorithm below makes the procedure explicit, with one simplification: it never tests whether a component is basic, which would require trying every vertex. Instead, it stops at a component of

maximum degree at most k (defined below), which it decomposes by a sweep, and at a component that becomes empty when the chosen vertex is deleted and reduction operations are applied.

Throughout, X is one of the six sequences of Theorem 9.2, the *seed*, with base degree k: $k = 3$ for C, $k = 4$ for M and F, and $k = 5$ for P, R and S. Let $(b_d)_{d\geq 0}$ be the bootstrap of X, with $b_0 = b_1 = b_2 = 0$ and $(b_3, \dots, b_k) = X$; it satisfies the hypotheses of Lemma 9.6 (proof of Theorem 9.2). For $X = \mathsf{S}$ it is denoted σ in Section 9.4; we keep the letter b of Section 9.1, because σ_v denotes standard deviations here. The sum $\psi(G) = \sum_{d\geq 3} b_d n_d(G)$, the potential ψ^- and the running minima H_d of the increments of b are as in Sections 9.1 and 9.3. The general case is the bootstrap case of S with the coefficients rounded up: by Proposition 9.13, $b_d \leq a_d$ for every $d \geq 3$ when $X = \mathsf{S}$, where a_d are the coefficients of Theorem 1.3.

Base layouts. Each seed comes with a layout for graphs J with $\delta(J) \geq 3$ and $\Delta(J) \leq k$, labeled with the parameters of some $n \geq \max\{|V(J)|, n_2\}$:

- C: the median-switched layout with the labels of the cubic case ($q \equiv 2$), as in Appendix A.5.8;
- M: the median-switched layout with the labels of the switched case ($q(x) = d(x) - 1$), as in Appendix A.6;
- F: the plain-sweep layout with the labels of the padded case ($q \equiv 3$);
- P, R, S: the plain-sweep layout with the labels of the mixed cases ($q(x) = d(x) - 1$) and the triple X.

For such a graph J, $\psi(J) = \sum_{d=3}^{k} b_d n_d(J)$, and this is the bound $\beta(J)$ of the case whose labels are used; for C and F because every vertex of J has degree 3, respectively degree 3 or 4. Each base layout has at most $12K + 3$ counts in the sense of Appendix A.5.5, and if every count is at most its mean plus some $\theta \geq 0$, then the layout is a path decomposition of J of width at most

$$\psi(J) + \rho_X(n)\,|V(J)| + 5\theta + 1, \tag{A.7}$$

where $\rho_X(n)$ is the function $\rho_0(n)$ of Appendix A.5.7 for the plain sweep (with the error term $\varepsilon(h)$ of the case) and $\rho_{\mathrm{med}}(n)$ of Appendix A.6 for the median-switched layout. This is Proposition A.32 for the plain sweep and Corollary A.39 for the median-switched layout.

The algorithm. Let G_0 be G without its isolated vertices, and $n_{\geq 1} := |V(G_0)|$. If $n_{\geq 1} \leq 2$, output one bag $V(G_0)$ and one bag for each isolated vertex. Otherwise compute all parameters from $n_{\geq 1}$, and an adjacency matrix of G_0.

(1) Apply reduction operations to G_0 until none applies, recording the removed vertices; let R_0 be the final graph.

(2) Keep a list of components, initially those of R_0, and a deleted set $D := \emptyset$. Repeat the following *round* while some component J in the list is not marked final:

- if $\Delta(J) \leq k$, mark J final (a *sweep component*);
- otherwise choose a vertex w of degree $\Delta(J)$ that has a neighbor of smaller degree, if one exists, and any vertex of degree $\Delta(J)$ otherwise. Apply reduction operations to $J - w$ until none applies; let J^* be the final graph.

– If J^* is empty, mark J final (a *thin component*). The deletion of w and these reductions are kept for step 3 only; they are not part of the global sequence. The reductions may have changed the shared adjacency matrix, so they are recorded in an undo log while they are applied and rolled back in time $O(|V(J)| + |E(J)|)$, which restores J.
– Otherwise *commit*: add w to D, record the deletion and the reductions as part of the global sequence, and replace J in the list by the components of J^*.

(3) Let R be the union of the final components. Decompose each sweep component J by the base layout of X, with the parameters of $n_{\geq 1}$ (randomized, Las Vegas or deterministic, as below). Decompose each thin component J by Lemma A.28, applied to $F = J$, the deleted set $\{w\}$ and the empty final graph. Concatenate these decompositions into a decomposition of R.

(4) Apply Lemma A.28 to $F = G_0$, the global sequence (deleted set D, final graph R) and the decomposition of R. Append one bag for each isolated vertex of G.

By (RP2), reductions inside a component of the current graph are reductions of the whole current graph, so the global sequence is a sequence of reductions and deletions from G_0 to R, as Lemma A.28 requires. Every component in the list has minimum degree at least 3 by (RP3), since it is a component of the final graph of a maximal sequence of reductions. Every final component is basic in the sense of Definition 9.8: a sweep component has $\Delta \leq k$, and a thin component has the witness w.

Lemma A.34 (the deletion process)**.** *The algorithm makes at most $n_{\geq 1}$ commits and at most $2n_{\geq 1}$ rounds, and $|D| + \psi^-(R) \leq \psi(G)$.*

Proof. Every commit deletes a vertex. Every other round marks a component final, and final components are disjoint, so there are at most $n_{\geq 1}$ of them.

Consider a commit at a component J. Then J is connected, $\delta(J) \geq 3$, $\Delta(J) \geq k+1$, the vertex w was chosen by the rule of Lemma 9.9, and J^* is the final graph of a maximal sequence of reductions applied to $J - w$, and is nonempty. So Lemma 9.9 gives $1 + \psi^-(J^*) \leq \psi^-(J)$. Hence the quantity $|D| + \sum_J \psi^-(J)$, summed over the components in the list, does not increase at a commit, and it does not change when a component is marked final. At the start it is $\psi^-(R_0)$ (by additivity), and at the end it is $|D| + \psi^-(R)$. Finally $\psi^-(R_0) \leq \psi(R_0) \leq \psi(G_0) = \psi(G)$, by (RP1), $H \geq 0$, and because isolated vertices have weight $b_0 = 0$. □

Proof of Theorem 10.1 in the bootstrap and general cases. Let $n_{\geq 1} \geq n_2$; otherwise the claims are trivial, because the output has width at most $n_{\geq 1} - 1$ and we put $\rho(N) := N$ and $\pi(N) := 0$ for $N < n_2$. Put

$$\rho_{\mathrm{sw}}(N) := \rho_X(N)\, N + 5 \max\{\theta_{\mathrm{D}}(N), \theta_{\mathrm{R}}(N)\} + 1,$$
$$\rho(N) := b_3 + \rho_{\mathrm{sw}}(N) + 6 \log_2 N + 2 \qquad (N \geq n_2),$$

with $\theta_{\mathrm{D}}(N) = \sqrt{8(12K+3)N}\, t^{1/4}$ and $\theta_{\mathrm{R}}(N) = N^{3/4}$ (parameters of N); we replace ρ by its running maximum (4.2). Then $\rho(N) = o(N)$, because $\rho_X(N) \to 0$.

We first bound the widths of the decompositions of the final components, of R and of G_0, assuming that every count of every sweep component is at most its mean plus some $\theta \leq \max\{\theta_{\mathrm{D}}(n_{\geq 1}), \theta_{\mathrm{R}}(n_{\geq 1})\}$. Then we show that the three algorithms ensure this assumption, for certain or with the stated probability.

Sweep components. Let J be a sweep component with N_J vertices, and suppose that all its counts are at most their means plus $\theta \leq \max\{\theta_{\mathrm{D}}(n_{\geq 1}), \theta_{\mathrm{R}}(n_{\geq 1})\}$. By (A.7), applied with the parameters

of $n_{\geq 1} \geq N_J$, its decomposition has width at most $\psi(J) + \rho_X(n_{\geq 1})N_J + 5\theta + 1 \leq \psi(J) + \rho_{\mathrm{sw}}(n_{\geq 1})$. Since J is basic, $\psi(J) \leq \psi^-(J) + b_3$ by Lemma 9.11.

Thin components. By Lemma A.28 with deleted set $\{w\}$ and R empty, the width is at most $1 + (-1) + 3\log_2 n_{\geq 1} + 1$.

The final graph. The width of the concatenation is the maximum over the components. Since ψ^- is nonnegative and additive, the decomposition of R has width at most $\psi^-(R) + b_3 + \rho_{\mathrm{sw}}(n_{\geq 1}) + 3\log_2 n_{\geq 1} + 1$.

Undoing. By Lemma A.28 and Lemma A.34, the decomposition of G_0 has width at most

$$|D| + \psi^-(R) + b_3 + \rho_{\mathrm{sw}}(n_{\geq 1}) + 6\log_2 n_{\geq 1} + 2 \leq \psi(G) + \rho(n_{\geq 1}) \leq \sum_{d\geq 3} b_d n_d(G) + \rho(n),$$

using $n_{\geq 1} \leq n$ and the monotonicity of ρ. The bags of isolated vertices do not increase the width. This is the width bound of Theorem 10.1 in the bootstrap case of X. In the general case, $X = \mathsf{S}$ and $b_d \leq a_d$ give the first width bound, $\sum_{d\geq 3} a_d n_d(G) + \rho(n)$.

Edge bound (general case). Every vertex of G_0 has degree at least 1, so $n_{\geq 1} \leq 2m$. By (9.9), $a_d \leq \frac{7}{48} \cdot \frac{d}{2}$ for every $d \geq 3$, and $b_0 = b_1 = b_2 = 0$; so $\psi(G) = \sum_d b_d n_d(G) \leq \frac{7}{48} \cdot \frac{1}{2} \sum_v d(v) = \frac{7}{48} m$. Hence the same decomposition also has width at most $\frac{7}{48} m + \rho(n_{\geq 1}) \leq \frac{7}{48} m + \rho(2m)$.

Probability. The randomized algorithm samples the variables of different sweep components independently. We apply Lemma A.27(b) to each sweep component J with the *global* threshold $\lambda = n_{\geq 1}^{3/4}$: the probability that some count of J exceeds its mean plus $n_{\geq 1}^{3/4}$ is at most $8\mu N_J t^{1/2} n_{\geq 1}^{-3/2}$, with $\mu \leq 12K + 3$. The sweep components are disjoint, so summing over them gives at most $8(12K+3)t^{1/2} n_{\geq 1}^{-1/2}$. With the sampling budget, which is $2n_{\geq 1}\lceil \log_2 n_{\geq 1} \rceil$ draws for the at most $n_{\geq 1}$ samples, the failure probability is at most $\pi(n_{\geq 1})$ with $\pi(N) := 8(12K+3)t^{1/2}N^{-1/2} + 1/\lceil \log_2 N \rceil$ (parameters of N). A union bound with each component's own threshold $N_J^{3/4}$ would not suffice, because there may be many components of moderate size.

Deterministic. Greedy fixing (Lemma A.27(c)) is run on each sweep component separately; the deviations are at most $\sqrt{8\mu N_J}\, t^{1/4} \leq \theta_{\mathrm{D}}(n_{\geq 1})$.

Las Vegas. Steps 1, 2 and 4 are deterministic. In step 3, each sweep component is sampled, and resampled until all its counts are at most their means plus $n_{\geq 1}^{3/4}$. By Lemma A.27(b) and (N3), a sample of J is accepted with probability at least $1 - 8\mu N_J t^{1/2} n_{\geq 1}^{-3/2} \geq \frac{1}{2}$, so the expected number of samples per component is at most 2. □

Remark A.35 (the rows and the sequences on all graphs). The bootstrap of a seed X is entrywise at most the row X of Table 4, which rounds it up (for $X \neq \mathsf{C}$; the bootstrap of C is given in the caption), and at most the sequence X of Theorem 9.2, which agrees with it up to degree k and has coefficient 1 above k (Lemma 9.3). So the bootstrap cases compute, for every graph, decompositions of width at most the bound of each row of Corollary 9.5 and of each sequence of Theorem 9.2, plus the remainder $\rho(n)$. No row is best for all graphs (Remark A.18); running the algorithm for each of the six seeds and keeping the narrowest decomposition realizes the minimum of the six bounds. For a graph of maximum degree at most k, the bootstrap case of X never commits a deletion, and its bound $\sum_{d=3}^{k} b_d n_d$ improves on the bound $\beta(G)$ of the cubic and padded cases when G has vertices of degree at most 2.

A.5.10 Running times

The width claims of Theorem 10.1, and the probability claims of Remark A.21, are proved in Appendices A.5.7 to A.5.9 and Appendix A.6. This subsection bounds the running times of the steps

of the algorithms, which completes the proof. Recall that $t = O(\log n)$, that $h, K = O(\log \log n)$, and that the number of counts is $\mu = O(K)$.

Cubic, padded, switched and mixed cases. Here $m \le \frac{5}{2}n$.

- *Reductions* (switched and mixed cases). Degrees never exceed 5, so each operation, including the adjacency test of (R3) and (R4), takes $O(1)$ time with a queue of vertices of degree at most 2; in total $O(n)$ operations on $O(\log n)$-bit integers, that is $O(n \log n)$ bit operations.
- *Coefficients and finite distribution.* $n(\log n)^{O(1)}$ and $(\log n)^{O(1)}$ bit operations (Appendices A.5.1 and A.5.2).
- *Labels.* Each $\tilde{Y}_v$ is a sum of at most $2t^{1/4}$ products of $O(\log t)$-bit rationals; in total $n\, t^{O(1)}$ bit operations.
- *Sampling* (randomized). At most $2n\lceil \log_2 n \rceil$ draws of $O(\log t)$ bits, each followed by a binary search over M prefix sums; in total $n(\log n)^{O(1)}$ bit operations in the worst case.
- *Sorting.* $O(n \log n)$ comparisons of labels, each in $(\log t)^{O(1)}$ bit operations.
- *Layout.* $O(n)$ operations for the plain sweep; for the median-switched layout, $O(n)$ operations by Lemma A.33 and the interval rules.
- *Undoing.* $O(n \log n)$ operations by Lemma A.28, since the layout of H has $s \le 2N + 2 \le 3n$ bags: N in the plain sweep, and $N + |W'| + 2$ in the median-switched layout.

So one sample (the randomized algorithm) uses $n(\log n)^{O(1)}$ bit operations. The Las Vegas algorithm additionally computes the means in $\mu n\, n^{o(1)} = n^{1+o(1)}$ bit operations (Lemma A.27(d)), checks the counts of each sample in $n(\log n)^{O(1)}$ bit operations, and takes at most two samples in expectation; so its expected running time is $n^{1+o(1)}$ (for $n < n_2$ it accepts the first sample). The deterministic algorithm replaces sampling by greedy fixing (Lemma A.27(c)), which uses $\mu n^2 n^{o(1)} = n^{2+o(1)}$ bit operations.

Bootstrap and general cases. Removing isolated vertices takes $O(n)$ operations, and building the adjacency matrix of G_0 takes $O(n_{\ge 1}^2)$.

- *Steps 1 and 2.* With adjacency lists that support deletion in $O(1)$ time, the adjacency matrix for the tests and insertions of (R3) and (R4), and a queue of vertices of degree at most 2, each reduction takes $O(1)$ operations. A component's reductions after the deletion of w take $O(|V(J)| + |E(J)|)$ operations, and so do the choice of w and the computation of the components of J^* by breadth-first search. By Lemma A.34 there are at most $2n_{\ge 1}$ rounds, so steps 1 and 2 take $O(n_{\ge 1}(n_{\ge 1} + m))$ operations.
- *Step 3.* The sweep components are disjoint and have maximum degree at most 5, and each is decomposed by the base layout of the seed, so the bounds of the previous paragraph, summed over them, give $n_{\ge 1}(\log n_{\ge 1})^{O(1)}$ bit operations (randomized), $n_{\ge 1}^{1+o(1)}$ in expectation (Las Vegas), and $\sum_J N_J^2 n_{\ge 1}^{o(1)} \le n_{\ge 1}^{2+o(1)}$ (deterministic). Each thin component takes $O(|E(J)| + |V(J)| \log n_{\ge 1})$ operations by Lemma A.28.
- *Step 4.* $O(m + n_{\ge 1} \log n_{\ge 1})$ operations by Lemma A.28, since the decomposition of R has at most $3n_{\ge 1}$ bags: a component J contributes at most $2|V(J)| + 2 \le 3|V(J)|$ of them, as $|V(J)| \ge 4$.

In total, the bootstrap and general cases use $O(n(n+m))$ operations on integers of $O(\log n)$ bits, plus $n(\log n)^{O(1)}$, expected $n^{1+o(1)}$, or $n^{2+o(1)}$ bit operations for one sample, the Las Vegas and the deterministic algorithm. This completes the proof of Theorem 10.1. □

A.5.11 Orderings of small cutwidth

Notation. We use the inner frontier $\partial^- X = \{u \in X : N(u) \not\subseteq X\}$ of Section 4.2. Let $e(X)$ be the number of edges with exactly one end in X. For an ordering $(u_1, \dots, u_n)$ of $V(G)$ with prefixes $U_i = \{u_1, \dots, u_i\}$, the *vertex separation* of the ordering is $\max_i |\partial^- U_i|$ and its *cutwidth* is $\max_i e(U_i)$.

Lemma A.36 (cutwidth from vertex separation)**.** *Let G be a graph with $\Delta(G) \le 3$ and $n \ge 1$ vertices, and let $(v_1, \dots, v_n)$ be an ordering of $V(G)$ of vertex separation w. Then an ordering of $V(G)$ of cutwidth at most $w+2$ can be computed in $O(n)$ operations on integers of $O(\log n)$ bits.*

Proof. The relations $s(G) \le \mathrm{vs}(G) + 2$ ([EST94, Lemma 2.2], half of their Theorem 2.1) and $\mathrm{cw}(G) \le s(G)$ for maximum degree 3 ([MS89, Theorem 3.2], which gives their Corollary 3.3), where s is the search number, have constructive proofs; we check that they apply to the given ordering and run in linear time. Both papers work with connected multigraphs. Cutwidth and vertex separation of an ordering are maxima over the components (restrict the ordering to each component, then concatenate), so let G be connected.

Reduction. Let G' be obtained from G by suppressing, while possible, a vertex of degree 2 with two distinct neighbors (deleting it and joining its neighbors by an edge). This creates no loops, keeps all degrees, and takes $O(n)$ operations. In G' every vertex of degree 2 is joined to a single vertex by two parallel edges, which is the form that the proof of [MS89, Theorem 3.2] brings the graph into (the reduction there, followed by subdividing every loop once). Restrict the given ordering to $V(G')$. If $u \in V(G')$ lies in a prefix and has a G'-neighbor outside it, then the path of G that the corresponding edge replaces leaves the prefix, and its last vertex in the prefix, counted from u, has a G-neighbor outside it. This vertex is u or an inner vertex of a path at u, so distinct such u give distinct vertices, and the restricted ordering of G' has vertex separation at most w.

Searching. The procedure `search1` in the proof of [EST94, Lemma 2.2] takes the vertices of G' in this order, places a searcher on each, clears its edges to earlier vertices with one further searcher, and then removes the searchers from vertices with no uncleared edge. By the two invariants stated in that proof it never uses more than $w+2$ searchers and never recontaminates an edge. This gives a search strategy S for G' with at most $k = w+2$ searchers, without recontamination, with $O(n)$ steps, and in which the step that clears each edge is known.

Ordering. The proof of [MS89, Theorem 3.2] orders the vertices of G' by the first step of S after which at least half of their incident edges are cleared, breaking ties by an explicit rule for vertices of degree 2, and shows that this ordering has cutwidth at most $\lfloor \Delta(G')/2 \rfloor (k-1) + 1 = k$. It is computed from the clearing steps by a bucket sort in $O(n)$ operations.

Back to G. Undo the suppressions in reverse order: an edge yz with $y \ne z$ that replaced a vertex x, with y before z, becomes the path yxz with x placed directly after y. The new cut between y and x is crossed by the same edges as the cut after y was, with yx in place of yz; every other cut that yz crossed is now crossed by xz instead. So the cutwidth does not change, and the resulting ordering of $V(G)$ has cutwidth at most $w+2$. □

Corollary 10.2 (computing orderings and bisections)**.** *For graphs G with $\Delta(G) \le 3$, an ordering of $V(G)$ of cutwidth at most $c_3 n + o(n)$, and a bisection of G with at most $c_3 n + o(n)$ edges between its two sides, can be computed in deterministic time $n^{2+o(1)}$.*

Proof. Theorem 10.1(b) computes a path decomposition $(W_1, \dots, W_r)$ of G of width at most $c_3 n + \rho(n)$, in interval form, in $n^{2+o(1)}$ bit operations. Order the vertices by $\mathrm{first}(v)$, breaking ties arbitrarily. Let L be a prefix of this ordering other than $V(G)$, let u' be the first vertex after L, and let $g = \mathrm{first}(u')$, which is the least value of $\mathrm{first}(y)$ over $y \notin L$. If $x \in L$ has a neighbor $y \notin L$, then $\mathrm{first}(x) \le g \le \mathrm{first}(y) \le j$ for a bag W_j that contains x and y, so $x \in W_g$. Hence $\partial^- L \subseteq W_g \setminus \{u'\}$; since $\partial^- V(G) = \emptyset$, the ordering has vertex separation at most $c_3 n + \rho(n)$. Lemma A.36 turns it into an ordering of cutwidth at most $c_3 n + \rho(n) + 2$ in $O(n \log n)$ bit operations. Splitting this ordering after its $\lfloor n/2 \rfloor$-th vertex gives a bisection with at most as many edges between its two sides. □

Theorem 10.3 (deterministic cut sweep)**.** *Let $p \in \{4, 5\}$. For graphs G with $\Delta(G) \le p$ and m edges, an ordering of $V(G)$ of cutwidth at most $\gamma_p m + o(n)$, and a bisection of G with at most that many edges between its two sides, can be computed in deterministic time $n^{2+o(1)}$. For $p = 4$ the bound is $\frac{m}{6} + o(n) \le \frac{n}{3} + o(n)$.*

Proof. We use the padded labels of Appendix A.5.1: the padded case ($q \equiv 3$) for $p = 4$ and the padded case of degree 5 ($q \equiv 4$) for $p = 5$, with the parameters of n; for $n < n_2$ every ordering has cutwidth at most $m \le \frac{p}{2} n$, which the remainder absorbs. Let $n \ge n_2$, let $v_1, \dots, v_n$ be the vertices in increasing order of $(\tilde{Y}_v, v)$, let $T_i = \{v_1, \dots, v_i\}$, and let $\tau_j = -3 + j/K$ for $0 \le j \le 6K$.

Counts. For $0 \le j < 6K$ and $1 \le k \le p$ let

$$X^{\mathrm{cut}}_{j,k} := \big|\{v : \ \tilde{Y}_v < \tau_j, \ |\{u \in N(v) : \tilde{Y}_u \ge \tau_j\}| \ge k\}\big|,$$
$$X^{\mathrm{win}}_{j,k} := \big|\{v : \ \tau_j \le \tilde{Y}_v < \tau_{j+1}, \ d(v) \ge k\}\big|,$$

and let X^{lo}_k and X^{hi}_k be the numbers of vertices v with $d(v) \ge k$ and $\tilde{Y}_v < -3$, respectively $\tilde{Y}_v \ge 3$. These $\mu = p(12K + 2)$ random variables are counts in the sense of Appendix A.5.5, since each indicator depends only on the labels in $N[v]$. Summing over k turns the conditions $c \ge k$ into the numbers c, so

$$C_j := \sum_k X^{\mathrm{cut}}_{j,k} = e(\tilde{S}_{\tau_j}), \qquad D_j := \sum_k X^{\mathrm{win}}_{j,k} = \sum_{v:\, \tau_j \le \tilde{Y}_v < \tau_{j+1}} d(v),$$

and $D^{\mathrm{lo}} := \sum_k X^{\mathrm{lo}}_k$ and $D^{\mathrm{hi}} := \sum_k X^{\mathrm{hi}}_k$ are the degree sums over the two tails. (Each edge leaving $\tilde{S}_{\tau_j}$ is counted once, at its end in $\tilde{S}_{\tau_j}$.) We split the cut and the degree sums into counts because Lemma A.27 needs indicators.

Grid. Every prefix satisfies

$$e(T_i) \le \max\Big\{D^{\mathrm{lo}}, \ D^{\mathrm{hi}}, \ \max_{0 \le j < 6K} (C_j + D_j)\Big\}. \tag{A.8}$$

Indeed, let $y = \tilde{Y}_{v_i}$ and let zu be an edge with $z \in T_i$ and $u \notin T_i$, so $\tilde{Y}_z \le y \le \tilde{Y}_u$. If $y < -3$, we charge the edge to z, which has $\tilde{Y}_z < -3$ and receives at most $d(z)$ edges; if $y \ge 3$, we charge it to u, which has $\tilde{Y}_u \ge 3$. Otherwise $\tau_j \le y < \tau_{j+1}$ for some j. If $\tilde{Y}_z < \tau_j$, the edge leaves $\tilde{S}_{\tau_j}$, because $\tilde{Y}_u \ge y \ge \tau_j$, and is counted by C_j; otherwise $\tau_j \le \tilde{Y}_z < \tau_{j+1}$, and we charge the edge to z, counted by D_j. This is the charging of the grid step in Section 3.4, for the discrete labels.

Means. Fix an edge uv and $\tau \in \mathbb{R}$, and couple the discrete and ideal labels as in Lemma A.24. Let $\mathcal{E}$ be the event that $|\tilde{Y}_x - Y_x| \le \xi$ for $x \in \{u, v\}$, which fails with probability at most p_t since $\{u, v\} \subseteq N[u]$, and let $\mathcal{W}$ be the event that neither Y_u nor Y_v lies in $[\tau - \xi, \tau + \xi]$, which fails with probability at most $4\xi\varphi(0)$. If $\tilde{Y}_u < \tau \le \tilde{Y}_v$, then on $\mathcal{E} \cap \mathcal{W}$ we have $Y_u < \tau + \xi$, hence $Y_u < \tau - \xi$, and $Y_v \ge \tau - \xi$, hence $Y_v > \tau + \xi$: the edge crosses τ for the ideal labels. The ideal labels have correlation at least $r_{pp}(1 - \frac{1}{h})$ (Lemma A.22), so by Lemma 3.3(b)

$$\mathbb{P}(uv \text{ leaves } \tilde{S}_\tau) \le \tfrac{1}{\pi} \arccos\big(r_{pp}(1 - \tfrac{1}{h})\big) + 4\xi\varphi(0) + p_t = \gamma_p + o(1),$$

uniformly in τ, because $h \to \infty$ and $\xi, p_t \to 0$. Summing over the edges, $\mathbb{E}C_j \le (\gamma_p + o(1))m$. By Corollary A.25(d) and (N2), a vertex lies in $[\tau_j, \tau_{j+1})$ with probability at most $\varphi(0)/K + o(1)$, and below -3 or at least 3 with probability at most 0.0014; as $\sum_v d(v) = 2m$, this gives $\mathbb{E}D_j \le (2\varphi(0)/K + o(1))m = o(m)$ and $\mathbb{E}D^{\mathrm{lo}}, \mathbb{E}D^{\mathrm{hi}} \le 0.0028\, m < \gamma_p m$.

Greedy fixing. By Lemma A.27(c), greedy fixing on these counts ends with values of the $\tilde{Z}_w$ for which every count is within $\theta = \sqrt{8\mu n}\, t^{1/4} = o(n)$ of its mean; so each of C_j, D_j, D^{lo} and D^{hi} is within $p\theta$ of its mean. By (A.8), the sorted order has cutwidth at most $\gamma_p m + o(n)$, and its first $\lfloor n/2 \rfloor$ vertices give the bisection. Since $\mu = O(K)$, greedy fixing takes $n^{2+o(1)}$ bit operations by Lemma A.27(d), and the other steps take $n(\log n)^{O(1)}$ (Appendices A.5.1 and A.5.2). Finally $\gamma_4 = \frac{1}{6}$ and $m \le 2n$ for $p = 4$. □

A.6 Computing the median-switched decompositions

This subsection completes the proof of Theorem 10.1 where the median-switched layout of Section 8.2 is used outside Appendix A.5.8: in the switched case, which makes Theorem 8.2 algorithmic, and in the base layouts of the bootstrap cases C and M (Appendix A.5.9). The proof follows Appendix A.5.8, and two points are new. First, the layout must work for graphs whose labels are built with the parameters of an integer n larger than their number of vertices, as the sweep components of the deletion process require. Second, the existence proof of Theorem 8.2 bounds $\mathbb{E}\hat{L}_0$ and $\mathbb{E}\hat{L}'_0$ only in the limit of a large radius, so we make the error term explicit; together with the bounds at nonzero thresholds, this gives expected frontier sizes that are bounded uniformly in the threshold (Lemma A.37). We use the notation and the parameters of Section 10.

Two settings. Let $n \ge n_2$, and let H be a graph with $N \le n$ vertices whose names are distinct positive integers. We consider two settings.

- *Cubic setting:* $\Delta(H) \le 3$, the labels are those of the cubic case (weight $q \equiv 2$), and $\beta(H) := c_3 N$.
- *Switched setting:* $\delta(H) \ge 3$ and $\Delta(H) \le 4$, the labels are those of the switched case (weight $q(x) = d(x) - 1$, as in the mixed cases), and $\beta(H) := c_3 n_3(H) + 0.2627\, n_4(H)$.

In both settings the labels are built with the parameters of n. The construction of Appendices A.5.1 and A.5.2 and Lemmas A.22 and A.24 and Corollary A.25 apply, since they require only $\Delta(H) \le 5$, and $\delta(H) \ge 3$ for the weights $q(x) = d(x) - 1$. In the switched setting the edge correlation bounds are $r_{uv} = r_{d(u)d(v)}$ with $\varepsilon_h = 8/(h-4)$. In both settings $\beta(H) \ge c_3 N$: in the cubic setting by definition, and in the switched setting because every vertex has degree 3 or 4 and both coefficients are at least c_3.

Layout and counts. Let $\hat{D}$ be the set of vertices with at least three neighbors on the other side of zero, as in Appendix A.3.1, and $\hat{L}_0 = |U| + |\hat{D} \cap W|$ and $\hat{L}'_0 = |W| + |\hat{D} \cap U|$; in the cubic setting $\hat{D}$ is the set D of Section 8.1, and $\hat{L}_0 = L_0$, $\hat{L}'_0 = L'_0$. The *median-switched layout* of H consists of the bags of Lemma 8.5, built as in Appendix A.5.8 from the order of $(\tilde{Y}_v, v)$, with $\hat{D}$ in place of D and a component order of $W' = W \setminus \hat{D}$ (Lemma A.33) for the items $u \in U \setminus \hat{D}$ with $e(u) = N(u) \cap W' \ne \emptyset$; each item has at most two endpoints by the definition of $\hat{D}$. Its interval form is computed in $O(N)$ time after sorting, as in Appendix A.5.8. The layout uses the grid $\tau_j = -2 + j/K$ $(0 \le j \le 4K)$ and the $\mu = 8K + 4$ counts of Appendix A.5.8: the inner counts X_j^{in} $(j < 2K)$, the outer counts X_j^{out} $(j > 2K)$, the windows X_j^{win}, the tails X^{lo}, X^{hi}, and $\hat{L}_0$, $\hat{L}'_0$ for the discrete labels. The last two

are counts in the sense of Appendix A.5.5: $\hat{L}_0 = \sum_v \mathbf{1}\{v \in U \cup (\hat{D} \cap W)\}$, a disjoint union because $U \cap W = \emptyset$, and the indicator depends only on the set $\{u \in N[v] : \tilde{Y}_u < 0\}$; similarly for $\hat{L}'_0$.

Error terms. Recall $r_h = r_*(1 - 1/h)$ and κ from (8.1). For $h \geq 13$ we have $\varepsilon_h = 8/(h-4) \leq \frac{8}{9}$; put

$$e_h := 205\,\varepsilon_h + \kappa\big(r_{33}(1-\varepsilon_h)\big) - c_3, \qquad \varepsilon_{\mathrm{sw}}(h) := \max\{e_h,\ 12\,\varepsilon_h^{1/3}\} \qquad (h \geq 13),$$

and $\varepsilon_{\mathrm{sw}}(h) := 1$ for $5 \leq h < 13$. Let $\varepsilon_{\mathrm{med}}(h)$ be $\kappa(r_h) - c_3$ in the cubic setting and $\varepsilon_{\mathrm{sw}}(h)$ in the switched setting. Since κ is continuous and $\varepsilon_h \to 0$, both tend to 0 as $h \to \infty$. Put

$$\rho_{\mathrm{med}}(n) := \varepsilon_{\mathrm{med}}(h) + \omega_4 + \frac{\varphi(0)}{K} + 2\xi\varphi(0) + p_t.$$

A.6.1 Uniform means in the switched setting

The proof of Theorem 8.2 bounds the expected frontiers at every nonzero threshold with the same error term e_R (Lemma A.14), so it is already uniform in the threshold. At threshold zero, Lemma A.11 makes the error term of Lemma A.12 explicit.

Lemma A.37 (uniform ideal means, switched setting)**.** *Let $h \geq 5$, and let (Y_v) be the ideal labels of radius h of a graph H with N vertices in the switched setting, and let S_τ, U, W, $\hat{D}$, $\hat{L}_0$ and $\hat{L}'_0$ be defined for the labeling Y. Then*

$$\mathbb{E}|\partial^+ S_\tau| \leq \beta(H) + \varepsilon_{\mathrm{sw}}(h)\,N \quad (\tau > 0), \qquad \mathbb{E}|\partial^- S_\tau| \leq \beta(H) + \varepsilon_{\mathrm{sw}}(h)\,N \quad (\tau < 0),$$
$$\mathbb{E}\hat{L}_0 \leq \beta(H) + \varepsilon_{\mathrm{sw}}(h)\,N, \qquad \mathbb{E}\hat{L}'_0 \leq \beta(H) + \varepsilon_{\mathrm{sw}}(h)\,N.$$

Proof. For $h < 13$ the claims are trivial, because each of the four quantities is at most N ($\hat{L}_0 \leq |U \cup W|$ since $U \cap W = \emptyset$, and similarly $\hat{L}'_0$). Let $h \geq 13$, so that $\varepsilon_h < 1$. By Lemma A.22, (Y_v) is a centered Gaussian vector with unit variances and $\mathrm{Corr}(Y_u, Y_v) \geq r_{d(u)d(v)}(1 - \varepsilon_h)$ for every edge uv. This is property (L1) of Section 7 with ε_h in place of ε_R. Moreover $-Y$ has the same distribution as Y, because Y is a linear function of the independent standard normal variables Z_w, and every Y_v has a continuous distribution, because Y_v has variance 1. The proofs of Lemmas A.12 and A.14 use the labels only through these properties: (L1), the sign symmetry and the continuity of the marginals.

Nonzero thresholds. Lemma A.14, with ε_h in place of ε_R, gives $\mathbb{E}|\partial^+ S_\tau| \leq \beta(H) + e_h N$ for $\tau > 0$ and $\mathbb{E}|\partial^- S_\tau| \leq \beta(H) + e_h N$ for $\tau < 0$, and $e_h \leq \varepsilon_{\mathrm{sw}}(h)$.

Threshold zero. Let v be a vertex of $\{3,4\}$-type T, of degree $d \leq 4$, with reference bounds r_T and transfer terms λ_T as in Definition 7.3. As in the proof of Lemma A.12, (A.2) and (L1) give $\mathbb{P}(v \in U) + \mathbb{P}(v \in \hat{D} \cap W) \leq M_d\big((1-\varepsilon_h)r_T\big)$, and Lemma A.11 with $\gamma = \varepsilon_h$ gives $M_d((1-\varepsilon_h)r_T) \leq M_d(r_T) + 3d\,\varepsilon_h^{1/3} \leq \beta_T + 12\,\varepsilon_h^{1/3}$, with β_T as in the table in that proof. By the second row of the table, with the transfer $s = 0.1529003 - c_3$ of Appendix A.3.1,

$$\mathbb{P}(v \in U) + \mathbb{P}(v \in \hat{D} \cap W) + \lambda_T(s,0,0) \leq \begin{cases} c_3 + 12\,\varepsilon_h^{1/3} & \text{if } d = 3, \\ 0.2627 + 12\,\varepsilon_h^{1/3} & \text{if } d = 4. \end{cases}$$

All degrees of H lie in $\{3,4\}$, so Lemma 7.4, with the coefficients c_3 and 0.2627 for the degrees 3 and 4 and with $e = 12\,\varepsilon_h^{1/3}$, gives $\mathbb{E}\hat{L}_0 \leq \beta(H) + 12\,\varepsilon_h^{1/3} N$. Up to null events, the events $\{v \in W\}$ and $\{v \in \hat{D} \cap U\}$ are the events $\{v \in U\}$ and $\{v \in \hat{D} \cap W\}$ for the labeling $-Y$, which has the same distribution; so the same bound holds for $\mathbb{E}\hat{L}'_0$. □

In the cubic setting, Lemma A.26(a) gives the analogous per-vertex bounds, with $\kappa(r_h)$ in place of $c_3 + \varepsilon_{\mathrm{sw}}(h)$.

A.6.2 The width of the median-switched layout

Proposition A.38 (median-switched sweep)**.** *Let $n \geq n_2$ and $\theta \geq 0$, and let H be a graph with $N \leq n$ vertices in the cubic or the switched setting. If every count of the median-switched layout is at most its mean plus θ, then this layout is a path decomposition of H of width at most $\beta(H) + \rho_{\mathrm{med}}(n)\, N + 2\theta + 1$.*

Proof. We first bound the means of the counts, then show that the layout is the path decomposition of Lemma 8.5 for an injective labeling that induces the order of $(\tilde{Y}_v, v)$, and finally check the hypotheses of that lemma.

Means. In the cubic setting the means are bounded in Appendix A.5.8, with N in place of n in the bounds and $\omega_3 \leq \omega_4$; the argument does not use $N = n$. In the switched setting we argue in the same way, with Lemma A.37 in place of Lemma A.26(a):

- for $j < 2K$, Corollary A.25(b) at the threshold τ_j, for which $\tau_j + \xi < 0$ by (N1), gives $\mathbb{E}X_j^{\mathrm{in}} \leq \mathbb{E}|\partial^- S_{\tau_j+\xi}| + \omega_4 N$, since every degree is at most 4 and ω_d is increasing in d;
- for $j > 2K$, similarly Corollary A.25(a) gives $\mathbb{E}X_j^{\mathrm{out}} \leq \mathbb{E}|\partial^+ S_{\tau_j-\xi}| + \omega_4 N$ with $\tau_j - \xi > 0$;
- Corollary A.25(c), applied to the indicator of $v \in U \cup (\hat{D} \cap W)$ as a function of the set of vertices of $N[v]$ with negative label (for the ideal labels, $Y_u < 0$ and $Y_u \leq 0$ differ by a null event), gives $\mathbb{E}\hat{L}_0 \leq \mathbb{E}\hat{L}_0^{\mathrm{id}} + \omega_4 N$, where $\hat{L}_0^{\mathrm{id}}$ is $\hat{L}_0$ for the ideal labels; similarly for $\hat{L}_0'$;
- the windows and tails are bounded by Corollary A.25(d) and (N2): $\mathbb{E}X_j^{\mathrm{win}} \leq (\varphi(0)/K + 2\xi\varphi(0) + p_t)N$ and $\mathbb{E}X^{\mathrm{lo}}, \mathbb{E}X^{\mathrm{hi}} < 0.023N$.

With Lemma A.37 (switched setting) and Lemma A.26(a) (cubic setting), in both settings

$$\mathbb{E}X_j^{\mathrm{in}},\ \mathbb{E}X_j^{\mathrm{out}},\ \mathbb{E}\hat{L}_0,\ \mathbb{E}\hat{L}_0' \leq \beta(H) + (\varepsilon_{\mathrm{med}}(h) + \omega_4)\, N.$$

An injective labeling. Let ν be the largest name of a vertex of H, let Σ be the finite set $\{\tilde{Y}_v : v \in V(H)\} \cup \{0\} \cup \{\tau_j : 0 \leq j \leq 4K\}$, let $\vartheta > 0$ be so small that $\vartheta\nu$ is less than every positive difference of two elements of Σ, and put $y_v := \tilde{Y}_v + \vartheta v$. Exactly as in Appendix A.5.8, with ν in place of n, the labeling y is injective, induces the order of $(\tilde{Y}_v, v)$, and has sublevel sets $S_\tau = \tilde{S}_\tau$ for every $\tau \in \Sigma$. In particular U, W and $\hat{D}$ are the same for y as for the discrete labels. By Lemma 8.5 in the cubic setting, and by the same lemma with $\hat{D}$ in place of D in the switched setting (which is established in the proof of Lemma A.9), the layout is a path decomposition of H of width at most w whenever w satisfies (i)–(iii) of Lemma 8.4 for y, with $\hat{L}_0$ and $\hat{L}_0'$ in (iii); the component order satisfies (8.3) by Lemma A.33.

The hypotheses. Put $w := \beta(H) + \rho_{\mathrm{med}}(n)N + 2\theta + 1$, and assume that every count is at most its mean plus θ. For a threshold $\tau < 0$ with $\tau \geq -2$, let $\tau \in [\tau_j, \tau_{j+1})$ with $j < 2K$. The argument of Appendix A.5.8 gives $|\partial^- S_\tau| \leq X_j^{\mathrm{in}} + X_j^{\mathrm{win}}$, which is at most

$$\beta(H) + (\varepsilon_{\mathrm{med}}(h) + \omega_4)N + (\varphi(0)/K + 2\xi\varphi(0) + p_t)N + 2\theta = w - 1.$$

For $0 < \tau < 2$, likewise $|\partial^+ S_\tau| \leq X_{j+1}^{\mathrm{out}} + X_j^{\mathrm{win}} \leq w - 1$, where $\tau \in [\tau_j, \tau_{j+1})$ with $2K \leq j < 4K$. For $\tau < -2$ and $\tau \geq 2$, the frontiers are bounded by the tails, which are at most $0.023N + \theta \leq c_3 N + \theta \leq \beta(H) + \theta \leq w - 1$. Finally, $\max(\hat{L}_0, \hat{L}_0') + 1 \leq \beta(H) + (\varepsilon_{\mathrm{med}}(h) + \omega_4)N + \theta + 1 \leq w$. So (i)–(iii) hold, and the width is at most w. □

For $H = G$ in the cubic setting, Proposition A.38 is the width bound proved in Appendix A.5.8, with ω_4 in place of ω_3.

A.6.3 The switched case

The algorithms. Let G be a graph with $\Delta(G) \le 4$ and $n \ge 3$ vertices (graphs with $n \le 2$ are decomposed by one bag). The randomized, Las Vegas and deterministic algorithms are those of Appendix A.5.7, with the following changes. In step 1, reduction operations are applied to G until none applies; the final graph H is empty or has $\delta(H) \ge 3$ by (RP3), and $\Delta(H) \le 4$ by (RP1), so it is in the switched setting. In step 2 the coefficients of H use the weights $q(x) = d_H(x) - 1$. In step 3 the algorithms build the median-switched layout of H and use its $8K + 4$ counts. Step 4 applies Lemma A.28 as in the mixed cases.

Proof of Theorem 10.1 in the switched case. Let $n \ge n_2$; for $n < n_2$ the claims are trivial (Appendix A.5.1). The coefficients $(0, 0, 0, c_3, 0.2627)$ are nonnegative and nondecreasing in the degree, and reductions remove vertices and increase no degree by (RP1), so $\beta(H) \le \beta(G)$. Put

$$\rho(n) := \rho_{\mathrm{med}}(n)\, n + 2 \max\{\theta_{\mathrm{D}}(n), \theta_{\mathrm{R}}(n)\} + 3 \log_2 n + 2 \qquad (n \ge n_2),$$

with $\theta_{\mathrm{D}}(n) = \sqrt{8(8K+4)n}\, t^{1/4}$ and $\theta_{\mathrm{R}}(n) = n^{3/4}$, and $\rho(n) := n$ for $n < n_2$; we replace ρ by its running maximum (4.2). Then $\rho(n) = o(n)$, because $\rho_{\mathrm{med}}(n) \to 0$. The three algorithms ensure that every count is at most its mean plus θ, exactly as in Appendices A.5.7 and A.5.8: deterministically with $\theta = \theta_{\mathrm{D}}(n)$ by Lemma A.27(c); in the randomized algorithm with $\theta = n^{3/4}$, except with probability at most $\pi(n) := 8(8K+4)t^{1/2}n^{-1/2} + 1/\lceil \log_2 n \rceil$; and in the Las Vegas algorithm with $\theta = n^{3/4}$, which accepts a sample with probability at least $\frac{1}{2}$ by Lemma A.27(b) and (N3). In each case Proposition A.38, applied with $N = |V(H)| \le n$, bounds the width of the layout of H by $\beta(G) + \rho_{\mathrm{med}}(n)n + 2\theta + 1$, and Lemma A.28 adds at most $3 \log_2 n + 1$. So the width is at most $\beta(G) + \rho(n)$. The running times are those of the cubic case (Appendix A.5.10): the reductions take $O(n)$ operations, since degrees never exceed 4, and the coefficients are computed as in the mixed cases. □

A.6.4 The sweep components of the bootstrap cases C and M

In the bootstrap case of a seed X (Appendix A.5.9), the deletion process stops at sweep components J with $\delta(J) \ge 3$ and $\Delta(J) \le k$, where $k = 3$ for $X = \mathsf{C}$ and $k = 4$ for $X = \mathsf{M}$, and decomposes them with the parameters of the number $n_{\ge 1} \ge |V(J)|$ of non-isolated vertices of the input graph.

Corollary A.39 (base layouts for C and M)**.** *Let $X \in \{\mathsf{C}, \mathsf{M}\}$ with base degree k, let $n \ge n_2$ and $\theta \ge 0$, and let J be a graph with $\delta(J) \ge 3$, $\Delta(J) \le k$ and $N_J \le n$ vertices, labeled in the cubic setting if $X = \mathsf{C}$ and in the switched setting if $X = \mathsf{M}$, with the parameters of n. If every count of the median-switched layout of J is at most its mean plus θ, then the layout is a path decomposition of J of width at most $\sum_{d=3}^{k} b_d n_d(J) + \rho_{\mathrm{med}}(n) N_J + 2\theta + 1$, where $(b_3, \ldots, b_k) = X$.*

Proof. For $X = \mathsf{C}$ the graph J is 3-regular, so $\beta(J) = c_3 N_J = c_3 n_3(J)$. For $X = \mathsf{M}$, $\beta(J) = c_3 n_3(J) + 0.2627\, n_4(J)$. In both cases Proposition A.38 gives the claim. □

A.7 Proofs for Part II

This subsection proves Lemma 11.2 (Appendix A.7.1), the statements on bucketing (Appendix A.7.2) and Theorem 14.1 (Appendix A.7.3). The analyses for Sections 12 and 13 are in Appendices C and D.

A.7.1 Pathwidth in terms of the vertices of degree three

Lemma 11.2 (pathwidth by the number of degree-3 vertices)**.** *There is a nondecreasing function* $\rho\colon \mathbb{N} \to \mathbb{R}$ *with* $\rho(k) = o(k)$ *such that every graph* H *with* $\Delta(H) \le 3$ *satisfies*

$$\mathrm{pw}(H) \le c_3\, n_3(H) + \rho\big(n_3(H)\big).$$

A path decomposition of this width can be computed in polynomial time.

Proof. Let $\rho_1\colon \mathbb{N} \to \mathbb{R}$ be the larger of the remainder of Theorem 8.1 and the remainder of the deterministic algorithm of Theorem 1.5 for that bound; replacing it by its running maximum (4.2), we take ρ_1 nonnegative and nondecreasing. We prove the lemma with

$$\rho(k) := \rho_1(k) + \log_2(2k+2) + 6 \qquad (k \in \mathbb{N}).$$

Write $n_3 := n_3(H)$. The proof has four steps: we remove the trees outside the 2-core of H, suppress the vertices of degree 2 of the core and decompose the simple graph underlying the resulting cubic multigraph by Theorem 8.1, and then put back first the suppressed paths and then the trees.

Step 1: the 2*-core.* Repeatedly delete a vertex of degree at most 1 until none is left, and call the resulting graph H' (possibly empty). No degree increases, so $n_3(H') \le n_3$. Every component T of $H - V(H')$ is a tree with at most one edge to H'. Otherwise T together with H' contains a cycle through a vertex of T, or a path between two vertices of H' whose interior lies in T; the first vertex of this cycle or path to be deleted had two neighbors at the time of its deletion, a contradiction.

Step 2: the core. Every vertex of H' has degree 2 or 3. Components of H' that are cycles have pathwidth at most 2. In every other component, every maximal path whose interior vertices have degree 2 joins two vertices of degree 3 (possibly the same one). Suppressing all vertices of degree 2 of these components therefore gives a multigraph M, possibly with loops and parallel edges, on the $n_3(H')$ vertices of degree 3 of H'. Its underlying simple graph G_M has maximum degree at most 3 and $n_3(H')$ vertices. By Theorem 8.1 and Theorem 1.5 we can compute a path decomposition of G_M of width at most $c_3 n_3(H') + \rho_1(n_3(H')) \le c_3 n_3 + \rho_1(n_3)$.

Step 3: restoring the subdivisions. An edge e of M with ends u, w (possibly $u = w$) stands for a path $u\, x_1 \cdots x_q\, w$ of H' with $q \ge 0$. If $q \ge 1$, choose a bag B of the current decomposition that contains u and w (it exists, because $uw \in E(G_M)$ or $u = w$). Directly after B insert the bags

$$B \cup \{x_1\},\ B \cup \{x_1, x_2\},\ B \cup \{x_2, x_3\},\ \ldots,\ B \cup \{x_{q-1}, x_q\}.$$

These cover the edges ux_1, $x_i x_{i+1}$ and $x_q w$. Each x_i occurs in consecutive bags, and every vertex of B still occurs in an interval of bags, since the inserted bags contain B. Paths that use the same bag B are inserted one after another. The width grows by at most 2. Finally append, for each cycle component of H', a decomposition of width 2. The result is a decomposition of H' of width at most $c_3 n_3 + \rho_1(n_3) + 2$.

Step 4: the trees. Let T be a component of $H - V(H')$ with N vertices, and let $t \le n_3$ be the number of vertices of degree 3 in T. If $N \ge 2$, then T has exactly $t + 2$ leaves (count $\sum_v (d_T(v) - 2) = -2$), and suppressing its vertices of degree 2 gives a tree T° with at most $2t + 2$ vertices; if $N = 1$, put $T^\circ = T$. A tree on $N^\circ \ge 1$ vertices has pathwidth at most $\log_2 N^\circ$: by induction, put a centroid into every bag of a concatenation of decompositions of the components of the tree minus the centroid, each of which has at most $N^\circ/2$ vertices. Restoring the subdivisions of T° as in Step 3 adds at most 2, so $\mathrm{pw}(T) \le \log_2(2t+2) + 2$. If T has an edge to a vertex v of H', take a bag $B \ni v$ of the decomposition of H' and insert directly after it the bags $B \cup Y$, for the bags Y of a decomposition of T in order; this covers the edge between T and v, and the width

grows by at most $\mathrm{pw}(T)+1$. Trees attached at the same bag are inserted one after another, so the growth is the maximum, not the sum, over the trees. Trees without an edge to H' are appended as separate segments.

Combining the steps,

$$\mathrm{pw}(H) \le c_3 n_3 + \rho_1(n_3) + 2 + \log_2(2n_3+2) + 3 \le c_3 n_3 + \rho(n_3),$$

and the bound also holds if H' is empty. Every step takes polynomial time; Step 2 computes the decomposition of G_M by the deterministic algorithm of Theorem 1.5. The function ρ is nondecreasing and $\rho(k) = o(k)$. □

A.7.2 Bucketing

We use $p(s)$ and P_k as defined before Lemma 11.6.

Lemma 11.6 (cycle cover plus path edges)**.** *There is a function $\rho(k) = o(k)$ such that the following holds. Let C be a multigraph on $[k]$, possibly with loops, of maximum degree at most 2, and let $A \subseteq P_k$ with $|A| = s$. Then the simple graph H underlying $C \cup A$ satisfies $\mathrm{pw}(H) \le p(s) + \rho(k)$, and a path decomposition of H of this width can be computed in polynomial time.*

Proof. Every vertex lies on at most two edges of P_k, so $d_H(v) \le d_C(v) + d_A(v) \le 4$. Let n_d be the number of vertices of degree d in H, which has k vertices; since $\Delta(H) \le 4$, $n_5 = 0$. Theorem 7.2 with the triple R and Theorem 1.5 give a path decomposition of H of width at most $0.1455\,n_3 + 0.2530\,n_4 + \rho(k)$, where ρ is the remainder of this algorithm. A vertex of degree $d \ge 3$ in H has $d_A(v) \ge d-2$, so $n_3 + 2n_4 \le \sum_v d_A(v) = 2s$; also $n_3 + n_4 \le k$. Hence

$$0.1455\,n_3 + 0.2530\,n_4 = 0.1455\,(n_3+2n_4) - 0.038\,n_4 \le 0.291\,s \quad \text{and}$$
$$0.1455\,n_3 + 0.2530\,n_4 = 0.038\,(n_3+n_4) + 0.1075\,(n_3+2n_4) \le 0.038\,k + 0.215\,s.$$ □

Lemma 11.7 (the trade-off)**.** *Let $\alpha = \frac{200}{243}$. For every $0 \le s \le k$,*

$$(1-\alpha)(k-s) + \alpha\,p(s) \le \tfrac{253}{1215}k.$$

Moreover, no $\alpha \in [0,1]$ gives a smaller bound than $\frac{253}{1215}k$ on $\max_{0\le s\le k}[(1-\alpha)(k-s)+\alpha p(s)]$.

Proof. Here $1-\alpha = \frac{215}{1215}$ and $\alpha = \frac{1000}{1215}$. Since $\alpha \ge 0$ and $p(s) \le 0.038\,k + 0.215\,s$,

$$(1-\alpha)(k-s) + \alpha\,p(s) \le \frac{215(k-s) + 38k + 215s}{1215} = \frac{253k}{1215}.$$

For the second statement, let $s^* := \frac{962}{1215}k$. Then $k - s^* = \frac{253}{1215}k$. The two terms in the definition of p are equal at $s = k/2$, and the second is the smaller one for $s \ge k/2$; since $s^* \ge k/2$, $p(s^*) = 0.038\,k + 0.215\,s^* = \frac{253}{1215}k$. So at $s = s^*$ the expression equals $\frac{253}{1215}k$ for every $\alpha \in [0,1]$. □

The three bounds of Theorem 7.1 and the bound of Theorem 8.2 also give upper bounds on $\mathrm{pw}(H)$ in Lemma 11.6: $F_4(n_3+n_4)$, $F_3 n_3 + 0.2575\,n_4$, $\frac{5}{32}n_3 + \frac{1}{4}n_4$ and $c_3 n_3 + 0.2627\,n_4$, up to $o(k)$. As long as only the constraints $n_3 + 2n_4 \le 2s$ and $n_3 + n_4 \le k$ of the proof of Lemma 11.6 are used, these four bounds do not lower the exponent $\frac{253}{1215}k$, even together with the bound of R. Indeed, let $s^* = \frac{962}{1215}k$ as in the proof above, and take $n_3 = \frac{506}{1215}k$ and $n_4 = \frac{709}{1215}k$, which satisfy both constraints at $s = s^*$ with equality. At these values $0.1455\,n_3 + 0.2530\,n_4 = \frac{253}{1215}k = k - s^* < 0.20824k$, while the four other bounds are at least $0.2457k$, $0.2086k$, $0.2109k$ and $0.2114k$, by the lower bounds on F_4 and F_3 of Proposition 6.9 and by $c_3 > 0.1397$. So the largest value, over all reals $n_3, n_4 \ge 0$ that

satisfy both constraints at $s = s^*$, of the minimum of the five bounds is $\frac{253}{1215}k = p(s^*)$: it is at least this value at the point above, and at most the bound of R, which is at most $p(s^*)$ by the proof of Lemma 11.6. With this value in place of $p(s^*)$, the expression of Lemma 11.7 at $s = s^*$ still equals $\frac{253}{1215}k$ for every $\alpha \in [0,1]$.

Theorem 11.8 (permutation patterns and k-opt)**.** *There are a function $\rho(k) = o(k)$ and numbers $\varepsilon_k \to 0$ such that:*

(a) *the decision and counting versions of Permutation Pattern Matching, for a pattern of length k and a text of length n, can be solved in time $2^{O(k)}\, n^{\frac{253}{1215}k+\rho(k)}$;*

(b) *for every fixed k, k-opt Optimization can be solved in time $O(n^{(253/1215+\varepsilon_k)k})$ on n-vertex edge-weighted graphs, and in time $O(n^{(7/48+\varepsilon_k)k})$ on those of bounded maximum degree.*

Proof. For (a) we follow Berendsohn, Kozma and Marx: we write the problem as a binary constraint satisfaction problem, split the text positions into intervals (strips), guess which entries of the pattern are leftmost in their strip and the strips of these entries, and solve each restricted instance by the dynamic program of Lemma 11.1; Lemmas 11.6 and 11.7 bound the total time. For (b) we substitute Lemmas 11.6 and 11.7 into the exponent of Cygan, Kowalik and Socała, and Remark 9.14 into the algorithm of Bonnet, Iwata, Jansen and Kowalik.

(a) *The constraint satisfaction formulation.* Following Berendsohn, Kozma and Marx [BKM21, §3], take variables $x_1, \dots, x_k$ with domain $[n]$; the value of x_i is the text position $f(i)$. Impose the *index constraints* $x_i < x_{i+1}$ for $i \in [k-1]$, and the *value constraints* $\tau(x_a) < \tau(x_b)$ for $a = \pi^{-1}(v)$, $b = \pi^{-1}(v+1)$ and $v \in [k-1]$. The solutions are exactly the occurrences: the index constraints say that f is strictly increasing, and the value constraints say that $\tau \circ f \circ \pi^{-1}$ is strictly increasing, which is equivalent to the order condition. The constraint graph is the union of the path P_k (index constraints) and the path Q on $[k]$ with edges $\{\pi^{-1}(v), \pi^{-1}(v+1)\}$ (value constraints).

Bucketing. Let $\alpha = \frac{200}{243}$, $q := \lceil n^\alpha \rceil$, and split $[n]$ into the $J := \lceil n/q \rceil \le n^{1-\alpha} + 1 \le 2n^{1-\alpha}$ intervals (*strips*) $I_j := \{(j-1)q+1, \dots, jq\} \cap [n]$, each of size at most $q \le 2n^\alpha$. (Berendsohn, Kozma and Marx use $n^{1/4}$ strips of width $n^{3/4}$; nothing else in their argument depends on the width.) For an occurrence f let $\sigma(i)$ be the index of the strip containing $f(i)$; since f is increasing, σ is nondecreasing. Let

$$X(f) := \{1\} \cup \{i \in \{2, \dots, k\} : \sigma(i-1) < \sigma(i)\},$$

the entries that are leftmost in their strip. For every i, $\sigma(i) = \sigma(i^*)$, where i^* is the largest element of $X(f)$ with $i^* \le i$, because σ does not increase between consecutive elements of $X(f)$.

The algorithm enumerates all *guesses*: sets $X \subseteq [k]$ with $1 \in X$ and strictly increasing maps $\sigma_X \colon X \to [J]$. For a guess it solves the *restricted instance*: the domain of x_i is $I_{\sigma_X(i^*)}$, with i^* the largest element of X that is at most i; all value constraints are kept; the index constraint between i and $i+1$ is kept if $i+1 \notin X$ and dropped if $i+1 \in X$. We claim that the solutions of the restricted instance are exactly the occurrences f with $X(f) = X$ and $\sigma|_X = \sigma_X$. Let f be a solution. A dropped constraint holds anyway: if $i+1 \in X$, then $f(i)$ lies in the strip $\sigma_X(i^*) < \sigma_X(i+1)$, so $f(i) < f(i+1)$. Hence f is an occurrence. Its strips are $\sigma(i) = \sigma_X(i^*)$, so for $i \in X \setminus \{1\}$ the strip strictly increases at i, and for $i \notin X$ the entries $i-1$ and i share a strip; thus $X(f) = X$ and $\sigma|_X = \sigma_X$. Conversely, an occurrence with $X(f) = X$ and $\sigma|_X = \sigma_X$ satisfies all constraints, and $f(i)$ lies in the strip $\sigma(i) = \sigma(i^*) = \sigma_X(i^*)$. So every occurrence is a solution of exactly one restricted instance. For the decision version we test each guess; for the counting version we add the numbers of solutions over all guesses.

Running time. The constraint graph of a restricted instance is the union of Q and the set $A_X := \{\{i, i+1\} : i+1 \notin X\}$ of $s := k - |X|$ edges of P_k (the $|X|-1$ elements of $X \setminus \{1\}$ each drop one index constraint). Q is a path, so Lemma 11.6 with $C = Q$ gives, in polynomial time, a path decomposition of width $w \le p(s) + \rho_1(k)$, where ρ_1 is the remainder there. Every domain has size at most $2n^\alpha$. We may assume $w \le k-1$ (otherwise use a single bag), so Lemma 11.1(a) over the counting semiring $(\mathbb{N}, +, \times)$, with scores in $\{0,1\}$ given by the constraints, takes $(2n^\alpha)^{w+1}\,\mathrm{poly}(n)$ steps; the restricted instance is built in polynomial time. (This is [BKM21, Lemmas 2 and 4].) There are at most $J^{|X|} \le (2n^{1-\alpha})^{|X|}$ maps σ_X for each X, and at most 2^k sets X. Since $|X| \le k$ and $w+1 \le k$, the total time is at most

$$\sum_X (2n^{1-\alpha})^{|X|}(2n^\alpha)^{p(k-|X|)+\rho_1(k)+1}\,\mathrm{poly}(n) \le 2^{3k}\, n^{\max_{0\le s\le k}[(1-\alpha)(k-s)+\alpha p(s)]+\rho_1(k)+1+O(1)},$$

and Lemma 11.7 bounds the maximum by $\frac{253}{1215}k$. With $\rho(k) := \rho_1(k) + c$, for ρ_1 taken nondecreasing by (4.2) and a suitable constant c, this is $2^{O(k)} n^{\frac{253}{1215}k+\rho(k)}$. Since $k \le n$, also $2^{O(k)} = n^{O(k/\log_2 n)} \le n^{O(k/\log_2 k)} = n^{o(k)}$, so the bound can be written $n^{\frac{253}{1215}k+o(k)}$, as Berendsohn, Kozma and Marx do for $n^{k/4}$.

(b) *General graphs.* Cygan, Kowalik and Socała [CKS19, Lemma 10] solve k-opt Optimization in time $2^{O(k\log k)}n^{c(k)}$, where

$$c(k) = \max_M \min_{\alpha\in[0,1]} \max_{A\subseteq P_k} \Big[(1-\alpha)(k-|A|) + \alpha\big(\mathrm{tw}(I_M \cup A) + 1\big)\Big],$$

the outer maximum ranges over the valid connection patterns M, and I_M is a graph on $[k]$ associated with M. Every component of I_M is a cycle or a single edge [CKS19, Proposition 3], and, as in the proof of [CKS19, Lemma 8], doubling every single-edge component gives a multigraph $I_M^* \supseteq I_M$ on $[k]$ in which every vertex has degree 2. Hence $I_M \cup A$ is a subgraph of the simple graph underlying $I_M^* \cup A$, and Lemma 11.6 with $C = I_M^*$ gives $\mathrm{tw}(I_M \cup A) \le p(|A|) + \rho_1(k)$. Taking $\alpha = \frac{200}{243}$ for every M, Lemma 11.7 gives $c(k) \le \frac{253}{1215}k + \rho_1(k) + 1$. Put $\varepsilon_k := (\rho_1(k)+1)/k \to 0$; for fixed k the factor $2^{O(k\log k)}$ is a constant, so the time is $O(n^{(253/1215+\varepsilon_k)k})$.

Bounded degree. Bonnet, Iwata, Jansen and Kowalik [BIJK19b, Thm. 3] solve k-opt Optimization on graphs of maximum degree at most a constant by running, for each of at most $(2k)!$ patterns, a dynamic program in time $O(n^{\mathrm{tw}(C_M)+1})$ for fixed k and maximum degree, where C_M is a simple graph that is a subgraph of an even multigraph with k edges. So C_M has at most k edges, and Remark 9.14 gives $\mathrm{tw}(C_M) \le \mathrm{pw}(C_M) \le \frac{7}{48}k + o(k)$. The time is $O(n^{(7/48+\varepsilon'_k)k})$ with $\varepsilon'_k \to 0$. The larger of ε_k and ε'_k serves for both bounds. □

A.7.3 Connectivity problems

Theorem 14.1 (connectivity problems)**.** *Polynomial factors are omitted, all algorithms use exponential space, and the Monte Carlo algorithms have one-sided error at most* $\frac{1}{2}$.

(a) *Feedback Vertex Set and Connected Vertex Cover can be solved by Monte Carlo algorithms in time* $2^{m/6+o(n)} \le 2^{n/3+o(n)} = O(1.2600^n)$ *on graphs with* m *edges and maximum degree* 4*, and in time* $2^{\gamma_5 m+o(n)} = O(1.4262^n)$ *on graphs of maximum degree* 5*, where* $\gamma_5 < 0.20484$ *is the constant of Theorem 3.6. Steiner Tree with unit weights can be solved by Monte Carlo algorithms in time* $3^{c_3 n+o(n)} = O(1.1661^n)$ *and* $3^{F_4 n+o(n)} = O(1.3100^n)$ *on graphs of maximum degree* 3 *and* 4*. Feedback Vertex Set can be solved on graphs of maximum degree* 6 *by a Monte Carlo algorithm in time* $3^{7n/16+o(n)} = O(1.6172^n)$*. Feedback Vertex Set can be solved on graphs of maximum degree* 4 *deterministically in time* $5^{F_4 n+o(n)} = O(1.4853^n)$.

(b) *Connected Dominating Set, Connected Feedback Vertex Set, Connected Odd Cycle Transversal, Exact k-Leaf Spanning Tree (in particular Minimum and Maximum Leaf Spanning Tree), Maximum Full Degree Spanning Tree, Graph Metric Traveling Salesperson, and the undirected versions of Min Cycle Cover and Longest Cycle can be solved by Monte Carlo algorithms in time* $4^{c_3 n+o(n)} = O(1.2139^n)$, $4^{F_4 n+o(n)} = O(1.4060^n)$ *and* $4^{11n/32+o(n)} = O(1.6105^n)$ *on graphs of maximum degree* 3, 4 *and* 5. *Minimum Leaf Spanning Tree (equivalently, Maximum Internal Spanning Tree) can be solved on graphs of maximum degree* 6 *by a Monte Carlo algorithm in time* $4^{7n/16+o(n)} = O(1.8341^n)$, *and Connected Dominating Set on graphs of maximum degree* 3 *by a Monte Carlo algorithm in time* $3^{c_3 n+o(n)} = O(1.1661^n)$. *Maximum Leaf Spanning Tree has the bounds of Connected Dominating Set.*

(c) *Hamiltonian Cycle can be solved by Monte Carlo algorithms in time* $(1+\sqrt{2})^{c_3 n+o(n)} = O(1.1312^n)$ *on graphs of maximum degree* 3, *in time* $(1+\sqrt{2})^{m/6+o(n)} \leq (1+\sqrt{2})^{n/3+o(n)} = O(1.3416^n)$ *on graphs with* m *edges and maximum degree* 4, *and in time* $(2+\sqrt{2})^{11n/32+o(n)} = O(1.5252^n)$ *on graphs of maximum degree* 5. *Undirected Longest Path can be solved by Monte Carlo algorithms in time* $(2+\sqrt{2})^{c_3 n+o(n)} = O(1.1874^n)$, $(2+\sqrt{2})^{F_4 n+o(n)} = O(1.3523^n)$ *and* $(2+\sqrt{2})^{11n/32+o(n)} = O(1.5252^n)$ *on graphs of maximum degree* 3, 4 *and* 5.

(d) *With* $\omega \leq \frac{9}{4}$ *[Ope26b, AvdBZ26], the traveling salesperson problem with nonnegative integer edge weights can be solved deterministically in time* $(1+2^{\omega/2})^{c_3 n+o(n)} = O(1.1757^n)$ *on graphs of maximum degree* 3, *and in time* $(2+2^{\omega/2})^{F_4 n+o(n)} = O(1.4214^n)$, $(2+2^{\omega/2})^{11n/32+o(n)} = O(1.6352^n)$ *and* $(2+2^{\omega/2})^{7n/16+o(n)} = O(1.8699^n)$ *on graphs of maximum degree* 4, 5 *and* 6.

(e) *The directed versions of Min Cycle Cover, Longest Path and Longest Cycle, and Exact k-Leaf Outbranching, can be solved by Monte Carlo algorithms in time* $6^{c_3 n+o(n)} = O(1.2848^n)$ *and* $6^{F_4 n+o(n)} = O(1.5533^n)$ *on directed graphs whose underlying undirected graph has maximum degree* 3 *and* 4.

(f) *Co-Path Packing (delete at most k vertices so that every component of the remaining graph is a path) can be solved by a Monte Carlo algorithm in time* $O^*(2.8192^k)$, *and k-Co-Path Set (delete at most k edges so that every component of the remaining graph is a path) by a Monte Carlo algorithm in time* $O^*(1.5423^k)$.

(g) *On graphs with* m *edges, Hamiltonian Cycle can be solved by a Monte Carlo algorithm in time* $(2+\sqrt{2})^{7m/48+o(m)} = O(1.1962^m)$, *and the traveling salesperson problem with nonnegative integer edge weights deterministically in time* $(2+2^{\omega/2})^{7m/48+o(m)} = O(1.2320^m)$, *with* ω *as in* (d). *For average degree at most* d, *that is* $m \leq \frac{d}{2}n$, *these are* $O(1.0937^{dn})$ *and* $O(1.1100^{dn})$.

Proof. For a graph of maximum degree 4, Theorem 10.1(b) computes a path decomposition of width at most $F_4 n + o(n)$ (the bound of Theorem 7.1(a)) in polynomial time, and for maximum degree 3 one of width at most $c_3 n + o(n)$ (the bound of Theorem 8.1); by Theorem 1.3 and Theorem 1.5, it computes one of width at most $\frac{11}{32}n + o(n)$ for maximum degree 5 and $\frac{7}{16}n + o(n)$ for maximum degree 6, since $a_d \leq a_5$ for $d \leq 5$ and $a_d \leq a_6$ for $d \leq 6$. A path decomposition is a tree decomposition of the same width. For maximum degree 3, Corollary 10.2 computes an ordering of cutwidth at most $c_3 n + o(n)$ in polynomial time. For a graph with m edges and maximum degree 4 or 5, Theorem 10.3 computes, in deterministic polynomial time, an ordering of cutwidth at most $\frac{m}{6} + o(n) \leq \frac{n}{3} + o(n)$ or $\gamma_5 m + o(n) \leq \frac{5}{2}\gamma_5 n + o(n)$, respectively. For a graph with m edges, Theorem 10.1 (general case) gives a path decomposition of width at most $\frac{7}{48}m + o(m)$ in polynomial time. We run the following algorithms on these decompositions and orderings.

(a) Bojikian, Chekan, Hegerfeld and Kratsch solve Connected Vertex Cover and Feedback Vertex Set by Monte Carlo algorithms with one-sided error in time $2^{\mathrm{ctw}}n^{O(1)}$ for a given ordering of cutwidth ctw [BCHK22, Theorems 12 and 43], on all graphs. We use them for maximum degree 4 and 5; the numerical bases follow from $2^{1/3} < 1.25993$ and $2^{5\gamma_5/2} < 1.42611$, where $\gamma_5 = \frac{1}{\pi}\arccos\frac{4}{5}$. Cygan et al. [CNP+11, Theorem 1.2] solve the three problems in time $3^{\mathrm{tw}}n^{O(1)}$, which we use for Steiner Tree and for Feedback Vertex Set at maximum degree 6, and Bodlaender et al. solve Feedback Vertex Set deterministically in time $5^{\mathrm{pw}}n^{O(1)}$ [BCKN15, Theorem 4.5].

(b) Cygan et al. solve these problems in time $4^{\mathrm{tw}}n^{O(1)}$ [CNP+11, Theorem 1.2]. For maximum degree 3, Bojikian et al. solve Connected Dominating Set in time $3^{\mathrm{ctw}}n^{O(1)}$ for a given ordering of cutwidth ctw [BCHK22, Theorem 23]. Maximum Leaf Spanning Tree has the bounds of Connected Dominating Set because a connected graph with $n \geq 3$ vertices has a spanning tree with at least ℓ leaves if and only if it has a connected dominating set of at most $n - \ell$ vertices. Indeed, the vertices that are not leaves of a spanning tree form a connected dominating set, and a spanning tree of the subgraph induced by a connected dominating set extends to a spanning tree in which every other vertex is a leaf. Finally, a spanning tree has $n - \ell$ internal vertices if it has ℓ leaves, so Minimum Leaf Spanning Tree and Maximum Internal Spanning Tree have the same optimal solutions.

(c) Bojikian, Chekan and Kratsch decide Hamiltonicity by a Monte Carlo algorithm with one-sided error in time $(1+\sqrt{2})^{\mathrm{ctw}}n^{O(1)}$ for a given ordering of cutwidth ctw, on all graphs [BCK25, Theorem 47]. We use it for maximum degree 3 and 4, where $(1+\sqrt{2})^{1/3} < 1.34151$. (The proof of [CKN12, Corollary 4.5] gives the same base for a path decomposition of a graph of maximum degree 3.) For maximum degree 5 we use the algorithm of [CKN12, Theorem 4.3]; there the ordering of Theorem 3.6 would give only $(1+\sqrt{2})^{5\gamma_5 n/2} > 1.570^n$. For Longest Path, Cygan, Kratsch and Nederlof decide whether a graph has a path of length k in time $(2+\sqrt{2})^{\mathrm{pw}}n^{O(1)}$ for a given path decomposition of width pw [CKN12, Theorem 4.6], with an algorithm that extends the Monte Carlo algorithm of their Theorem 4.3 by a counter; we run it for each k.

(d) For maximum degree 4, 5 and 6 we use the algorithm of [BCKN15, Theorem 3.18]. Its exponent ω enters through minimum weight row bases over GF(2) [BCKN15, Lemma 3.15], which reduce to products of matrices over GF(2). The bound $\omega \leq \frac{9}{4}$ is proved over the complex numbers in [Ope26b], and Amanov, van den Berg and Zuiddam prove it over every field [AvdBZ26, Theorem 1.1], in particular over GF(2). For cubic graphs, Bodlaender et al. obtain their bound $1.2186^n n^{O(1)}$ [BCKN15, Corollary 3.19] by running that algorithm on a path decomposition with fewer states. By an observation of Cut&Count [CNP+11], they discard the states in which the degree of a vertex in a partial solution is incompatible with the number of its edges introduced so far. In a graph of maximum degree 3, a vertex of a bag that has at most one introduced edge then has degree 0 or 1, and one with two introduced edges has degree 1 or 2. So besides degree 1 each vertex has one possible degree instead of two, and the count of states in the proof of [BCKN15, Theorem 3.18] gives the base $1+2^{\omega/2}$ instead of $2+2^{\omega/2}$. With the width $\frac{n}{6}+o(n)$ this is $(1+2^{\omega/2})^{n/6+o(n)} = O(1.2187^n)$ with $\omega < 2.3727$, which they state as 1.2186^n, and $O(1.2128^n)$ with $\omega \leq \frac{9}{4}$. On a path decomposition of width $c_3 n + o(n)$, the same algorithm runs in time $(1+2^{\omega/2})^{c_3 n+o(n)}$.

(e) Cygan et al. solve these problems in time $6^{\mathrm{tw}}n^{O(1)}$ for a tree decomposition of width tw of the underlying undirected graph [CNP+11, Theorem 1.2].

(g) We run the algorithms of [CKN12, Theorem 4.3] and [BCKN15, Theorem 3.18] on a path decomposition of width $\frac{7}{48}m+o(m)$. Numerically $(2+\sqrt{2})^{7/48} < 1.19612$ and $(2+2^{\omega/2})^{7/48} < 1.23198$ for $\omega \leq \frac{9}{4}$, and the bounds in dn are their square roots.

(f) *Co-Path Packing.* Liu and Xiao [LX26] branch, with branching factor 2.8192, until the remaining instance (G, k'), with $k' \leq k$, is a *proper* graph [LX26, Section 3.1]: its maximum degree is at most 4, and if it has a solution of size at most k', then $|V(G)| \leq 100k'$ and $\frac{n_3}{6} + \frac{n_4}{3} \leq \frac{2k'}{3}$ [LX26,

Lemma 2], where n_d is the number of vertices of degree d. On a proper graph, if these two conditions hold, they run a Monte Carlo algorithm in time $5^{\mathrm{pw}} n^{O(1)}$ [LX26, Theorem 4] on a path decomposition of width $\frac{2k'}{3} + o(k')$, which the bound of Fomin, Gaspers, Saurabh and Stepanov gives; otherwise they reject the instance. If these two conditions hold, Theorem 1.3 and Theorem 1.5 give instead a path decomposition of width at most

$$\tfrac{5}{32} n_3 + \tfrac{1}{4} n_4 + o(n) = \tfrac{15}{16} \cdot \tfrac{n_3}{6} + \tfrac{3}{4} \cdot \tfrac{n_4}{3} + o(n) \leq \tfrac{15}{16} \cdot \tfrac{2k'}{3} + o(k') = \tfrac{5}{8} k' + o(k'),$$

and the dynamic program takes time $5^{5k'/8 + o(k')} < 2.7344^{k'}$ for large k'. Since 2.7344 is below the branching factor 2.8192, the whole algorithm runs in time $O^*(2.8192^k)$.

k-Co-Path Set. Sullivan and van der Poel [SvdP16, §4.3, Algorithm 2] first reduce the instance, in polynomial time, to an equivalent instance (G', k') with $k' \leq k$ and at most $6k'$ vertices (a kernel of Feng, Zhou and Wang that they cite). They then branch on the vertices of degree greater than 10 and solve each resulting instance by a Monte Carlo algorithm in time $4^{\mathrm{tw}} n^{O(1)}$ on a tree decomposition of width tw [SvdP16, Theorem 4], which applies to every graph. With our width bound the branching is not needed. A yes-instance (G', k') satisfies $\sum_{d \geq 3} (d-2) n_d \leq 2k'$, since at least $d-2$ edges at each vertex of degree $d \geq 3$ are deleted [SvdP16, Lemma 10]; we reject (G', k') if this inequality fails. Otherwise, since $a_d \leq \frac{5}{32}(d-2)$ for every $d \geq 3$, Theorem 1.3 and Theorem 1.5 give a path decomposition of G' of width at most $\frac{5}{32} \cdot 2k' + o(k') = \frac{5}{16} k' + o(k')$, in place of the bound $\frac{k}{3} + o(n)$ that their Lemma 11 gives for the instances after branching, and their algorithm runs on it in time $4^{5k'/16 + o(k')} n^{O(1)} \leq 2^{5k/8 + o(k)} n^{O(1)}$. Since $2^{5/8} < 1.5423$, this is $O^*(1.5423^k)$. □

B Formal verification

Several results of Part I, among them Theorem 3.1, have formal proofs in the Lean 4 proof assistant (version 4.34.0) with the Mathlib library (commit `5ed2965`). In addition, a Lean port of the implementation of the algorithms of Theorem 1.5, for the cases described in Appendix B.1, is proved to output path decompositions that attain the bounds of Theorems 1.3, 7.1, 7.2 and 8.1 and Corollary 1.4 for suitable labelings, under the same hypotheses. The development is part of the accompanying code, and the driver script of Appendix A.2.6 has a third mode, *lean*, that rebuilds the formal proof of Theorem 8.1. Table 8 lists the formal proofs; the text after the table states the hypotheses they assume, describes the formal statements and what is not formalized, and records where the formal proofs follow other routes than the written ones. Appendix B.1 describes the implementation, its tests and its Lean port. No result of Part II or of Appendices A.7, C and D is formalized.

What each result rests on. The results rest on three kinds of evidence, and we say for each headline result which kind it needs.

(a) Theorem 1.1(a), Corollary 1.2 and the switch lemma behind them (Lemma 8.9) use no computation; their written proof uses the case $d = 3$ of Lemma 6.5, which is a case of Mulgund's theorem [Mul26] and also has an elementary proof (Remark 6.6). Theorem 8.1, which is Theorem 1.1(a), has a complete formal proof in Lean, without hypotheses. Its written proof takes pages 45–56 and lemmas from Sections 3 to 7; the formal proof follows it with the differences listed under "Routes" below.

(b) Theorem 1.1(b) uses two inputs. The first is, for vertices of degree 4, the extremality of the regular simplex among four Gaussian variables, which is a case of Mulgund's theorem [Mul26]

and also follows from the four-cap theorem of Fejes Tóth [FT50] ([Bör04, Theorem 5.6.2]). The second is a certified enclosure of F_4 (Proposition 6.9).

(c) Theorem 1.3 and Corollary 1.4 are derived from the triple $(\frac{5}{32}, \frac{1}{4}, \frac{11}{32})$ for maximum degree 5 (Theorem 7.2). This triple rests on a computer-assisted charging certificate, verified in ball arithmetic and exact rational arithmetic (Proposition 7.5), and on the case of five variables of the simplex theorem. That case is Mulgund's theorem [Mul26], proved in a preprint with a Lean formalization that we compiled and checked against the inequality we use (see H1 below). Theorem 7.1(b),(c) are restrictions of triples of Theorem 7.2 and rest on the same certificate.

The bootstrap from degree 5 to all degrees (Lemma 9.3) is a combinatorial induction without computation. Theorem 8.2 rests on two further computer-assisted certificates. Theorem 1.1(b), Theorem 1.3, Corollary 1.4 and Theorems 7.2 and 8.2 also have formal proofs in Lean, conditional on the certified and external inputs named above.

For every entry of Table 8, the final Lean theorem has no hypotheses beyond those of the written statement and those listed, and Lean's `#print axioms` reports only the three standard axioms `propext`, `Classical.choice` and `Quot.sound`. The development contains no unproved steps (`sorry`), no additional axioms and no trusted native computation. The last column of the table says what each formal statement was checked against: the written statement it formalizes, and for the implementation also the code it ports.

Since Mathlib has no notion of pathwidth, the development defines path decompositions directly: a path decomposition is a nonempty finite sequence of bags that covers every vertex and every edge, such that the bags containing any vertex are consecutive. Graphs are Mathlib's simple graphs on the vertex set $\{0, \dots, n-1\}$, and a bound $\mathrm{pw}(G) \le w$ means that G has a path decomposition whose bags have at most $w+1$ vertices. Treewidth is not formalized. Unless said otherwise below, a formal statement is the uniform ε–N form of the written one: for every real $\varepsilon > 0$ there is N such that every graph of the class with $n \ge N$ vertices satisfies $\mathrm{pw}(G) \le \beta(G) + \varepsilon n$, where β is the bound without its remainder. All theorem names below live in the Lean namespace `CubicPathwidth`.

The hypotheses and their status. The five hypotheses are mathematics established outside the development, stated as explicit hypotheses of the final theorems rather than as axioms.

H1 (`SimplexFourClaim`) The case $d = 4$ of Lemma 6.5, written exactly as the theorem `FSC.cdf_simplex_le` of Mulgund's formalization [Mul26] for four variables. Status: proved in that formalization, by the author of [Mul26]. We compiled it at the commit its author designates as certified, and checked that its main theorem depends only on the three standard axioms and that, for $n \in \{3,4,5\}$ variables and after the substitution $X = -U$, its statement is exactly the inequality of Lemma 6.5. The distribution in that statement is Mathlib's multivariate Gaussian measure, which for a positive semidefinite matrix is the centered Gaussian distribution with that covariance, singular matrices included. The reproducibility package does not rebuild this formalization. It uses Lean 4.31.0 and another version of Mathlib, so our development cannot import it; H1 therefore restates the theorem word for word. The written proof uses the same theorem.

H2 (`AdmissiblePair`, at $(F_3, 0.2575)$ and at $(\frac{5}{32}, \frac{1}{4})$) Admissibility for $D = \{3,4\}$ as in Definition 7.3, with the exact center bounds; these pairs are the restrictions of P and S noted after Proposition 7.5. Status: certified by the computer-assisted proof of Proposition 7.5 in ball arithmetic and exact rational arithmetic (Appendix A.2). The computation is not repeated inside Lean, because that would require verified enclosures of Φ, of the simplex functions S_d and of Owen's T function on about a thousand cells.

Table 8: The formal proofs. "Axioms" lists what `#print axioms` reports beyond the three standard axioms, which every entry uses. The hypotheses H1, H2, H3, H6 and H7 are stated, with their status, after the table.

Result	Lean theorem	Hypotheses	Axioms	Checked against
Theorem 8.1, that is Theorem 1.1(a)	`C3.c3MainClaim`, proving the proposition `C3MainClaim`	none	none	Theorem 8.1
Theorem 3.1	`CutSweep.cutSweep`, proving the proposition `CutSweepClaim`	none	none	Theorem 3.1 and Lemma 3.2
Lemma 5.4 for $p = 3$	`Locality.` `graph_correlation_lower`	none	none	a lemma inside the development of Theorem 8.1; not separately checked
Theorem 7.1(a)	`T2.target2a`	H1	none	Theorem 7.1
Theorem 7.1(b), (c)	`T2.target2b`, `T2.target2c`	H2 at the pair	none	Theorem 7.1
Theorem 7.2, triples P, R, S	`T3.target3P`, `T3.target3R`, `T3.target3S`	H3 at the triple	none	Theorem 7.2
Theorem 1.3; Corollary 1.4	`T4.target4`; `T4.target5`	H3 at S	none	Theorem 1.3; Corollary 1.4
Definition 9.1 (its ε–N form is equivalent to its $\rho(n) = o(n)$ form); Lemma 9.3: its first claim; the whole lemma when $(a_3, \ldots, a_k)$ is a pathwidth sequence	`PwSeq.targetRhoForm`; `PwSeq.targetBootstrapLemma`; `PwSeq.targetLemma2`	none	none	Definition 9.1 and Lemma 9.3
Theorem 9.2 for C, F, P, R, S	`PwSeq.targetThm1C`, `targetThm1F`, `targetThm1P`, `targetThm1R`, `targetThm1S`	none for C; H1 for F; H3 at the triple for P, R, S	none	Theorem 9.2
Theorem 8.2; Theorem 9.2 for M	`Sw4.targetThm62`; `Sw4.targetThm1M`	H6 and H7	none	Theorem 8.2; Theorem 9.2
Implementation of Appendix B.1: every output is a path decomposition	`Alg.algoT1Valid`, `algoT2aValid`, `algoMixedValid`, `algoT4Valid`	none	none	the Lean port against the implementation, and the statements against Appendix B.1
Implementation: no fallback on the inputs of these cases of Theorem 10.1; the width bounds of Lemmas A.28 and A.33 and of Lemma A.34 for the bootstrap of S; the bounds of Part I attained for some labeling	`Alg.algoT1RealizesTheorem1`, ..., `Alg.` `algoT4RealizesCorollary5`, with 29 statements in all	those of the attained theorem (none, H1, H2, H3)	none	the statements against Appendix B.1 and the theorems they realize

H3 (`AdmissibleTriple`, at P, R and S) Proposition 7.5 as stated. Status: as for H2. The written method of this certificate uses the cases $d = 4$ and $d = 5$ of Lemma 6.5, which rest on Mulgund's preprint (Remark 6.6). The simplex theorem for five variables enters the formal proofs only through this certificate.

H6 (`SwitchMixedClaim`) Lemma A.10 as stated, for centered Gaussian vectors described by their correlation matrices and for the bounds in every order, with Ψ written as the expectation $\mathbb{E}[\mathbf{1}\{Y_0 \le 0\}\lceil N/2\rceil]$, where N counts the neighbors with positive label. Status: proved in the paper by a coupling reduction and the box certificates of Appendix A.3.2 in ball arithmetic; not formalized.

H7 (`SwitchedChargingClaim`) Proposition A.13 with c_3 exact and the center bounds of H2 and H3. Status: a further problem of the certificate of Proposition 7.5 (Appendix A.2); not formalized.

The formal statements. `C3MainClaim` is the uniform ε–N form of Theorem 8.1 with c_3 written as the exact real expression $\frac{3}{4\pi}\arccos\frac{5}{6}$: for every real $\varepsilon > 0$ there is N such that for every $n \ge N$, every simple graph on the vertex set $\{0, \dots, n-1\}$ all of whose degrees are at most 3 has a path decomposition in which every bag has at most $(c_3 + \varepsilon)n + 1$ vertices. The proposition also contains the inequality $c_3 < 2c_*/(1+6c_*) \approx 0.16443$, where $c_* = \frac{3}{2\pi}\arccos\frac{2\sqrt{2}}{3}$ is the constant of Theorem 3.1; this compares c_3 with a weaker coefficient proved earlier in the same development and plays no role here. The statement `graph_correlation_lower` is the inequality $\mathrm{Corr}(Y_u, Y_v) \ge r_*(1 - \frac{1}{R})$ for every graph of maximum degree at most 3, every radius $R \ge 1$ and every edge uv.

In the statements for Theorem 7.1 and Theorem 7.2, F_4 and F_3 are defined as the suprema over the threshold of the ideal star functions (2.2), and Theorem 7.1(a) is stated in the form $\mathrm{pw}(G) \le F_4 n + o(n)$. The formal target for Corollary 1.4 is the ε–M form of Remark 9.14, which implies Corollary 1.4: for every $\varepsilon > 0$ there is M such that every graph with $m \ge M$ edges, on any number of vertices, satisfies $\mathrm{pw}(G) \le (\frac{7}{48} + \varepsilon)m$. It implies the $o(n)$ form: Lemma 3.2 gives $\mathrm{pw}(G) \le m$, as at most m edges leave any prefix of an ordering, and $\mathrm{pw}(G) \le n - 1 < \frac{7}{48}m$ if $m > \frac{48}{7}n$, so $\mathrm{pw}(G) \le \frac{7}{48}m + \max\{M, \frac{48}{7}\varepsilon n\}$ for every graph.

The formal definition of a pathwidth sequence is the finite-tuple phrasing given after Definition 9.1, for tuples $(a_3, \dots, a_k)$ with $k \in \{3, 4, \dots\} \cup \{\infty\}$, with the remainder in uniform ε–N form: for every $\varepsilon > 0$ there is N such that every graph with $n \ge N$ vertices satisfies $\mathrm{pw}(G) \le \sum_{d=3}^{k} a_d n_d + \sum_{d>k} n_d + \varepsilon n$. The theorem `PwSeq.targetRhoForm` proves this equivalent to the form of Definition 9.1 with one remainder function $\rho(n) = o(n)$. The statements for Theorem 9.2 bound every graph, with coefficient 1 for each vertex of degree greater than k. `CutSweepClaim` states Theorem 3.1 in four parts: $c_* = \frac{3}{2\pi}\arcsin\frac{1}{3}$ for c_* written as $\frac{3}{2\pi}\arccos\frac{2\sqrt{2}}{3}$; $c_* < \frac{1}{6}$; $\mathrm{pw}(G) \le w$ whenever some ordering of $V(G)$ has at most w edges leaving each of its prefixes, for every graph (the generality of Lemma 3.2); and the uniform ε–N form of $\mathrm{cw}(G) \le c_* n + o(n)$ for maximum degree at most 3.

What is not formalized. The following are not formalized: Corollary 1.2; Theorem 3.6; Corollary 9.5; Corollary 9.15; the enclosures of Proposition 6.9; the certificates behind H2, H3, H6 and H7; the explicit remainders of Lemma 4.5 and (9.4); the explicit threshold $n \ge 2^{46}$ and the explicit constants of the proof of Theorem 3.1; the computation of the labels (Appendices A.5.2 to A.5.5) and the running times of Theorem 10.1; Corollary 10.2 and Theorem 10.3; and treewidth.

Routes. Each formal proof follows the written proof, with the following differences. The formal proof of Theorem 8.1 proves the switch lemma (Lemma 8.9) by a different deformation than Lemma 8.11: it rotates one vector of the residual configuration at a time, controls each rotation

by a Plackett-type derivative computation, and reaches a planar configuration after finitely many rotations by a potential-function argument; this route exists only in the formal development. It also proves the case $d = 3$ of Lemma 6.5 directly, by the elementary argument of Remark 6.6, instead of taking it from Mulgund's theorem. All formal proofs prove only $o(n)$ remainders, so they choose their parameters from ε, replace explicit rates by continuity or elementary estimates (for Theorems 7.1 and 7.2, they also prove the inequalities $5\Phi(-3) \le \frac{1}{50} \le F_3 \le \frac{103}{400}$ by elementary estimates instead of the certified values of Proposition 6.9), and, except for Theorem 3.1, bound the expected total absolute deviation of all counts instead of applying Chebyshev's inequality to each count. For vertices of degree at most 2, the formal proofs of Theorems 1.3, 7.1, 7.2 and 8.2 do not follow Lemma 4.5: they keep a smallest set of vertices to which every component of the rest, of treewidth at most 2, attaches at most twice, and lift the decomposition by a torso construction, with the remainder $12\sqrt{n} + 1$. The formal bootstrap proves $\sigma_d \le a_d$ through the conserved deficit $d(1 - \sigma_d) + (1 - \sigma_{d-1}) = \frac{129}{32}$ instead of Lemma 9.12; with base degree $k \ge 3$ in place of 5 it proves Lemma 9.3, and with it Theorem 9.2 for C, F, P, R and S, and for M under H6 and H7 with the base of Theorem 8.2.

B.1 The implementation

The accompanying code implements the randomized algorithms of the cubic, padded, mixed and general cases of Theorem 10.1, with floating-point labels, and its outputs were valid path decompositions on all test inputs; for the sweeps, the widths were between 1.19 and 1.63 times the bounds without remainder at $n \le 2000$. A Lean port of these algorithms, which agrees with the implementation on 16,650 test calls, is proved to output path decompositions that, for suitable labelings, attain Theorems 1.3, 7.1, 7.2 and 8.1 and Corollary 1.4 under the hypotheses of their formal proofs.

C Separate, Measure and Conquer: the analyses

This appendix contains the analyses behind Section 12. Appendix C.1 explains the technique operationally and shows where the separator constant enters, and Appendix C.2 gives the conventions for the computer-assisted steps. Appendix C.3 derives balanced separators from our decompositions. Appendix C.4 reruns the linear program of Gaspers and Sorkin for cubic Max 2-CSP; we show that the optimum of their linear program is exactly $\frac{2c}{1+4c}$ as a function of the separator constant c. Appendices C.5 to C.9 prove the bounds for counting independent sets and coloring.

Throughout this appendix $r \ge 2$ denotes the domain size of Max $(r, 2)$-CSP, as in Section 11.1. All logarithms without a base are binary.

C.1 The technique

This subsection describes Separate, Measure and Conquer in the generality needed below, following [GS17, Sections 1.1 and 4], and isolates the one constraint in which the separator constant appears. Section, lemma and equation numbers of [GS17] refer to its arXiv version [GS14]. The algorithm of Appendix C.4, and the subcubic part of the algorithm of Appendix C.5, are instances of this scheme.

A *separation* of a graph G is a partition (L, S, R) of $V(G)$ such that no edge joins L and R; S is the *separator*, L and R are the left and right sides.

Branching on the separator. Consider a branching algorithm that removes a vertex in every branch, as most algorithms for Max 2-CSP, counting independent sets or counting dominating sets do. Suppose it is given a separation (L, S, R) of the input with $|S| \le cn$ and $|L|, |R| \le \frac{1-c}{2}n$, and that it always branches on a vertex of S. Used on its own, the separator gives a weak bound: if branching on each separator vertex multiplies the running time by r and nothing else is gained, the running time satisfies $t(n) = r^{cn} \cdot 2\, t(\frac{1-c}{2}n)$, whose solution is $r^{\frac{2c}{1+c}n}$, that is $r^{2n/7}$ for $c = \frac{1}{6}$. This is slower than the simple $r^{n/4}$ algorithm for cubic Max 2-CSP [SS07], which branches on any vertex and uses that its three neighbors become contractible.

A better bound comes from combining the two effects. Locally, branching on a separator vertex simplifies the instance as much as branching on any other vertex, and the split at the end comes in addition. In the cubic Max 2-CSP setting, each branching on a separator vertex removes it and makes its three neighbors contractible. If these neighbors fall on the two sides evenly, each side shrinks by $\frac{3}{2}$ per separator vertex, and the *optimistic recurrence* of [GS17, Section 1.1] is

$$t(n) = r^{cn} \cdot 2\, t\Big(\frac{1-c}{2}n - \frac{3}{2}\,cn\Big).$$

With $t(n) = r^{xn}$ this reads $x = c + x \cdot \frac{1-4c}{2}$, so $x = \frac{2c}{1+4c}$: for $c = \frac{1}{6}$ this is the $\frac{1}{5}$ of Gaspers and Sorkin, and for $c = c_3$ it is 0.179341. Edwards [Edw16], whose polynomial-space algorithm is also separator-based, remarks that an improved coefficient $c < \frac{1}{6}$ for cubic graphs would give the exponent $\frac{2c}{1+4c}$ by his argument.

The measure. The optimistic recurrence is not a proof: the branchings need not split evenly between the sides, the algorithm must sometimes move vertices between L, S and R to keep branching on the separator, and the split takes effect only at the end, all at once. Gaspers and Sorkin make it rigorous with a measure that assigns weight only to the separator and to the larger side. Let μ_s and μ_r assign nonnegative weights to vertex sets ("separator" and "right" weights, depending on the current labeled graph), and let the sides be named so that $\mu_r(R) \ge \mu_r(L)$. The measure is [GS17, (26)]

$$\mu(L, S, R) = \mu_s(S) + \mu_r(R) + \max\Big\{0,\, B - \frac{\mu_r(R) - \mu_r(L)}{2}\Big\} + (1+B)\log_{1+\eta}\big(\mu_r(R) + \mu_s(S)\big), \quad \text{(C.1)}$$

where B is a constant larger than the change of the imbalance $\mu_r(R) - \mu_r(L)$ in any single transformation except the separation step below, and $\eta > 0$ is a small constant. The vertices of L carry no weight: when the separator is exhausted, the remaining running time is at most twice that for the larger side R, and the smaller side is accounted for by the halving of the weight of R. The third term is a penalty for balanced instances. Gaspers and Sorkin combine it with a condition on the transformations [GS17, condition (30)]: if $\mu_r(R) - \mu_r(L) > B$, a transformation must decrease $\mu_r(R)$ at least as much as $\mu_r(L)$. Under this condition every transformation decreases $\mu_r(R)$ plus the penalty by at least half of its total progress on the two sides. The logarithmic term offsets the penalty B that may appear each time a new separation is computed, and it contributes only a polynomial factor to the running time.

Where the separator constant enters. Every transformation of the algorithm must satisfy the usual measure-and-conquer condition (Lemma C.1 below). For the local branching and simplification rules this gives constraints that do not involve the separator. The only transformation that does is the *separation step*: when $S = \emptyset$, the instance on $L \cup R$ is solved as two independent instances, and for each of them (say the one on R) a new separation (L', S', R') is computed. The measure must

not increase; the logarithmic term offsets the new penalty provided that (the case $S = \emptyset$ of [GS17, (29)])

$$\mu_r(R) \geq (1+\eta)(\mu_r(R') + \mu_s(S')). \tag{C.2}$$

Suppose that every vertex has right weight w_r and separator weight at most w_s, where $w_s \geq \frac{w_r}{2}$, and that the new separation has $|S'| \leq c|R|$ and $|L'| = |R'|$, so that $|R'| = \frac{|R|-|S'|}{2}$. Then $\mu_r(R') + \mu_s(S') \leq w_s|S'| + w_r\frac{|R|-|S'|}{2}$, which is nondecreasing in $|S'|$ because $w_s \geq \frac{w_r}{2}$, so it is at most $(c\,w_s + \frac{1-c}{2}w_r)|R|$ (this monotonicity argument is carried out in Lemma C.4). Hence (C.2) holds as soon as

$$c\,w_s + \frac{1-c}{2}\,w_r < w_r, \qquad \text{equivalently} \qquad w_s < \frac{1+c}{2c}\,w_r. \tag{C.3}$$

The separations of Appendix C.3 have $|S'| \leq c|R|$ only up to a lower-order term; this is absorbed by taking c slightly larger than the separator constant and $|R|$ large.

This is the mechanism by which the separator constant controls the running time. A branching on a separator vertex removes that vertex from S, so its large weight w_s contributes to the decrease of the measure; the branching constraints therefore need w_s large compared with the right weights. The time bound is governed by w_r, the weight of a vertex in the initial instance $(\emptyset, \emptyset, V)$. Constraint (C.3) caps the ratio w_s/w_r by $\frac{1+c}{2c}$, which is 3.5 for $c = \frac{1}{6}$ and 4.076 for $c = c_3$. A smaller separator constant relaxes the cap, and w_r can decrease. Everything else in the analysis is unchanged. Gaspers and Sorkin note that any improvement of their separator lemma improves their running times [GS17, Section 3.3]; Lemma C.5 quantifies this exactly for cubic Max 2-CSP.

The analyses rest on the following lemma, which bounds the running time of a branching algorithm by a measure that decreases suitably along every branching, and allows a hand-over to a subroutine with its own measure. In it, A is an algorithm for a problem, $\mathcal{B}$ is an algorithm for a class $\mathcal{C}$ of instances, and μ, $\mu_{\mathcal{B}}$ and ν are nonnegative measures on instances.

Lemma C.1 ([GS17, Lemma 4.2], [GL23, Lemma 1])**.** *Let $c_0 \geq 0$ and $\lambda > 1$ be constants, and let $\mu_{\mathcal{B}}(I) \leq \mu(I)$ for every $I \in \mathcal{C}$. Suppose that on every instance I, either $I \in \mathcal{C}$ and A solves I by calling $\mathcal{B}$ in time $O(\nu(I)^{c_0+1}\lambda^{\mu_{\mathcal{B}}(I)})$, or A reduces I to instances $I_1, \ldots, I_k$, solves them recursively and combines the solutions, using time $O(\nu(I)^{c_0})$ outside the recursive calls, where $\nu(I_i) \leq \nu(I) - 1$ for every i and*

$$\sum_{i=1}^{k} \lambda^{\mu(I_i)} \leq \lambda^{\mu(I)}.$$

Then A solves every instance I in time $O(\nu(I)^{c_0+1}\lambda^{\mu(I)})$.

The case without a subroutine $\mathcal{B}$ is [GS17, Lemma 2.3].

For a two-way branching with $\lambda = 2$ whose branches decrease the measure by at least a and b, the condition reads $2^{-a} + 2^{-b} \leq 1$; we call (a, b) a *branching vector*. For Max $(r, 2)$-CSP the branching creates r instances with the same measure μ', and with $\lambda = r$ the condition reads $1 + \mu' - \mu \leq 0$.

C.2 Computer-assisted steps

The analyses of Appendix C.4 and of Appendices C.5 to C.9 end in finitely many explicit inequalities between explicit rational weights. We list every inequality and every weight, so that each can be checked by hand or by machine. The accompanying code checks them as follows. Linear inequalities are checked in exact rational arithmetic (Python `fractions`). Every branching vector (a, b) is checked to satisfy $2^{-a} + 2^{-b} \leq 1$ in interval arithmetic (`mpmath.iv`, 40 decimal digits),

with the weights entered as exact rationals, except in the finite enumerations of configurations of Appendix C.9 (`g2_enumeration.py`): there the vectors are computed in exact rational arithmetic and the condition is checked in floating point, with 40-digit decimal arithmetic for sums close to 1. Floating-point optimization (`scipy`) is used only to find the weights and to report optima that the theorems do not use; nothing in a proof depends on it.

C.3 Balanced separators from the decompositions

The previous analyses obtain their separators from the bounds of Fomin and Høie and of Fomin, Gaspers, Saurabh and Stepanov [FH06, FGSS09]: Gaspers and Sorkin use separators of size $\alpha_\Delta|V| + o(|V|)$ with $\alpha_3, \alpha_4, \alpha_5, \alpha_6 = \frac{1}{6}, \frac{1}{3}, \frac{13}{30}, \frac{23}{45}$ [GS17, Lemmas 1.2 and 4.1], and Gaspers and Lee use $\frac{n_3}{6}$ [GL23]. The following lemma provides the same kind of separators from our decompositions. Part (a) is used for Max 2-CSP, part (b) for counting independent sets.

Let $a_3^* = c_3$, $a_4^* = F_4$, $a_5^* = \frac{11}{32}$ and $a_6^* = \frac{7}{16}$. For a graph G and $\theta > 0$, a function μ_r from the subsets of $V(G)$ to $[0, \infty)$ is *θ-Lipschitz* if $\mu_r(\emptyset) = 0$ and $|\mu_r(X \cup \{v\}) - \mu_r(X)| \le \theta$ for all $X \subseteq V(G)$ and $v \in V(G)$. We write V_3 for the set of vertices of degree 3 of G, so that $n_3 = |V_3|$.

Lemma C.2 (balanced separators)**.** *Let $\theta > 0$.*

(a) *Let $3 \le \Delta \le 6$. There is a function $\rho'_\Delta(x) = o(x)$ such that for every graph G with $\Delta(G) \le \Delta$ and every θ-Lipschitz μ_r computable in polynomial time, a separation (L, S, R) of G with $|S| \le a_\Delta^*|V(G)| + \rho'_\Delta(|V(G)|)$ and $|\mu_r(L) - \mu_r(R)| \le \theta$ can be computed deterministically in polynomial time; if $\mu_r(X) = |X|$, it can moreover have $|L| = |R|$.*

(b) *There is a function $\rho_3(x) = o(x)$ such that for every graph G with $\Delta(G) \le 3$ and $\delta(G) \ge 2$ and every θ-Lipschitz μ_r computable in polynomial time with $\mu_r(X) = \mu_r(X \cap V_3)$ for all X, a separation (L, S, R) of G with $S \subseteq V_3$, $|S| \le c_3 n_3 + \rho_3(n_3)$ and $|\mu_r(L) - \mu_r(R)| \le \theta$ can be computed deterministically in polynomial time.*

For comparison, $c_3 < \frac{1}{6}$, $F_4 < 0.2458 < \frac{1}{3}$, $\frac{11}{32} < \frac{13}{30}$ and $\frac{7}{16} < \frac{23}{45}$. The proof is the argument of [GS17, Lemma 4.1]: a nice path decomposition gives a sequence of separations whose separators are its bags, and the difference of μ_r between the two sides changes by at most θ from one separation to the next. Part (b) applies this to the cubic multigraph obtained by suppressing the vertices of degree 2.

Proof. We first show how a path decomposition yields the separation. Let H be a graph and $(W_1, \dots, W_t)$ a path decomposition of H of width w in which consecutive bags differ in exactly one vertex (a *nice* path decomposition). Every path decomposition, given for instance in the interval form of Theorem 10.1, can be turned into a nice one of the same width with at most $2|V(H)|$ bags in polynomial time: list the bags, and between two consecutive bags first remove the vertices of $W_i \setminus W_{i+1}$ one at a time, then add those of $W_{i+1} \setminus W_i$ one at a time; start with the empty bag and end by removing the last vertices one at a time (empty bags are discarded at the end). For $1 \le i \le t$ let

$$L_i = \Big(\bigcup_{j<i} W_j\Big) \setminus W_i, \qquad R_i = V(H) \setminus (L_i \cup W_i).$$

Then (L_i, W_i, R_i) is a separation: a vertex of L_i occurs only in bags before W_i and a vertex of R_i only in bags after W_i, so by the interval property no bag contains both, and they are not adjacent. Let $f(i) = \mu_r(L_i) - \mu_r(R_i)$. If $W_{i+1} = W_i \cup \{v\}$, then $v \in R_i$, $L_{i+1} = L_i$ and $R_{i+1} = R_i \setminus \{v\}$; if $W_{i+1} = W_i \setminus \{v\}$, then $L_{i+1} = L_i \cup \{v\}$ and $R_{i+1} = R_i$. In both cases $|f(i+1) - f(i)| \le \theta$, and in both cases $|L_i| - |R_i|$ increases by exactly one. Moreover $L_1 = \emptyset$ and

$R_t = \emptyset$, so $f(1) = -\mu_r(R_1) \le 0 \le \mu_r(L_t) = f(t)$. Let i be the largest index with $f(i) \le 0$. If $i = t$ then $f(t) = 0$; otherwise $f(i+1) > 0$ and $f(i) \ge f(i+1) - \theta > -\theta$. Either way $|f(i)| \le \theta$. For $\mu_r(X) = |X|$ the integer $|L_i| - |R_i|$ increases by exactly one at each step, from at most 0 to at least 0, so it vanishes at some i. All of this takes polynomial time, and the separator W_i has at most $w + 1$ vertices.

(a) Let $n = |V(G)|$. By Theorem 10.1(b) a path decomposition of G of width at most $W(G)+\rho(n)$, with $\rho(n) = o(n)$, is computed deterministically in polynomial time, where $W(G)$ is the respective bound: $c_3 n$ by Theorem 8.1 if $\Delta \le 3$; $F_4 n$ by Theorem 7.1(a) if $\Delta \le 4$; $\frac{5}{32}n_3 + \frac{1}{4}n_4 + \frac{11}{32}n_5 \le \frac{11}{32}n$ by Theorem 7.2 with the triple S if $\Delta \le 5$; and $\sum_{d \le 6} a_d n_d \le a_6 n = \frac{7}{16}n$ by Theorem 1.3 if $\Delta \le 6$. (Vertices of degree at most 2 have coefficient 0 in these bounds, and the coefficients are nondecreasing in d.) So $W(G) \le a^*_\Delta n$, and the construction above gives a separation with $|S| \le a^*_\Delta n + \rho(n) + 1$. This is one algorithm for each Δ, and ρ depends only on Δ (through the case of Theorem 10.1 used); take $\rho'_\Delta(x) = \rho(x) + 1$.

(b) Suppress the vertices of degree 2: every component of G that is a cycle contains no vertex of V_3, and in every other component each vertex of degree 2 lies on a unique path $u\, x_1 \cdots x_q\, w$ whose ends u, w (possibly equal) are in V_3 and whose inner vertices have degree 2. Replacing every such path by an edge uw gives a multigraph M (possibly with loops and parallel edges) on V_3 in which every vertex has degree 3. Let G_M be the underlying simple graph, without loops; $\Delta(G_M) \le 3$. By Theorem 8.1 and Theorem 10.1(b) a path decomposition of G_M of width at most $c_3 n_3 + \rho(n_3)$ is computed deterministically in polynomial time, and the construction above, applied with μ_r restricted to V_3, yields a separation (L_M, S, R_M) of G_M with $|S| \le c_3 n_3 + \rho(n_3) + 1$ and $|\mu_r(L_M) - \mu_r(R_M)| \le \theta$. Extend it to G: put every cycle component and every suppressed path whose ends lie in $L_M \cup S$ into L, and every other suppressed path into R; the ends of a suppressed path are adjacent or equal in M, so they never lie in L_M and R_M respectively. Then (L, S, R) with $L \supseteq L_M$, $R \supseteq R_M$ is a separation of G with $S \subseteq V_3$, and $\mu_r(L) = \mu_r(L_M)$, $\mu_r(R) = \mu_r(R_M)$ because μ_r depends only on degree-3 vertices. Take $\rho_3(x) = \rho(x) + 1$. □

C.4 Max 2-CSP

This subsection reruns the analysis of Gaspers and Sorkin [GS17, Section 3] for cubic Max $(r, 2)$-CSP with the separators of Lemma C.2(a), and lifts the result to all instances with the theorem of Scott and Sorkin. For cubic instances the exponent drops from the $\frac{n}{5}$ of Gaspers and Sorkin to $\frac{2c_3}{1+4c_3}n + o(n) < 0.179341n + o(n)$ (Theorem C.3). The algorithm is that of Gaspers and Sorkin, rules R1–R9, with the separations of Lemma C.2(a). Only one constraint of their linear program involves the separator constant (Lemma C.4), and we solve the program exactly as a function of that constant c: its optimum is $\frac{2c}{1+4c}$, with explicit weights and a three-line dual certificate (Lemma C.5). At $c = \frac{1}{6}$ these are the weights of Gaspers and Sorkin.

Put

$$\gamma = \frac{2c_3}{1 + 4c_3} \qquad \text{and} \qquad \alpha = \frac{2\gamma}{3}.$$

RCSP and PCSP are the ring and polynomial versions of Section 11.1, and polynomially bounded PCSP is meant in the sense of Scott and Sorkin [SS09].

Theorem C.3 (polynomial-space Max 2-CSP)**.** *Max $(r, 2)$-CSP with n variables and m constraints can be solved in polynomial space and time*

(a) $r^{\gamma n + o(n)} = r^{\alpha m + o(m)}$ *if the constraint graph is cubic;*

(b) $r^{(\frac{1}{8} + \frac{3}{8}\alpha)m + o(m)}$ *if the constraint graph has maximum degree at most* 4*;*

(c) $r^{(\frac{7}{50}+\frac{3}{10}\alpha)m+o(m)}$ *for all instances.*

The same bounds hold for the number of ring operations of RCSP and for the running time of polynomially bounded PCSP. The numbers of dominating sets of each size of an n-vertex graph of maximum degree at most 3 can be computed in polynomial space and time $3^{\gamma n+o(n)}$.

Here $\gamma < 0.179341$, $\alpha < 0.119561$, $\frac{1}{8}+\frac{3}{8}\alpha < 0.169836$, $\frac{7}{50}+\frac{3}{10}\alpha < 0.175869$ and $3^\gamma < 1.21778$, so the last bound is $O(1.2178^n)$. The polynomially bounded PCSPs include counting cuts by size and the Ising partition function.

The previous bounds are $r^{n/5}$, $r^{7m/40}$, $r^{9m/50}$ and $3^{n/5} = O(1.2458^n)$ [GS17, Theorems 3.1–3.3 and 5.1]; Edwards [Edw16] also obtains $r^{9m/50+o(m)}$ in polynomial space. In exponential space, Theorem 11.3 is faster. In polynomial space, this route does not improve Max 2-SAT, which is solved in time $O^*(2^{m/6.321})$ [GS12, Theorem 4], nor counting the dominating sets of general graphs, which takes time $O(1.5183^n)$ [GS17, Theorem 5.2]: a separator constraint of that analysis is not tight, so the new separators do not lower this bound.

Reductions. Instances are as in Section 11.1. Scott and Sorkin [SS07] use four reductions, each computable in polynomial time: a vertex of degree 0 is deleted after maximizing its score (Reduction 0); a vertex of degree 1 is folded into its neighbor (Reduction 1); a vertex of degree 2 is replaced by an edge between its two neighbors, merged with an existing edge if there is one (Reduction 2); and a vertex v of degree at least 3 is *branched on* (Reduction 3): for each color $C \in [r]$, v is deleted and its score functions are absorbed into those of its neighbors, which gives r instances with the common constraint graph $G - v$. An optimal solution of the instance is obtained in polynomial time from optimal solutions of the subinstances.

The algorithm. The algorithm maintains a separation (L, S, R) of the current constraint graph G, initially $(\emptyset, \emptyset, V)$. Let L_3, S_3, R_3 denote the vertices of degree 3 in L, S, R, let S_2 denote the vertices of degree 2 in S, and name the sides so that $|L_3| \le |R_3|$. The measure is [GS17, (6)]

$$\mu(L,S,R) = w_s|S_3| + w_{s,2}|S_2| + w_r|R_3| + w_b\mathbf{1}(|R_3| = |L_3|) \\ + w_c\mathbf{1}(|R_3| = |L_3|+1) + w_d \log_{3/2}(|R_3| + |S_3|), \quad \text{(C.4)}$$

a special case of (C.1) with integral imbalance: w_b and w_c play the role of the penalty, and the logarithm that of the last term. The algorithm applies the first applicable rule of the following list; when a rule is described, the earlier rules are assumed not to apply. The constraint numbers refer to Table 9.

R1. *Degree* 0. Apply Reduction 0. The measure does not change.

R2. *Half-edge deletion.* If some vertex has degree 1, or a vertex of $L \cup R$ has degree 2, apply Reduction 1 or 2. Viewing every edge deletion as two half-edge deletions, each lowering one degree, the measure does not increase under constraints (7)–(11). From now on the graph is cubic except for vertices of degree 2 in S.

R3. *Separation.* If $S = \emptyset$: the instance is the disjoint union of the instances on L and R, which are solved independently. For each of them a new separation is computed by Lemma C.2(a) with $\mu_r(X) = |X|$. Constraint (12_c) below.

R4. *Degree* 2 *in* S. If $s \in S$ has degree 2, apply Reduction 2. If its neighbors are $u_1 \in L$ and $u_2 \in R$, the new edge u_1u_2 would join the sides, so u_2 is moved into S. Constraints (13)–(15).

Table 9: The constraints of the cubic Max 2-CSP analysis [GS14, Section 3.2], each in the form "expression ≤ 0". In (12_c) the previous value $c = \frac{1}{6}$ is replaced by the separator constant c. The last column gives the expression for the weights of Lemma C.5, multiplied by $D = 1 + 4c$; a constant term 1 becomes D.

No.	Rule (case)	Constraint	$D \cdot$ (value at Lemma C.5)
(7)	measure nonnegative	$w_s, w_{s,2}, w_r, w_b, w_c, w_d \ge 0$	holds
(8)	R2, degree $3 \to 2$ in L	$-w_b + w_c$	$-c$
(9)	R2, degree $3 \to 2$ in S	$-w_s + w_{s,2}$	$-c$
(10)	R2, degree $3 \to 2$ in R	$-w_r + w_c$	$-c$
(11)	R2, degree $3 \to 2$ in R	$-w_r + w_b - w_c$	$-c$
(12_c)	R3, separation	$c\,w_s + \frac{1-c}{2} w_r - w_r$	0
(13)	R4, balanced	$-w_{s,2} + w_s - w_b + w_c$	0
(14)	R4, imbalanced	$-w_{s,2} + w_s - w_r + w_c$	0
(15)	R4, imbalanced	$-w_{s,2} + w_s - w_r + w_b - w_c$	0
(16)	R5	$-w_s + w_r$	$c - 1$
(17)	R6	$-w_s + w_c$	-1
(18)	R6	$-w_s + w_b - w_c$	-1
(19)	R7, branch	$1 - 2w_s + w_{s,2} - w_r$	0
(20)	R8, branch	$1 - w_s - w_r - w_b + w_c$	0
(21)	R8, branch	$1 - w_s - w_r - w_c$	0
(22)	R8, drag	$-w_r + w_b$	0
(23)	R9, branch	$1 - w_s - 2w_r - w_c + w_b$	0
(24)	R9, branch	$1 - w_s - 2w_r + w_c$	0

R5. *No neighbor in L.* If $s \in S$ has no neighbor in L, move it into R ("drag" it). Constraint (16).

R6. *No neighbor in R.* If $s \in S$ has no neighbor in R, move it into L (swapping the names of the sides if $|R_3| = |L_3|$). Constraints (17), (18).

R7. *One neighbor in each of L, S, R.* Branch on such an s. Constraint (19).

R8. *Two neighbors in L.* Let $s \in S$ have two neighbors in L and one, u, in R. If $|R_3| \le |L_3| + 1$, branch on s; constraints (20), (21). Otherwise move s into L and u into S; constraint (22).

R9. *Two neighbors in R.* Otherwise every $s \in S$ has two neighbors in R and one in L; branch on such an s. Constraints (23), (24), and (20) after swapping the sides when $|R_3| = |L_3|$.

Every rule takes polynomial time. The rules that only move vertices between L, S and R decrease the secondary measure $\nu = 3|S_3| + 2|R_3| + |L_3| + 2|E(G)|$ [GS17, end of Section 3.2], and so do all rules that delete vertices; Lemma C.4 treats the separation step R3.

Table 9 lists the constraints. They are derived in [GS17, Section 3.2] from the condition of Lemma C.1: for the rules that create one instance the measure must not increase, and for a branching on a vertex, which creates r instances of the same measure μ', the condition $r \cdot r^{\mu'} \le r^{\mu}$ is $1 + \mu' - \mu \le 0$. For example, in R7 the separator loses the degree-3 vertex s, its neighbor in S drops to degree 2, and one degree-3 vertex of R and one of L drop to degree 2; the balance is unchanged, which gives (19). We do not repeat the other derivations; the only one that involves the separator is that of (12_c). In the following lemma, an instance with fewer than N_0 vertices that arises in the Separation rule R3 is counted as a leaf of measure 0.

Lemma C.4 (the separation constraint)**.** *Let $0 < c < \frac{1}{3}$, and suppose that R3 uses separations with $|S'| \le c|V|$ and $|L'| = |R'|$ on every cubic graph with $|V| \ge N_0$ to which it applies. If the weights*

satisfy (7), (12_c) and (16) of Table 9 and $w_d = w_b + 1$, then on every instance with $|R| \geq N_0$, R3 satisfies the condition of Lemma C.1 and decreases the secondary measure ν.

Proof. We follow [GS17, Section 3.2, Separation]. Let the instance be $(L, \emptyset, R)$ with $N = |R| \geq |L|$. Since R1 and R2 do not apply and S is empty, every vertex has degree 3, so $N = |R_3|$, and $\mu \geq w_r N + w_d \log_{3/2} N$ because the penalty terms are nonnegative. The graph $G[R]$ is cubic, since $S = \emptyset$ and no edge joins L and R. Let (L', S', R') be the new separation of $G[R]$ and $s = |S'| \leq cN$. It is balanced ($|L'_3| = |R'_3|$), so its measure is

$$\mu' = w_s s + w_r \frac{N - s}{2} + w_b + w_d \log_{3/2} \frac{N + s}{2}.$$

The derivative of $w_s s + w_r \frac{N-s}{2}$ in s is $w_s - \frac{w_r}{2} \geq 0$ by (16) and (7), so this part is at most $(c\, w_s + \frac{1-c}{2} w_r)N$. Since $c < \frac{1}{3}$ we have $\frac{N+s}{2} \leq \frac{1+c}{2} N \leq \frac{2}{3} N$, hence $w_d \log_{3/2} \frac{N+s}{2} \leq w_d \log_{3/2} N - w_d$. With $w_d = w_b + 1$,

$$\mu' \leq \Big(c\, w_s + \frac{1-c}{2} w_r\Big) N + w_d \log_{3/2} N - 1 \leq w_r N + w_d \log_{3/2} N - 1 \leq \mu - 1$$

by (12_c). The instance on L has measure $\mu'' \leq \mu - 1$ by the same bound with $|L| \leq N$ in place of N, since the bound is nondecreasing in N; if $|L| < N_0$, it is a leaf of measure $\mu'' = 0 \leq \mu - 1$. Since $r \geq 2$, $r^{\mu'} + r^{\mu''} \leq 2r^{\mu - 1} \leq r^{\mu}$. For the secondary measure: if $L \neq \emptyset$, splitting $(L, \emptyset, R)$ into $(\emptyset, \emptyset, R)$ and $(\emptyset, \emptyset, L)$ decreases ν because $|L| \leq |R|$ and both sides are nonempty; if $L = \emptyset$, as for the initial instance, this first step is void. Replacing $(\emptyset, \emptyset, R)$ by (L', S', R') changes ν by $3s + 2|R'| + |L'| - 2N = s - |L'| = s - \frac{N-s}{2} < 0$, because $s \leq cN < \frac{N}{3}$. □

Thus the separator constant enters the linear program only through (12_c), and the proof once more through $c < \frac{1}{3}$ (the logarithmic estimate and the secondary measure). The previous derivation estimates $|R'_3| \leq \frac{5}{12}|R_3|$ and $|S'_3| \leq \frac{1}{6}|R_3| + o(|R_3|)$ separately; the monotonicity in s used above shows that the combined form (12_c) is what is needed for every separator fraction $c < \frac{1}{3}$.

Lemma C.5 (the linear program)**.** *For every $c \in (0, 1]$ the weights*

$$w_s = \frac{1 + c}{1 + 4c}, \quad w_{s,2} = \frac{1}{1 + 4c}, \quad w_r = w_b = \frac{2c}{1 + 4c}, \quad w_c = \frac{c}{1 + 4c}, \quad w_d = w_b + 1$$

satisfy every constraint of Table 9. Conversely, every solution of (12_c), (21) and (24) has $w_r \geq \frac{2c}{1+4c}$. Hence the minimum of w_r subject to the constraints of Table 9 is exactly $\frac{2c}{1+4c}$.

Proof. Put $D = 1 + 4c$ and multiply all weights by D: the scaled weights are $w_s = 1 + c$, $w_{s,2} = 1$, $w_r = w_b = 2c$ and $w_c = c$, and the constant 1 in (19)–(21), (23) and (24) becomes D. Substituting gives the last column of Table 9: (8)–(11) each equal $-c$; (12_c) equals $c(1 + c) + \frac{1-c}{2} \cdot 2c - 2c = 0$; (13) equals $-1 + (1 + c) - 2c + c = 0$, (14) equals (13) because $w_r = w_b$, and (15) equals $-1 + (1 + c) - 2c + 2c - c = 0$; (16) equals $c - 1 \leq 0$; (17) and (18) equal -1; (19) equals $D - 2(1+c) + 1 - 2c = 0$; (20) equals $D - (1+c) - 2c - 2c + c = 0$; (21) equals $D - (1+c) - 2c - c = 0$; (22) equals 0; (23) equals $D - (1 + c) - 4c - c + 2c = 0$; and (24) equals $D - (1 + c) - 4c + c = 0$. All weights are nonnegative, so (7) holds.

For the converse, (12_c) says $w_s \leq \frac{1+c}{2c} w_r$, and the sum of (21) and (24) says $2 \leq 2w_s + 3w_r$. Together, $2 \leq (\frac{1+c}{c} + 3) w_r = \frac{1+4c}{c} w_r$. □

At $c = \frac{1}{6}$ these are the previous weights $(w_s, w_{s,2}, w_r, w_b, w_c) = (0.7, 0.6, 0.2, 0.2, 0.1)$ [GS17, Section 3.2], and the tight constraints are those listed there, (12)–(15) and (19)–(24). Gaspers

and Sorkin solved the program numerically. The value $\frac{2c}{1+4c}$ is the exponent of the optimistic recurrence, and Edwards notes that his argument would give it for an improved separator coefficient (Appendix C.1); Lemma C.5 adds that it is the exact optimum of the linear program of Gaspers and Sorkin for every c, with the explicit weights and the certificate $(12_c)+(21)+(24)$. The accompanying code confirms both (exact check of the weights for several rational c, and numerical optima at $c = \frac{1}{6}, c_3, 0.1, 0.3$).

We now prove Theorem C.3.

Proof. We prove (a) with the weights of Lemma C.5 and Lemma C.1, and obtain (b) and (c) from (a) by the lifting theorem of Scott and Sorkin. The ring versions reuse the search tree of (a), and the count of dominating sets reuses its analysis with $r = 3$.

(a) The algorithm is that of Gaspers and Sorkin, rules R1–R9, with the separations of Lemma C.2(a) for $\Delta = 3$ and $\mu_r(X) = |X|$. Since that lemma gives one algorithm, this algorithm has no parameter; ε and N_0 below enter only the analysis. Fix $\varepsilon \in (0, \frac{1}{10})$, put $c = c_3 + \varepsilon < \frac{1}{3}$, and choose N_0 such that $\rho'_3(x) \le \varepsilon x$ for $x \ge N_0$. For the analysis, an instance with fewer than N_0 vertices that arises in the Separation rule is a leaf: the algorithm runs on it as usual, and its running time is bounded by a constant C_ε. By Lemma C.2(a) the separations computed on larger instances satisfy $|S'| \le c|V|$ and $|L'| = |R'|$. By Lemma C.5 the weights of the lemma for this c satisfy every constraint of Table 9, and $w_d = w_b + 1$. By Lemma C.4 and the derivations of [GS17, Section 3.2] for the other rules, every rule satisfies the condition of Lemma C.1 with $\lambda = r$, and ν decreases in every step. So the number of leaves of the search tree is at most r^μ for the initial measure

$$\mu(\emptyset, \emptyset, V) = w_r n + w_d \log_{3/2} n = \frac{2c}{1+4c}\, n + O(\log n),$$

and, since ν is polynomial in n and each step takes polynomial time, the running time is $C_\varepsilon\, r^{\frac{2c}{1+4c} n}\, \mathrm{poly}(n)$. The function $c \mapsto \frac{2c}{1+4c}$ is continuous, so letting $\varepsilon \to 0$ gives $r^{\gamma n + o(n)}$ for the one algorithm. The recursion depth is polynomial and each call uses polynomial space, so the space is polynomial. For a cubic graph $m = \frac{3}{2}n$, so $\gamma n = \alpha m$.

(b), (c) Scott and Sorkin [SS07, Theorem 22] show that an algorithm solving cubic instances in time $O^*(r^{\alpha' m})$ with $\frac{1}{9} \le \alpha' \le \frac{1}{5}$ yields one solving instances of maximum degree 4 in time $O^*(r^{(\frac{1}{8}+\frac{3}{8}\alpha')m})$ and arbitrary instances in time $O^*(r^{(\frac{7}{50}+\frac{3}{10}\alpha')m})$; the new algorithm uses the cubic one as a subroutine and is in polynomial space if the subroutine is [GS17, Section 3.3]. This is how Gaspers and Sorkin obtain $\frac{7}{40}$ and $\frac{9}{50}$ from $\alpha' = \frac{2}{15}$. Our cubic algorithm runs in time $O^*(r^{(\alpha+\varepsilon)m})$ for every $\varepsilon > 0$, and $\alpha \approx 0.1195606 \in [\frac{1}{9}, \frac{1}{5}]$, so the lifted algorithm runs in time $O^*(r^{(\frac{1}{8}+\frac{3}{8}(\alpha+\varepsilon))m})$ and $O^*(r^{(\frac{7}{50}+\frac{3}{10}(\alpha+\varepsilon))m})$ for every small $\varepsilon > 0$; again the algorithm does not depend on ε.

RCSP and PCSP: Gaspers and Sorkin [GS17, Theorems 3.2 and 3.3] obtain their bounds by running the same algorithm with the ring versions of Reductions 0–3 of Scott and Sorkin [SS09], which change the constraint graph exactly as the original reductions and cost $O(r^3)$ ring operations each. The search tree, and hence the analysis, is that of (a); the lift to maximum degree 4 and to all instances is the same.

Dominating sets: Gaspers and Sorkin [GS17, Section 5.1, Theorem 5.1] count the dominating sets of every size of a subcubic graph G by branching, with a three-way branching ("in", "optional", "forbidden"), on the vertices of the cubic structure of G, the cubic multigraph obtained by contracting the edges at vertices of degree at most 2. The pivot is chosen exactly as in the Max 2-CSP algorithm, the analysis is that of cubic Max $(3, 2)$-CSP, and instances whose cubic structure has at most a constant number A of vertices are solved in polynomial time by dynamic programming over a tree decomposition of width at most A. We use the separations of Lemma C.2(a) of the underlying

simple graph of the cubic structure (a separation of it is one of the multigraph) and take $A \geq N_0$. The analysis of (a) with $r = 3$ gives $3^{\gamma n + o(n)}$. □

C.5 Counting independent sets, and graph coloring

Gaspers and Lee [GL23] count the independent sets of a graph in polynomial space with a branching algorithm whose subcubic part is a Separate, Measure and Conquer subroutine. It branches on the vertices of a balanced separator, computed from a path decomposition (Lemma C.2), until the graph falls apart into two halves that are counted independently; a measure that gives the separator vertices extra weight (Appendix C.1) makes the measure drop at each branching on a separator vertex by enough to cover the extra weight of the separator. They extend the measure to maximum degree 4 and to all graphs with Wahlström's compound measures [Wah08], and they compute the chromatic number from these counts with the inclusion–exclusion method of Björklund, Husfeldt and Koivisto [BHK09]. The separator constant $\frac{1}{6}$ enters their subcubic analysis in two places, the separator constraint and the function $g(d)$ of [GL23, Lemma 3]. A better pathwidth bound shortens the separator phase, and the gain carries over to the larger degrees and to the chromatic number. With c_3 in place of $\frac{1}{6}$ (Lemma C.2(b)) and the changes to the algorithm described below, their analysis, adapted, gives the following bounds.

Theorem C.6 (counting independent sets, and coloring)**.** *Let G be a graph with n vertices. The number of independent sets of G can be computed*

(a) *in polynomial space and time $O(1.1304^n)$ if $\Delta(G) \leq 3$;*

(b) *in polynomial space and time $O(1.2037^n)$ if $\Delta(G) \leq 4$;*

(c) *in polynomial space and time $O(1.2341^n)$, and the chromatic number of G in polynomial space and time $O(2.2341^n)$;*

(d) *in time and space $O(1.2287^n)$.*

The previous bounds are $O(1.1389^n)$, $O(1.2070^n)$, $O(1.2356^n)$, $O(2.2356^n)$ and $O(1.2330^n)$ [GL23, Lemmas 5 and 6, Theorems 1 and 2]. The proof of Theorem C.6, in the rest of this appendix, is computer-assisted. It runs the analysis of Gaspers and Lee with c_3 in place of $\frac{1}{6}$, with three changes. First, the analysis of their subcubic subroutine lacks condition (30) of Gaspers and Sorkin [GS17] (Appendix C.1), without which a step that crosses the balance threshold of the measure can decrease the measure by less than its branching vector counts. Their local analysis also has some further minor gaps. Fixing condition (30) increases the running time given by the analysis; for maximum degree 3 we obtain $O(1.1304^n)$. In exchange, the new pathwidth bound makes the special treatment of so-called spider vertices (the separator vertices to which their subroutine gives a larger weight) unnecessary: in our subroutine every separator vertex has the same weight, which gives a simpler algorithm and analysis. Second, above maximum degree 3 their branching on lazy 2-separators need not satisfy the measure condition of Lemma C.1 (Proposition C.26); our rule tests that condition in exact rational arithmetic. Third, we prove two steps that their paper leaves unproved, the branching vectors of their Lemma 8 for degrees 5 and 6 and a vector for the lazy rule (Proposition C.12).

A recent claim of $O^*(1.2321^n)$ for counting independent sets [DMRBC26] would beat part (c) of Theorem C.6, but not part (d). Its proof is not valid. For instance, its Lemma 3 asserts $|E(G)| \leq c\,|V(G)|$ for every graph G and vertex v, with $c = \max\{\delta(N[v])/(2|N[v]|),\ \Delta(G - N[v])/2\}$ and $\delta(N[v])$ the sum of the degrees of the vertices of $N[v]$; for $G = K_{2,3}$ and v on the side of size 2, $c = \frac{9}{8}$ and $|E(G)| = 6 > \frac{45}{8}$. Moreover, the analysis takes the maximum of bounds obtained in two different measures.

We prove part (a) of Theorem C.6 in Appendix C.6, and parts (b)–(d) in Appendix C.7, using two steps proved in Appendices C.8 and C.9. Numbering refers to the journal version of Gaspers and Lee [GL23]. The computer-assisted steps follow the paragraph conventions of Appendix C.2; the scripts named below are in `experiments/algorithmic-consequences` of the accompanying code.

The algorithm of Gaspers and Lee. The algorithm #IS [GL23, Section 3] computes a weighted count of independent sets, the weights (a *cardinality function*) recording earlier steps. It solves graphs of maximum degree at most 2 in polynomial time and splits a disconnected graph into its components. Otherwise it picks a vertex x of maximum degree that maximizes the associated average degree of Wahlström [Wah08], and applies, in this order: the *multiplier reduction*, which folds a part of constant size attached at a cut vertex v into the cardinality function of v and deletes it; a branching for a *lazy* 2*-separator* (defined in Appendix C.7); if $\Delta(G) = 3$ and no vertex of degree 3 has all neighbors of degree 3 (a $(3,3,3)$ *vertex*), a Separate, Measure and Conquer subroutine; otherwise it *branches* on x, counting the independent sets of $G - x$ and of $G - N[x]$. All rules only delete vertices, so degrees never increase, and for $\Delta(G) \le 3$ no rule creates a $(3,3,3)$ vertex (it would have been one before).

C.6 Maximum degree 3

Skeletons. Let G be a graph with $\Delta(G) \le 3$ and $\delta(G) \ge 2$, and let V_3 be its set of vertices of degree 3. A 2*-path* is a path of length at least 2 whose inner vertices have degree 2 and whose ends lie in V_3. The *skeleton* of G is the multigraph on V_3 with one edge for every edge of G inside V_3 and one edge for every 2-path. Every vertex of the skeleton has degree 3, a loop counted twice; a loop at v makes v a cut vertex. The far end of a skeleton edge at v is a *skeleton neighbor* of v. For a separation (L, S, R) with $S \subseteq V_3$, the inner vertices of a 2-path lie all in L or all in R, so no skeleton edge joins L and R. For $v \in S$ let $e_L(v)$, $e_R(v)$ and $e_S(v)$ be the numbers of skeleton edges at v whose far end lies in L, R and S.

The measure. Let $r_3, s_3 > 0$ be weights, $B = 42r_3$ and $J = 40r_3$. For a separation (L, S, R) let $a = r_3|R \cap V_3|$, $b = r_3|L \cap V_3|$, $\Xi(a,b) = \max\{a, b, \frac{a+b}{2} + B\}$, let $\eta > 0$, and let $\Lambda(x) = (1+B)\log_{1+\eta}\max\{1, x\}$. The measure of a separated instance is

$$\mu_{\mathrm{sep}}(L,S,R) = s_3|S| + \Xi(a,b) + \Lambda\big(\max\{a,b\} + s_3|S|\big). \tag{C.5}$$

This is (C.1) with $\mu_r(X) = r_3|X \cap V_3|$ and $\mu_s(X) = s_3|X|$, written symmetrically in the sides. It is nondecreasing in a, b and $|S|$. We name the sides so that $a \ge b$ and call $t = a - b$ the *imbalance*. Each of the three linear *pieces* a, b, $\frac{a+b}{2} + B$ of Ξ is at most Ξ, and one of them attains it. Hence, if (a,b) changes to (a',b') and π is a piece that attains $\Xi(a',b')$,

$$\Xi(a',b') - \Xi(a,b) \le \pi(a',b') - \pi(a,b). \tag{C.6}$$

This inequality, together with the pull rule (S6) below, replaces condition (30) of Gaspers and Sorkin [GS17] (Appendix C.1).

The algorithm. Let N_0 be a constant and $K_0 = 8$. To *count directly* a graph of maximum degree at most 3 with at most N_0 vertices of degree 3 means to run the standard dynamic program over a path decomposition of width $N_0 + 2$ (deleting the vertices of degree 3 leaves paths and cycles), in time $2^{N_0+3}n^{O(1)}$ and polynomial space. The multiplier reduction is applied only to parts with

at most K_0 vertices of degree 3; such a part with its cut vertex has pathwidth at most $K_0 + 3$, so the reduction takes polynomial time. An instance is a graph G with $\Delta(G) \leq 3$ and a cardinality function, and either no separation (a *fresh* instance) or a separation (L, S, R) with $S \subseteq V_3$ (a *separated* instance). The input is fresh. The algorithm applies the first rule that applies; for a fresh instance:

(F1) if $\Delta(G) \leq 2$ or $n_3 \leq N_0$, count directly;

(F2) if G is disconnected, solve its components as fresh instances and multiply the results;

(F3) apply the multiplier reduction;

(F4) if G has a $(3,3,3)$ vertex x (chosen as by Gaspers and Lee): if exactly two skeleton edges join $N(x)$ to vertices outside $N[x]$, with far ends y and z, branch on y, and in each branch fold what remains of the part cut off by $\{y, z\}$ that contains x into z by the multiplier reduction (or count it directly if it has become a component); otherwise branch on x;

(F5) otherwise compute a separation by Lemma C.2(b) with $\mu_r(X) = r_3|X \cap V_3|$, which is r_3-Lipschitz, and continue with the separated instance.

For a separated instance, with the sides named so that $a \geq b$:

(S1) if $\Delta(G) \leq 2$ or $n_3 \leq N_0$, count directly;

(S2) if $S = \emptyset$, solve the components of G as fresh instances and multiply the results;

(S3) if some component of G has at most N_0 vertices of degree 3, count it directly and delete it;

(S4) apply the multiplier reduction;

(S5) *(move)* if some $v \in S$ has $e_L(v) = 0$, move v into R, together with the inner vertices of its 2-paths that lie in L; symmetrically if $e_R(v) = 0$;

(S6) *(pull)* if $t > 2B - J$ and some $v \in S$ has $e_L(v) > e_R(v)$, let q be the far end of the skeleton edge at v that ends in R; move v and the inner vertices of that skeleton edge into L, and q into S;

(S7) otherwise branch on any $v \in S$, keeping the separation restricted to the remaining vertices.

In (F2) and (S2), components with at most N_0 vertices of degree 3 are counted directly. Every step that deletes vertices is followed, as part of the same step, by a *clean-up*: vertices of degree 1 are folded into their neighbors by the multiplier reduction and vertices of degree 0 are counted, as long as there are any; then, in a separated instance, the separator is *normalized*:

(N) while some $u \in S$ has degree 2, let P be the maximal path or cycle through u whose inner vertices have degree 2, with ends $f_1, f_2 \in V_3$ if it is a path; if no end lies in R, move the inner vertices of P into L; else, if no end lies in L, move them into R; else, say $f_1 \in L$ and $f_2 \in R$, move them into L and f_2 into S.

Each operation of (N) gives a separation and lowers the number of vertices of degree 2 in S, so afterwards $S \subseteq V_3$. Every rule computes the weighted count exactly; the separation only steers the choice of vertices.

Lemma C.7 (moves and deletions)**.** *Let $s_3 > r_3$, and let I' be obtained from a separated instance I by* (S3)*,* (S4)*,* (S5) *or* (S6)*, including the clean-up. Then $\mu_{\mathrm{sep}}(I') \le \mu_{\mathrm{sep}}(I)$, and the argument $\max\{a, b\} + s_3|S|$ of Λ does not increase.*

Proof. (S3), (S4) and the removal of vertices of degree at most 1 delete vertices and lower degrees, so none of a, b, $|S|$ increases. An operation of (N) moves vertices of weight 0 and either lowers $|S|$ or keeps $|S|$ and lowers a by r_3. In (S5) with $e_L(v) = 0$, the result is a separation, $|S|$ drops by one, a grows by r_3 and b is unchanged; by (C.6) Ξ grows by at most r_3, and the argument of Λ changes by at most $r_3 - s_3 < 0$. In (S6), $e_R(v) \ge 1$ and $e_L(v) \ge 2$ force $(e_L, e_R, e_S)(v) = (2, 1, 0)$, so q is unique and the result is a separation; $|S|$ is unchanged, $a' = a - r_3$, $b' = b + r_3$ and $a' - b' = t - 2r_3 > 2B - J - 2r_3 > 0$. So the piece b is not active at (a', b'), and by (C.6) Ξ does not increase; the argument of Λ drops by r_3. □

For G with $\Delta(G) \le 3$ and $\delta(G) \ge 2$, the *closure* of $X_0 \subseteq V_3$ is the smallest $Z \supseteq X_0$ that contains every vertex of V_3 joined to Z by at least two skeleton edges. Let ∂Z be the set of skeleton edges with exactly one end in Z, and Y the set of vertices of $V_3 \setminus Z$ joined to Z by exactly one skeleton edge.

Lemma C.8 (the clean-up after a deletion)**.** *Let $\Delta(G) \le 3$, $\delta(G) \ge 2$, let $\emptyset \ne X_0 \subseteq V_3$ have closure Z, and let G' be obtained from G by deleting X_0 and some neighbors of degree 2 of vertices of X_0, and then removing vertices of degree at most 1 as long as there are any. Then G' is G minus Z and the inner vertices of the skeleton edges with an end in Z; in G' the vertices of Y have degree 2 and those of $V_3 \setminus (Z \cup Y)$ degree 3. Moreover $|Y| = |\partial Z| \le |\partial X_0|$, $|Z| \le |X_0| + |\partial X_0|$, and $\partial Z \ne \emptyset$ if some component of the skeleton that meets X_0 has more than $|X_0| + |\partial X_0|$ vertices.*

Proof. Let W be Z with the inner vertices of the skeleton edges at Z. The removals leave the largest subgraph of minimum degree at least 2, whatever their order. In $G - W$ a vertex of $V_3 \setminus Z$ has lost one edge per skeleton edge to Z, so it has degree 2 if it is in Y and 3 otherwise; all deleted vertices lie in W, so no vertex outside W is removed. Conversely, Z arises from X_0 by adding, one at a time, vertices joined to the current set by two or more skeleton edges; once the current set is gone, the inner vertices of its skeleton edges are removed one after another, and then the next added vertex has degree at most 1. So all of W disappears. Adding a vertex joined to the set by $k \ge 2$ skeleton edges changes the number of leaving skeleton edges by $(3 - k) - k \le -1$, so $|Y| = |\partial Z| \le |\partial X_0| - |Z \setminus X_0|$. If $\partial Z = \emptyset$, then Z contains every component of the skeleton that meets X_0. □

Lemma C.9 (branchings)**.** *Let $N_0 \ge 17$, $s_3 + \frac{3}{2}r_3 \ge 1$, $2s_3 \ge 1$ and $s_3 \ge \frac{r_3}{2}$. Let I be a separated instance to which none of* (S1)–(S6) *applies, and let I' be either branch of* (S7)*, including the clean-up. Then $\mu_{\mathrm{sep}}(I') \le \mu_{\mathrm{sep}}(I) - 1$, and the argument of Λ does not increase.*

Proof. Let $v \in S$ be the branching vertex. As (S1)–(S4) do not apply, $\delta(G) \ge 2$, every component of G meeting V_3 has more than $N_0 \ge 17$ vertices of degree 3, and the skeleton has no loop; as (S5) does not apply, $e_L(v), e_R(v) \ge 1$. Let $X_0 = \{v\}$ in the branch $G - v$, and let X_0 be v with its neighbors of degree 3 in the branch $G - N[v]$. With Z, Y as in Lemma C.8, $|\partial X_0| \le 12$, so $|Y| \le 12$, $|Z| \le 16$ and $\partial Z \ne \emptyset$.

We have $|Z \setminus \{v\}| + |Y| \ge 3$: if $Z = \{v\}$, the three skeleton neighbors of v are distinct and form Y; if $Z = \{v, u\}$, at most two skeleton edges join v and u, so $|\partial Z| \ge 2$; if $|Z| \ge 3$, use $|Y| \ge 1$. After the clean-up the vertices of degree 2 in S are those of $Y \cap S$. Each path or cycle P of (N) through $k \ge 1$ of them takes these k vertices out of S and puts at most one end $f \in R$ into S; each such f serves one P only. Give a *credit* to every vertex of $(Z \setminus \{v\}) \cup Y$: an *a-credit* if it lies in R, a *b-credit*

if it lies in L, an *s-credit* if it lies in $(Z \cap S) \setminus \{v\}$; on each P, $k-1$ of its vertices of $Y \cap S$ get an s-credit and the last gets an a-credit (for f) if an end enters S and an s-credit otherwise. With $\#_a, \#_b, \#_s$ the numbers of credits, $\#_a + \#_b + \#_s \geq 3$, and, as no vertex of degree 3 enters L or R,

$$|S'| \leq |S| - 1 - \#_s, \qquad a' \leq a - r_3\#_a, \qquad b' \leq b - r_3\#_b.$$

So the argument of Λ does not increase, and the imbalance changes by at most $r_3(|Z| + 2|Y|) \leq 40r_3 = J$.

If $t \leq 2B - J$, then $|a' - b'| \leq 2B$, $\Xi = \frac{a+b}{2} + B$ before and after, and the measure drops by at least $s_3(1 + \#_s) + \frac{r_3}{2}(\#_a + \#_b) \geq s_3 + \frac{3}{2}r_3 \geq 1$. If $t > 2B - J$, then, as (S6) does not apply, $(e_R, e_L, e_S)(v)$ is $(2,1,0)$ or $(1,1,1)$. We claim $\#_s \geq 1$ or $\#_a \geq 2$. Suppose $\#_s = 0$; then $Z \cap S = \{v\}$ and every vertex of $Y \cap S$ has an a-credit, so every skeleton edge from $Z_R \cup \{v\}$, where $Z_R = Z \cap R$, to a vertex of $(R \cup S) \setminus Z$ yields its own a-credit. If $Z_R = \emptyset$, the $e_R(v)$ edges into R give $e_R(v)$ credits, and if $e_R(v) = 1$ the skeleton neighbor of v in S gives another. If $|Z_R| \geq 2$ we are done. If $Z_R = \{z\}$, at most two of the three skeleton edges at z go to v, and the others end in $(R \cup S) \setminus Z$. Now if $\#_s \geq 1$, $|S|$ drops by two and the measure by $2s_3 \geq 1$. Otherwise $\#_a \geq 2$, $\#_a + \#_b \geq 3$ and $a' - b' \geq t - J > 2B - 2J > 0$, so b is not active after the step and, by (C.6), Ξ drops by at least $\min\{2r_3, \frac{3}{2}r_3\}$; the measure drops by at least $s_3 + \frac{3}{2}r_3 \geq 1$. □

Lemma C.10 (initial measure)**.** *Let $c \in (c_3, \frac{1}{3})$, $s_3 \geq \frac{r_3}{2}$, $\kappa = c\,s_3 + \frac{1-c}{2}r_3$ and $\kappa < r_3$. There are N_0' and $\eta' > 0$ such that for every $\eta > 0$ and every connected G with $\Delta(G) \leq 3$, $\delta(G) \geq 2$ and $n_3 \geq N_0'$, the separation of* (F5) *satisfies $\mu_{\mathrm{sep}}(L,S,R) \leq \kappa n_3 + 2B + \Lambda(r_3 n_3/(1+\eta'))$, and the argument of Λ in it is at most $r_3 n_3/(1+\eta')$.*

Proof. By Lemma C.2(b), $k = |S| \leq c_3 n_3 + \rho_3(n_3) \leq c n_3$ for large n_3, $|a - b| \leq r_3$ and $a + b = r_3(n_3 - k)$. As $\max\{a,b\} = \frac{a+b}{2} + \frac{|a-b|}{2}$, both $s_3 k + \Xi(a,b) - B$ and $\max\{a,b\} + s_3 k$ are at most $s_3 k + \frac{r_3(n_3-k)}{2} + \frac{r_3}{2}$, which is nondecreasing in k; so the first is at most $\kappa n_3 + B$ (as $\frac{r_3}{2} < B$), and the second at most $\kappa n_3 + \frac{r_3}{2} \leq r_3 n_3/(1+\eta')$ for a suitable η' and large n_3, as $\kappa < r_3$. The bound on the measure follows as Λ is nondecreasing. □

Measures of all instances. Graphs with a $(3,3,3)$ vertex are measured with weights w_2', w_3' for the vertices of degree 2 and 3, as in [GL23, Lemma 5]. Let $\ell(G) = w_2' n_2 + w_3' n_3$, $m(G) = \min\{\ell(G), \kappa n_3\}$, and let η' be as in Lemma C.10. The measure $\mu(I)$ of an instance I on G is 0 if $\Delta(G) \leq 2$ or $n_3 \leq N_0$; otherwise it is $\mu_3(G) = \ell(G) + \Lambda(r_3 n_3)$ for a fresh instance with a $(3,3,3)$ vertex, $\mu_f(G) = m(G) + 2B + \Lambda(r_3 n_3/(1+\eta'))$ for a fresh instance without one, and $\mu_{\mathrm{sep}}(L,S,R)$ for a separated instance.

Weights and constants. Let $c = 0.13985$, so $c_3 < 0.1398222 < c < \frac{1}{3}$, and

$$r_3 = 0.179364, \quad s_3 = 0.730954 = 1 - \tfrac{3}{2}r_3, \qquad w_2' = 0.0052182, \quad w_3' = 0.176755, \tag{C.7}$$

exact decimals. They satisfy the constraints of Table 10, checked by `is3_uniform_repair.py` (part 1); in particular $w_2' \leq w_3'$ and $\kappa < r_3$. Let N_0', η' be as in Lemma C.10. We fix $\eta > 0$ with $(1+B)\log_{1+\eta}(1+\eta') \geq 2B+1$, and $N_0 \geq \max\{N_0', 17\}$ with $w_3' N_0 \geq 1$, $\kappa N_0 \geq 1$ and $r_3 N_0 \geq 1 + \eta'$. Then every graph with $n_3 > N_0$ satisfies

$$\mu_f(G) \leq \mu_3(G) - 1, \tag{C.8}$$

as $m \leq \ell$ and $\Lambda(r_3 n_3/(1+\eta')) = \Lambda(r_3 n_3) - (1+B)\log_{1+\eta}(1+\eta')$. For reals $x_1, \ldots, x_p \geq 1$, induction on $2^{x+y} - 2^x - 2^y = (2^x - 1)(2^y - 1) - 1 \geq 0$ gives

$$\textstyle\sum_{j=1}^{p} 2^{x_j} \leq 2^{x_1 + \cdots + x_p}. \tag{C.9}$$

Table 10: The constraints of the subcubic analysis and their values at the weights (C.7), with $c = 0.13985$ and $\kappa = c\,s_3 + \frac{1-c}{2} r_3$. Linear constraints read "expression ≥ 0" and are checked exactly; for the branching vectors (a, b) the last column is an upper bound on $2^{-a} + 2^{-b}$ in interval arithmetic. Decimals are rounded.

No.	Origin	Expression	value
(U1)	Lemma C.9	$s_3 + \frac{3}{2} r_3 - 1$	0
(U2)	Lemmas C.7 and C.9	$s_3 - r_3$; $2s_3 - 1$	0.5516; 0.4619
(U3)	Lemma C.10, (C.3)	$r_3 - \kappa$	$1.108 \cdot 10^{-7}$
(U4)	hand-over (F5)	w'_2; $2w'_3 + w'_2 - 2\kappa$	0.0052; $4.22 \cdot 10^{-7}$
(B11)	(F4), branch on x	$(4w'_3 - 3w'_2,\ 8w'_3 - 4w'_2)$	0.9999956
(B12)	(F4), branch on y	$(6w'_3 - w'_2,\ 6w'_3 - w'_2)$	0.9623865

Proof of Theorem C.6(a). We apply Lemma C.1 with $\lambda = 2$, no subroutine, the measure μ and the secondary measure ν defined below, and check its condition rule by rule. (F1), (S1) and the direct counts make no recursive call. (S3)–(S6): Lemma C.7. (F3) deletes vertices and lowers degrees; as $0 \leq w'_2 \leq w'_3$, neither ℓ nor n_3 grows, no $(3,3,3)$ vertex appears, and if the $(3,3,3)$ vertices disappear the measure changes from μ_3 to $\mu_f \leq \mu_3 - 1$. (S7): Lemma C.9, with the vector $(1, 1)$.

(F2): let $C_1, \ldots, C_p$ be the components with more than N_0 vertices of degree 3. If G has a $(3,3,3)$ vertex, $\mu(C_j) \leq \ell(C_j) + \Lambda(r_3 n_3(G))$ by (C.8) and $\sum_j \ell(C_j) \leq \ell(G)$. Otherwise $\mu(C_j) \leq m(C_j) + 2B + \Lambda(r_3 n_3(G)/(1+\eta'))$ and $\sum_j m(C_j) \leq m(G)$. As $\ell(C_j) \geq m(C_j) \geq 1$, (C.9) gives $\sum_j 2^{\mu(C_j)} \leq 2^{\mu(G)}$.

(F4): G is connected with $n_3 > N_0$ and $\delta(G) \geq 2$, its skeleton has no loop, the new instances are fresh with measure at most $\ell + \Lambda(r_3 n_3(G))$ by (C.8), so it suffices that ℓ drops by the components of (B11) or (B12). Let e be the number of skeleton edges inside $N(x)$; then $6 - 2e$ skeleton edges leave $N[x]$, and $e = 3$ would make $n_3 = 4$. If $6 - 2e \geq 4$ we branch on x. In $G - x$, ℓ drops by $4w'_3 - 3w'_2$. In $G - N[x]$ let Z, Y be as in Lemma C.8 for $X_0 = N[x]$; as $n_3 > 17 > |X_0| + |\partial X_0|$, ℓ drops by at least $|Z| w'_3 + |Y|(w'_3 - w'_2)$ with $|Z| \geq 4$ and $|Y| \geq 1$, which is at least $8w'_3 - 4w'_2$ if $|Z| + |Y| \geq 8$. If $Z = N[x]$, then $|Y| = 6 - 2e \geq 4$. Otherwise $|Z| \geq 5$ and $|Y| = |\partial Z| \equiv 3|Z| \pmod 2$, so $|Z| + |Y| \geq 8$ unless $|Z| = 5$, $|Y| = 1$; then the vertex of Y cuts off a part with five vertices of degree 3, and (F3) would apply as $K_0 \geq 5$. If $6 - 2e = 2$, let H be $N[x]$ with the inner vertices of the skeleton edges at $N(x)$; it has four vertices of degree 3, and all edges leaving H end in y or z. If $y = z$, (F3) would apply at y; so $y \neq z$ and $d(y) = d(z) = 3$. In both branches y and the four vertices of degree 3 of H are deleted and z loses an edge or is deleted, so ℓ drops by at least $5w'_3 + (w'_3 - w'_2)$, which is (B12).

(F5): G is connected, $\delta(G) \geq 2$, $n_3 > N_0 \geq N'_0$, and every vertex of degree 3 has a neighbor of degree 2, so $n_3 \leq 2n_2$. If $w'_3 \geq \kappa$, then $\ell(G) \geq \kappa n_3$; otherwise $\ell(G) - \kappa n_3 \geq (\frac{w'_2}{2} + w'_3 - \kappa) n_3 \geq 0$ by (U4). So $m(G) = \kappa n_3$, and Lemma C.10 gives $\mu(I') \leq \mu_f(G)$.

(S2): let the instance be $(L, \emptyset, R)$ with $a \geq b$, so $\mu \geq a + \Lambda(a)$. A component C with $n_3(C) > N_0$ lies in L or in R, has no $(3,3,3)$ vertex, and $1 + \eta' \leq r_3 n_3(C) \leq a$; by $\kappa < r_3$ and the choice of η,

$$\mu(C) \leq \kappa n_3(C) + 2B + \Lambda(r_3 n_3(C)) - (1+B)\log_{1+\eta}(1+\eta') \leq r_3 n_3(C) + \Lambda(a) - 1.$$

As $\sum_{C \subseteq R} r_3 n_3(C) \leq a$ and $\sum_{C \subseteq L} r_3 n_3(C) \leq b \leq a$, (C.9) gives $\sum_C 2^{\mu(C)} \leq 2^{\Lambda(a)-1}(2^a + 2^b) \leq 2^\mu$. If G is connected, the new instance is G itself, with measure at most $\mu - 1$.

Secondary measure. For an instance on a graph with n vertices and m edges let $p = n + m$, $T = 3(n+1)^2$, $M = \lfloor n + 2B + \Lambda(n) \rfloor \geq \lfloor \mu \rfloor$ (as $s_3 + r_3 < 1$ and $w'_3 < 1$), $\tau = (2n+2)|S| + |a-b|/r_3 < T$ for a separated and $\tau = T$ for a fresh instance, and $\nu = (M+1)(T+1)p + (T+1)\lfloor \mu \rfloor + \tau$. A step

that deletes a vertex or an edge lowers p and so ν. Of the steps that keep the graph, none raises μ; (F5) lowers τ from T; (S2) on a connected graph lowers $\lfloor\mu\rfloor$ and raises τ by at most T; (S5) lowers τ by at least $2n+1$, and (S6) by 2. So ν drops in every step, and the rules terminate.

Conclusion. The input has $\mu \le \mu_3(G) \le w_3' n + \Lambda(n)$. By Lemma C.1 the running time is $2^{w_3' n} n^{O(1)}$, and $2^{0.176755} < 1.130339$. Every step takes polynomial time and space, and the recursion depth is at most the initial ν, so the space is polynomial. □

C.7 Maximum degree at least 4

The algorithm. On graphs of maximum degree at most 3 the algorithm is that of Appendix C.6. Fix constants $K \ge 65$, $K_m \ge 6K+1$ and $N_1 = 7K+2$. In a graph of maximum degree at least 4 the algorithm splits a disconnected graph into its components, solves every component with at most N_1 vertices directly in constant time, and otherwise selects a vertex x of maximum degree: as by Gaspers and Lee if $\Delta(G) = 4$ or the component is regular, and otherwise, if $\Delta(G) \ge 5$, a vertex of maximum degree with a neighbor of smaller degree (one exists in every connected non-regular graph). It then applies the first of the following rules that applies.

- *Multiplier reduction* at a cut vertex v such that some component of $G - v$ has at most K_m vertices.
- *Hybrid rule*, if x has a *lazy* 2*-separator*: distinct vertices y, z with $d(y), d(z) \ge 3$ such that $G - \{y, z\}$ is disconnected and $N[x] \setminus \{y, z\}$ lies in one component of $G - \{y, z\}$ with at most K vertices (x may belong to the pair). For each $v \in \{y, z, x\}$ in this order it computes the two instances of the branching on v, and it branches on the first v whose instances I_1, I_2 satisfy $2^{\mu(I_1)} + 2^{\mu(I_2)} \le 2^{\mu(G)}$ for the measures below without their sublinear term Λ, and on y if there is none. The weights are rational, so this is a polynomial-time test in exact arithmetic. As $\Lambda(I_j) \le \Lambda(G)$, a branching that passes the test satisfies the same inequality for the full measures.
- Otherwise branch on x.

The *instances* of a branching on v are $G - v$ and $G - N[v]$, each taken after the reductions that make no recursive call, applied as long as one applies: the removal of vertices of degree at most 1, the multiplier reduction, and the direct solution of components with at most N_1 vertices. Our lazy 2-separator is a reading of the rule of Gaspers and Lee [GL23, Algorithm #IS], which branches on y. Proposition C.26 below shows that under this reading their branching alone can violate the condition of Lemma C.1.

The measures. We follow [GL23, Lemmas 6 and 8, Theorem 1]. All measures have the form $\sum_i w_i n_i$ with $w_0 = w_1 = 0$, plus, for maximum degree at least 4, the sublinear term $\Lambda(r_3|V(G)|)$, which no step increases and which is at least the term $\Lambda(r_3 n_3)$ of μ_3 of every graph of maximum degree at most 3 that arises from G; we omit it below. We write $\Delta w_i = w_i - w_{i-1}$. From now on a 2-path is a path of length at least 2 whose inner vertices have degree 2 and whose ends have degree at least 3; its *far end* from one end is the other.

Maximum degree 4. The pivots $p_2, \dots, p_5 = 3, \frac{16}{5}, \frac{7}{2}, \frac{15}{4}$ cut the interval $[2, 4]$ of average degrees $d(G) = 2|E(G)|/|V(G)|$ into five *sections*. Section 1, $[2, 3]$, has the weights $(w_2^{(1)}, w_3^{(1)}) = (w_2', w_3')$ of (C.7) and a weight $w_4^{(1)}$, and its measure $\mu^{(1)}$ has the additional term $\psi \cdot \mathrm{ind}(G)$, where $\psi \ge 0$ and $\mathrm{ind}(G) = 1$ if all degrees of G are 2 or 4 and no two vertices of degree 4 are adjacent, and

$\mathrm{ind}(G) = 0$ otherwise. Section $k \geq 2$ has $w_i^{(k)} = w_i^{(k-1)} - (i - p_k)\alpha_k$ $(i = 2, 3, 4)$ with $\alpha_k \geq 0$, so that

$$\mu_{\mathrm{lin}}^{(k)}(G) - \mu_{\mathrm{lin}}^{(k-1)}(G) = -\alpha_k |V(G)|\,(d(G) - p_k), \tag{C.10}$$

where $\mu_{\mathrm{lin}}^{(k)} = \sum_i w_i^{(k)} n_i$ and $\mu^{(k)} = \mu_{\mathrm{lin}}^{(k)}$ for $k \geq 2$. A graph G of maximum degree 4 has the measure $\mu_4(G) = \min_k \mu^{(k)}(G)$, and its *minimizing section* is the smallest k with $\mu^{(k)}(G) = \mu_4(G)$. Let τ_k be the maximum of $\sum_i x_i w_i^{(k)}$ over the degree distributions x on $\{2, 3, 4\}$ with average degree in section k. The degree-4 branchings are those of [GL23, Table 1], listed by the degrees of the neighbors of x (the *type* of x), each with a branching vector and a *highest average degree* h, the largest $d(G)$ of a graph in which the selection rule can pick a vertex of this type (Wahlström's averaging argument [Wah08], which we use as Gaspers and Lee do). Its *group* is the section that contains h, the lower of the two sections if h is a common end of two sections. Two entries of that table are misprinted (as in Figure 4 of [GL16]): the first component for type $(2, 2, 2, 3)$ is $4w_4 - 2w_3 + 2w_2$, and in cases (a) and (b) of type $(2, 2, 2, 2)$ both components are $5w_4 - 4w_3 + 4w_2 + \psi$ and $4w_4 - 2w_3 + 3w_2 - \psi$, respectively (the signs of ψ are swapped in print). Case (a) is a vertex of type $(2, 2, 2, 2)$ all of whose 2-paths end at vertices of degree 4; case (b) has a 2-path to a vertex of degree 3.

Degrees 5 *and* 6. One measure $\sum_i w_i n_i$ with weights $w_2, \ldots, w_6$, at least those of section 5 on degrees $2, 3, 4$, increasing and concave. Branching on a vertex v with $d(v) \in \{5, 6\}$ has the vector [GL23, Lemma 8]

$$\Big(w_{d(v)} + \sum_{u \in N(v)} w_{d(u)} + \mathrm{out}(v)\,\Delta w_{d(v)},\quad w_{d(v)} + \sum_{u \in N(v)} \Delta w_{d(u)} + \deg_2(v)\,\Delta w_{d(v)}\Big), \tag{C.11}$$

where $\deg_2(v)$ is the number of neighbors of degree 2, $\mathrm{out}(v) = 5$ for the neighbor degrees $(2, 2, 2, 2, 2)$ and $(2, 2, 2, 2, 2, 3)$, $\mathrm{out}(v) = 6$ for $(2, \ldots, 2)$ with $d(v) = 6$, and otherwise $\mathrm{out}(v) = 3$ if $\sum_{u \in N(v)} d(u) - d(v)$ is odd and 4 if it is even.

Degree at least 7. The measure is $w_6|V(G)|$; a branching deletes one vertex or at least eight.

Cross-section cases. A type with highest average degree h can be selected in every graph with $d(G) \leq h$. By (C.10), a graph whose average degree lies inside section k has minimizing section k, and a graph with $d(G) = p_k$ has minimizing section at most $k - 1$. So in a graph with minimizing section k the selection rule picks only types whose group is at least k; we call these types *admissible* for section k, and we check the vector of every type under the weights of its own section and of every earlier section. A graph with $\mathrm{ind}(G) = 1$ has $4n_4 \leq 2n_2$, so $d(G) \leq \frac{8}{3}$, and its minimizing section is 1, since $\mu^{(2)} - \mu^{(1)} \geq \alpha_2 |V(G)|(3 - d(G)) - \psi \geq \frac{\alpha_2}{3} - \psi > 0$ for our weights.

The potential. The printed vectors account for ind only in cases (a) and (b). Case (a) is selected only when every vertex of degree 4 has type $(2, 2, 2, 2)$, the type of smallest associated average degree, and no vertex has degree 3 (otherwise #IS would branch on a case-(b) vertex), so $\mathrm{ind}(G) = 1$; case (b) occurs only when $\mathrm{ind}(G) = 0$. Other section-1 branchings may switch the indicator from 0 to 1. The indicator is evaluated after the removal of vertices of degree at most 1; each further reduction in the definition of the instances deletes at least two vertices of degree at least 2, so it lowers the linear part of $\mu^{(1)}$ by at least $2w_2^{(1)} = 0.0104364 > \psi$, and no $\mu^{(k)}$ increases even if the indicator switches. Subtracting ψ from both components keeps every section-1 vector other than (a), (b) and $(2, 2, 2, 3)$ valid (sums at most 0.90706), and every later-group vector under section-1 weights (sums at most 0.89110). For $(2, 2, 2, 3)$, Lemma C.11 gives the excess $\varepsilon = (w_3^{(1)} - w_2^{(1)}) - (w_4^{(1)} - w_3^{(1)}) = 0.0122914 > \psi$ in each branch where the indicator switches; with

$\varepsilon - \psi$ added in the first, the second or both components, the sums are at most 0.99543, 0.99544 and 0.99331. These checks are part 2 of `gaspers_lee_general_uniform.py`.

In the next lemma $w_i = w_i^{(1)}$, the instance G' of a branch is taken after the removal of vertices of degree at most 1, and the *printed component* of a branch is the corresponding component of the vector of $(2, 2, 2, 3)$ in [GL23, Table 1]. The *decrease* $\sum_v (w_{d_G(v)} - w_{d_{G'}(v)})$, with $w_0 = w_1 = 0$ and deleted vertices of degree 0, is a sum of nonnegative terms.

Lemma C.11 (the potential in case $(2, 2, 2, 3)$)**.** *Let G be a connected graph with more than N_1 vertices, minimizing section 1, $\Delta(G) \leq 4$ and $\delta(G) \geq 2$, to which neither the multiplier reduction nor the hybrid rule applies, and let x have type $(2, 2, 2, 3)$. If the instance G' of a branch on x has indicator 1, its decrease exceeds the printed component by at least ε.*

Proof. Let u be the neighbor of degree 3 of x, y_1, y_2, y_3 the others, and t_i the far end of the 2-path that starts with xy_i. The printed components charge w_4 to x, w_2 to each y_i, $w_4 - w_3$ to each t_i (the least drop of a vertex of degree at least 3 that loses an edge, as $w_4 - w_3 < w_3 - w_2$), and $w_3 - w_2$ (branch $G - x$) or w_3 (branch $G - N[x]$) to u. In G' no vertex has degree 3.

Closed parts. Call $P \supseteq \{x, y_1, y_2, y_3\}$ *closed* if it consists otherwise of vertices of degree 3 and inner vertices of 2-paths, and deleting x and then removing vertices of degree at most 1 removes all of P. Below, P is a component of G minus at most two vertices, so $|P| > b$ with $b = K$, K_m or N_1 (otherwise the hybrid rule, the multiplier reduction or the rule for small components would apply). Then P disappears in both branches, and the excess is at least $\sum_{v \in P \setminus \{x,y_1,y_2,y_3,u\}} w_{d(v)} - 3(w_4 - w_3)$, which is at least $387w_2 - 3(w_4 - w_3) = 1.5417072 > \varepsilon$ if $b \geq K_m \geq 391$, and at least $2w_3 + 60w_2 - 3(w_4 - w_3) = 0.1888658 > \varepsilon$ if $b = K \geq 65$, $u \notin P$ and P has two vertices of degree 3.

G has an even number of vertices of degree 3, so one, z, besides u; choose it with the largest number k of indices i with $t_i = z$. The charges at z are $k(w_4 - w_3)$, and z is deleted or loses an edge. The printed vectors use the drop of one vertex to cover a charge at another only for a far end of degree 3 at which two 2-paths end (the drop along its third skeleton edge), which is then z with $k = 2$, and when u is a far end, treated last. If $k = 3$, x, z and the three 2-paths form a closed part, a component of $G - u$. If $k = 0$, z drops by at least $w_3 - w_2 > \varepsilon$. If $k = 1$, z drops by $w_3 - w_2 = (w_4 - w_3) + \varepsilon$ or by $w_3 > (w_4 - w_3) + \varepsilon$. If $k = 2$, z is deleted, and the far end z'' of its third 2-path loses an edge and has degree 2 or is deleted in G'. If z'' is not charged, the excess at z, z'' is at least $w_3 + (w_3 - w_2) - 2(w_4 - w_3) = 0.0298010$. If z'' is the third far end and $d(z'') = 4$, it is at least $w_3 + (w_4 - w_2) - 3(w_4 - w_3) = 0.0298010$. If z'' is the third far end, $d(z'') = 3$ and f is the far end of its third 2-path, then x, z, z'' and the 2-paths between them and to f form a closed part, a component of $G - \{u, f\}$ (of $G - u$ if $f = u$), with two vertices of degree 3. If $z'' = u$: in $G - x$, u is deleted and the far end q of its third 2-path loses an edge; in $G - N[x]$ the drop at q is not charged; arguing for q as for z'' (if q is the third far end and $d(q) = 3$, x, u, z and their 2-paths form a closed part, a component of $G - q$), the excess at z, u, q is at least $2w_3 - 2(w_4 - w_3) = 0.0350192$ in $G - x$ and $w_3 + (w_3 - w_2) - 2(w_4 - w_3)$ in $G - N[x]$.

If $u = t_i$, the printed vectors use the drop at the third skeleton neighbor q of u for the charges at u. If $z \neq q$, the cases above apply, with $z'' \neq q$ for $k = 2$ (otherwise x, u, z and their 2-paths form a closed part, a component of $G - q$). If $z = q$ and $k = 2$, then G is a closed part with $b = N_1$. If $z = q$ and $k \leq 1$, the excess at u and z is at least $w_3 - (w_4 - w_3)$ and ε in the two branches if $k = 0$, and $2w_3 - 2(w_4 - w_3)$ and $w_3 + (w_3 - w_2) - 2(w_4 - w_3)$ if $k = 1$. The excesses of the cases $k \leq 2$ and $u = t_i$ are checked exactly in part 2 of `gaspers_lee_general_uniform.py`. □

Regular neighborhoods. The vectors (C.11) are not needed for the types $(5, 5, 5, 5, 5)$ and $(6, \ldots, 6)$: a vertex of maximum degree all of whose neighbors have this degree is selected only in

Table 11: The weights of the five sections at maximum degree 4; section 1 also has $\psi = 0.0061458$. All decimals are exact, and τ_k is computed exactly.

k	section	α_k	$w_2^{(k)}$	$w_3^{(k)}$	$w_4^{(k)}$	τ_k
1	$[2, 3]$	–	0.0052182	0.176755	0.3360004	0.176755
2	$[3, \frac{16}{5}]$	0.0484291	0.0536473	0.176755	0.2875713	0.19891826
3	$[\frac{16}{5}, \frac{7}{2}]$	0.0142955	0.0708019	0.1796141	0.2761349	0.2278745
4	$[\frac{7}{2}, \frac{15}{4}]$	0.0122574	0.089188	0.1857428	0.2700062	0.24894035
5	$[\frac{15}{4}, 4]$	0.0105598	0.10766765	0.19366265	0.26736625	0.26736625

a regular component. A Δ-regular component with $\Delta \in \{5, 6, 7\}$ arises at most once for each Δ on a root-to-leaf path of the search tree, since degrees never increase and components are solved separately; so these branchings cost a constant factor in the running time.

Weights. The polynomial-space algorithms use the weights of Table 11 and, for degrees 5 and 6,

$$(w_2, \ldots, w_6) = (0.1136655,\ 0.1986263,\ 0.2673669,\ 0.295498,\ 0.3034362). \tag{C.12}$$

The exponential-space algorithm of part (d) uses one measure for degrees 5 to 7, with

$$(w_2, \ldots, w_7) = (0.1103289,\ 0.1932144,\ 0.2500011,\ 0.2867833,\ 0.296828,\ 0.2970601), \tag{C.13}$$

and $w_7|V(G)|$ for maximum degree at least 8. All decimals are exact.

The analysis needs two steps that the paper of Gaspers and Lee does not prove: the vectors of their Lemma 8 rest on a bound on the number of edges leaving $N[v]$ (their Lemma 7) stated for 3-connected graphs, and the lazy 2-separator branching has no analyzed vector. We prove them for our algorithm in Appendices C.8 and C.9.

Proposition C.12. *For the algorithm and the constants above, the following hold.*

(G1) *Let v be a vertex of maximum degree $d(v) \in \{5, 6\}$ with the weights* (C.12)*, or $d(v) \in \{5, 6, 7\}$ with the weights* (C.13)*, on which the algorithm branches other than in the hybrid rule. Then* (C.11) *is a branching vector of this branching for $\sum_i w_i n_i$, the instances being taken after the removal of vertices of degree at most* 1*.*

(G2) *Every branching of the hybrid rule in a graph of maximum degree $\Delta \geq 4$ whose component, if $\Delta \in \{5, 6\}$, is not Δ-regular, satisfies the condition of Lemma C.1 for the measures above.*

Proof of Theorem C.6(b)–(d). (b), (c): we apply Lemma C.1 with $\lambda = 2$ to the algorithm above with the measures above. The following are checked by part 1 of `gaspers_lee_general_uniform.py`: every vector of [GL23, Table 1], corrected as above, under the weights of its own section and of every earlier section (interval arithmetic; largest sums 0.9999997 and 0.9671); every vector (C.11) for $d(v) \in \{5, 6\}$ and all neighbor degrees in $\{2, \ldots, d(v)\}$ except the regular neighborhoods (interval arithmetic); that the weights (C.12) are at least the section-5 weights on degrees $2, 3, 4$, increasing and concave, and that $\alpha_k, \psi \geq 0$ (exact); and the values τ_k, by exact enumeration of the vertices of the polytope of degree distributions with average degree in the section. By Lemma C.19 below and the paragraphs "Cross-section cases" and "The potential", the degree-4 branchings satisfy the condition of Lemma C.1, and the hand-overs between the measures are covered by Lemma C.19; so do those at degrees 5 and 6 by Proposition C.12(G1) and Lemma C.19, the reductions, which only delete vertices, and the branchings of the hybrid rule by Proposition C.12(G2), except in

regular components, which cost a constant factor. For maximum degree at least 7, the measure $w_6|V(G)|$ dominates $\sum_i w_i n_i$, and the vector $(w_6, 8w_6)$ is valid because $2^{w_6} > 1.23408$ exceeds the positive root $1.232055\ldots$ of $\lambda^8 = \lambda^7 + 1$. So graphs of maximum degree at most 4 are solved in time $2^{\max_k \tau_k n} n^{O(1)}$ and $2^{0.26736625} < 1.203609$, and all graphs in time $2^{w_6 n} n^{O(1)}$ and $2^{0.3034362} < 1.234081$, in polynomial space.

For the chromatic number, Björklund, Husfeldt and Koivisto [BHK09] show that G is k-colorable if and only if $\sum_{X \subseteq V} (-1)^{|V \setminus X|} i(G[X])^k > 0$, where $i(H)$ is the number of independent sets of H. Computing every $i(G[X])$ in polynomial space takes time $\sum_X O(1.2341^{|X|}) = O(2.2341^n)$, and binary search over k adds a polynomial factor.

(d): if the current graph has maximum degree at most 4, compute a path decomposition of width $\frac{5}{32} n_3 + \frac{1}{4} n_4 + \rho(|V|)$ by Theorem 7.1(c) and Theorem 10.1(b), with $|V|$ the number of vertices of the current graph, and count by the standard dynamic program in time $2^{\frac{5}{32} n_3 + \frac{1}{4} n_4 + \rho(|V|)} \operatorname{poly}(n)$. Above maximum degree 4 use the weights (C.13). They are at least $(0, \frac{5}{32}, \frac{1}{4})$ on degrees $2, 3, 4$, increasing and concave, and satisfy the vectors (C.11) for degrees 5 to 7 (largest sum 0.9999990, interval arithmetic; part 4 of `gaspers_lee_general.py`). These vectors are branching vectors by Proposition C.12(G1), and the hybrid rule is covered by Proposition C.12(G2), except in regular components. For maximum degree at least 8, $(w_7, 9w_7)$ is valid since the positive root of $\lambda^9 = \lambda^8 + 1$ is below $1.21315 < 2^{w_7}$. We apply Lemma C.1 with the dynamic program as the subroutine $\mathcal{B}$ and the measures $\mu + \rho(|V|)$ and $\mu_{\mathcal{B}} + \rho(|V|)$, where $\mu_{\mathcal{B}} = \frac{5}{32} n_3 + \frac{1}{4} n_4 \leq \mu$. No step increases $|V|$ and ρ is nonnegative and nondecreasing (Theorem 10.1), so every branching condition for μ remains valid. The time is $2^{w_7 n + \rho(n)} \operatorname{poly}(n)$, and $2^{0.2970601} < 1.228639$. □

C.8 Proof of (G1)

Throughout Appendices C.8 and C.9, $\mu(G) = \sum_{u \in V(G)} w_{d(u)}$ for weights $w_0 = w_1 = 0 < w_2 \leq \cdots \leq w_D$, which are *concave from* 2 if $\Delta w_2 \geq \Delta w_3 \geq \cdots \geq \Delta w_D > 0$. For $X \subseteq V(G)$ let $G \ominus X$ be the largest subgraph of $G - X$ of minimum degree at least 2, which the removal of vertices of degree at most 1, in any order, leaves; let $W(X) = V(G) \setminus (X \cup V(G \ominus X))$. For $A, B \subseteq V(G)$, $e(A, B)$ is the number of edges between A and B, and $e(A)$ the number inside A.

Lemma C.13 (decrease of the measure)**.** *Let $X \subseteq V(G)$, $G' = G \ominus X$ and $W = W(X)$. Then*

$$\mu(G) - \mu(G') = \sum_{u \in X \cup W} w_{d(u)} + \sum_{s \in V(G')} \left(w_{d(s)} - w_{d_{G'}(s)} \right),$$

with nonnegative terms, and $2 \leq d_{G'}(s) \leq d(s) - e(\{s\}, X \cup W)$ for $s \in V(G')$. If $X \subseteq X'$, then $\mu(G \ominus X') \leq \mu(G \ominus X)$.

Proof. The identity sums the weights over X, W and $V(G')$; the terms are nonnegative as the weights are nondecreasing. $G \ominus X'$ is a subgraph of $G - X$ of minimum degree at least 2, hence of $G \ominus X$. □

Let $\theta(v) = e(N[v], V(G) \setminus N[v])$. Counting edges at $N(v)$,

$$\textstyle\sum_{u \in N(v)} d(u) = d(v) + 2e(N(v)) + \theta(v), \tag{C.14}$$

so in (C.11) the parity cases have $\theta(v) \equiv \operatorname{out}(v) \pmod 2$.

Lemma C.14 (the situation of a branching)**.** *Let the algorithm branch on its selected vertex v, other than in the hybrid rule, with $5 \leq d(v) = \Delta(G) \leq 7$. Then G is connected, $|V(G)| > N_1$, $\delta(G) \geq 2$, and:*

(1) *if $P \subseteq V(G) \setminus \{v\}$ is a nonempty union of components of $G - v$ with $|P| \le K_m$, then v has no neighbor outside P; and if every edge leaving $N[v]$ ends at one vertex c, then c has no neighbor outside $N[v]$;*

(2) *there are no distinct $y, z \notin N[v]$ with $d(y), d(z) \ge 3$ such that every edge leaving $N[v]$ ends in $\{y, z\}$ and $G - \{y, z\}$ has a vertex outside $N[v]$.*

Proof. Small components are solved directly and vertices of degree at most 1 are removed. In (1), a neighbor of v outside P, or of c outside $N[v]$, would make v or c a cut vertex with a part of at most K_m vertices, and the multiplier reduction would apply. In (2), $\{y, z\}$ would be a lazy 2-separator for v, as $|N[v]| \le 8 \le K$. □

Lemma C.15 (the branch $G-v$)**.** *Let the weights be concave from 2 with $w_2 \ge 2\Delta w_D$, and let G, v be as in Lemma C.14 with $d(v) = D$. Then $\mu(G) - \mu(G \ominus \{v\}) \ge w_D + \sum_{u \in N(v)} \Delta w_{d(u)} + \deg_2(v)\Delta w_D$.*

Proof. Use Lemma C.13 with $X = \{v\}$. Let U_2 and U_3 be the neighbors of degree 2 and of degree at least 3. Each $u \in U_2$ lies in W and contributes $w_2 = \Delta w_{d(u)}$; let $t(u)$ be its other neighbor. By Lemma C.14(1), $t(u) \notin U_2$. For $s \ne v$ let $m(s) = |\{u \in U_2 : t(u) = s\}|$; it suffices that every s contributes at least $[s \in U_3]\Delta w_{d(s)} + m(s)\Delta w_D$. If s survives, it loses at least $[s \in U_3] + m(s)$ edges, and its term is a sum of that many increments Δw_i with $3 \le i \le d(s)$, the largest being $\Delta w_{d(s)}$. If $s \in W$, its term is $w_{d(s)}$, and $w_{d(s)} - \Delta w_{d(s)} \ge (d(s) - 1)\Delta w_D \ge m(s)\Delta w_D$ for $s \in U_3$, while $w_{d(s)} \ge w_2 + (d(s) - 2)\Delta w_D \ge d(s)\Delta w_D \ge m(s)\Delta w_D$ otherwise. □

Lemma C.16 (the branch $G - N[v]$)**.** *Let the weights and G, v, D be as in Lemma C.15. Then $\mu(G) - \mu(G \ominus N[v]) \ge w_D + \sum_{u \in N(v)} w_{d(u)} + \mathrm{out}(v)\,\Delta w_D$.*

Proof. Let $X = N[v]$, $W = W(X)$, $\theta = \theta(v)$ and $k_X(t) = e(\{t\}, X)$. By Lemma C.13 the decrease is $w_D + \sum_u w_{d(u)} + Q$, where, as every surviving s loses $e(\{s\}, X \cup W)$ edges and each edge leaving X ends in W or at a survivor,

$$Q \ge \sum_{t \in W} w_{d(t)} + e(X \cup W, V(G \ominus X))\,\Delta w_D \ge \theta\,\Delta w_D + \sum_{t \in W} \big(w_{d(t)} - k_X(t)\Delta w_D\big),$$

and every summand $w_{d(t)} - k_X(t)\Delta w_D$ is at least $w_2 - 2\Delta w_D \ge 0$, and at least $w_2 - \Delta w_D$ if $k_X(t) \le 1$. We show $Q \ge \mathrm{out}(v)\Delta w_D$. As $|V(G)| > N_1 \ge D + 3$ and G is connected, $\theta \ge 1$.

$\mathrm{out}(v) \in \{5, 6\}$. By Lemma C.14(1) there is no edge inside the set U_2 of neighbors of degree 2, and no two edges from U_2 to a neighbor of degree 3. So $e(N(v)) = 0$, or $e(N(v)) \le 1$ for $(2, 2, 2, 2, 2, 3)$, and (C.14) gives $\theta \ge \mathrm{out}(v)$.

$\mathrm{out}(v) \in \{3, 4\}$. If $\theta \ge \mathrm{out}(v)$ we are done; by parity, otherwise $(\theta, \mathrm{out}(v))$ is $(1, 3)$ or $(2, 4)$. If a leaving edge ab ($a \in N(v)$) has $d(b) = 2$, the vertices of degree 2 from b away from a, $b = t_1, \ldots, t_p$, end at a vertex b' of degree at least 3; if no other leaving edge ends on this walk, then $t_1, \ldots, t_p \in W$. If $\theta = 1$, with leaving edge ab, then $d(b) = 2$ by Lemma C.14(1), and $Q \ge w_2 + \Delta w_D \ge 3\Delta w_D$ (b' is removed or loses an edge). If $\theta = 2$, with leaving edges a_1b_1, a_2b_2: if $b_1 = b_2 = b$, then $d(b) \ge 3$ contradicts Lemma C.14(1), and $d(b) = 2$ makes $X \cup \{b\}$ a component of $D + 2 < |V(G)|$ vertices; if $b_1 \ne b_2$ have degree at least 3, Lemma C.14(2) is contradicted; if $d(b_1) = d(b_2) = 2$, then $b_1, b_2 \in W$ with $k_X(b_i) = 1$ and $Q \ge 2\Delta w_D + 2(w_2 - \Delta w_D) \ge 4\Delta w_D$; if $d(b_1) = 2 < d(b_2)$, the walk from b_1 ends at b' ($b' = b_2$ possible), and $Q \ge pw_2 + w_3 \ge 4\Delta w_D$ if $b_2 \in W$, $Q \ge w_2 + 2\Delta w_D$ if $b_2 \notin W$ and $b' = b_2$, and $Q \ge w_2 + \Delta w_D + \min\{w_3, \Delta w_D\}$ otherwise. □

Proof of (G1). By Lemmas C.15 and C.16 it suffices that the weights are concave from 2 with $w_2 \ge 2\Delta w_D$ for $D = d(v)$. This is checked exactly by part A of `g1_check.py`: both weight vectors are increasing and concave from 2, $w_2 = 0.1136655 \ge 4\Delta w_5 \ge 4\Delta w_6$ for (C.12), and $w_2 = 0.1103289 \ge 2\Delta w_5 = 0.0735644$ with $\Delta w_6, \Delta w_7 < \Delta w_5$ for (C.13). □

C.9 Proof of (G2)

A *lazy configuration* is the current graph G with $\Delta = \Delta(G) \geq 4$ when the hybrid rule applies, with the selected vertex x and its lazy 2-separator $\{y, z\}$, whose component $C \supseteq N[x] \setminus \{y, z\}$ of $G - \{y, z\}$ has $|C| \leq K$. Let C^+ be the union of the components of $G - \{y, z\}$ with at most K vertices, $R = V(G) \setminus (C^+ \cup \{y, z\})$, $o_y = |N(y) \cap R|$, $o_z = |N(z) \cap R|$, $H = G[C^+ \cup \{y, z\}]$ and $\bar{e} = |C^+ \setminus N[x]|$.

Lemma C.17 (the configuration)**.** *In a lazy configuration, $\delta(G) \geq 2$; the neighbors of C^+ lie in $C^+ \cup \{y, z\}$; every component of $G[R]$ has more than K vertices; every component of $G[C^+]$ contains a neighbor of y and one of z; and $|C^+| \leq \Delta K$. If $\Delta \leq 7$, then $o_y, o_z \geq 1$ and $x \notin \{y, z\}$.*

Proof. A component of $G[C^+]$ without a neighbor of z would be a component of $G - y$ with at most $K \leq K_m$ vertices; so $G[C^+]$ has at most $\min\{d(y), d(z)\}$ components. The rest of the first sentence follows from the definitions. Let $\Delta \leq 7$. G is connected with $|V(G)| > N_1 = 7K + 2 \geq \Delta K + 2$, so $R \neq \emptyset$ and not both o_y, o_z vanish. If $o_z = 0 < o_y$, then $C^+ \cup \{z\}$ is a union of components of $G - y$ with at most $(d(y) - 1)K + 1 \leq K_m$ vertices, and the multiplier reduction would apply; similarly if $o_y = 0 < o_z$. If $x = y$, all neighbors of x but z lie in C, so $o_y = 0$. □

In the branch $G - y$ the vertex z cuts off every component of $G[C^+]$, which the multiplier reduction deletes; in $G - N[y]$ the components of $G[C^+ \setminus N[y]]$ are deleted by the multiplier reduction at z or, if $yz \in E(G)$, solved directly. So the instances arise from $I_1 = G \ominus (\{y\} \cup C^+)$ and $I_2 = G \ominus (N[y] \cup C^+)$ by further reductions, which by Lemma C.13 do not increase μ; the same holds for z, and for x with $G \ominus \{x\}$ and $G \ominus N[x]$.

Lemma C.18 (credit from a large part)**.** *Let $w_2 > 0$ and $\Delta w_3 \geq \cdots \geq \Delta w_D > 0$ with $D \geq \Delta(G)$. Let $X_0 \subseteq V(G)$, $S \subseteq V(G \ominus X_0)$, and let P be a connected set of vertices of $G - X_0$, disjoint from S, of vertices of degree at least 2, with $N(P) \subseteq P \cup X_0 \cup S$. Let $W_P = P \cap W(X_0)$. Then the terms of the vertices of P in Lemma C.13 add up to at least $e(P, X_0)\Delta w_D + |W_P|w_2 - 2[W_P = P]\Delta w_D$, and so to at least $e(P, X_0)\Delta w_D$ if $|P| > K$ and $Kw_2 \geq 2\Delta w_D$.*

Proof. Let $P' = P \setminus W_P$. When $t \in W_P$ is removed it has at most one remaining neighbor, its *out-neighbor*, in W_P, P' or S; let $N_{\text{out}}, N_S, N_{\text{iso}}$ count the t whose out-neighbor lies in P', lies in S, or does not exist. Counting the edges at W_P at their later endpoints, $e(X_0, W_P) = \sum_{t \in W_P}(d(t) - 2) + N_{\text{out}} + N_S + 2N_{\text{iso}}$. The vertices of P' contribute at least $(N_{\text{out}} + e(X_0, P'))\Delta w_D$, and $t \in W_P$ contributes $w_{d(t)} \geq w_2 + (d(t) - 2)\Delta w_D$; so P contributes at least $|W_P|w_2 + (e(P, X_0) - N_S - 2N_{\text{iso}})\Delta w_D$. A component A of $G[W_P]$ has $e(A) = |A| - N_{\text{out}}(A) - N_S(A) - N_{\text{iso}}(A) \geq |A| - 1$; if $W_P \neq P$, its edge to P' is an out-edge, so $N_S(A) = N_{\text{iso}}(A) = 0$. If $W_P = P$, then $G[W_P]$ is connected and $N_S + 2N_{\text{iso}} \leq 2$. □

Lemma C.19 (one section suffices)**.** *Let I_1, I_2 be the instances of a branching of G. If $\Delta(G) = 4$, G has minimizing section k, and $a_j = \mu^{(k)}_{\text{lin}}(G) - \mu^{(k)}_{\text{lin}}(I_j) - [k = 1]\psi$ satisfy $2^{-a_1} + 2^{-a_2} \leq 1$, then $2^{\mu(I_1)} + 2^{\mu(I_2)} \leq 2^{\mu(G)}$. The same holds for $\Delta(G) \geq 5$ with $a_j = \sum_i w_i n_i(G) - \sum_i w_i n_i(I_j)$ for the weights* (C.12) *or* (C.13)*.*

Proof. It suffices that $\mu(G) - \mu(I_j) \geq a_j$. For $\Delta(G) = 4$, $\mu_4(G) = \mu^{(k)}(G) \geq \mu^{(k)}_{\text{lin}}(G)$, and $\mu(I_j) \leq \mu^{(k)}_{\text{lin}}(I_j) + [k = 1]\psi$: for $\Delta(I_j) = 4$ as $\mu_4 \leq \mu^{(k)}$, and for $\Delta(I_j) \leq 3$ as $\mu(I_j) \leq \mu^{(1)}_{\text{lin}}(I_j) \leq \mu^{(k)}_{\text{lin}}(I_j)$ by (C.10). For $\Delta(G) \geq 5$, an instance of maximum degree at most 4 has measure at most $\mu^{(5)}_{\text{lin}}$, or at most $\frac{5}{32}n_3 + \frac{1}{4}n_4$, and both are at most $\sum_i w_i n_i$. □

The three candidates. For $X_0 \subseteq V(G)$, the *H-clean-up* deletes $X_0 \cap V(H)$ from H and then removes vertices of degree at most 1, where y and z count o_y and o_z extra edges to vertices never removed. Let W' be the removed vertices, $d'(v)$ the final degree of a surviving $v \in V(H)$, and ℓ_R the number of edges between R and $\{y,z\} \cap (X_0 \cup W')$. Let

$$\beta(X_0) = \sum_{v \in (X_0 \cap V(H)) \cup W'} w_{d(v)} + \sum_{v \in V(H) \setminus (X_0 \cup W')} \left(w_{d(v)} - w_{d'(v)}\right) + \ell_R \, \kappa',$$

with $\kappa' = \Delta w_D$ if $X_0 \subseteq V(H)$ and $\kappa' = \min\{w_2, \Delta w_D\}$ if $X_0 = N[v] \cup C^+$ with $v \in \{y,z\}$. The *candidate vectors* are $\left(\beta(\{v\} \cup C^+), \max\{\beta(\{v\} \cup C^+) + o_v w_2, \beta(N[v] \cup C^+)\}\right)$ for $v \in \{y,z\}$ and $(\beta(\{x\}), \beta(N[x]))$, each component reduced by ψ in section 1.

Lemma C.20 (the candidates)**.** *In a lazy configuration, let the weights be nondecreasing with $w_2 > 0$ and $\Delta w_3 \geq \cdots \geq \Delta w_D > 0$, $D \geq \Delta$, and $(K+1-\Delta)w_2 \geq 2\Delta w_D$. Then each candidate vector is a lower bound on the decreases of $\mu_{\text{lin}} = \sum_i w_i n_i$ in the two branches of its branching. Consequently, by Lemma C.19, if one candidate vector (a,b), computed with the weights of the minimizing section for $\Delta = 4$ and with* (C.12) *or* (C.13) *for $\Delta \geq 5$, has $2^{-a} + 2^{-b} \leq 1$, the hybrid rule branches validly.*

The hypothesis holds for $K \geq 65$ and all our weights: $2\Delta w_D / w_2$ is largest for $\Delta = D = 4$ in section 1, where it is $61.03 < 62 \leq K + 1 - \Delta$.

Proof. By Lemma C.13 the decrease from G to $G \ominus X_0$ is a sum of nonnegative terms. *The H-side:* by induction on the removal order, $W' \subseteq W(X_0)$, since a vertex of W' has at most one neighbor left in $G - X_0$ when it is removed in H (vertices of C^+ have no neighbor in R). So every $v \in (X_0 \cap V(H)) \cup W'$ contributes $w_{d(v)}$, and every other $v \in V(H)$ either lies in $W(X_0)$ and contributes $w_{d(v)} \geq w_{d(v)} - w_{d'(v)}$, or survives with $d_{G \ominus X_0}(v) \leq d'(v)$, as it loses its edges to $X_0 \cup W'$; so H contributes at least the first two sums of $\beta(X_0)$. *The R-side, $X_0 \subseteq V(H)$:* with $X_0' = X_0 \cup (W(X_0) \cap V(H))$, $G \ominus X_0' = G \ominus X_0$, and each component P of $G[R]$ satisfies Lemma C.18 with X_0' and $S = \{y,z\} \setminus X_0'$; so R contributes at least $\ell_R \, \Delta w_D$. *The R-side, $X_0 = N[y] \cup C^+$:* each $r \in N(y) \cap R$ is deleted and contributes $w_{d(r)}$; if $z \in X_0 \cup W'$, each $s \in (N(z) \cap R) \setminus N(y)$ contributes w_2 if removed and $\Delta w_{d(s)}$ otherwise, and $s \in N(y) \cap N(z) \cap R$ has $d(s) \geq 3$ (else $\{s\} \subseteq C^+$) and contributes $w_{d(s)} \geq 2 \min\{w_2, \Delta w_D\}$; so R contributes at least $\ell_R \kappa'$. *The term $o_y w_2$:* with $X = \{y\} \cup C^+$ and $X' = N[y] \cup C^+$, $G \ominus X' = (G \ominus X) \ominus (X' \cap V(G \ominus X))$, so by Lemma C.13 the second branch gains at least w_2 over the first for each $s \in N(y) \cap R$ that survives in $G \ominus X$. Each other $s \in N(y) \cap R$ lies in W_P for a component P of $G[R]$, and by Lemma C.18, applied in the first branch with $X^\circ = X \cup (W(X) \cap V(H))$ in place of X_0, P contributes at least $e(P, X^\circ)\Delta w_D + |W_P \cap N(y)| w_2$: trivially if $W_P \neq P$, and if $W_P = P$ because $|P \setminus N(y)| \geq K + 1 - \Delta$ and $(K+1-\Delta)w_2 \geq 2\Delta w_D$. The case of z is symmetric. □

Lemma C.21 (the branching on y in closed form)**.** *In a lazy configuration, with the weights as in Lemma C.20, let $\zeta = w_{d(z)} - w_{o_z}$ if $o_z \geq 2$ and $\zeta = w_{d(z)} + \Delta w_D$ if $o_z = 1$, and*

$$L = w_{d(y)} + \sum_{c \in C^+} w_{d(c)} + \zeta + o_y \, \Delta w_D.$$

Then the candidate vector of y is at least $(L, L + o_y w_2)$. L does not decrease when vertices are added to C^+ or degrees grow, with o_y, o_z fixed. If $L - [k=1]\psi \geq 1$, with k the minimizing section when $\Delta = 4$, the branching on y satisfies the condition of Lemma C.1.

Proof. For $X_0 = \{y\} \cup C^+$, $W' \subseteq \{z\}$; if $o_z \geq 2$, z survives with $d'(z) = o_z$ and $\ell_R = o_y$, and if $o_z = 1$, z is removed and $\ell_R = o_y + 1$. So $\beta(X_0) = L$. The monotonicity holds as the weights are nondecreasing and $w_2 > 0$, and the last claim by Lemma C.20, as $2 \cdot 2^{-1} = 1$. □

Corollary C.22 (maximum degree at least 7)**.** *The branching on y of a lazy configuration satisfies the condition of Lemma C.1 if $\Delta = 7$ with the weights* (C.13)*, and for the measures $w_6|V|$ ($\Delta \geq 7$) and $w_7|V|$ ($\Delta \geq 8$).*

Proof. For $\Delta = 7$ and (C.13), Lemma C.17 gives $x \in C^+$ and $|N(x) \cap C^+| \geq 5$, so $L > w_3 + w_7 + 5w_2 = 1.041919$. For the measures $w_6|V|$ and $w_7|V|$, both branches delete $\{y\} \cup C^+$, which contains y, x and $N(x) \setminus \{y, z\}$, at least Δ vertices, and for $\Delta = 7$ the branch $G - N[y]$ also deletes the $o_y \geq 1$ neighbors of y in R. With $\lambda < 2^{w_6}$ the root of $\lambda^8 = \lambda^7 + 1$, $2^{-7w_6} + 2^{-8w_6} < \lambda^{-1} + \lambda^{-8} = 1$, and $2 \cdot 2^{-8w_6} < 0.38$, $2 \cdot 2^{-8w_7} < 0.39$. □

For $\Delta \geq 5$ a type is *admissible* if it contains a degree smaller than Δ, which holds in non-regular components.

Lemma C.23 (thresholds for the branching on y)**.** *The branching on y satisfies the condition of Lemma C.1 in every lazy configuration with an admissible type of x in which*

(1) *$\Delta = 4$ and the minimizing section is 4 or 5;*

(2) *$\Delta = 4$, the minimizing section is 2 or 3, and $\bar{e} \geq 2$ or some vertex of $C^+ \setminus N[x]$ has degree at least 3;*

(3) *$\Delta = 4$, the minimizing section is 1, and some vertex of $C^+ \setminus \{x\}$ has degree at least 3;*

(4) *$\bar{e} \geq 1$, and $\Delta \in \{5, 6\}$ with* (C.12) *or $\Delta = 6$ with* (C.13)*;*

(5) *$\bar{e} \geq 2$, $\Delta = 5$ and* (C.13)*.*

Proof. By Lemma C.21 with $D = \Delta$ it suffices that $L - [k = 1]\psi \geq 1$. By Lemma C.17, $d(x) = \Delta$, the degrees of $N(x) \setminus \{y, z\}$ are those of the type, $3 \leq d(y), d(z) \leq \Delta$, $1 \leq o_y \leq d(y) - 1$, $1 \leq o_z \leq d(z) - 1$, and the $\bar{e}$ vertices of $C^+ \setminus N[x]$ contribute at least $\bar{e}w_2$, or w_3 for one of degree at least 3. As L is monotone, the remaining parameters (type, whether $y, z \in N(x)$, $d(y), d(z), o_y, o_z$) take finitely many values, and `g2_thresholds.py` (parts A and B) computes the minima of L exactly. They are 1.0800248 and 1.2168722 in (1); at least 1.0236558 in (2); $L - \psi \geq 1.0070736$ in (3); at least 1.0050485 in (4); and 1.1052066 in (5). □

Theorem C.24 (the hybrid rule on bounded configurations)**.** *In every lazy configuration of the following classes, one of the three candidate vectors (a, b) satisfies $2^{-a} + 2^{-b} \leq 1$:*

(1) *$\Delta = 4$, any minimizing section k, an admissible type for section k, and $C^+ = (N[x] \setminus \{y, z\}) \cup E$ with $|E| \leq 2$;*

(2) *$\Delta \in \{5, 6\}$, an admissible type, and $C^+ = N[x] \setminus \{y, z\}$;*

(3) *$\Delta = 5$, an admissible type, and $C^+ = (N[x] \setminus \{y, z\}) \cup E$ with $|E| \leq 1$;*

in classes (2) *and* (3) *with both weight vectors* (C.12) *and* (C.13)*, and in each class optionally with one further vertex $p \in C^+$ with $N(p) = \{y, z\}$.*

Proof. The proof is computer-assisted (`g2_enumeration.py`, parts A and B). For each class the program generates every graph H on $C^+ \cup \{y, z\}$ with the prescribed degrees (all at most Δ, with $d(y), d(z) \in [3, \Delta]$ enumerated independently of the type), and every pair $o_y, o_z \geq 1$. It skips H if some $c \in C^+$ leaves a component of $H - c$ without y and z (a part of $G - c$ with at most $K \leq K_m$ vertices, so the multiplier reduction would apply), or if a component of $G[C^+]$ misses y or z (Lemma C.17). For the others it computes the three candidate vectors in exact rational arithmetic and tests $2^{-a} + 2^{-b} \leq 1$ in floating point with margin 10^{-9}, and with 40 decimal digits in the doubtful band. About $8.7 \cdot 10^6$ configurations are tested in class (1), $4.1 \cdot 10^7$ in class (2) and $2.1 \cdot 10^7$ in class (3) (the last two for each weight vector); none fails. □

Lemma C.25 (the 2-path family)**.** *Let a lazy configuration with $\Delta = 4$ have all vertices of $C^+ \setminus \{x\}$ of degree 2. Replace every maximal path of vertices of degree 2 in C that starts at a neighbor u of x by u alone, joined to the end (y or z) of the path, and every other component of $G[C^+]$, which is a path from a neighbor of y to a neighbor of z, by one vertex p with $N(p) = \{y, z\}$. The result H', with the same o_y, o_z, satisfies all conditions that the proof of Theorem C.24 tests for class (1), with the same type, $d(y), d(z), o_y, o_z$, and each candidate vector of H is componentwise at least that of H'.*

Proof. From $u \in N(x) \setminus \{y, z\}$ the vertices of degree 2 lead to x, y or z; to x would give a component of $G - x$ of at most K_m vertices. Two such paths are disjoint, and they cover C. A component of $G[C^+]$ other than C is not a cycle and has neighbors of y and z (Lemma C.17), so it is a path from y to z; there are at most $d(y) - o_y - 1 \leq 2$ of them. So H' has the same degrees at x, y, z, at most two vertices outside $N[x]$, no vertex skipped by the proof of Theorem C.24, and its part $\{x\} \cup (N(x) \setminus \{y, z\})$ keeps a neighbor of y and one of z, as the paths keep their ends. A path of degree-2 vertices is removed by a clean-up as soon as an attachment vertex in $\{x, y, z\}$ or a deleted vertex of the path goes, and keeps all degrees 2 otherwise; so the H- and H'-clean-ups remove the same vertices among x, y, z and the same paths, with the same ℓ_R, and β for H exceeds β for H' by the weights $w_2 \geq 0$ of the extra removed vertices. □

Proof of (G2). By Lemma C.20 it suffices that one candidate vector satisfies $2^{-a} + 2^{-b} \leq 1$; for $\Delta \leq 7$, $x \notin \{y, z\}$ and $C^+ \supseteq N[x] \setminus \{y, z\}$ (Lemma C.17). For $\Delta = 4$ let k be the minimizing section; the type of x is admissible for k (paragraph "Cross-section cases"). Sections 4, 5: Lemma C.23(1). Sections 2, 3: Lemma C.23(2), or class (1) of Theorem C.24. Section 1: Lemma C.23(3), or Lemma C.25 and Theorem C.24. For $\Delta \in \{5, 6\}$ the component is not regular, so x has an admissible type: Lemma C.23(4),(5) or classes (2), (3). For $\Delta \geq 7$: Corollary C.22. □

The hybrid rule cannot be replaced by the branching on y alone:

Proposition C.26 (the branching on y alone fails)**.** *For $K = 65$ there is a connected graph G with 488 vertices, $\Delta(G) = 4$, $\delta(G) = 2$, no cut vertex and minimizing section 2, all of whose vertices of degree 4 have type $(2, 3, 3, 4)$, with a vertex x and a lazy 2-separator $\{y, z\} \subseteq N(x)$, $d(y) = 3$, for which both instances of the branching on y are $G - \{x, u_1, u_2, y, r\}$ and $2^{\mu_4(I_1)} + 2^{\mu_4(I_2)} = 1.0210061 \cdot 2^{\mu_4(G)}$.*

Proof. Take $m = 80$ and the vertices $a_i, a'_i, b_{i,1}, b_{i,2}, b_{i,3}$ ($i \in \mathbb{Z}_m$), c_i ($i \in \mathbb{Z}_m \setminus \{0\}$) and $x, u_1, u_2, y, z, r, s_1, s_2, s_3$, with the edges $a_i a'_i$, $a_i b_{i,1}$, $a_i b_{i,2}$, $a'_i b_{i,2}$, $a'_i b_{i,3}$, $b_{i,1} b_{i,3}$, $b_{i,2} b_{i,3}$ for all i; $a'_i c_i$, $c_i a_{i+1}$ for $i \neq 0$; $b_{i,1} b_{i+1,1}$ for even $i \neq 0$; and $a'_0 r$, ry, yx, yu_2, xu_1, xu_2, $u_1 u_2$, xz, zs_1, $s_1 a_1$, zs_2, zs_3, $s_2 s_3$, $s_2 b_{0,1}$, $s_3 b_{1,1}$. In $G - y$, z cuts off the part $\{x, u_1, u_2\}$, which the multiplier reduction deletes, and r becomes pendant; in $G - N[y] = G - \{y, x, u_2, r\}$, u_1 loses both neighbors. So in both branches the vertices y, x, u_1, u_2, r disappear and z, a'_0 drop from degree 4 to 3, so the decrease is $2w_2^{(2)} + 2w_3^{(2)} + w_4^{(2)} + 2\Delta w_4^{(2)} = 0.9700085$. The remaining properties are checked by part C of `g2_enumeration.py`. □

D Counting dominating sets: the analysis

This appendix carries out the running-time analysis outlined in Section 13.2, in the notation of Section 13.1, and proves Theorem 13.1:

Theorem 13.1 (counting dominating sets)**.** *The numbers of dominating sets of each size of an n-vertex graph can be computed deterministically in time and space $O(1.4815^n)$. If the vertices carry real weights with $n^{O(1)}$ distinct subset sums, the same holds for the numbers of dominating sets of each size and each weight.*

Appendix D.1 proves the three steps of the analysis for a general threshold and list of exceptional types, Appendix D.2 applies them to $\mathcal{A}_{4,\mathcal{X}_{30}}$, and Appendix D.3 treats the weighted counts. Appendix D.5 applies the analysis to the previous algorithm.

D.1 The three steps of the analysis

The analysis follows the measure-and-conquer scheme of the thesis, in three steps. First, a set of branching recurrences bounds the number of instances of each measure on which the dynamic program is called [vR11, Lemma 8.11] (Lemmas D.1 to D.3). Next, a linear program bounds the pathwidth of these instances [vR11, Lemma 8.12] (Lemmas D.4 and D.5). Finally, the two bounds are combined [vR11, Theorem 8.13] (Lemma D.6). We prove the three steps for both thresholds at once.

For $a, b > 0$ let $\tau(a, b)$ be the unique $x > 1$ with $x^{-a} + x^{-b} = 1$, the *branching factor* of (a, b). A vector (a, b) *holds at* α if $\alpha^{-a} + \alpha^{-b} \leq 1$, that is, if $\tau(a, b) \leq \alpha$.

Branching vectors. For $d \geq 5$ and $r = (r_1, \ldots, r_d)$ with $\sum_i r_i = d$, $r_1 \leq 1$, and $(r_1, \ldots, r_5) \notin \mathcal{X}$ when $d = 5$, the *element vector* $E_{d,r}$ is

$$\Big(v(d) + \sum_i r_i\, \Delta w(i), \quad v(d) + \sum_i r_i\, w(i) + \Delta v(d) \sum_i (i-1) r_i\Big). \tag{D.1}$$

For $d \geq \sigma + 1$, $f(d) = \max\{d - 1, 4\}$ and $r = (r_2, \ldots, r_{f(d)})$ with $\sum_i r_i = d$, the *set vector* $S_{d,r}$ is

$$\Big(w(d) + \sum_i r_i\, \Delta v(i) + r_2\, w(1), \quad w(d) + \sum_i r_i\, v(i) + \Delta w(d) \sum_i (i-1) r_i\Big). \tag{D.2}$$

These are the recurrences of [vR11, p. 152] and [NvRvD14, Lemma 29, pp. 728–729], with the range of cases that the algorithm actually produces; Appendix D.5 shows that the previous range misses some of these cases.

Lemma D.1 (branching recurrences)**.** *Let $\sigma \in \{3, 4\}$, let $\mathcal{X}$ satisfy* (13.1)*, and let the weights satisfy (W1)–(W4). Whenever $\mathcal{A}_{\sigma,\mathcal{X}}$ branches on an instance of measure k, there is an element vector $(a, b) = E_{d,r}$ or a set vector $(a, b) = S_{d,r}$ such that the two child instances, after the reduction rules have been applied exhaustively, have measures at most $k - a$ and $k - b$.*

Proof. Since the reduction rules do not increase the measure, it suffices to bound the measure right after the branching step, except for the term $r_2 w(1)$ in the set vector, which comes from reductions and is treated separately. When the algorithm branches, no reduction rule applies, so every element has frequency at least 2, distinct sets have distinct elements, and no two elements satisfy $N_I(e_1) \subseteq N_I(e_2)$. We treat the two branching rules in turn. For each, we first bound the

sizes and frequencies that can occur, which shows that the vector is one of those defined above, and then compute the decrease of the measure in the two branches.

Branching on an element e (rule (C)). Rule (B) did not apply, so $d := d(e) \geq 5$ and $d(s) \leq d$. Every set containing e has size at most d: non-exceptional sets have size at most $d(s)$, and exceptional ones have size $\sigma + 1 \leq 5$. Every element has frequency at most d: non-exceptional ones by the choice of e, exceptional ones have frequency 5. At most one set containing e has size 1, since two such sets would both equal $\{e\}$. If $d = 5$, the type of e is not in $\mathcal{X}$, since e is not exceptional. Let r_i be the number of sets of size i containing e.

In the optional branch, e is deleted, which removes $v(d)$, and each set containing e shrinks by one, which removes $\Delta w(i)$ for a set of size i (a set of size 1 becomes empty and loses $w(1) = \Delta w(1)$). This is the first entry of (D.1). In the forbidden branch, e and the sets containing it are deleted, which removes $v(d) + \sum_i r_i w(i)$, and every other element f of these sets loses frequency. The sets containing e have $\sum_i (i-1) r_i$ incidences with other elements in total. We claim that an element of frequency j that loses $t \geq 1$ of its sets loses at least $t\,\Delta v(d)$ in weight; summing over the elements gives the second entry. We have $2 \leq j \leq d$ and $1 \leq t \leq j$. If $j - t \geq 1$, then $v(j) - v(j-t) = \sum_{l=j-t+1}^{j} \Delta v(l)$ with $2 \leq l \leq d$ (recall $v(1) = 0$), and each term is at least $\Delta v(d)$ by (W3). If $j - t = 0$, then $t = j$ and $v(j) - v(0) = v(2) + \sum_{l=3}^{j} \Delta v(l) \geq 2\Delta v(d) + (j-2)\Delta v(d) = t\,\Delta v(d)$, using $v(2) \geq 2\Delta v(5) \geq 2\Delta v(d)$ from (W4), (W3) and $d \geq 5$.

Branching on a set s (rule (B)). Let $d = d(s)$. First, $d \geq \sigma + 1$: if $d \leq \sigma$, then either $d(e) \leq 4$, and the switch (A) applies, or $d(e) \geq 5 > d(s)$, and the algorithm branches on e. Next, s contains no exceptional element: if $d \leq 5$, a set of size in $\{\sigma + 1, \ldots, 5\}$ containing an exceptional element would be exceptional, and s is not; if $d \geq 6$, s cannot contain an exceptional element, because by (13.1) exceptional elements lie only in sets of size at most $\sigma + 1 \leq 5$. Hence every element of s is non-exceptional and has frequency at most $d(e)$. Since rule (B) applies, $d(e) < d$ or $d(e) \leq 4$, so every element of s has frequency in $\{2, \ldots, f(d)\}$. Every set has size at most d: non-exceptional sets by the choice of s, exceptional sets have size $\sigma + 1 \leq d$. Let r_i be the number of elements of s of frequency i.

In the take branch, s and its elements are deleted, which removes $w(d) + \sum_i r_i v(i)$, and every other set loses the elements it shares with s; these are $\sum_i (i-1) r_i$ incidences. A set of size $j \leq d$ that loses $t \geq 1$ elements loses $w(j) - w(j-t) = \sum_{l=j-t+1}^{j} \Delta w(l) \geq t\,\Delta w(d)$, since $1 \leq l \leq d$ and (W3). This is the second entry of (D.2).

In the discard branch, s is deleted, which removes $w(d)$, and each element of s of frequency i loses $\Delta v(i)$ (for $i = 2$ this is $v(2)$). This gives $w(d) + \sum_i r_i \Delta v(i)$ before reductions. Let E_2 be the elements of s of frequency 2, and for $e \in E_2$ let $s_e \neq s$ be the other set containing e. The sets s_e are distinct: if $s_e = s_{e'}$ for $e \neq e'$, then $N_I(e) = \{s, s_e\} = N_I(e')$ before branching, and (R3) would have applied. We show that every s_e is deleted during the reductions that follow. Then the final measure is at most the measure before reductions minus $\sum_{e \in E_2} w(d(s_e)) \geq r_2 w(1)$, because no weight increases during reductions, each s_e is a distinct vertex of weight at least $w(1)$ before reductions, and it is absent at the end.

During these reductions no element reaches frequency 0: elements of s have frequency at least 1 after the discard; (R2) and (R3) delete elements but do not lower the frequency of the remaining ones; and (R1) lowers the frequency of an element only if it lies in two identical sets, leaving frequency at least 1. Suppose some s_e is never deleted. We may assume that it is never removed by (R1) either, by following the merged representative, which has the same elements. Then e is not deleted by (R2), since that would delete its only set s_e. If e survives to the end, it has frequency 1 there, and (R2) would apply, a contradiction. So e is deleted by (R3), because of an element e_1 with $\emptyset \neq N_I(e_1) \subseteq N_I(e) = \{s_e\}$; then $e_1 \in s_e$ has frequency 1, and we repeat the argument with e_1 in place of e. As there are finitely many elements, this is a contradiction. So every s_e is deleted,

which completes the proof of the first entry of (D.2). □

Lemma D.2 (high degrees)**.** *Assume (W1)–(W3). For every element or set vector of degree $d \geq 10$ there is a vector of the same kind and of degree $d-1$ that is componentwise at most the given one. Hence, if every vector of degree at most 9 holds at α, then every vector holds at α.*

Proof. By (W1), $\Delta v(i) = \Delta w(i) = 0$ and $v(i) = v(7)$, $w(i) = w(7)$ for $i \geq 8$.

Element vectors. Let $d \geq 10$. Delete one of the d neighbors from r, and replace every remaining neighbor of size d by one of size $d-1$; the result is a neighborhood r' of $d-1$ sets of sizes at most $d-1$ with $r'_1 \leq r_1 \leq 1$, so $E_{d-1,r'}$ is an element vector ($d-1 \geq 9 > 5$). Since $d-1 \geq 9$, the weights $v(d) = v(d-1)$, $\Delta v(d) = \Delta v(d-1) = 0$, $w(d) = w(d-1)$ and $\Delta w(d) = \Delta w(d-1) = 0$ do not change, and deleting a set of size i lowers the two entries by $\Delta w(i) \geq 0$ and $w(i) \geq 0$.

Set vectors. Let $d \geq 10$, so $f(d) = d-1$ and $f(d-1) = d-2$. Delete one element from r and replace every remaining element of frequency $d-1$ by one of frequency $d-2$; this gives a set vector $S_{d-1,r'}$. Since $d-2 \geq 8$, we have $v(d-1) = v(d-2)$ and $\Delta v(d-1) = \Delta v(d-2) = 0$; also $w(d) = w(d-1)$ and $\Delta w(d) = \Delta w(d-1) = 0$. Deleting an element of frequency i lowers the entries by $\Delta v(i)$ (plus $w(1)$ if $i = 2$) and by $v(i)$, both nonnegative.

Since τ is decreasing in each argument, a vector componentwise at least a vector that holds at α also holds at α; induction on d gives the last statement. □

The bound $d \geq 10$ cannot be lowered to $d \geq 9$ for set vectors: at $d = 9$, elements of frequency 8 would be replaced by elements of frequency 7, and $\Delta v(7) > 0 = \Delta v(8)$, so the discard entry may increase. Vectors of degree 9 are therefore checked directly.

Lemma D.3 (leaves)**.** *Let σ, $\mathcal{X}$ and the weights be as in Lemma D.1, and let every element and set vector hold at $\alpha > 1$. For every instance J of measure k with incidence graph I, the recursion tree of J has at most α^{k-h} calls of the dynamic program on instances of measure h, for every real $h \geq 0$, and at most $2(1 + |V(I)|)$ times as many nodes as leaves.*

Proof. Let N_h be the number of calls of the dynamic program on instances of measure exactly h in the recursion tree of an instance J of measure k. We show $N_h \leq \alpha^{k-h}$ by induction on the size of the recursion tree. If J is itself a leaf, $N_h \leq 1$, and $N_h = 0$ unless $k = h$. If J is reduced, its child has measure at most k, and N_h is the same. If J branches, its two children, after reductions, have measures at most $k-a$ and $k-b$ by Lemma D.1, so $N_h \leq \alpha^{k-a-h} + \alpha^{k-b-h} \leq \alpha^{k-h}$. For the second statement, every branching node has two children, so there are fewer branching nodes than leaves. The reduction steps form chains, each ending at a branching node or a leaf, so there are at most twice as many chains as leaves, and each chain has length at most $|V(I)|$, since each step deletes a vertex. □

The instances handed to the dynamic program. Let J be an instance on which the switch (A) applies, and let I be its incidence graph. Then every non-exceptional set has size at most σ, every non-exceptional element has frequency at most 4, every exceptional element has frequency 5 and a type in $\mathcal{X}$, and every set of size larger than σ is exceptional and has size $\sigma + 1$. In particular $\Delta(I) \leq 5$. Let I' be obtained from I by deleting the sets of size 1 and the isolated vertices.

Lemma D.4 (pendant vertices)**.** *With I and I' as above,* $\mathrm{pw}(I) \leq \mathrm{pw}(I') + 1$.

Proof. Isolated vertices can be added as one-vertex bags. A set of size 1 is a vertex of degree 1 adjacent to its element f. Take a path decomposition of I', and for each such vertex s choose a bag $B \ni f$; insert directly after B the bag $B \cup \{s\}$, and insert the bags for vertices attached at the same

B one after another. Every new bag has at most $|B|+1$ vertices, every edge sf is covered, and every vertex still occurs in consecutive bags. This is the argument of [NvRvD14, Proposition 28] and [vR11, Remark 8.1]. □

We now bound $\mathrm{pw}(I')$ by *pathwidth forms*. A form is a nondecreasing function $\varphi : \{1,\dots,5\} \to [0,1]$ such that every graph H without isolated vertices and with $\Delta(H) \le 5$ satisfies

$$\mathrm{pw}(H) \le \sum_{x \in V(H)} \varphi(d_H(x)) + \rho_\varphi(|V(H)|) \tag{D.3}$$

with a remainder $\rho_\varphi(t) = o(t)$, which we take nondecreasing (Section 4.1). A form is *computable* if for every such H a path decomposition of H of width at most the right side of (D.3) can be computed in polynomial time. We use the following forms, written as $(\varphi(1),\dots,\varphi(5))$; the decimal values are rational upper bounds for the constants, as used in the certificates below.

- $\varphi_{\mathsf{S}} = (0, 0, \frac{5}{32}, \frac{1}{4}, \frac{11}{32})$, by Theorem 1.3 (or Theorem 7.2 with S);
- $\varphi_{\mathsf{R}} = (0, 0, 0.1455, 0.2530, 0.3600)$ and $\varphi_{\mathsf{P}} = (0, 0, 0.140159, 0.2575, 0.3667)$, by Theorem 7.2, using $F_3 \le 0.140159$ (Proposition 6.9);
- $\varphi_a = (F_4, F_4, F_4, F_4, 1)$ with $F_4 \le 0.24577197475$ (Proposition 6.9): by Theorem 7.1(a) applied to $H - V_5$, where V_5 is the set of vertices of degree 5, and $\mathrm{pw}(H) \le \mathrm{pw}(H - V_5) + |V_5|$ (Section 4.1); Theorem 7.1(a) counts every vertex of $H - V_5$, whatever its degree;
- $\varphi_b = (0, 0, 0.140159, 0.2575, 1)$: by Theorem 7.1(b) applied to $H - V_5$;
- $\varphi_3 = (0, 0, 0.139823, 1, 1)$: by Lemma 11.2 applied to $H - V_{\ge 4}$, where $V_{\ge 4}$ is the set of vertices of degree at least 4, with $c_3 \le 0.139823$.

In the last three items, let X be the deleted set, V_5 or $V_{\ge 4}$. The degrees in $H - X$ are at most the degrees in H, and the cited bounds have coefficients nondecreasing in the degree, so (D.3) holds with the degrees of H. All these forms are computable: a decomposition of the stated width is computed deterministically in polynomial time by Theorem 10.1(b) or by Lemma 11.2, and in the last three items X is then added to all bags.

Assign to each vertex of I' a *nominal degree*: a set gets its size, a non-exceptional element its frequency, and an exceptional element of type c gets $5 - [c_1 > 0]$, its degree after the deletion of its set of size 1 (it has at most one). The nominal degree is at least the degree in I', since an element lies in at most one set of size 1 (two such sets would be identical); the two differ only for elements in a set of size 1. Since each form is nondecreasing and I' has no isolated vertices and maximum degree at most 5, $\mathrm{pw}(I')$ is at most the sum of the form over the nominal degrees of the vertices of I', plus the remainder.

The linear program. For a leaf instance of measure $h > 0$, divide the numbers of vertices of each kind by h: let x_i ($1 \le i \le \sigma + 1$) be the scaled number of sets of size i, y_i ($2 \le i \le 4$) that of non-exceptional elements of frequency i, and e_c ($c \in \mathcal{X}$) that of exceptional elements of type c.

These numbers satisfy

$$\sum_{i=1}^{\sigma+1} w(i)\, x_i + \sum_{i=2}^{4} v(i)\, y_i + v(5) \sum_c e_c = 1, \tag{D.4}$$

$$\sum_{i=1}^{\sigma+1} i\, x_i = \sum_{i=2}^{4} i\, y_i + 5 \sum_c e_c, \tag{D.5}$$

$$\sum_c c_j\, e_c \le j\, x_j \qquad (1 \le j \le \sigma+1), \tag{D.6}$$

$$x_{\sigma+1} \le \sum_c c_{\sigma+1}\, e_c, \tag{D.7}$$

and all variables are nonnegative. Here (D.4) is the measure (sets of size 0 and elements of frequency 0 have weight 0, and elements of frequency 1 do not occur), and (D.5) counts the edges of I from both sides. Inequality (D.6) holds because the $\sum_c c_j e_c$ incidences between exceptional elements and sets of size j are at most j per such set, and (D.7) holds because every set of size $\sigma+1$ is exceptional and so contains an exceptional element.

For a form φ let $\Phi_\varphi(x, y, e)$ be the sum of φ over the nominal degrees, scaled:

$$\Phi_\varphi = \sum_{i=2}^{\sigma+1} \varphi(i)\, x_i + \sum_{i=2}^{4} \varphi(i)\, y_i + \sum_c \varphi(5 - [c_1 > 0])\, e_c,$$

where sets of size 1 contribute nothing. Let $L(\sigma, \mathcal{X}, \Psi)$ be the maximum of $\min_{\varphi\in\Psi} \Phi_\varphi$ over the polytope (D.4)–(D.7), for a set Ψ of forms. This is the linear program of [vR11, Lemma 8.12] and [NvRvD14, Lemma 30, pp. 729–731], with several forms and with the constraints (D.6)–(D.7) for a general threshold.

Lemma D.5 (pathwidth at the switch)**.** *Let σ, $\mathcal{X}$ and the weights be as in Lemma D.1, let Ψ be a finite set of computable forms, and let $L \ge L(\sigma, \mathcal{X}, \Psi)$. Then for every instance of measure h on which $\mathcal{A}_{\sigma,\mathcal{X}}$ calls the dynamic program, a path decomposition of its incidence graph of width at most $L\,h + o(h)$ can be computed in polynomial time, where the $o(h)$ depends only on the weights and on Ψ.*

Proof. The scaled counts of a leaf instance are feasible, so $\min_\varphi \Phi_\varphi \le L$, that is, $\min_\varphi \sum_x \varphi(\cdot) \le Lh$ before scaling. By the discussion of nominal degrees and (D.3), $\mathrm{pw}(I') \le Lh + \max_\varphi \rho_\varphi(|V(I')|)$, and $\mathrm{pw}(I) \le \mathrm{pw}(I') + 1$ by Lemma D.4. Every vertex of I' has weight at least $\min\{v(2), w(1)\} > 0$ in the measure, so $|V(I')| \le h/\min\{v(2), w(1)\}$, and the remainder is $o(h)$. The decomposition is the narrowest of those computed for the forms, followed by the insertions of Lemma D.4. □

Certificates for L. Write the program in terms of the vector $z = (x, y, e)$: (D.4) as $\mathbf{m}^\top z = 1$, (D.5) as $\mathbf{a}^\top z = 0$, (D.6)–(D.7) as $Bz \le 0$, and $\Phi_\varphi(z) = \mathbf{f}_\varphi^\top z$. By weak duality, a bound $L \ge L(\sigma, \mathcal{X}, \Psi)$ is certified by $\lambda_\varphi \ge 0$ ($\varphi \in \Psi$) with $\sum_\varphi \lambda_\varphi = 1$, a real μ and a vector $\nu \ge 0$ such that, componentwise,

$$\sum_{\varphi\in\Psi} \lambda_\varphi\, \mathbf{f}_\varphi \le L\,\mathbf{m} + \mu\,\mathbf{a} + B^\top \nu. \tag{D.8}$$

Indeed, the inner product of (D.8) with a feasible point $z \ge 0$ gives $\min_\varphi \Phi_\varphi(z) \le \sum_\varphi \lambda_\varphi \Phi_\varphi(z) \le L\,\mathbf{m}^\top z + \mu\,\mathbf{a}^\top z + \nu^\top B z \le L$, since $\nu \ge 0$ and $Bz \le 0$.

Table 12: Weights of the measure (13.3): previous ([vR11, p. 151], [NvRvD14, p. 728]), certified for the previous algorithm with all cases (Proposition D.10), and certified for Theorem 13.1. All values are exact decimals; $v(i) = v(7)$ and $w(i) = w(7)$ for $i \geq 7$.

i	1	2	3	4	5	6	7
previous $v(i)$	0	0.275603	0.551206	0.658484	0.719009	0.745509	0.749538
previous $w(i)$	0.354816	0.625964	0.785112	0.923325	0.986711	1	1
repaired $v(i)$	0	0.2351251	0.4702502	0.5842398	0.6447407	0.6729014	0.6766715
repaired $w(i)$	0.3126782	0.5477033	0.6838767	0.8200501	0.8626895	0.8695487	0.8695487
Theorem 13.1 $v(i)$	0	0.2175171	0.4350342	0.5562242	0.6220752	0.6539177	0.6655099
Theorem 13.1 $w(i)$	0.3062246	0.5338365	0.6754745	0.7616247	0.8178813	0.8324929	0.8324929

Lemma D.6 (running time)**.** *Let σ, $\mathcal{X}$ and the weights be as in Lemma D.1, let Ψ be a finite set of computable forms, and let $\alpha > 1$ with $2^{L(\sigma,\mathcal{X},\Psi)} < \alpha$. If every branching vector of degree at most 9 holds at α, then $\mathcal{A}_{\sigma,\mathcal{X}}$, with the dynamic program run on the narrowest decomposition of Lemma D.5, counts the dominating sets of each size of an n-vertex graph in time and space $O\big(\alpha^{(v(7)+w(7))n}\, n^{O(1)}\big)$.*

Proof. Fix a real $L \geq L(\sigma, \mathcal{X}, \Psi)$ with $2^L < \alpha$. By Lemma D.2 every branching vector holds at α, so by Lemma D.3 at most α^{k-h} leaves have measure h, where $k \leq (v(7) + w(7))n$ is the initial measure. By Lemma D.5 and the dynamic program, a leaf of measure h takes time $2^{Lh+o(h)}n^{O(1)}$. Since $2^L < \alpha$ and the remainder is $o(h)$, there is a constant C with $2^{Lh+o(h)} \leq C\alpha^h$ for all $h \geq 0$. The measure is determined by the numbers of sets of each size $1, \ldots, 7$ (sizes at least 7 together) and of elements of each frequency $2, \ldots, 7$; these are at most $2n$, so at most $(2n+1)^{13}$ values of h occur. Every vertex of the recursion tree costs polynomial time besides the dynamic program, and there are polynomially many nodes per leaf by Lemma D.3. Summing,

$$T \leq n^{O(1)} \sum_h \alpha^{k-h}\, C\, \alpha^h \leq n^{O(1)}\, \alpha^{(v(7)+w(7))n}.$$

The space is dominated by the tables of the dynamic program, at most $2^{Lh+o(h)}n^{O(1)} \leq C\alpha^h n^{O(1)}$ at a leaf of measure $h \leq k$. □

D.2 Proof of Theorem 13.1

The algorithm of Theorem 13.1 is $\mathcal{A}_{4,\mathcal{X}_{30}}$. Compared with the previous one, it switches to dynamic programming when every non-exceptional set has size at most 4 (rather than 3) and every non-exceptional element has frequency at most 4. Its exceptional elements are those with a type in $\mathcal{X}_{30}$, and its exceptional sets are the sets of size 5 that contain one. The dynamic program runs on the narrowest of the path decompositions for the forms $\Psi = \{\varphi_{\mathsf{S}}, \varphi_{\mathsf{R}}, \varphi_{\mathsf{P}}, \varphi_a, \varphi_b, \varphi_3\}$ listed before Lemma D.5, computed by Theorem 10.1(b) and Lemma 11.2, instead of the decomposition of Fomin, Gaspers, Saurabh and Stepanov [FGSS09, Lemma 1]. Correctness is Lemma 13.2.

Proof of Theorem 13.1 for unweighted counting. We apply Lemma D.6 with $\sigma = 4$, $\mathcal{X} = \mathcal{X}_{30}$, $\alpha = \frac{13}{10}$, the weights of Table 12 and the set Ψ of six computable forms above. Each step is checked by the accompanying code, as follows.

Weights. The weights are rational and satisfy (W1)–(W4), checked in exact rational arithmetic. Explicitly, the increments

$$(\Delta v(2),\dots,\Delta v(7)) = (0.2175171, 0.2175171, 0.1211900, 0.0658510, 0.0318425, 0.0115922),$$
$$(\Delta w(1),\dots,\Delta w(7)) = (0.3062246, 0.2276119, 0.1416380, 0.0861502, 0.0562566, 0.0146116, 0)$$

are nonnegative and nonincreasing, and $2\Delta v(5) = 0.131702 \le v(2)$. The list $\mathcal{X}_{30}$ satisfies (13.1).

Branching. By Lemma D.1 the branching vectors are the element vectors $E_{d,r}$ ($d \ge 5$) and the set vectors $S_{d,r}$ ($d \ge 5$, frequencies at most $d-1$). There are 30,972 of them of degree at most 9. Their entries are exact rationals; the code evaluates $\alpha^{-a} + \alpha^{-b}$ in interval arithmetic (mpmath, 120-bit precision), and the largest upper end is $0.999999610 < 1$. By Lemma D.2, every branching vector holds at α.

Dynamic programming. The code computes an exact rational solution of (D.8) for the program with $\sigma = 4$, $\mathcal{X}_{30}$ and Ψ, which certifies $L(4, \mathcal{X}_{30}, \Psi) \le L$ with $L \approx 0.378511274$, and $2^L < 1.2999997 < \alpha$, by interval arithmetic. In floating point, the optimum is attained at $x_4 = y_4 \approx 0.7492$ with small amounts of sets of size 2 and 3 and of exceptional elements of type $(0, 1, 4, 0, 0)$; so the worst instances are close to incidence graphs of cubic graphs. The exact certificate puts weight only on φ_{S} (Theorem 1.3) and φ_a (Theorem 7.1(a) after deleting the vertices of degree 5), with $\lambda_{\varphi_{\mathsf{S}}} = \frac{460610214}{535261133} \approx 0.8605$ and $\lambda_{\varphi_a} = \frac{74650919}{535261133} \approx 0.1395$, so it also certifies the bound for these two forms alone; with φ_{S} alone the value is $0.3794062 > \log_2 \alpha$ in floating point.

Combination. By Lemma D.6 the running time is $O(\alpha^{(v(7)+w(7))n} n^{O(1)})$ with $v(7) + w(7) = 1.4980028$, and $\alpha^{1.4980028} < 1.481452 < 1.4815$, which absorbs the polynomial factor. The algorithm is deterministic, since Theorem 10.1(b) and Lemma 11.2 are. □

Since the exact certificate puts weight only on φ_{S} and φ_a, it also certifies the bound with $\Psi = \{\varphi_{\mathsf{S}}, \varphi_a\}$; with this Ψ the proof uses, among the results of this paper, only Theorem 1.3, Theorem 7.1(a) and Theorem 10.1. The other forms are kept because adding a form to Ψ never raises $L(\sigma, \mathcal{X}, \Psi)$, and its decomposition is computed in polynomial time.

D.3 Weighted versions

Let the vertices of G have real weights, and let the set $\Sigma = \{\sum_{v\in X} \mathrm{wt}(v) : X \subseteq V(G)\}$ of subset sums have size polynomial in n. We count dominating sets by size and weight; the least weight with a nonzero count, and that count, solve Minimum Weight Dominating Set and #Minimum Weight Dominating Set. The algorithm and its measure stay the same; only the bookkeeping changes.

The difficulty is the identical-sets rule (R1): true twins of different weights give identical sets of different weights, so after merging a multiplicity no longer describes the copies. We therefore replace the multiplicity of a set s by the generating polynomial

$$P_s(x, z) = \prod_c (1 + x\, z^{\mathrm{wt}(c)}) - 1$$

of the nonempty selections of its copies c (vertices of G), by number (x) and weight (z); initially $P_{N[v]} = x\, z^{\mathrm{wt}(v)}$. All copies have the same elements, so every nonempty selection covers the same elements. The algorithm returns a polynomial $\sum_{\ell,\omega} c_{\ell,\omega} x^\ell z^\omega$, and the rules become:

- (R1): the merged set gets $(1 + P_{s_1})(1 + P_{s_2}) - 1$;
- (R2), the take branch of (B), and the dynamic program when it puts s into the solution: multiply by P_s in place of (13.2); discarding s multiplies by 1;

- (R3) and (C) involve no weights and are unchanged.

Correctness. Different sets have disjoint sets of copies, since the rules only merge or remove sets. Call a set Y of copies a *cover* if the sets of its copies have union $\mathcal{U}$; its size is $|Y|$ and its weight is $\sum_{c\in Y} \mathrm{wt}(c)$. By induction on $|V(I)|$, the algorithm returns the generating polynomial of the covers by size and weight; the recursion terminates, since every rule produces instances with fewer vertices of I. The nonempty selections of copies of a set s have generating polynomial P_s, and each covers exactly the elements of s. (R1): a selection of copies of the merged set is a pair of selections of copies of s_1 and of s_2. (R2) and the take branch of (B): a cover that contains a copy of s is a nonempty selection of copies of s together with a cover of $\mathcal{U} \setminus s$ by the other sets, and in (R2) every cover contains a copy of s. The discard branch of (B): the covers without a copy of s are the covers by $\mathcal{S} \setminus \{s\}$. (R3): every cover contains a copy of a set containing e_1, which contains e_2, so a set of copies covers $\mathcal{U}$ if and only if it covers $\mathcal{U} \setminus \{e_2\}$. (C): a set of copies that covers $\mathcal{U} \setminus \{e\}$ either covers e or has no copy of a set containing e, and the latter are the covers of $\mathcal{U} \setminus \{e\}$ by $\mathcal{S} \setminus N_I(e)$. (A): the dynamic program of [NvRvD14, Proposition 27] handles each set with its polynomial when the set is processed, multiplying by P_s when it puts s into the solution; an empty set s contributes the factor $1 + P_s$ (any selection of its copies, in place of $(1 + x)^{\mathrm{mult}(s)}$), and an element of frequency 0 gives 0. Every polynomial that occurs is the generating function of selections of pairwise disjoint sets of vertices of G (products combine disjoint sets of copies), so its exponents lie in $\{0, \ldots, n\} \times \Sigma$, and every arithmetic operation on these polynomials takes time polynomial in n and $|\Sigma|$. The measure, the recursion tree and the decompositions are those of the unweighted case, so the running time is again $O(1.4815^n)$. This proves the weighted part of Theorem 13.1.

Nederlof, van Rooij and van Dijk state only Minimum Weight Dominating Set [NvRvD14, Corollary 32], with a proof that refers to their polynomial-space algorithm, which has no merged sets; van Rooij's thesis also states the counting version [vR11, Corollary 8.14].

The accompanying code checks this bookkeeping (not the running time) against brute force on small random graphs with true twins of different weights.

D.4 Remarks

Remark D.7 (variants). The following values come from the model of Appendix D.1 (recurrences and linear program), adapted to variants of the algorithm. Certified values are marked; the others are floating-point optima of the weights at $\alpha = 1.3$ (the objective $\alpha^{v(7)+w(7)}$ does not depend on the scale of the weights).

- With the previous list $\mathcal{X}_{14}$ and $\sigma = 4$, the same proof gives $O(1.4835^n)$ (certified: base < 1.483408, with $L \approx 0.378511246$ and 30,988 vectors). With the form φ_{S} of Theorem 1.3 alone, the bases are 1.4837 ($\mathcal{X}_{14}$) and 1.4818 ($\mathcal{X}_{30}$).

- The improvement comes from handing sets of size 4 to the dynamic program. With the previous $\sigma = 3$, the new width bounds do not lower the running-time bound (as noted after Proposition D.10); with the coefficients of Fomin, Gaspers, Saurabh and Stepanov [FGSS09, Lemma 1], the switch at $\sigma = 4$ gives only 1.5997^n.

- Other thresholds are worse: switching when sets have size at most 3 and elements frequency at most 5 gives 1.4930^n; sets at most 4 and elements at most 5, 1.5117^n; sets at most 5 and elements at most 4, 1.5240^n (all without exceptional elements). With weights that are constant only from degree 9 on, the optimum for $\mathcal{A}_{4,\mathcal{X}_{30}}$ is unchanged.

The list $\mathcal{X}_{30}$ was found by a greedy search; we did not try to optimize it further.

Remark D.8 (status of the computations). The floating-point optimizer only produces candidate weights; every constant in the proofs of Proposition D.10 and Theorem 13.1 is certified as described there. A separate recheck, with code written apart from the main script, of all branching vectors up to degree 12 (1,747,700 vectors) found the largest value 0.99999961 of $\alpha^{-a} + \alpha^{-b}$, consistent with Lemma D.2.

D.5 A gap in the previous analysis

This subsection describes a gap in the previous running-time analysis of the algorithm of Nederlof, van Rooij and van Dijk for counting dominating sets [NvRvD14], which also appears in van Rooij's thesis [vR11, §8.5]. In the notation of Section 13, it is the algorithm $\mathcal{A}_{3,\mathcal{X}_{14}}$, with switching threshold $\sigma = 3$ and the previous list $\mathcal{X}_{14}$ of exceptional types (Section 13.1). The previous recurrences do not cover branchings on a set of size 4 that contains an element of frequency 4, and with the previous weights some of these cases violate the claimed branching factor (Proposition D.9). We show that the algorithm does branch on such sets, and that with all cases included it runs in time $O(1.5004^n)$ (Proposition D.10). The analysis of Theorem 13.1 in Appendix D.2 uses the corrected set of recurrences.

Throughout, the notation is that of Section 13.1 and Appendix D.1: the measure k of (13.3) with weights v and w and their increments Δv and Δw, the set vectors $S_{d,r}$ and element vectors $E_{d,r}$ of (D.2) and (D.1), the branching factor $\tau(a, b)$, and the linear program for the pathwidth at the switch.

The previous analysis. The thesis [vR11, Lemma 8.11, p. 152] and the journal version [NvRvD14, Lemma 29, pp. 728–729] state that, with the previous weights of Table 12 ([vR11, p. 151], [NvRvD14, p. 728]), the branching produces at most 1.26089^{k-h} subproblems of measure h, where k is the measure of the input instance. The linear program of [vR11, Lemma 8.12] and [NvRvD14, Lemma 30] has value 0.325782, so that the dynamic program on a subproblem of measure h runs in time $O(1.25334^h)$. From these, [vR11, Theorem 8.13] and [NvRvD14, Theorem 31] conclude the running time $O(1.26089^{(0.749538+1)n}) = O(1.5002^n)$.

We reproduced the previous analysis exactly, including the value 0.325782 of the linear program. With the previous weights, every branching vector of Lemma D.1 for $\sigma = 3$ has $\tau \le 1.260893$, except four set vectors $S_{4,r}$ of degree 4, where r_i is the number of elements of the set that have frequency i.

Proposition D.9 (the uncovered case)**.** *For $\sigma = 3$ and the previous weights, the set vector $S_{4,r}$ with $r_4 = 4$ is*

$$\big(w(4) + 4\Delta v(4),\ w(4) + 4v(4) + 12\Delta w(4)\big) = (1.352437,\ 5.215817),$$

and $\tau(1.352437, 5.215817) = 1.27583 > 1.26089$. Also $\tau > 1.26089$ for the set vectors $S_{4,r}$ with (r_3, r_4) equal to $(3, 1)$, $(2, 2)$ or $(1, 3)$.

Proof. With the previous weights, $\Delta v(4) = 0.658484 - 0.551206 = 0.107278$ and $\Delta w(4) = 0.923325 - 0.785112 = 0.138213$, so the entries are $0.923325 + 0.429112 = 1.352437$ and $0.923325 + 2.633936 + 1.658556 = 5.215817$. The values of τ were computed numerically by the accompanying code; $\tau(1.352437, 5.215817)$ lies in $(1.2758, 1.2759)$, since $1.2758^{-1.352437} + 1.2758^{-5.215817} > 1$ and $1.2759^{-1.352437} + 1.2759^{-5.215817} < 1$. The largest value of τ for the other three vectors is 1.27004. □

The algorithm branches on these sets. The switch of [vR11, Algorithm 8.2] requires $d(s) \le 3$, so when $d(s) = 4$ and $d(e) \le 4$ the algorithm branches on the set (rule (B)), as the journal version

says [NvRvD14, pp. 726–727]. The case already occurs at the root for every cubic graph without true twins, where every set has size 4, every element has frequency 4 and no reduction rule applies. The coefficient $12\,\Delta w(4)$ of the printed recurrence is the correct one, since no exceptional type in $\mathcal{X}_{14}$ has a set of size 5.

A repair. We do not know whether the algorithm is slower than $O(1.5002^n)$; the gap concerns the proof. Including all cases of Lemma D.1 and re-optimizing the weights gives the following repair.

Proposition D.10 (the previous algorithm, repaired)**.** *The algorithm $\mathcal{A}_{3,\mathcal{X}_{14}}$ of Nederlof, van Rooij and van Dijk, with the decompositions of [FGSS09, Lemma 1], counts the dominating sets of each size of an n-vertex graph in time and space $O(1.5004^n)$.*

Proof. We apply Lemma D.6 with $\sigma = 3$, $\mathcal{X} = \mathcal{X}_{14}$ (which satisfies (13.1)), $\alpha = \frac{13}{10}$, the repaired weights of Table 12, and $\Psi = \{\varphi_\varepsilon\}$, and check its hypotheses.

The form. Let $\varphi_{\mathrm{FGSS}} = (0, 0, \frac{1}{6}, \frac{1}{3}, \frac{13}{30})$ be the coefficients of [FGSS09, Lemma 1] for degrees $1, \ldots, 5$, and let $\varphi_\varepsilon = \varphi_{\mathrm{FGSS}} + \varepsilon$ on every degree, for a fixed $\varepsilon > 0$ chosen below. For fixed ε, Lemma 1 of [FGSS09] gives a path decomposition of width at most $\sum_x \varphi_{\mathrm{FGSS}}(d_H(x)) + \varepsilon|V(H)| + O(1)$ in polynomial time, so φ_ε is a computable form.

The branching vectors. The accompanying code checks (W1)–(W4) in exact rational arithmetic, and all 31,003 branching vectors of degree at most 9 in interval arithmetic (mpmath, 120-bit precision); the largest upper end of $\alpha^{-a} + \alpha^{-b}$ is 0.999999551.

The linear program. The code also computes an exact rational solution of (D.8) for $\{\varphi_{\mathrm{FGSS}}\}$, which certifies $L(3, \mathcal{X}_{14}, \{\varphi_{\mathrm{FGSS}}\}) \le L$ with $L \approx 0.371341949$ and $2^L < 1.2935555 < \alpha$. Every vertex counted in Φ_φ has weight at least $m_0 = \min\{v(2), w(1)\}$ in (D.4), so $\Phi_{\varphi_\varepsilon} \le \Phi_{\varphi_{\mathrm{FGSS}}} + \varepsilon/m_0$ on the polytope and $L(3, \mathcal{X}_{14}, \Psi) \le L + \varepsilon/m_0$. A fixed $\varepsilon > 0$ with $2^{L+\varepsilon/m_0} < \alpha$ exists because $2^L < 1.2936 < \alpha$.

The running time. Finally $v(7) + w(7) = 1.5462202$ and $\alpha^{1.5462202} < 1.500312$. □

The pathwidth bounds of this paper do not improve this running time. With the forms $\varphi_{\mathsf{S}}, \ldots, \varphi_3$ listed before Lemma D.5 in place of φ_{FGSS}, the optimum for $\sigma = 3$ is still 1.500312 in floating point, because the dynamic program is not the bottleneck for $\sigma = 3$.

The source of the gap. The case range of the previous recurrences appears to come from the polynomial-space analysis of the same authors [vR11, Theorem 8.5], [NvRvD14, Theorem 11], where set branchings are needed only for elements of frequency at most $d(s) - 1$; re-optimizing the weights with that range reproduces the previous 1.50016^n in floating point. The journal version has the same switching rule, weights and recurrences, so the gap is there too. Two printing slips, which also occur in the thesis [vR11], do not affect the results: constraints 2–4 on the weights [NvRvD14, p. 728] are meant for $i \ge 1$, $i \ge 1$ and $i \ge 2$, and the subsumption rule [NvRvD14, Algorithm 4, line 7] is printed with closed neighborhoods, $N[e_1] \subseteq N[e_2]$, where (R3) uses open ones.

References

[Alo97] Noga Alon. On the edge-expansion of graphs. *Combinatorics, Probability and Computing*, 6(2):145–152, 1997. doi:10.1017/S096354839700299X.

[AT07] Robert J. Adler and Jonathan E. Taylor. *Random Fields and Geometry*. Springer Monographs in Mathematics. Springer, 2007. doi:10.1007/978-0-387-48116-6.

[AvdBZ26] Alimzhan Amanov, Maxim van den Berg, and Jeroen Zuiddam. Multiplying matrices via polynomials: exposition and barriers. arXiv:2610.30054, 2026. URL: https://arxiv.org/abs/2610.30054.

[BBC+18] Cristina Bazgan, Ljiljana Brankovic, Katrin Casel, Henning Fernau, Klaus Jansen, Kim-Manuel Klein, Michael Lampis, Mathieu Liedloff, Jérôme Monnot, and Vangelis Th. Paschos. The many facets of upper domination. *Theoretical Computer Science*, 717:2–25, 2018. doi:10.1016/J.TCS.2017.05.042.

[BCEP13] Nicolas Bourgeois, Federico Della Croce, Bruno Escoffier, and Vangelis Th. Paschos. Fast algorithms for min independent dominating set. *Discrete Applied Mathematics*, 161(4-5):558–572, 2013. doi:10.1016/J.DAM.2012.01.003.

[BCHK22] Narek Bojikian, Vera Chekan, Falko Hegerfeld, and Stefan Kratsch. Tight bounds for connectivity problems parameterized by cutwidth, 2022. Version 1. URL: https://arxiv.org/abs/2212.12385v1.

[BCK25] Narek Bojikian, Vera Chekan, and Stefan Kratsch. Tight bounds for some classical problems parameterized by cutwidth, 2025. Version 1. URL: https://arxiv.org/abs/2502.15884v1.

[BCKN15] Hans L. Bodlaender, Marek Cygan, Stefan Kratsch, and Jesper Nederlof. Deterministic single exponential time algorithms for connectivity problems parameterized by treewidth. *Information and Computation*, 243:86–111, 2015. doi:10.1016/J.IC.2014.12.008.

[BF12] Daniel Binkele-Raible and Henning Fernau. An exact exponential-time algorithm for the directed maximum leaf spanning tree problem. *Journal of Discrete Algorithms*, 15:43–55, 2012. doi:10.1016/J.JDA.2012.03.006.

[BHK09] Andreas Björklund, Thore Husfeldt, and Mikko Koivisto. Set partitioning via inclusion-exclusion. *SIAM Journal on Computing*, 39(2):546–563, 2009. doi:10.1137/070683933.

[BHKK12] Andreas Björklund, Thore Husfeldt, Petteri Kaski, and Mikko Koivisto. The traveling salesman problem in bounded degree graphs. *ACM Transactions on Algorithms*, 8(2):18:1–18:13, 2012. doi:10.1145/2151171.2151181.

[BIJK19a] Édouard Bonnet, Yoichi Iwata, Bart M. P. Jansen, and Łukasz Kowalik. Fine-grained complexity of k-OPT in bounded-degree graphs for solving TSP, 2019. Version 2. URL: https://arxiv.org/abs/1908.09325v2.

[BIJK19b] Édouard Bonnet, Yoichi Iwata, Bart M. P. Jansen, and Łukasz Kowalik. Fine-grained complexity of k-OPT in bounded-degree graphs for solving TSP. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms, ESA 2019, Munich/Garching, Germany, September 9-11, 2019*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICS.ESA.2019.23.

[BKKZ15] Andreas Björklund, Vikram Kamat, Łukasz Kowalik, and Meirav Zehavi. Spotting trees with few leaves, 2015. Version 2. URL: https://arxiv.org/abs/1501.00563v2.

[BKM21] Benjamin Aram Berendsohn, László Kozma, and Dániel Marx. Finding and counting permutations via CSPs. *Algorithmica*, 83(8):2552–2577, 2021. doi:10.1007/S00453-021-00812-Z.

[BKT17] Alexey Balitskiy, Roman Karasev, and Alexander Tsigler. Optimality of codes with respect to error probability in Gaussian noise, 2017. Version 1. URL: https://arxiv.org/abs/1701.07986v1.

[Bod98] Hans L. Bodlaender. A partial k-arboretum of graphs with bounded treewidth. *Theoretical Computer Science*, 209(1-2):1–45, 1998. doi:10.1016/S0304-3975(97)00228-4.

[Bör04] Károly Böröczky, Jr. *Finite Packing and Covering*, volume 154 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2004. doi:10.1017/CBO9780511546587.

[BSV15] Ágnes Backhausz, Balázs Szegedy, and Bálint Virág. Ramanujan graphings and correlation decay in local algorithms. *Random Structures & Algorithms*, 47(3):424–435, 2015. doi:10.1002/RSA.20562.

[CDM17] Radu Curticapean, Holger Dell, and Dániel Marx. Homomorphisms are a good basis for counting small subgraphs. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 210–223. ACM, 2017. doi:10.1145/3055399.3055502.

[CGHV15] Endre Csóka, Balázs Gerencsér, Viktor Harangi, and Bálint Virág. Invariant Gaussian processes and independent sets on regular graphs of large girth. *Random Structures & Algorithms*, 47(2):284–303, 2015. doi:10.1002/RSA.20547.

[CKN12] Marek Cygan, Stefan Kratsch, and Jesper Nederlof. Fast Hamiltonicity checking via bases of perfect matchings, 2012. Version 1. URL: https://arxiv.org/abs/1211.1506v1.

[CKS19] Marek Cygan, Łukasz Kowalik, and Arkadiusz Socała. Improving TSP tours using dynamic programming over tree decompositions. *ACM Transactions on Algorithms*, 15(4):54:1–54:19, 2019. doi:10.1145/3341730.

[Cla26] Claude. A $2^{o(n)}$ time algorithm for 3-SAT. Anthropic preprint, dated 26 October 2026, with a Lean 4 formalization, in the folder eth of the repository, 2026. URL: https://github.com/anthropics/formal-math.

[CNP+11] Marek Cygan, Jesper Nederlof, Marcin Pilipczuk, Michał Pilipczuk, Johan M. M. van Rooij, and Jakub Onufry Wojtaszczyk. Solving connectivity problems parameterized by treewidth in single exponential time, 2011. Version 1. URL: https://arxiv.org/abs/1103.0534v1.

[CP12] Federico Della Croce and Vangelis Th. Paschos. Efficient algorithms for the max k-vertex cover problem. In Jos C. M. Baeten, Thomas Ball, and Frank S. de Boer, editors, *Theoretical Computer Science - 7th IFIP TC 1/WG 2.2 International Conference, TCS 2012, Amsterdam, The Netherlands, September 26-28, 2012. Proceedings*, volume 7604 of *Lecture Notes in Computer Science*, pages 295–309. Springer, 2012. doi:10.1007/978-3-642-33475-7_21.

[DLP22] Louis Dublois, Michael Lampis, and Vangelis Th. Paschos. Upper dominating set: Tight algorithms for pathwidth and sub-exponential approximation. *Theoretical Computer Science*, 923:271–291, 2022. doi:10.1016/J.TCS.2022.05.013.

[DM63] Joaquin B. Diaz and Frederic T. Metcalf. Stronger forms of a class of inequalities of G. Pólya–G. Szegő, and L. V. Kantorovich. *Bulletin of the American Mathematical Society*, 69(3):415–418, 1963. doi:10.1090/S0002-9904-1963-10953-2.

[DMRBC26] Guillermo De Ita Luna, J. Raymundo Marcial-Romero, Pedro Bello López, and Meliza Contreras González. An exact algorithm for counting the number of independent sets of a graph. *Mathematics*, 14(2):275, 2026. doi:10.3390/math14020275.

[DPS02] Josep Díaz, Jordi Petit, and Maria J. Serna. A survey of graph layout problems. *ACM Computing Surveys*, 34(3):313–356, 2002. doi:10.1145/568522.568523.

[Edw16] Keith J. Edwards. A faster polynomial-space algorithm for Max 2-CSP. *Journal of Computer and System Sciences*, 82(3):536–550, 2016. doi:10.1016/J.JCSS.2015.11.013.

[EM15] Keith Edwards and Eric McDermid. A general reduction theorem with applications to pathwidth and the complexity of MAX 2-CSP. *Algorithmica*, 72(4):940–968, 2015. doi:10.1007/s00453-014-9883-7.

[ES73] Paul Erdős and John L. Selfridge. On a combinatorial game. *Journal of Combinatorial Theory, Series A*, 14(3):298–301, 1973. doi:10.1016/0097-3165(73)90005-8.

[EST94] John A. Ellis, Ivan Hal Sudborough, and Jonathan S. Turner. The vertex separation and search number of a graph. *Information and Computation*, 113(1):50–79, 1994. doi:10.1006/INCO.1994.1064.

[Fei00] Uriel Feige. Approximating the bandwidth via volume respecting embeddings. *Journal of Computer and System Sciences*, 60(3):510–539, 2000. doi:10.1006/JCSS.1999.1682.

[FGK05] Fedor V. Fomin, Fabrizio Grandoni, and Dieter Kratsch. Some new techniques in design and analysis of exact (exponential) algorithms. *Bulletin of the EATCS*, 87:47–77, 2005.

[FGK+13] Fedor V. Fomin, Fabrizio Grandoni, Dieter Kratsch, Daniel Lokshtanov, and Saket Saurabh. Computing optimal Steiner trees in polynomial space. *Algorithmica*, 65(3):584–604, 2013. doi:10.1007/s00453-012-9612-z.

[FGSS09] Fedor V. Fomin, Serge Gaspers, Saket Saurabh, and Alexey A. Stepanov. On two techniques of combining branching and treewidth. *Algorithmica*, 54(2):181–207, 2009. doi:10.1007/s00453-007-9133-3.

[FH06] Fedor V. Fomin and Kjartan Høie. Pathwidth of cubic graphs and exact algorithms. *Information Processing Letters*, 97(5):191–196, 2006. doi:10.1016/j.ipl.2005.10.012.

[FIK+08] Fedor V. Fomin, Kazuo Iwama, Dieter Kratsch, Petteri Kaski, Mikko Koivisto, Lukasz Kowalik, Yoshio Okamoto, Johan M. M. van Rooij, and Ryan Williams. 08431 open

problems – moderately exponential time algorithms. In Fedor V. Fomin, Kazuo Iwama, and Dieter Kratsch, editors, *Moderately Exponential Time Algorithms, 19.10.–24.10.2008*, volume 08431 of *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany, 2008. doi:10.4230/DagSemProc.08431.3.

[FK10] Fedor V. Fomin and Dieter Kratsch. *Exact Exponential Algorithms*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2010. doi:10.1007/978-3-642-16533-7.

[FT50] László Fejes Tóth. Extremum properties of the regular polyhedra. *Canadian Journal of Mathematics*, 2:22–31, 1950. doi:10.4153/CJM-1950-003-x.

[Gas10] Serge Gaspers. *Exponential Time Algorithms: Structures, Measures, and Bounds*. VDM Verlag Dr. Müller, 2010.

[Gas26] Serge Gaspers. Path decompositions of sparse graphs via Gaussian sweeps and faster exponential algorithms: Code and formal proofs. Companion website, 2026. doi:10.5281/zenodo.23883489.

[GdHKS16] Robert Ganian, Ronald de Haan, Iyad A. Kanj, and Stefan Szeider. On existential MSO and its relation to ETH. In *41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016*, volume 58 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 42:1–42:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPIcs.MFCS.2016.42.

[GK14] Alexander Golovnev and Konstantin Kutzkov. New exact algorithms for the 2-constraint satisfaction problem. *Theoretical Computer Science*, 526:18–27, 2014. doi:10.1016/J.TCS.2014.01.010.

[GL16] Serge Gaspers and Edward J. Lee. Faster graph coloring in polynomial space, 2016. Version 2. URL: https://arxiv.org/abs/1607.06201v2.

[GL17] Serge Gaspers and Edward J. Lee. Exact algorithms via multivariate subroutines, 2017. Version 1. URL: https://arxiv.org/abs/1704.07982v1.

[GL23] Serge Gaspers and Edward J. Lee. Faster graph coloring in polynomial space. *Algorithmica*, 85(2):584–609, 2023. doi:10.1007/S00453-022-01034-7.

[GLP+24] Kishen N. Gowda, Aditya Lonkar, Fahad Panolan, Vraj Patel, and Saket Saurabh. Improved FPT algorithms for deletion to forest-like structures. *Algorithmica*, 86(5):1657–1699, 2024. doi:10.1007/s00453-023-01206-z.

[GS12] Serge Gaspers and Gregory B. Sorkin. A universally fastest algorithm for Max 2-Sat, Max 2-CSP, and everything in between. *Journal of Computer and System Sciences*, 78(1):305–335, 2012. doi:10.1016/J.JCSS.2011.05.010.

[GS14] Serge Gaspers and Gregory B. Sorkin. Separate, Measure and Conquer: Faster polynomial-space algorithms for Max 2-CSP and counting dominating sets, 2014. Version 2. URL: https://arxiv.org/abs/1404.0753v2.

[GS17] Serge Gaspers and Gregory B. Sorkin. Separate, Measure and Conquer: Faster polynomial-space algorithms for Max 2-CSP and counting dominating sets. *ACM Transactions on Algorithms*, 13(4):44:1–44:36, 2017. doi:10.1145/3111499.

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995. doi:10.1145/227683.227684.

[HV15] Viktor Harangi and Bálint Virág. Independence ratio and random eigenvectors in transitive graphs. *The Annals of Probability*, 43(5):2810–2840, 2015. doi:10.1214/14-AOP952.

[IP01] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2):367–375, 2001. doi:10.1006/JCSS.2000.1727.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001. doi:10.1006/JCSS.2001.1774.

[Joh17] Fredrik Johansson. Arb: Efficient arbitrary-precision midpoint-radius interval arithmetic. *IEEE Transactions on Computers*, 66(8):1281–1292, 2017. doi:10.1109/TC.2017.2690633.

[Kes59] Harry Kesten. Symmetric random walks on groups. *Transactions of the American Mathematical Society*, 92(2):336–354, 1959. doi:10.1090/S0002-9947-1959-0109367-6.

[Kin92] Nancy G. Kinnersley. The vertex separation number of a graph equals its path-width. *Information Processing Letters*, 42(6):345–350, 1992. doi:10.1016/0020-0190(92)90234-M.

[KM20] Łukasz Kowalik and Konrad Majewski. The asymmetric travelling salesman problem in sparse digraphs, 2020. Version 2. URL: https://arxiv.org/abs/2007.12120v2.

[KMRR09] Joachim Kneis, Daniel Mölle, Stefan Richter, and Peter Rossmanith. A bound on the pathwidth of sparse graphs with applications to exact algorithms. *SIAM Journal on Discrete Mathematics*, 23(1):407–427, 2009. doi:10.1137/080715482.

[LM23] Lyuben Lichev and Dieter Mitsche. On the minimum bisection of random 3-regular graphs. *The Electronic Journal of Combinatorics*, 30(2):P2.40, 2023. doi:10.37236/11085.

[LN19] Jason Li and Jesper Nederlof. Detecting feedback vertex sets of size k in $O^\star(2.7^k)$ time. *CoRR*, abs/1906.12298, 2019. URL: http://arxiv.org/abs/1906.12298.

[LRV13] Anand Louis, Prasad Raghavendra, and Santosh S. Vempala. The complexity of approximating vertex expansion. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, Berkeley, CA, USA, October, 26-29, 2013*, pages 360–369. IEEE Computer Society, 2013. doi:10.1109/FOCS.2013.46.

[LS01] W. V. Li and Q.-M. Shao. Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes: Theory and Methods*, volume 19 of *Handbook of Statistics*, pages 533–597. Elsevier, 2001. doi:10.1016/S0169-7161(01)19019-X.

[LX26] Yuxi Liu and Mingyu Xiao. Solving Co-Path/Cycle Packing and Co-Path Packing faster than 3^k. *Information and Computation*, 309:105406, 2026. doi:10.1016/J.IC.2026.105406.

[Lyo17] Russell Lyons. Factors of IID on trees. *Combinatorics, Probability and Computing*, 26(2):285–300, 2017. doi:10.1017/S096354831600033X.

[MP06] Burkhard Monien and Robert Preis. Upper bounds on the bisection width of 3- and 4-regular graphs. *Journal of Discrete Algorithms*, 4(3):475–498, 2006. doi:10.1016/J.JDA.2005.12.009.

[MS89] Fillia Makedon and Ivan Hal Sudborough. On minimizing width in linear layouts. *Discrete Applied Mathematics*, 23(3):243–265, 1989. doi:10.1016/0166-218X(89)90016-4.

[Mul26] Abhijeet Mulgund. Stochastic domination of Gaussian maxima by the regular simplex, 2026. Version 2. URL: https://arxiv.org/abs/2609.28452v2.

[Nil91] A. Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991. doi:10.1016/0012-365X(91)90112-F.

[NvRvD14] Jesper Nederlof, Johan M. M. van Rooij, and Thomas C. van Dijk. Inclusion/Exclusion meets Measure and Conquer. *Algorithmica*, 69(3):685–740, 2014. doi:10.1007/S00453-013-9759-2.

[Ope26a] OpenAI. The Exponential Time Hypothesis is false. OpenAI Math Release preprint, dated 17 October 2026, revised 25 October 2026, in the folder preprints/The-Exponential-Time-Hypothesis-is-False-October-25-2026 of the repository, 2026. URL: https://github.com/openai/math.

[Ope26b] OpenAI. An upper bound of 9/4 for the matrix multiplication exponent. OpenAI Math Release preprint, dated 2 October 2026, in the folder preprints/Matrix-Multiplication-Nine-Fourths-October-2-2026 of the repository, 2026. URL: https://github.com/openai/math.

[Owe56] Donald B. Owen. Tables for computing bivariate normal probabilities. *The Annals of Mathematical Statistics*, 27(4):1075–1090, 1956. doi:10.1214/aoms/1177728074.

[Pas26] Rafael Pass. Matrix multiplication from polynomial packings: A constructive and elementary proof of $\omega \leq 9/4$. arXiv:2610.28627, 2026. URL: https://arxiv.org/abs/2610.28627.

[Pla54] R. L. Plackett. A reduction formula for normal multivariate integrals. *Biometrika*, 41(3-4):351–360, 1954. doi:10.1093/biomet/41.3-4.351.

[PVC10] PVC (Research Infrastructure), UNSW Sydney. Katana. UNSW, Sydney, 2010. doi:10.26190/669X-A286.

[Rag88] Prabhakar Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *Journal of Computer and System Sciences*, 37(2):130–143, 1988. doi:10.1016/0022-0000(88)90003-7.

[She99] W. F. Sheppard. On the application of the theory of error to cases of normal distribution and normal correlation. *Philosophical Transactions of the Royal Society of London. Series A*, 192:101–167, 1899. doi:10.1098/rsta.1899.0003.

[Sle62] David Slepian. The one-sided barrier problem for Gaussian noise. *Bell System Technical Journal*, 41(2):463–501, 1962. doi:10.1002/j.1538-7305.1962.tb02419.x.

[SS07] Alexander D. Scott and Gregory B. Sorkin. Linear-programming design and analysis of fast algorithms for Max 2-CSP. *Discrete Optimization*, 4(3-4):260–287, 2007. doi:10.1016/J.DISOPT.2007.08.001.

[SS09] Alexander D. Scott and Gregory B. Sorkin. Polynomial constraint satisfaction problems, graph bisection, and the Ising partition function. *ACM Transactions on Algorithms*, 5(4):45:1–45:27, 2009. doi:10.1145/1597036.1597049.

[SvdP16] Blair D. Sullivan and Andrew van der Poel. A fast parameterized algorithm for co-path set. In Jiong Guo and Danny Hermelin, editors, *11th International Symposium on Parameterized and Exact Computation, IPEC 2016, Aarhus, Denmark, August 24-26, 2016*, volume 63 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28:1–28:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPICS.IPEC.2016.28.

[VGO+20] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020. doi:10.1038/s41592-019-0686-2.

[vR11] Johan M. M. van Rooij. *Exact Exponential-Time Algorithms for Domination Problems in Graphs*. PhD thesis, Utrecht University, 2011. URL: https://dspace.library.uu.nl/handle/1874/205442.

[vRB12] Johan M. M. van Rooij and Hans L. Bodlaender. Exact algorithms for edge domination. *Algorithmica*, 64(4):535–563, 2012. doi:10.1007/S00453-011-9546-X.

[vRBR09] Johan M. M. van Rooij, Hans L. Bodlaender, and Peter Rossmanith. Dynamic programming on tree decompositions using generalised fast subset convolution. In Amos Fiat and Peter Sanders, editors, *Algorithms - ESA 2009, 17th Annual European Symposium, Copenhagen, Denmark, September 7-9, 2009. Proceedings*, volume 5757 of *Lecture Notes in Computer Science*, pages 566–577. Springer, 2009. doi:10.1007/978-3-642-04128-0_51.

[Wah08] Magnus Wahlström. A tighter bound for counting max-weight solutions to 2SAT instances. In Martin Grohe and Rolf Niedermeier, editors, *Parameterized and Exact Computation, Third International Workshop, IWPEC 2008, Victoria, Canada, May 14-16, 2008. Proceedings*, volume 5018 of *Lecture Notes in Computer Science*, pages 202–213. Springer, 2008. doi:10.1007/978-3-540-79723-4_19.

[Wik26a] Wikipedia contributors. Cubic graph. Wikipedia, The Free Encyclopedia, revision 1346173041, 30 March 2026. URL: https://en.wikipedia.org/w/index.php?oldid=1346173041.

[Wik26b] Wikipedia contributors. List of unsolved problems in mathematics. Wikipedia, The Free Encyclopedia, revision 1382281920, 27 October 2026. URL: https://en.wikipedia.org/w/index.php?oldid=1382281920.

[Wik26c] Wikipedia contributors. Pathwidth. Wikipedia, The Free Encyclopedia, revision 1369812266, 17 August 2026. URL: https://en.wikipedia.org/w/index.php?oldid=1369812266.

[Wil05] Ryan Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theoretical Computer Science*, 348(2-3):357–365, 2005. doi:10.1016/j.tcs.2005.09.023.

[Woe00] Wolfgang Woess. *Random Walks on Infinite Graphs and Groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2000. doi:10.1017/CBO9780511470967.

[YG80] Mihalis Yannakakis and Fanica Gavril. Edge dominating sets in graphs. *SIAM Journal on Applied Mathematics*, 38(3):364–372, 1980. doi:10.1137/0138030.