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

Short version

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 $0.2458n + o(n)$, 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.

This short version states the main results, proves the bound for maximum degree 3, and sketches the other proofs. The full version [34] contains all proofs, the computer-assisted certificates, the algorithms and the derivations of the running times.

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 [31] 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 [56, 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 [28] and of Edwards and McDermid [23] 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 [64], as well as the bound $\frac{m}{5.769}$ of Kneis, Mölle, Richter and Rossmanith [48], 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.

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 $\mathrm{pw}(G) \le c_3\, n + o(n)$.*

(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* (4.2)*.*

Part (a) is Theorem 5.1 in Section 5, and part (b) is Theorem 4.4 in Section 4. 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 [22, Lemma 2.4], cutwidth equals the search number for maximum degree 3 [55, Corollary 3.3], the search number is at most the vertex separation number plus 2 [24, Theorem 2.1], and vertex separation equals pathwidth [47, Theorem 3.1]. For bisection width, Corollary 1.2 improves the bound $\frac{n}{6} + o(n)$ of Monien and Preis [56, Theorem 1] for cubic graphs. The constant c_3 is not new as a bisection constant: Lichev and Mitsche [51] proved that *random* cubic graphs have bisection width at most $c_3 n + o(n)$ asymptotically almost surely, 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 simplest form of our method, which counts the edges that cross a threshold, bounds the cutwidth for larger maximum degree (Theorem 3.1): every graph 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 bound $\frac{2n}{5} + o(n)$ of Monien and Preis [56, Theorem 2] for the bisection width of 4-regular graphs; for $\Delta(G) \le 5$ it is at most $0.5121n + o(n)$.

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: [31] for $\Delta \le 3$; [28] for $\Delta \le 4$ and for degrees 3 to 6; the coefficients $g_{1/6}(d)$ of [23] for $d \ge 7$; and $\frac{13}{75}m$ from [64] ([48] 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	$1-\frac{4}{d+1}$	$\frac{7}{48}$

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*

$$a_d = \frac{3d-4}{32} \quad (3 \le d \le 6), \qquad a_d = 1 - \frac{111}{28\,(d+1)} \quad (7 \le d \le 10), \qquad a_d = 1 - \frac{4}{d+1} \quad (d \ge 11).$$

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

Both are proved in Section 6. 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 $\frac{13}{75} \approx 0.1733$, and for ours it is $d = 6$, with $\frac{2a_6}{6} = \frac{7}{48} \approx 0.1458$. Theorem 1.3 is one of several bounds of this form, none best for all graphs (Theorem 6.4).

1.2 Algorithms and consequences

The proofs of the width bounds are probabilistic. They can be made algorithmic: the Gaussian variables are approximated by a finite rational distribution and the construction is derandomized by the method of conditional expectations (Section 7).

Theorem 1.5 (computing the decompositions)**.** *Let $\beta(G)$ be any of the bounds of Theorems 1.1, 1.3 and 6.2 to 6.4 and Corollary 1.4 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 $n^{1+o(1)} + O(n(n+m))$.*

Many exact algorithms run in time $c^w \,\mathrm{poly}(n)$ on a path decomposition of width w, so the new widths give new running times. Section 8 lists them, one table line per problem and graph class, together with the previous bounds. Among them are Minimum Weight Dominating Set in time $O(1.4815^n)$ for vertex weights with polynomially many distinct subset sums (previously $O(1.5002^n)$), Max $(r,2)$-CSP in time $r^{7m/48+o(m)}$ (previously $r^{13m/75+o(m)}$), counting k-edge subgraphs in time $k^{O(k)}n^{7k/48+o(k)}$, and, on graphs of maximum degree 4, Feedback Vertex Set in time $O(1.2600^n)$ by a Monte Carlo algorithm and the traveling salesperson problem deterministically in time $O(1.4214^n)$. To our knowledge, each bound in the first eight groups of Table 3, the bounds in exponential space that two recent preprints claiming to refute the Exponential Time Hypothesis would not supersede (Section 8), is the fastest known for its problem and graph class (for a deterministic bound, among deterministic algorithms), subject to the table notes.

1.3 Evidence and verification

Theorem 5.1, which is Theorem 1.1(a), and Corollary 1.2 use no computation, and Theorem 5.1 has a complete formal proof in the Lean proof assistant, without hypotheses. Their written proofs use the simplex theorem for three Gaussian variables (Lemma 4.2), which is a case of Mulgund's theorem [57], a preprint whose Lean formalization we checked; for three variables it also has an elementary proof, which the formal proof of Theorem 5.1 uses. This short version contains the complete written proof of Theorem 5.1 (Sections 2 and 5), except for Plackett's identity and the simplex theorem for three variables, which it cites. 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 [12, 27]. They rest in addition on computer-assisted certificates in ball arithmetic and exact rational arithmetic (Propositions 4.3 and 6.1, and two certificates for Theorem 6.3). Theorem 1.1(b), Theorems 1.3, 6.2 and 6.3, Corollary 1.4 and Theorem 6.4 for its tuples other than C also have formal proofs in Lean, conditional on these inputs; for the tuple C, and for Theorem 3.1 with $p = 3$, the formal proof is unconditional. The programs of the computer-assisted steps and the Lean proofs are published on the companion website [35]; the full version [34] states which result needs which input.

1.4 Related work

Width bounds. Fomin, Grandoni and Kratsch [29, §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 [31, §5] and Fomin and Kratsch [33, notes to Chapter 5, p. 100] asked to reduce the gap between the upper bound $\frac{n}{6}$ and a lower bound $0.082n$. Van Rooij stated 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 [68, Open Problem 14.2]; see also the Dagstuhl seminar report [32]. Theorems 1.1 and 1.5 solve this problem 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 [16, 42, 54]. The cut density $\frac{3}{2\pi} \arccos \frac{2\sqrt{2}}{3}$ of a single threshold appears in Lyons [54], and the correlation sequence of our labels on the 3-regular tree is the extremal one of Backhausz, Szegedy and Virág [2] for factor of i.i.d. processes. Lichev and Mitsche [51] 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 4 takes the place of the tree structure. Orderings by a Gaussian projection with bounds on prefix boundaries appear in the approximation algorithms of Feige [26] and of Louis, Raghavendra and Vempala [53], with non-constant-factor losses. We are not aware of 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.

Organization. Section 2 gives the layout lemma, the Gaussian tools, the labels and the concentration step, and Section 3 the simplest sweep, which counts crossing edges. Section 4 introduces the star comparison and the plain sweep, Section 5 the median switch and the complete proof of the cubic bound, and Section 6 the mixed degrees and the bootstrap to all graphs. Section 7 outlines the algorithms, Section 8 lists the algorithmic applications, and Section 9 concludes.

2 Preliminaries

2.1 Layouts from labelings

A *labeling* of a graph 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*

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

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 $S \subseteq V(G)$, and $e(S)$ is the number of edges with exactly one end in S.

Lemma 2.1 (layout)**.** *Let G be a nonempty graph, let $v_1, \dots, v_n$ be an ordering of $V(G)$, 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 $\max_i |\partial^+ T_i| \le \max_i e(T_i)$; in particular $\mathrm{pw}(G) \le \mathrm{cw}(G)$. If y is an injective labeling, then $\mathrm{pw}(G) \le \max_{t\in\mathbb{R}} |\partial^+ S_t|$ and $\mathrm{pw}(G) \le \max_{t\in\mathbb{R}} |\partial^- S_t|$.*

Proof. 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. 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, of width $\max_i |\partial^+ T_i|$, and $|\partial^+ T_i| \le e(T_i)$ because every vertex of $\partial^+ T_i$ is the end of an edge leaving T_i. For an injective labeling, order the vertices by increasing label; then $T_i = S_{y_{v_i}}$. For the bound by inner frontiers, apply the bound by outer frontiers to $-y$: its sublevel sets are the complements $V(G) \setminus A$ of the sets $A = \{v : y_v < -t\}$, each empty or a sublevel set of y, and $\partial^+(V(G) \setminus A) = \partial^- A$. □

The quantity $\max_i |\partial^- T_i|$, minimized over all orderings, is the vertex separation number, which equals the pathwidth by a theorem of Kinnersley [47, Theorem 3.1]; Lemma 2.1 is the easy direction of that theorem. We need labels for which every frontier is small at every threshold. For random labels, $\mathbb{E}|\partial^+ S_t| = \sum_v \mathbb{P}(v \in \partial^+ S_t)$, so everything rests on the probability that one vertex lies on the frontier, together with concentration and a finite grid of thresholds (Section 2.4).

2.2 Gaussian tools

We write φ and Φ for the standard normal density and distribution function, $s_r = \sqrt{1 - r^2}$ for $r \in [-1, 1]$, and $\varphi_2(x, y; \rho)$ for the standard bivariate normal density with correlation $\rho \in (-1, 1)$. A random vector is *centered Gaussian* if every linear combination of its coordinates is a centered normal variable (possibly degenerate); a *correlation matrix* is a positive semidefinite matrix with unit diagonal. Every $k \times k$ correlation matrix C is the Gram matrix of unit vectors $v_1, \dots, v_k$, and $U_i = \langle v_i, \Gamma \rangle$ with $\Gamma \sim N(0, I_k)$ is a centered Gaussian vector with correlation matrix C. If all coordinates of a centered Gaussian vector have positive variance, ties with constants have probability zero, so strict and weak inequalities may be exchanged below.

Lemma 2.2 (two Gaussians)**.** *Let (U, V) be centered Gaussian with unit variances and correlation $\rho > -1$. 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* $t \in \mathbb{R}$.

Proof. If $\rho = 1$, then $U = V$ and both claims are trivial. Let $\theta = \arccos\rho \in (0, \pi)$. (a) Write $U = G_1$ and $V = \cos\theta\, G_1 + \sin\theta\, G_2$ with independent standard normal G_1, G_2, and $(G_1, G_2) = r(\cos\phi, \sin\phi)$. The angle ϕ is uniform, $U > 0$ exactly when ϕ lies in the open half-circle centered at 0, and $V > 0$ exactly when it lies in that half-circle rotated by θ. The two half-circles differ by two arcs of length θ, so $\mathbb{P}(UV \le 0) = 2\theta/(2\pi)$. (b) Let $S = \frac{U+V}{2}$ and $D = \frac{U-V}{2}$. They are uncorrelated, hence independent, and S has a density f that is symmetric and decreasing in $|s|$. The event of (b) implies $(U-t)(V-t) = (S-t)^2 - D^2 \le 0$. An interval of given length has the largest mass under f when it is centered at 0, so, conditioning on D, $\mathbb{P}(|S-t| \le |D|) \le \mathbb{P}(|S| \le |D|) = \mathbb{P}(UV \le 0)$, and (a) finishes the proof. □

Lemma 2.3 (Sheppard's formula [65])**.** *Let (X_1, X_2, X_3) be centered Gaussian with unit variances and put $\vartheta_{ij} = \arccos\mathrm{Corr}(X_i, X_j)$. Then $\mathbb{P}(X_1 > 0, X_2 > 0) = \frac{1}{2} - \frac{\vartheta_{12}}{2\pi}$ 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. The first formula follows from Lemma 2.2(a) and the symmetry $X \mapsto -X$. Three nonzero signs have either 0 or 2 disagreeing pairs, so $\mathbb{P}(\text{not all signs equal}) = \frac{1}{2}\sum_{i<j}\vartheta_{ij}/\pi$ by Lemma 2.2(a), and symmetry halves the probability that all signs are equal. □

Lemma 2.4 (Plackett's identity [61])**.** *Let $k \ge 2$ 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, 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 equals 1 if $k = 2$.

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

$$\text{mean}\ \ h\,\frac{\Sigma_{il} + \Sigma_{jl}}{1 + \Sigma_{ij}} \qquad \text{and variance}\ \ \frac{\det\Sigma}{1 - \Sigma_{ij}^2} > 0. \tag{2.1}$$

Plackett's identity implies Slepian's inequality [66]: 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 given thresholds with at least the probability that all coordinates of the second do; this holds also for singular covariance matrices.

2.3 Gaussian labels with critical decay

A vertex lies on a frontier only if one of its edges crosses the threshold. By Lemma 2.2, two standard normal labels with correlation ρ are separated by a given threshold with probability at most $\arccos(\rho)/\pi$. So the labels should be as correlated as possible across every edge, while each label depends only on a bounded neighborhood. Let $(Z_w)_{w \in V(G)}$ be independent standard normal variables, and let $B_R(v)$ be the ball of radius R around v.

Definition 2.5 (labels)**.** (a) *Padded labels of degree p*, for integers $p \ge 3$, $R \ge 1$ and graphs with $\Delta(G) \le p$: $X_v = \sum_{w \in B_R(v)} (p-1)^{-\operatorname{dist}(v,w)/2} Z_w$ and $Y_v = X_v/\sigma_v$ with $\sigma_v^2 = \operatorname{Var} X_v$. For $p = 3$ the coefficients are $2^{-\operatorname{dist}(v,w)/2}$.

(b) *Mixed labels*, for $R > 0$ and graphs with $\delta(G) \geq 3$: every arc $x \to y$ has length $\frac{1}{2}\ln(d(x)-1)$, $D(v,w)$ is the least length of a walk from v to w, and $X_v = \sum_{w:\, D(v,w)\leq R} e^{-D(v,w)} Z_w$, $Y_v = X_v/\sqrt{\operatorname{Var} X_v}$.

In both cases the decay is *critical*: the arcs leaving a vertex x, except any one of them, carry total squared weight at most 1, that is, $\sum_{y\in N(x)\setminus\{y_0\}} e^{-2\ell(x\to y)} \leq 1$, where $\ell(x\to y)$ is $\frac{1}{2}\ln(p-1)$ for padded and $\frac{1}{2}\ln(d(x)-1)$ for mixed labels. On a tree every sphere then carries the same squared mass, so the squared mass is spread evenly over the R layers of the ball, and the outer shell, where the balls of two neighbors differ, carries only a fraction $O(1/R)$ of the variance. Faster decay would make each label close to Z_v alone; slower decay would let the outermost sphere dominate. For integers $i, j \geq 3$ put

$$r_{ij} = \operatorname{sech}\Big(\tfrac{1}{4}\ln\big((i-1)(j-1)\big)\Big), \qquad r_* = r_{33} = \tfrac{2\sqrt{2}}{3}, \quad r_{44} = \tfrac{\sqrt{3}}{2}, \quad r_{55} = \tfrac{4}{5},$$

so that $r_{pp} = \frac{2\sqrt{p-1}}{p}$.

Lemma 2.6 (edge correlations)**.** (a) *For padded labels of degree p at radius R, every edge uv satisfies* $\operatorname{Corr}(Y_u, Y_v) \geq r_{pp}(1 - \frac{1}{R})$.

(b) *For mixed labels on a graph with $\delta(G) \geq 3$ and $\Delta(G) \leq \Delta_0$, every edge uv satisfies* $\operatorname{Corr}(Y_u, Y_v) \geq r_{d(u)d(v)}(1-\varepsilon_R)$, *where $\varepsilon_R \to 0$ as $R \to \infty$ for fixed Δ_0.*

Proof of (a). Write $q = p-1$. First, if $a_w, b_w > 0$ satisfy $b_w/\alpha \leq a_w \leq \beta b_w$ on a finite set I, then with $A = \sum_I a_w^2$ and $B = \sum_I b_w^2$,

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

the first inequality by summing $(a_w - b_w/\alpha)(\beta b_w - a_w) \geq 0$ (Diaz and Metcalf [21]), the second by the inequality of arithmetic and geometric means; so $\sum_I a_w b_w \geq \operatorname{sech}(\frac{1}{2}\ln(\alpha\beta))\sqrt{AB}$. Let uv be an edge, and apply this to the coefficients $a_w = q^{-\operatorname{dist}(u,w)/2}$ and $b_w = q^{-\operatorname{dist}(v,w)/2}$ on $I = B_R(u) \cap B_R(v)$. Since $\operatorname{dist}(u,w)$ and $\operatorname{dist}(v,w)$ differ by at most one, $\alpha = \beta = \sqrt{q}$ will do, and $\operatorname{sech}(\frac{1}{2}\ln q) = 2\sqrt{q}/(q+1) = r_{pp}$. So $\operatorname{Cov}(X_u, X_v) \geq r_{pp}\sqrt{AB}$. It remains to show $A \geq (1-\frac{1}{R})\sigma_u^2$, and likewise for B. Let s_j be the number of vertices at distance j from u. A vertex at distance $j \geq 1$ has at most q neighbors at distance $j+1$, so $s_{j+1} \leq q s_j$ for $j \geq 1$, and the sphere masses $s_j q^{-j}$ are nonincreasing in $j \geq 1$. Hence 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, which proves the claim. □

No step uses regularity, girth, connectivity or a lower bound on the degrees. For mixed labels the sphere argument is replaced by a cut argument in a shortest-path tree of the directed distance [34]. The constant r_{pp} is best possible for labels whose weights depend only on the distance: on the infinite p-regular tree the adjacency operator has norm $2\sqrt{p-1}$ (Kesten [46]; [70, Lemma 1.24]), so no such *radial* Gaussian labeling has a neighbor correlation above $\frac{2\sqrt{p-1}}{p} = r_{pp}$, and our labels approximate the Gaussian wave, which attains this value in the limit [34].

Lemma 2.7 (distinctness and locality)**.** *For padded labels at radius R, put $\varrho_R = R$; for mixed labels, $\varrho_R = \lfloor 2R/\ln 2 \rfloor$. Then the labels are pairwise distinct with probability* 1, *each Y_v is a function of $(Z_w)_{w\in B_{\varrho_R}(v)}$, and the events $\{v \in \partial^+ S_t\}$ and $\{v \in \partial^- S_t\}$ are determined by $(Z_w)_{w\in B_{\varrho_R+1}(v)}$.*

Proof. Let $u \neq v$. If the coefficients of Z_u and of Z_v in $Y_u - Y_v$ both vanished, then $\sigma_v = \sigma_u e^{-D(v,u)}$ and $\sigma_v = \sigma_u e^{D(u,v)}$, so $D(u,v) + D(v,u) = 0$, which is impossible. So $Y_u - Y_v$ has positive variance and $\mathbb{P}(Y_u = Y_v) = 0$. Every arc has length at least $\frac{1}{2}\ln 2$, which gives the ball $B_{\varrho_R}(v)$, and a frontier event of v depends on the labels of $N[v]$. □

2.4 Reductions and concentration

The *reduction operations* delete a vertex of degree 0 or 1, suppress a vertex of degree 2 whose neighbors are not adjacent (delete it and join its neighbors), or delete a vertex of degree 2 whose neighbors are adjacent. They do not increase any degree, and a maximal sequence of them ends with an empty graph or a graph of minimum degree at least 3.

Lemma 2.8 (undoing reductions [34])**.** *Apply to a nonempty graph F any finite sequence of steps, each a reduction operation or the deletion of a vertex. Let D be the set of deleted vertices and R the final graph. Then* $\mathrm{pw}(F) \le |D| + \mathrm{pw}(R) + 3\log_2 |V(F)| + 1$.

The proof shows that every component C of $(F - D) - V(R)$ has $\mathrm{tw}(F[C]) \le 2$ and attaches to R at a single vertex or at the two ends of an edge of R, and uses that a graph of treewidth k on N vertices has pathwidth at most $(k+1)\lfloor \log_2 N \rfloor$. With no deletions, it shows that a bound $\sum_{d\ge 3} a_d n_d + o(n)$ with nondecreasing coefficients needs to be proved only for graphs of minimum degree at least 3. The next three lemmas turn bounds on expectations at finitely many thresholds into a bound for one realization at all thresholds.

Lemma 2.9 (local variance)**.** *Let $\Delta(G) \le \Delta$ with $\Delta \ge 2$, let (Z_w) be independent, and let $W = \sum_v I_v$, where each I_v takes values in $[0,1]$ and is a function of $(Z_w)_{w \in B_\varrho(v)}$. Then* $\mathrm{Var}\, W \le \frac{1}{2} n \Delta^{2\varrho}$.

Proof. $\mathrm{Cov}(I_u, I_v) = 0$ unless $B_\varrho(u) \cap B_\varrho(v) \ne \emptyset$, that is, $v \in B_{2\varrho}(u)$, and $|\mathrm{Cov}(I_u, I_v)| \le \frac{1}{4}$ always. There are at most $n \max_u |B_{2\varrho}(u)| \le 2n\Delta^{2\varrho}$ such pairs, since $|B_k(u)| \le \sum_{i=0}^{k} \Delta^i \le 2\Delta^k$. □

Lemma 2.10 (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 $v \in \partial^+ S_\tau$, with a neighbor u such that $y_u \le \tau$. (i) If $y_v > \tau_{j+1}$, then $v \in \partial^+ S_{\tau_{j+1}}$ with the same neighbor u; otherwise 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 $u \in \partial^- S_{\tau_j}$; otherwise u is counted by N_j. (ii) Map v to u, whose label is at most τ_0; each u is the image of at most $\Delta(G)$ vertices. Also $\partial^- S_\tau \subseteq S_{\tau_0}$. (iii) $y_v > \tau \ge \tau_K$. □

Lemma 2.11 (concentration of local counts)**.** *Let G have N vertices and $\Delta(G) \le \Delta$, where $\Delta \ge 2$, let (Y_v) be labels as in Definition 2.5, and let $W_1, \dots, W_\mu$ be sums over $v \in V(G)$ of $[0,1]$-valued 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 2.9 and Chebyshev's inequality, $\mathbb{P}(|W_i - \mathbb{E}W_i| \ge \lambda) \le N\Delta^{2\varrho}/(2\lambda^2)$; a union bound over the μ counts and Lemma 2.7 finish the proof. □

3 The cut sweep

The simplest form of the method counts edges instead of frontier vertices. Every prefix of the order by labels is a set S_t, so the cutwidth of that order is the largest number of edges that *cross* a

threshold, that is, have one end labeled at most t and the other above t. For an integer $p \geq 3$ put $\gamma_p = \frac{1}{\pi} \arccos r_{pp}$; then $\gamma_4 = \frac{1}{6}$, $\gamma_5 < 0.20484$, and

$$\tfrac{3}{2}\gamma_3 = c_* := \frac{3}{2\pi} \arcsin \frac{1}{3} = \frac{3}{2\pi} \arccos \frac{2\sqrt{2}}{3} \approx 0.16226 < \frac{1}{6}.$$

The inequality $c_* < \frac{1}{6}$ means $\frac{1}{3} < s$ for $s = \sin \frac{\pi}{9}$; it holds because $h(x) = 3x - 4x^3$ is increasing on $[0, \frac{1}{2}]$, $h(s) = \sin \frac{\pi}{3} = \frac{\sqrt{3}}{2}$, and $h(\frac{1}{3}) = \frac{23}{27} < \frac{\sqrt{3}}{2}$.

Theorem 3.1 (cut sweep)**.** *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* $\mathrm{pw}(G) \leq \mathrm{cw}(G) \leq c_* n + o(n)$ *if* $\Delta(G) \leq 3$, *and* $\mathrm{bw}(G) \leq \mathrm{cw}(G) \leq \frac{m}{6} + o(n) \leq \frac{n}{3} + o(n)$ *if* $\Delta(G) \leq 4$.

Proof sketch. Take the padded labels of degree p at radius R. By Lemma 2.6(a) and Lemma 2.2(b), every edge crosses every threshold with probability at most $\frac{1}{\pi} \arccos(r_{pp}(1 - \frac{1}{R}))$, so by linearity of expectation every threshold is crossed by at most $\gamma_p m + O_p(n/R)$ edges in expectation. Let $\tau_j = -3 + j/K$ for $0 \leq j \leq 6K$. An edge crossing $\tau \in [\tau_j, \tau_{j+1})$ either crosses τ_j or has its lower end among the N_j vertices with labels in $(\tau_j, \tau_{j+1}]$, each the end of at most p edges; for $\tau \notin [-3, 3)$ the crossing edges are bounded by the degree sums over the vertices with $|Y_v| > 3$, whose means are $2m\Phi(-3)$. All these counts are sums of $[0, p]$-valued local terms, so with $K = R = \lfloor \frac{1}{8} \log_{p-1} n \rfloor$, Chebyshev's inequality and a union bound (as in Lemma 2.11) give a realization in which every count is within $n^{7/8}$ of its mean. Ordering the vertices by its labels gives $\mathrm{cw}(G) \leq \gamma_p m + O_p(n/R) + (p+1)n^{7/8}$, and Lemma 2.1 gives $\mathrm{pw}(G) \leq \mathrm{cw}(G)$. For $\Delta(G) \leq 3$, $\gamma_3 m \leq \frac{3}{2}\gamma_3 n = c_* n$; for the bisection width, take the first $\lfloor n/2 \rfloor$ vertices of the ordering. □

The full version gives a complete, self-contained proof of the case $p = 3$, using only Sheppard's formula (the random-hyperplane lemma of Goemans and Williamson [40]) and Chebyshev's inequality. The argument loses in two places. First, the angle: on cubic graphs of large girth no radial labels do better than $\arcsin \frac{1}{3}$, so counting crossing edges cannot give a smaller constant with such labels. Second, the count: a vertex of $\partial^+ S_t$ 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 the next section handles them.

4 The star comparison and the plain sweep

4.1 Stars and the reduction to the center

Fix a vertex v of degree d with neighbors $u_1, \ldots, u_d$. The vector $(Y_v, Y_{u_1}, \ldots, Y_{u_d})$ is a centered Gaussian vector with unit variances whose *incident* correlations $\mathrm{Corr}(Y_v, Y_{u_i})$ are at least known lower bounds r_i; we call it a *d-star with lower bounds r*, with center $Y_0 = Y_v$ and neighbors $Y_i = Y_{u_i}$. 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* $\{Y_0 > \tau,\ \min_i Y_i \leq \tau\}$ occurs. We bound its probability from above in terms of r alone.

Lemma 4.1 (reduction to the center)**.** *Let $(Y_0, \ldots, Y_d)$ be a d-star with lower bounds $r \in (0,1)^d$, and let $P_{\min}(a)$ be the infimum of $\mathbb{P}(U_i > a_i\ \forall i)$ over all centered Gaussian vectors $U = (U_1, \ldots, U_d)$ with unit variances. Then for every $\tau \in \mathbb{R}$,*

$$\mathbb{P}\big(Y_0 > \tau,\ \min_i Y_i \leq \tau\big) \ \leq\ \Phi(-\tau) - \int_\tau^\infty \varphi(z)\, P_{\min}\big(a(z)\big)\, dz, \qquad a_i(z) = \frac{\tau - r_i z}{s_{r_i}}. \tag{4.1}$$

Proof sketch. The event $\{Y_0 > \tau\}$ is the disjoint union of the frontier event and the event that the whole star lies above τ, so it suffices to bound the latter from below. With $\rho_i = \operatorname{Corr}(Y_0, Y_i)$ and fresh independent standard normal ξ_i, put $Y_0' = Y_0$ and $Y_i' = Y_i - (\rho_i - r_i)Y_0 + \sqrt{\rho_i^2 - r_i^2}\,\xi_i$. Then $\operatorname{Var} Y_i' = 1$, $\operatorname{Corr}(Y_0', Y_i') = r_i$, and $\operatorname{Cov}(Y_i', Y_j') = \operatorname{Cov}(Y_i, Y_j) - \rho_i\rho_j + r_i r_j \le \operatorname{Cov}(Y_i, Y_j)$, so by Slepian's inequality the new star lies above τ with at most the probability of the old one. In the new star $Y_i' = r_i Y_0 + s_{r_i} U_i$ with U independent of Y_0, and given $Y_0 = z$ the neighbor Y_i' lies above τ exactly when $U_i > a_i(z)$. Integrating over $z > \tau$ and taking the least favorable correlation matrix of U for each z gives (4.1). □

For equal thresholds the infimum $P_{\min}$ 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. Write $S_d(a) = \mathbb{P}(U_i^* > a\ \forall i)$.

Lemma 4.2 (simplex theorem for $d \le 5$)**.** *Let $d \in \{3, 4, 5\}$ and let U be a centered Gaussian vector in $\mathbb{R}^d$ with unit variances, possibly with a singular correlation matrix. Then $\mathbb{P}(U_i > a\ \forall i) \ge S_d(a)$ for every $a \in \mathbb{R}$.*

This is a case of a recent theorem of Mulgund [57, Theorem 1.1], which proves a conjecture of Balitskiy, Karasev and Tsigler [3]. For $d = 3$ it also has an elementary proof: condition on one coordinate, apply the Fréchet bound $\mathbb{P}(A \cap B) \ge \mathbb{P}A + \mathbb{P}B - 1$ to the other two, note that the bound is attained by a planar configuration with the same conditional marginals, and compare planar configurations by convexity. For $d = 4$ it also follows from the four-cap theorem of L. Fejes Tóth [12, 27].

4.2 Ideal stars and the plain sweep

Suppose that all lower bounds are equal to some r, as on a d-regular graph. Then all thresholds $a_i(z)$ equal $(\tau - rz)/s_r$, and by the simplex theorem the bound of Lemma 4.1 becomes an explicit integral. It is the frontier probability of an actual star, the *ideal d-star*: $Y_0^* = Z$ and $Y_i^* = rZ + s_r U_i^*$, with Z standard normal and independent of U^*. Its incident correlations are exactly r, its neighbors sum to drZ, and any two neighbors have correlation $(dr^2 - 1)/(d - 1)$, the smallest common value possible. Given $Z = z$, all neighbors lie above τ with probability exactly $S_d((\tau - rz)/s_r)$, so its frontier probability, the *ideal star function*, is

$$f_{d,r}(\tau) = \mathbb{P}(Y_0^* > \tau,\ \min_i Y_i^* \le \tau) = \Phi(-\tau) - \int_\tau^\infty \varphi(z)\, S_d\Big(\frac{\tau - rz}{s_r}\Big) dz. \tag{4.2}$$

By Lemmas 4.1 and 4.2, for $3 \le d \le 5$ every star with at most d neighbors whose lower bounds are all at least r has frontier probability at most $f_{d,r}(\tau)$ at every threshold. Since the ideal star is itself such a star, 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.

Proposition 4.3 (enclosures of F_3 and F_4)**.** *The suprema $F_3 = \sup_\tau f_{3,r_{33}}(\tau)$ and $F_4 = \sup_\tau f_{4,r_{44}}(\tau)$ 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 [45], from one-dimensional integral formulas for the ideal star functions [34]. The first supremum is attained near $\tau \approx -0.068$, the second near $\tau \approx -0.164$.

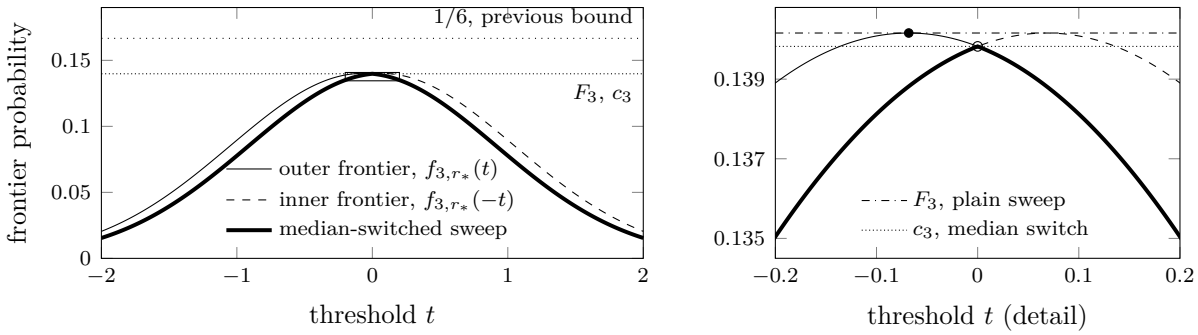


Figure 1: The ideal 3-star function f_{3,r_*} at the cubic correlation $r_* = \frac{2\sqrt{2}}{3}$ and its reflection; the right panel enlarges the box of the left panel, and the dotted line at $1/6$ is the previous bound for maximum degree 3 [31]. The plain sweep has width at most $F_3 n + o(n) \approx 0.1402n$, 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 $c_3 n + o(n)$, the maximum of the thick envelope, attained at $t = 0$ (open dot): by the switch lemma (Lemma 5.7), the passage from inner to outer frontiers at $t = 0$ adds nothing at leading order.

Theorem 4.4 (plain sweep)**.** *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* $\mathrm{pw}(G) \le F_3 n + o(n)$*.*

Proof sketch. Take padded labels of degree 4 at radius R. By Lemma 2.6, every star has lower bounds $r_{44}(1 - \frac{1}{R})$, so every vertex lies in $\partial^+ S_\tau$ with probability at most $f_{4,r_{44}(1-1/R)}(\tau)$, which is at most $F_4 + o_R(1)$ uniformly in τ by an explicit continuity estimate of the center bound in the lower bounds [34]. By linearity of expectation $\mathbb{E}|\partial^+ S_\tau| \le (F_4 + o_R(1))n$ at every threshold. With the grid $\tau_j = -3 + j/K$, Lemmas 2.10 and 2.11 give a realization in which every outer frontier has at most $(F_4 + o_R(1) + \varphi(0)/K)n + 4n^{7/8}$ vertices (outside $[-3, 3]$ there are at most $4\Phi(-3)n + 4n^{7/8}$), and Lemma 2.1 finishes the proof as $R, K \to \infty$ slowly with n. The same argument with padded labels of degree 3 gives F_3. □

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

5 The median switch: maximum degree 3

This section proves Theorem 1.1(a). The proof is complete here, except for the three-variable case of the simplex theorem (Lemma 4.2) and Plackett's identity (Lemma 2.4), which we cite.

Theorem 5.1 (maximum degree 3)**.** *Every graph* G *with* $\Delta(G) \le 3$ *satisfies* $\mathrm{pw}(G) \le c_3\, n + o(n)$*.*

Compared with the plain sweep, only the layout changes: vertices still enter in increasing order of their 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*. The change is suggested by the shape of $f_{3,r}$ (Fig. 1): it exceeds its value at the median only below the median.

5.1 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 $r_* = \frac{2\sqrt{2}}{3}$, and put

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

For $r \in (0,1)$ we have $b(r) \in (-\frac{1}{2}, 1)$, so κ is well defined and continuous on $(0,1)$, and $b(r_*) = \frac{5}{6}$, so $\kappa(r_*) = c_3$.

Labels. Fix an integer $R \ge 2$ and let (Y_v) be the padded labels of degree 3 at radius R, and $r_R = r_*(1 - \frac{1}{R}) \in (0, r_*)$. By Lemmas 2.6 and 2.7:

(L1) $(Y_v)_v$ is centered Gaussian with unit variances, and every edge uv satisfies $\mathrm{Corr}(Y_u, Y_v) \ge r_R$;

(L2) with probability 1, the labels are pairwise distinct;

(L3) Y_v is a function of $(Z_w)_{w \in B_R(v)}$.

Since $-Y$ is again centered Gaussian with the same covariance, $-Y$ has the same distribution as Y; we call this *sign symmetry*.

Threshold zero. For a labeling y, 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 $y_v \le 0 < \min_{x \in N(v)} y_x$ or $y_v > 0 \ge \max_{x \in N(v)} y_x$. Then $D \subseteq U \cup W$. Put $U' = U \setminus D$, $W' = W \setminus D$ and

$$L_0 = |U| + |D \cap W| = |U'| + |D|, \qquad L_0' = |W| + |D \cap U| = |W'| + |D|. \tag{5.1}$$

5.2 The median-switched layout

The connector at zero needs an ordering lemma. 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 5.2 (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 $q_1, \dots, q_k$ of Q and $0 \le j \le k$, let $c(j)$ be the number of items p with $e(p) \subseteq \{q_1, \dots, q_j\}$. Then some ordering satisfies*

$$j - c(j) \le \max(0,\, k - |P|) + 1 \qquad \textit{for all } 0 \le j \le k. \tag{5.2}$$

Proof. We use induction on k. If $k = 0$, then $P = \emptyset$ and (5.2) 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 $\mathrm{ex}(A) = |A| - |P_A|$; thus $\mathrm{ex}(Q) = k - |P|$.

Case 1: some nonempty proper subset Y of Q is closed. Then $Q \setminus Y$ is closed as well, and 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 and B the other, so that $\mathrm{ex}(Q) = \mathrm{ex}(A) + \mathrm{ex}(B) \ge 2\,\mathrm{ex}(A)$. By induction, (A, P_A) and (B, P_B) have orderings satisfying (5.2) with $\mathrm{ex}(A)$, respectively $\mathrm{ex}(B)$, in place of $k - |P|$; order Q as A followed by B. For $j \le |A|$, $c(j) = c_A(j)$, so $j - c(j) \le \max(0, \mathrm{ex}(A)) + 1$, which is at most 1 if $\mathrm{ex}(A) \le 0$,

and otherwise at most $\mathrm{ex}(Q)+1$ because $\mathrm{ex}(Q) \geq 2\,\mathrm{ex}(A)$. For $j = |A| + j'$ with $1 \leq j' \leq |B|$, $c(j) = |P_A| + c_B(j')$ and $j - c(j) \leq \mathrm{ex}(A) + \max(0, \mathrm{ex}(B)) + 1$. If $\mathrm{ex}(B) \geq 0$, this is $\mathrm{ex}(Q)+1$; if $\mathrm{ex}(B) < 0$, then $\mathrm{ex}(A) \leq \mathrm{ex}(B) < 0$ and it is less than 1.

Case 2: no nonempty proper subset of Q is closed. Choose q_1 arbitrarily. Given $X = \{q_1, \dots, q_j\}$ with $1 \leq j < k$, the set X is not closed, so some item p has $e(p) = \{x, y\}$ with $x \in X$ and $y \notin X$; put $q_{j+1} = y$. Then $c(j+1) \geq c(j)+1$, so $c(j) \geq j-1$ for every $j \geq 1$, and $j - c(j) \leq 1$ for all j. □

Lemma 5.3 (median-switched layout)**.** *Let G be a graph with $\Delta(G) \leq 3$, let $y\colon V(G) \to \mathbb{R}$ be injective, and let $w \geq 0$ be a real number such that*

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

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

(iii) $\max(L_0, L_0') + 1 \leq w$*, with* L_0, L_0' *as in* (5.1)*.*

Then $\mathrm{pw}(G) \leq w$*.*

Proof. Let $v_1, \dots, v_n$ be the vertices in increasing order of y, $\ell = |S_0|$, and $T_i = \{v_1, \dots, v_i\}$. A vertex $u \in U'$ is not in D, so either $d(u) \leq 2$ or u has a neighbor at or below zero; in both cases u has at most two neighbors above zero. Let P be the set of $u \in U'$ for which $e(u) := N(u) \cap W'$ is nonempty, so that $|e(u)| \leq 2$, and let $q_1, \dots, q_k$ be an ordering of W' given by Lemma 5.2 for $Q = W'$ and the items P. The *inner bags* are $A_i = \{v_i\} \cup \partial^- T_{i-1}$ for $1 \leq i \leq \ell$, the *outer bags* are $A_i = \{v_i\} \cup \partial^+ T_i$ for $\ell < i \leq n$, and the *connector bags* are

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

in the order $A_1, \dots, A_\ell, C_{-1}, C_0, \dots, C_k, A_{\ell+1}, \dots, A_n$.

Sizes. For $i \leq \ell$, $T_{i-1} = S_t$ for some $t < 0$, so $|A_i| \leq w+1$ by (i); for $i > \ell$, $T_i = S_t$ for some $t > 0$, so $|A_i| \leq w+1$ by (ii). Next $|C_{-1}| = L_0$ and $|C_0| \leq |D| + |U'| = L_0$. For $1 \leq j \leq k$, the number of items in C_j is $|P| - c(j-1)$, so by (5.2) at $j-1$, $|C_j| \leq |D| + |P| + \max(0, k - |P|) + 2 = |D| + \max(|P|, |W'|) + 2 \leq \max(L_0, L_0') + 2 \leq w + 1$.

Intervals. A vertex of D lies in all connector bags, q_j in $C_j, \dots, C_k$, an item $u \in P$ in $C_{-1}, \dots, C_J$ for the largest position J of an element of $e(u)$, and every other vertex of U in C_{-1} only. Let $u = v_a$ with $a \leq \ell$, and let b be the largest index of a neighbor of u. The inner bags containing u are $A_a, \dots, A_{\min(\max(a,b),\ell)}$, and no outer bag contains u. If $u \notin U$ then $b \leq \ell$ and u is in no connector bag; if $u \in U$ its bags are $A_a, \dots, A_\ell$ followed by a prefix of the connector bags. Symmetrically, a vertex $x = v_c$ with $c > \ell$ lies in no inner bag, in consecutive outer bags ending with A_c and starting with $A_{\ell+1}$ if $x \in W$, and, if $x \in W$, in a suffix of the connector bags.

Edges. Let $v_a v_b$ be an edge with $a < b$. If $b \leq \ell$, both ends lie in A_b; if $a > \ell$, 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 lie in C_k. Otherwise $u \in P$ and $x = q_j \in e(u)$ for some j, so $u, x \in C_j$. Hence the bags form a path decomposition of width at most w. □

5.3 The star comparison at nonnegative thresholds

The ideal 3-star at correlation $r \in (0,1)$ is Y_0^* standard normal and $Y_i^* = rY_0^* + s_r U_i^*$, where U^* is the regular simplex vector in three variables, realized as the projections of a standard planar Gaussian vector onto three unit vectors at mutual angles $\frac{2\pi}{3}$. Its neighbor pairs have correlation $b(r)$, ties have probability zero, and

$$Y_1^* + Y_2^* + Y_3^* = 3rY_0^*. \tag{5.3}$$

Lemma 5.4. *For every $r \in (0,1)$, $f_{3,r}(0) = \kappa(r)$.*

Proof. If $Y_i^* > 0$ for all i, then $Y_0^* > 0$ by (5.3). Hence $f_{3,r}(0) = \mathbb{P}(Y_0^* \geq 0) - \mathbb{P}(\min_i Y_i^* > 0) = \frac{1}{2} - \big(\frac{1}{2} - \frac{3\arccos b(r)}{4\pi}\big) = \kappa(r)$ by Lemma 2.3. □

Lemma 5.5. *For every $r \in (0,1)$ and every $t \geq 0$, $f_{3,r}(t) \leq f_{3,r}(0)$.*

Proof. Let $M = \min_i Y_i^*$. If $M > t \geq 0$, then $Y_0^* > t/r \geq t$ by (5.3), so $f_{3,r}(t) = \mathbb{P}(Y_0^* \geq t) - \mathbb{P}(M > t)$ for all $t \geq 0$, and $f_{3,r}(0) - f_{3,r}(t) = \mathbb{P}(0 \leq Y_0^* < t) - \mathbb{P}(0 < M \leq t)$. The event $\{0 < M \leq t\}$ is contained in the union of the events $A_k = \{0 < Y_k^* \leq t,\ Y_k^* \leq Y_j^*\ \forall j\}$. Fix k, write $b = b(r)$, and put $R_j = Y_j^* - bY_k^*$ for $j \neq k$. The pair $(R_j)_{j\neq k}$ is uncorrelated with, hence independent of, Y_k^*, and on A_k both $R_j \geq (1-b)Y_k^* > 0$. The two residuals have correlation $\frac{b-b^2}{1-b^2} = \frac{b}{1+b} < \frac{1}{2}$, so by Lemma 2.3 both are positive with probability less than $\frac{1}{3}$. Hence $\mathbb{P}(A_k) \leq \frac{1}{3}\mathbb{P}(0 < Y_0^* \leq t)$, and summing over k gives the claim. □

Proposition 5.6 (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 5.4 and 5.5. Put $a_i = \mathrm{Corr}(Y_0, Y_i) \geq r$. Indices with $a_i = 1$ contribute only null events and are discarded; if none remains, the probability is 0. For the remaining i put $\eta_i = (Y_i - a_i Y_0)/s_{a_i}$, which are standard normal and independent of Y_0; repeat one of them if fewer than three remain. On $\{Y_0 \geq t\}$ the function $g(a) = (t - aY_0)/\sqrt{1-a^2}$ satisfies $g'(a) = (at - Y_0)(1-a^2)^{-3/2} \leq 0$ on $[r,1)$, because $at \leq t \leq Y_0$. Hence $Y_i \leq t$, that is, $\eta_i \leq g(a_i)$, implies $\eta_i \leq g(r) = (t - rY_0)/s_r$. By independence and Lemma 4.2 with $d = 3$,

$$\mathbb{P}(Y_0 \geq t,\ \exists i : Y_i \leq t) \leq \int_t^\infty \varphi(z)\Big(1 - S_3\Big(\frac{t - rz}{s_r}\Big)\Big)dz = f_{3,r}(t). \qquad \square$$

For $t < 0$ the monotonicity of g fails on part of the event $\{Y_0 \geq t\}$; the median-switched layout uses outer frontiers only at $t > 0$ and inner frontiers, which by sign symmetry are outer frontiers of $-Y$, only at $t < 0$.

5.4 Level zero: the switch lemma

For a vertex v of degree d with neighbors $x_1, \ldots, x_d$, the events $v \in U$ and $v \in D \cap W$ are $\{Y_v \leq 0 < \max_i Y_{x_i}\}$ and $\{d = 3,\ Y_v > 0 \geq \max_i Y_{x_i}\}$, so the following lemma bounds the contribution of v to $\mathbb{E}L_0$ by the same $\kappa(r)$ as at nonzero thresholds.

Lemma 5.7 (switch lemma)**.** *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*

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

At the ideal star the second event is impossible, by (5.3), 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, it remains to show that the probability that at least two of three standard normal residuals exceed a level $h > 0$ is smallest for U^*. This is not an orthant event, so the simplex theorem does not apply; we use a deformation monitored by Plackett's identity.

Lemma 5.8 (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 $\det C = \big(\cos\alpha - \cos(\beta+\gamma)\big)\big(\cos(\beta-\gamma) - \cos\alpha\big)$, and 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 of both sides. 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 $[0,\pi]$, both factors are nonnegative, and both are positive exactly when the four inequalities are strict. The 2×2 principal minors are $\sin^2 \ge 0$, so C is positive semidefinite, and positive definite exactly when $\det C > 0$. Conversely, a positive semidefinite C is the Gram matrix of unit vectors v_1, v_2, v_3 with these angles; the triangle inequality for angles gives the first three inequalities, and applied to $v_1, -v_3, v_2$ it gives $\gamma \le (\pi-\beta)+(\pi-\alpha)$. □

For a correlation matrix C and $h > 0$ let $\Pi_C(h) = \mathbb{P}_C(\text{at least two of } \eta_1, \eta_2, \eta_3 \text{ exceed } h)$ for $\eta \sim N(0, C)$, 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, by weak convergence and because the boundary of the event is null.

Lemma 5.9 (two exceedances)**.** *For every 3×3 correlation matrix C and every $h > 0$, $\Pi_C(h) \ge \Pi_{C_{\mathrm{eq}}}(h)$.*

Proof. By continuity and $(1-\xi)C + \xi I \to C$, we may assume that C is positive definite.

The water-filling path. Write $C_{ij} = \cos\theta_{ij}$ with $\theta_{ij} \in [0,\pi]$; by Lemma 5.8 the four inequalities hold strictly. For $s \ge 0$ let $\theta_{ij}(s) = \max(\theta_{ij}, s)$ and $\Sigma(s) = \sum_{i<j}\theta_{ij}(s)$, which is continuous and nondecreasing with $\Sigma(0) < 2\pi \le \Sigma(\frac{2\pi}{3})$. Let $s^* \in (0, \frac{2\pi}{3}]$ be the smallest s with $\Sigma(s) = 2\pi$, and $C(s)$ the matrix with entries $\cos\theta_{ij}(s)$. For $s < s^*$, $C(s)$ is positive definite: $\Sigma(s) < 2\pi$, and a triangle inequality $\theta_{ij}(s) < \theta_{ik}(s) + \theta_{jk}(s)$ follows from the original one if $\theta_{ij}(s) = \theta_{ij}$, and otherwise from $s < 2s$. By continuity $C(s^*)$ is positive semidefinite with angle sum 2π.

Monotonicity. On each interval between the breakpoints θ_{ij} the set of *moving* pairs, those with $\theta_{ij} < s$, is fixed, and they have $C_{ij}(s) = \cos s$. By inclusion and exclusion, $\Pi_C(h) = P_{12}+P_{13}+P_{23}-2P_{123}$, with $P_{ij} = \mathbb{P}(\eta_i > h, \eta_j > h)$ and $P_{123} = \mathbb{P}(\eta_1,\eta_2,\eta_3 > h)$. Since P_{ik}, P_{jk} do not depend on C_{ij}, Lemma 2.4 gives, for each pair $\{i,j\}$ with third index k,

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

For a moving pair, $C_{ik}(s), C_{jk}(s) \le C_{ij}(s) = \cos s$ and $\cos s \ge -\frac{1}{2}$, so by (2.1) the conditional mean is at most $h \cdot 2C_{ij}(s)/(1+C_{ij}(s)) \le h$, and the conditional probability is at most $\frac{1}{2}$. So every such partial derivative is nonnegative, and $\frac{d}{ds}\Pi_{C(s)}(h) = -\sin s \sum_{\text{moving}} \partial\Pi/\partial C_{ij} \le 0$. By continuity, $\Pi_C(h) \ge \Pi_{C(s^*)}(h)$.

The planar endpoint. With angle sum 2π, $\det C(s^*) = 0$, so $C(s^*)$ is the Gram matrix of three unit vectors in $\mathbb{R}^2$. The three circular gaps g_k between consecutive directions sum to 2π, and since the angles $\theta_{ij}(s^*)$ also sum to 2π and $\theta_{ij} \le g_k$ for the gap g_k between v_i and v_j not containing v_k, equality holds and every $g_k \le \pi$. Write the planar standard Gaussian vector as $\varrho(\cos\omega, \sin\omega)$ with ω uniform. Given $\varrho > h$, the projection on v_i exceeds h exactly when ω is within $\lambda = \arccos(h/\varrho) < \frac{\pi}{2}$ of the direction of v_i. No point lies in all three arcs, since the complementary gap would exceed π, and two arcs separated by the gap g_k overlap in length $(2\lambda - g_k)_+$. Therefore

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

by Jensen's inequality for the convex function $g \mapsto (2\lambda - g)_+$, since the configuration of U^* has three equal gaps $\frac{2\pi}{3}$. □

Proof of Lemma 5.7. Reduction to $d = 3$. If $d < 3$, add coordinates $Y_j = rY_0 + s_r\eta_j$ with fresh independent standard normal η_j; the first event can only grow, and the second term is nonnegative.

The functional. Let $d = 3$ and $N = \#\{i : Y_i > 0\}$. By sign symmetry, $\mathbb{P}(Y_0 > 0,\ Y_i \le 0\ \forall i) = \mathbb{P}(Y_0 \le 0,\ N = 3)$. Since $\lceil N/2 \rceil = \mathbf{1}\{N \ge 1\} + \mathbf{1}\{N = 3\}$, the left side of the lemma equals $\Psi = \mathbb{E}[\mathbf{1}\{Y_0 \le 0\}\lceil N/2 \rceil]$.

Coupling. Put $a_i = \mathrm{Corr}(Y_0, Y_i) \ge r$ and $s_i = \sqrt{1 - a_i^2}$; for $s_i > 0$ let $\eta_i = (Y_i - a_iY_0)/s_i$, and for $s_i = 0$ an auxiliary independent standard normal variable. Then η is a centered Gaussian vector with unit variances, independent of Y_0; let C be its correlation matrix. On $\{Y_0 < 0\}$, if $s_i = 0$ then $Y_i = Y_0 < 0$, and if $s_i > 0$ then $Y_i > 0$ if and only if $\eta_i > |Y_0|a_i/s_i \ge |Y_0|r/s_r$, as $a \mapsto a/\sqrt{1 - a^2}$ is increasing. Hence $N \le N_h := \#\{i : \eta_i > h\}$ with $h = h(Y_0) = r|Y_0|/s_r$. Since $\lceil n/2 \rceil$ is nondecreasing and equals $n - \mathbf{1}\{n \ge 2\}$ on $\{0, 1, 2, 3\}$, independence gives

$$\Psi \le \int_{-\infty}^{0} \varphi(z)\Big(3\Phi(-h(z)) - \Pi_C(h(z))\Big)dz \le \int_{-\infty}^{0} \varphi(z)\Big(3\Phi(-h(z)) - \Pi_{C_{\mathrm{eq}}}(h(z))\Big)dz,$$

the second step by Lemma 5.9.

The ideal value. For the ideal star the coupling is an identity: its residuals are U^* and its incident correlations are r. So the right side equals Ψ for the ideal star, which is $\mathbb{P}(Y_0^* \le 0 < \max_i Y_i^*) + \mathbb{P}(Y_0^* \le 0 < \min_i Y_i^*)$. The second probability is 0 by (5.3), and by sign symmetry the first equals $\mathbb{P}(Y_0^* \ge 0,\ \min_i Y_i^* \le 0) = f_{3,r}(0) = \kappa(r)$ by Lemma 5.4. □

5.5 Concentration and the proof of Theorem 5.1

Proof of Theorem 5.1. It suffices to show that 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 $\mathrm{pw}(G) \le (c_3 + \varepsilon)n$: then the maximum $\rho(n)$ of $\max(0, \mathrm{pw}(G) - c_3 n)$ over such graphs on n vertices is $o(n)$.

Fix $\varepsilon > 0$. Choose an integer $R \ge 2$ with $\kappa(r_R) \le c_3 + \frac{\varepsilon}{3}$, an integer $K_0 \ge 1$ with $\varphi(0)/K_0 \le \frac{\varepsilon}{3}$, and the grid $\tau_j = -2 + j/K_0$ for $0 \le j \le 4K_0$, which contains $\tau_{2K_0} = 0$. Consider the following $8K_0 + 4$ counts, each a sum over v of indicators determined, by (L3), by $(Z_w)_{w \in B_{R+1}(v)}$: $|\partial^- S_{\tau_j}|$ for $j < 2K_0$; $|\partial^+ S_{\tau_j}|$ for $2K_0 < j \le 4K_0$; L_0 and L_0' (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); the window counts $M_j = \#\{v : \tau_j < Y_v \le \tau_{j+1}\}$; and $\#\{v : Y_v \le -2\}$ and $\#\{v : Y_v > 2\}$.

Means. Fix v with neighbors $x_1, \dots, x_d$. By (L1), $(Y_v, Y_{x_1}, \dots, Y_{x_d})$ and its negative satisfy the hypotheses of Proposition 5.6 and Lemma 5.7 with $r = r_R$. The event $v \in \partial^+ S_\tau$ for $\tau > 0$ is contained in the event of Proposition 5.6 at τ, and the event $v \in \partial^- S_\tau$ for $\tau < 0$ in that event for the negated vector at $-\tau$. Up to null events, the indicators of L_0 and L_0' are the events of Lemma 5.7 for the vector and for its negative. So 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 the two tail counts have mean $\Phi(-2)n < 0.023n$.

Concentration. Choose $N(\varepsilon)$ so that for $n \ge N(\varepsilon)$, $(8K_0 + 4)\, 9^{R+1} < 2n^{1/2}$, $2n^{3/4} \le \frac{\varepsilon}{3}n$, $n^{3/4} + 1 \le \frac{\varepsilon}{3}n$ and $0.023n + n^{3/4} \le c_3 n$. The first condition is that of Lemma 2.11 with $\Delta = 3$, $\varrho = R + 1$ and $\lambda = n^{3/4}$, so some realization has pairwise distinct labels and every count within $n^{3/4}$ of its mean; fix it.

All thresholds. Let $w = (c_3 + \varepsilon)n$. For $t \in [-2, 0)$ with $t \in [\tau_j, \tau_{j+1})$, Lemma 2.10(i) gives $|\partial^- S_t| \le |\partial^- S_{\tau_j}| + M_j \le \kappa(r_R)n + \frac{\varphi(0)}{K_0}n + 2n^{3/4} \le w$. For $t \in (0, 2)$ the same computation with

the outer count at $\tau_{j+1} > 0$ gives $|\partial^+ S_t| \le w$. For $t < -2$ and $t \ge 2$, Lemma 2.10(ii),(iii) give inner, respectively outer, frontiers of size at most $0.023n + n^{3/4} \le c_3 n \le w$. Finally $\max(L_0, L_0') + 1 \le \kappa(r_R)n + n^{3/4} + 1 \le w$. So Lemma 5.3 gives $\mathrm{pw}(G) \le w$. □

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 9).

6 Mixed degrees and all graphs

6.1 Degrees 3 to 5: the charging certificate

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 center bound of Lemma 4.1 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 it $C_T(\tau)$. There are 46 types for degrees 3 to 5. 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 $j > i$. These transfers sum to zero over the graph, since each edge contributes $-t_{ij}$ at one end and $+t_{ij}$ at the other. Coefficients (a_3, a_4, a_5) are *admissible* if for every $\tau \in [-3, 3]$ some transfer vector (t_{34}, t_{35}, t_{45}) brings $C_T(\tau)$ plus the transfers of every type below the coefficient of its degree; then $\mathbb{E}|\partial^+ S_\tau| \le \sum_d a_d n_d + o(n)$ at every threshold. (Beyond $|\tau| = 3$ the frontier probabilities are below $5\Phi(-3) < 0.007$.)

Proposition 6.1 (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.

The proof is computer-assisted. The interval $[-3, 3]$ of thresholds is partitioned into cells; each cell has one transfer vector and, for every type, a rational lower bound on a Riemann sum of the integral in $C_T(\tau)$, in which $P_{\min}$ is replaced by an explicit lower bound: a truncation of the simplex bound or, for the seven types of degree 3 with unequal lower bounds, the larger of this and an arc bound. All values are enclosures in Arb ball arithmetic [45], rounded outward, and a separate program rechecks the final inequalities in exact rational arithmetic with positive slack [34]. With concentration as in Theorem 4.4, and the reductions of Section 2.4 for vertices of degree at most 2, this gives:

Theorem 6.2 (maximum degree 4 and 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* $\mathsf{P}, \mathsf{R}, \mathsf{S}$*. In particular, every graph with $\Delta(G) \le 4$ satisfies* $\mathrm{pw}(G) \le F_3 n_3 + 0.2575\, n_4 + o(n)$ *and* $\mathrm{pw}(G) \le \frac{5}{32} n_3 + \frac{1}{4} n_4 + o(n)$*.*

The triple P is best when few vertices have degree 4 or 5, and S when many do. The median switch also combines with mixed labels:

Theorem 6.3 (switched sweep, maximum degree 4)**.** *Every graph G with $\Delta(G) \le 4$ satisfies* $\mathrm{pw}(G) \le c_3\, n_3 + 0.2627\, n_4 + o(n)$*.*

The proof keeps the layout of Lemma 5.3, with the set D replaced by the vertices with at least three neighbors on the other side of zero, and uses mixed labels with the charging above. At threshold zero, stars with unequal correlation bounds need a switch lemma of their own, proved with a computer-assisted certificate, and each vertex of degree 3 sends the same amount $s \approx 0.01308$ to each neighbor of degree 4. At nonzero thresholds, a vertex of degree 3 whose neighbors all have degree 3 is bounded by c_3 through Proposition 5.6, the other vertices by the star bounds of Section 4, and the charging is certified as in Proposition 6.1, with c_3 in place of F_3 [34].

6.2 Pathwidth sequences and the bootstrap

A sequence $a = (a_d)_{d\ge 3}$ 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 $d > k$, which is what putting a vertex into every bag adds to the width.

Theorem 6.4 (pathwidth sequences)**.** *Each of the following six tuples is a pathwidth sequence:* $\mathsf{C} = (c_3)$, $\mathsf{M} = (c_3, 0.2627)$, $\mathsf{F} = (F_4, F_4)$, S, P *and* R. *So is the bootstrap* (6.1) *of each of them; the bootstraps, rounded up, are the rows of Table 2.*

For graphs of maximum degree at most k, the six tuples are Theorems 4.4, 5.1, 6.2 and 6.3. The extension to all graphs is the following bootstrap. Let $0 \le b_3 \le \cdots \le b_k$ and $b_2 := 0$, and put $h_d = b_d - b_{d-1}$ and $H_d = \min_{3\le i\le d} h_i$. The *bootstrap* defines, for $d = k+1, k+2, \dots$ in turn,

$$b_d := \max\Big\{ \frac{1 + (d-1)\, b_{d-1} - H_{d-1}}{d},\ 1 - d\, H_{d-1} \Big\}. \tag{6.1}$$

Lemma 6.5 (bootstrap)**.** *Let $k \ge 3$ and $0 \le b_3 \le \cdots \le b_k$ with $b_k + H_k \le 1$, and suppose that every graph J with $\delta(J) \ge 3$ and $\Delta(J) \le k$ satisfies $\mathrm{pw}(J) \le \sum_{d=3}^{k} b_d n_d(J) + o(|V(J)|)$. Then the bootstrap of $(b_3, \dots, b_k)$ is a pathwidth sequence; it is nondecreasing, with all entries at most* 1.

Proof sketch. No probability is involved. As in the work of Fomin, Gaspers, Saurabh and Stepanov [28, Lemma 1] and of Edwards and McDermid [23, Theorem 3.2], we delete vertices of maximum degree $d > k$ one at a time, and reduce vertices of degree at most 2 after every deletion. Putting a deleted vertex back into every bag adds at most one to the width, and we compare this with the decrease of the potential $\psi(G) = \sum_v b_{d(v)}$. Deleting a vertex of degree d removes its own term b_d, and lowers each neighbor's term by an increment $h_{d'} \ge H_{d'}$. By (6.1) the entries satisfy $1 - b_d = (d-1)H_d + H_{d-1}$, so a deletion that has a neighbor of smaller degree lowers the potential by at least 1; regular components are handled by a modified potential. The reductions cost $O(\log n)$ in total (Section 2.4), and the remaining graph has maximum degree at most k. □

Three things are new compared with the reduction theorem of Edwards and McDermid: the seeds are arbitrary, with the running minimum H_d in place of their increment, so that a cubic coefficient below $\frac{1}{6}$ is admissible; the base is a per-degree bound for minimum degree 3 and maximum degree k, not a bound for cubic graphs; and the reductions are accounted for once. The recurrence also explains why a better cubic constant alone improves nothing at higher degrees: with $k = 3$ and $b_3 = c_3$ the second term wins, and $b_4 = 1 - 4c_3 \approx 0.4407 > \frac{1}{3}$. This is why Theorem 1.3 is bootstrapped from the base S of maximum degree 5.

Proof sketch of Theorem 1.3 and Corollary 1.4. The bootstrap of S satisfies the hypotheses of Lemma 6.5 with $k = 5$, by Theorem 6.2. Its entries are $\frac{3d-4}{32}$ up to degree 6, and from degree 7 on they satisfy the recurrence $d\, b_d = (d-2)\, b_{d-1} + b_{d-2} + 1$ of Edwards and McDermid. For the

Table 2: The bootstraps of the sequences of Theorem 6.4, rounded up to four decimals; in the last column, the bootstrap is bounded by a closed form. Entries in bold are the smallest of their column. The bootstrap of C is entrywise at least row M and is omitted.

	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	$\frac{5}{32}$	$\frac{1}{4}$	$\mathbf{\frac{11}{32}}$	**0.4375**	**0.5045**	**0.5581**	**0.6012**	**0.6368**	$\mathbf{1 - \frac{4}{d+1}}$

deficits $u_d = 1 - b_d$ this recurrence has the conserved quantity $d\,u_d + u_{d-1}$, which keeps $(d+1)\,u_d$ in a narrow interval and gives the closed forms of Theorem 1.3 as upper bounds [34]. For Corollary 1.4, $2a_d/d \leq \frac{7}{48}$ for every $d \geq 3$, with equality only at $d = 6$, so $\sum_d a_d n_d \leq \frac{7}{48} \sum_d \frac{d}{2} n_d \leq \frac{7}{48} m$. □

7 Computing the decompositions

Theorem 1.5 is proved in the full version [34], for each case in two forms: a Las Vegas algorithm in expected $n^{1+o(1)}$ bit operations and a deterministic one in $n^{2+o(1)}$, with $O(n(n+m))$ further operations on integers of $O(\log n)$ bits in the bootstrap cases. The decomposition is output by the indices of the first and the last bag of every vertex. Each algorithm computes discrete labels, sorts the vertices by them and builds the layout of the case. Four ingredients bound the counts that control the width.

(1) *Exact local labels.* The coefficients are rounded to rationals, and the Gaussian variables are replaced by a finite rational distribution coupled with the standard normal distribution.

(2) *Perturbation.* In this coupling every per-vertex probability used above changes by $o(1)$, uniformly in the threshold.

(3) *Uniform means.* The grid refines as n grows, so the bounds on the expected frontier sizes are needed uniformly in the threshold, with explicit error terms.

(4) *Derandomization.* The method of conditional expectations (Erdős and Selfridge [25]; Raghavan [63]), applied to the sum of the squared deviations of the counts, fixes the variables one at a time. The locality radius grows only like $\frac{1}{8} \log_2 \log_2 n$, so every conditional expectation is a sum of local terms, computable exactly in $n^{o(1)}$ time.

The bootstrap cases run the deletion process of Lemma 6.5 and undo the reductions explicitly. For maximum degree 3, 4 and 5 the same method computes, in deterministic time $n^{2+o(1)}$, linear orderings of cutwidth at most $c_3 n + o(n)$, $\frac{m}{6} + o(n)$ and $\gamma_5 m + o(n)$, and bisections with at most as many edges between their sides. The remainders $o(n)$ of all our bounds become small compared with n only for astronomically large n; the bounds, like the records they improve, are worst-case asymptotic statements.

8 Algorithmic applications

Part II of the full version [34] follows the new constants through the literature of exact algorithms. The new bounds arise 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 $c^{W(G)+o(n)}\,\mathrm{poly}(n)$, and the decomposition is computed by Theorem 1.5; this gives most rows, including, through Cut&Count, the rank-based approach and algorithms over orderings of small cutwidth, those for connectivity problems. *Rerunning a measure-and-conquer analysis*: in polynomial space, the widths enter Separate, Measure and Conquer [39] through balanced separators of size cn, and the separator constant enters its mathematical program. *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 [58] 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.

Table 3 lists every application, one line each. Polynomial factors are omitted and every exponent carries an additive $o(\cdot)$ term; r is the domain size, n the number of variables or vertices, m the number of constraints or edges, k the pattern size (for Co-Path Packing and k-Co-Path Set, the number of deleted vertices or edges), and $\bar{d}$ the average degree. "Previous" is the fastest bound that we are aware of. Numerical bases are rounded up in the last printed digit. The new bounds for feedback vertex sets, the connected and connectivity problems, cycles and paths, Co-Path Packing, k-Co-Path Set and the directed problems are achieved by Monte Carlo algorithms with one-sided error, except those marked det.

The Exponential Time Hypothesis. Two recent preprints, by OpenAI [59] and Anthropic [14], claim to refute the Exponential Time Hypothesis [43] by a subexponential-time algorithm for #3-SAT. Through the subexponential reduction families of Impagliazzo, Paturi and Zane [44] and direct counting algorithms, they would supersede the seventeen bounds of the groups marked ‡ in Table 3. As far as we know, the other problems lie outside this reach: they have real weights, a global connectivity or acyclicity constraint, 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.

Table 3: The algorithmic applications, one line each: problem and graph class, previous bound, new bound, and the work behind the previous bound. Letters refer to the notes after the table; groups marked ‡ would be superseded if the preprints [14, 59] are correct.

Problem	Previous	New	Previous work
Feedback vertex sets			
Feedback Vertex Set, $\Delta \leq 4$	1.443^n	$2^{n/3} < 1.2600^n$	[20][f]
Feedback Vertex Set, $\Delta \leq 4$, det.	1.7117^n	$5^{F_4 n} < 1.4853^n$	[36]
Feedback Vertex Set, $\Delta \leq 5$	1.61^n	$2^{0.5121n} < 1.4262^n$	[20][f]
Feedback Vertex Set, $\Delta \leq 6$	1.6297^n	$3^{7n/16} < 1.6172^n$	[50][f]
Connected FVS and others[e], $\Delta \leq 3$	1.26^n	$4^{c_3 n} < 1.2139^n$	[20][f]
Connected FVS and others[e], $\Delta \leq 4$	1.588^n	$4^{F_4 n} < 1.4060^n$	[20][f]
Connected FVS and others[e], $\Delta \leq 5$	1.824^n	$4^{11n/32} < 1.6105^n$	[20][f]
Domination			
(#)Min. Weight Dominating Set[c]	1.5002^n	1.4815^n	[58][a]
(#)Min. Weight Dominating Set, $\Delta \leq 3$	$3^{n/6} < 1.2010^n$	$3^{c_3 n} < 1.1661^n$	[31]
Connected Dominating Set, $\Delta \leq 3$	1.26^n	$3^{c_3 n} < 1.1661^n$	[20][f]
Connected Dominating Set, $\Delta \leq 4$	1.588^n	$4^{F_4 n} < 1.4060^n$	[20][f]
Connected Dominating Set, $\Delta \leq 5$	1.8139^n	$4^{11n/32} < 1.6105^n$	[6]
Independence and vertex covers			
(#)Max. Weight Indep. Set, $\Delta \leq 3$	$2^{n/6} < 1.1225^n$	$2^{c_3 n} < 1.1018^n$	[31]

Problem	Previous	New	Previous work
(#)Max. Weight Indep. Set, $\Delta \leq 4$	$2^{n/3} < 1.2600^n$	$2^{F_4 n} < 1.1858^n$	[28]
Connected Vertex Cover, $\Delta \leq 4$	1.443^n	$2^{n/3} < 1.2600^n$	[20][f]
Connected Vertex Cover, $\Delta \leq 5$	1.61^n	$2^{0.5121n} < 1.4262^n$	[20][f]
Max 2-CSP			
Max $(r,2)$-CSP, all instances	$r^{13m/75} < r^{m/5.769}$	$r^{7m/48} < r^{m/6.857}$	[64]
Max $(r,2)$-CSP, $\Delta \leq 3$	$r^{n/6}$	$r^{c_3 n} < r^{0.13983n}$	[31]
Max $(r,2)$-CSP, $\Delta \leq 4$	$r^{n/3}$, $r^{m/6}$	$r^{F_4 n}$, $r^{\frac{3F_4}{2+16F_4}m}$	[28], [64]
Max $(r,2)$-CSP[b], avg. degree $\bar{d} \geq 7$	$r^{n(1-3.5/(\bar{d}+1)+O(\bar{d}^{-3}))}$	$r^{n(1-3.964/(\bar{d}+1))}$	[23]
Max $(r,2)$-CSP[b], avg. degree $\bar{d} \geq 11$	$r^{n(1-3.5/(\bar{d}+1)+O(\bar{d}^{-3}))}$	$r^{n(1-4/(\bar{d}+1))}$	[23]
Patterns and packings			
Counting k-edge subgraphs[d]	$n^{0.174k}$	$n^{(7/48)k} < n^{0.1459k}$	[17]
Counting k-paths, k-cycles[d]	$n^{0.1704k}$	$n^{0.1459k}$	[10]
Permutation Pattern Matching	$n^{k/4}$	$2^{O(k)} n^{(253/1215)k}$	[5]
k-opt for TSP, general graphs	$n^{(1/4+\varepsilon_k)k}$	$n^{(253/1215+\varepsilon_k)k}$	[18]
k-opt for TSP, bounded degree	$n^{(23/135+\varepsilon_k)k}$	$n^{(7/48+\varepsilon_k)k}$	[11]
Co-Path Packing	2.9241^k	2.8192^k	[52]
k-Co-Path Set	1.588^k	1.5423^k	[67]
Cycles and paths			
Hamiltonian Cycle, $\Delta \leq 3$	1.1583^n	1.1312^n	[19][f]
Hamiltonian Cycle, $\Delta \leq 4$	1.5058^n	1.3416^n	[19][f,g]
Hamiltonian Cycle, $\Delta \leq 5$	1.6246^n	1.5252^n	[8][f]
Hamiltonian Cycle, avg. degree d	1.12^{dn}	1.0937^{dn}	[49]
TSP, $\Delta \leq 3$, det.	1.2128^n	1.1757^n	[9]
TSP, $\Delta \leq 4$, det.	1.6110^n	1.4214^n	[9][g]
TSP, $\Delta \leq 5$, det.	1.8588^n	1.6352^n	[9][g]
TSP, $\Delta \leq 6$, det.	1.9672^n	1.8699^n	[7]
TSP, avg. degree d, det.	1.1321^{dn}	1.1100^{dn}	[49]
Longest Path, $\Delta \leq 3$	1.26^n	1.1874^n	[20][f]
Longest Path, $\Delta \leq 4$	1.5875^n	1.3523^n	[8][f]
Longest Path, $\Delta \leq 5$	1.6246^n	1.5252^n	[8][f]
Other connectivity problems			
Steiner Tree, $\Delta \leq 3$	1.201^n	$3^{c_3 n} < 1.1661^n$	[20][f]
Steiner Tree, $\Delta \leq 4$	1.3533^n	$3^{F_4 n} < 1.3100^n$	[30]
Max Internal Spanning Tree, $\Delta \leq 6$	1.8423^n	$4^{7n/16} < 1.8341^n$	[8][f]
Directed problems			
Directed problems[h], $\Delta \leq 3$	1.349^n	$6^{c_3 n} < 1.2848^n$	[20][f]
Directed problems[h], $\Delta \leq 4$	1.818^n	$6^{F_4 n} < 1.5533^n$	[20][f]
Polynomial space			
Max $(r,2)$-CSP, all instances	$r^{9m/50}$	$r^{0.17587m}$	[39]
Max $(r,2)$-CSP, $\Delta \leq 4$	$r^{7m/40}$	$r^{0.16984m}$	[39]
Max $(r,2)$-CSP, cubic instances	$r^{n/5}$	$r^{0.17935n}$	[39]
Max 2-SAT, counting dominating sets and independent sets ‡			
Max 2-SAT	$2^{m/6.321}$	$2^{m/6.857}$	[38]
#Dominating Set, all graphs	1.5002^n	1.4815^n	[58][a]
#Dominating Set (every size), $\Delta \leq 3$	$3^{n/6} < 1.2010^n$	$3^{c_3 n} < 1.1661^n$	[31]
#Independent Sets, $\Delta \leq 3$	$2^{n/6} < 1.1225^n$	$2^{c_3 n} < 1.1018^n$	[31]
#Independent Sets, all graphs	1.2330^n	1.2287^n	[37]
Local problems on graphs of bounded degree (deterministic) ‡			
Max k-Vertex Cover, $\Delta \leq 3$	1.123^n	$2^{c_3 n} < 1.1018^n$	[15]
Min. Independent Dominating Set, $\Delta \leq 3$	1.3351^n	$3^{c_3 n} < 1.1661^n$	[13][i]
Min. Independent Dominating Set, $\Delta \leq 4$	1.3351^n	$3^{F_4 n} < 1.3100^n$	[13][i]
Upper Domination, $\Delta \leq 3$	1.3481^n	$6^{c_3 n} < 1.2848^n$	[4]

Problem	Previous	New	Previous work
Edge Dominating Set[j], $\Delta \le 3$	$3^{n/6} < 1.2010^n$	$3^{c_3 n} < 1.1661^n$	[31]
Odd Cycle Transversal, $\Delta \le 4$	$3^{n/3} < 1.4423^n$	$2^{n/3} < 1.2600^n$	[28]
Odd Cycle Transversal, $\Delta \le 5$	$3^{13n/30} < 1.6098^n$	$2^{0.5121n} < 1.4262^n$	[28]
Polynomial space ‡			
#Dominating Sets of each size, $\Delta \le 3$	1.2458^n	1.2178^n	[39]
#Independent Sets, $\Delta \le 3$	1.1389^n	1.1304^n	[37][k]
#Independent Sets, $\Delta \le 4$	1.2070^n	1.2037^n	[37]
#Independent Sets, all graphs	1.2356^n	1.2341^n	[37]
Chromatic Number	2.2356^n	2.2341^n	[37]

Notes to Table 3. 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 $O(1.5004^n)$ [34]. [b] A record only for moderate $\bar{d}$; for large $\bar{d}$, the algorithms of Williams [69], for integer scores of polynomial size, and of Golovnev and Kutzkov [41], for Boolean domains, are faster. [c] When the set of possible weight sums has polynomial size, as in [58, Cor. 32] (minimum weight) and [68, Cor. 8.14] (counting). [d] Both bounds carry a further factor $k^{O(k)}$. [e] 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 [28], and for TSP with $\omega \le \frac{9}{4}$. [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. [i] A bound for all graphs. [j] Also Minimum Maximal Matching. [k] The previous analysis has open points; the new bound includes the necessary corrections [34]. The bounds for TSP, previous and new, are evaluated with the matrix multiplication exponent $\omega \le \frac{9}{4}$ [1, 60].

9 Conclusion and open problems

We replaced the bisection behind the bound $\frac{n}{6} + o(n)$ of Fomin and Høie [31] 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 $c_3 n + o(n)$; a charging scheme certified by computer and a bootstrap to large degrees extended the method to all graphs. This solves the open problem of van Rooij [68, Open Problem 14.2].

The constant for maximum degree 3. The smallest constant c such that every graph of maximum degree 3 has pathwidth at most $cn + o(n)$ lies in $[0.1032, c_3]$, and the lower end comes from random cubic graphs [51]. 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 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 two children of the root both have the root as their only neighbor above zero.

Conjecture 9.1 (ideal-wave bound)**.** *Every cubic graph G satisfies* $\mathrm{pw}(G) \le (c_3 - 3p_{\mathrm{ch}})n + o(n)$*, where* $c_3 - 3p_{\mathrm{ch}} = 0.13224778941\ldots$.

Maximum degree 4, degree 6 and treewidth. For maximum degree 4 the ideal 4-star function has value 0.2423 at threshold zero, below $F_4 \approx 0.2458$; 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. For degrees 6 and more all our coefficients come from the bootstrap, and a better coefficient at degree 6 alone would lower the bound $\frac{7}{48}m$ per edge. Finally, is the worst-case coefficient of

treewidth smaller than that of pathwidth for any of these classes? The Gaussian sweep has given us no gain specific to treewidth at leading order.

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. Claude also drafted this short version from the full version.
- 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.

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 [62]. 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.

References

[1] 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`.

[2] Ágnes Backhausz, Balázs Szegedy, and Bálint Virág. Ramanujan graphings and correlation decay in local algorithms. *Random Struct. Algorithms*, 47(3):424–435, 2015. `doi:10.1002/RSA.20562`.

[3] 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`.

[4] 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. *Theor. Comput. Sci.*, 717:2–25, 2018. doi:10.1016/J.TCS.2017.05.042.

[5] 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.

[6] Daniel Binkele-Raible and Henning Fernau. An exact exponential-time algorithm for the directed maximum leaf spanning tree problem. *J. Discrete Algorithms*, 15:43–55, 2012. doi:10.1016/J.JDA.2012.03.006.

[7] Andreas Björklund, Thore Husfeldt, Petteri Kaski, and Mikko Koivisto. The traveling salesman problem in bounded degree graphs. *ACM Trans. Algorithms*, 8(2):18:1–18:13, 2012. doi:10.1145/2151171.2151181.

[8] Andreas Björklund, Vikram Kamat, Łukasz Kowalik, and Meirav Zehavi. Spotting trees with few leaves, 2015. Version 2. URL: https://arxiv.org/abs/1501.00563v2.

[9] Hans L. Bodlaender, Marek Cygan, Stefan Kratsch, and Jesper Nederlof. Deterministic single exponential time algorithms for connectivity problems parameterized by treewidth. *Inf. Comput.*, 243:86–111, 2015. doi:10.1016/J.IC.2014.12.008.

[10] 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.

[11] 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 *ESA 2019*, volume 144 of *LIPIcs*, pages 23:1–23:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICS.ESA.2019.23.

[12] Károly Böröczky, Jr. *Finite Packing and Covering*, volume 154 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2004. doi:10.1017/CBO9780511546587.

[13] Nicolas Bourgeois, Federico Della Croce, Bruno Escoffier, and Vangelis Th. Paschos. Fast algorithms for min independent dominating set. *Discrete Appl. Math.*, 161(4-5):558–572, 2013. doi:10.1016/J.DAM.2012.01.003.

[14] 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.

[15] Federico Della Croce and Vangelis Th. Paschos. Efficient algorithms for the max k-vertex cover problem. In *IFIP TCS 2012*, volume 7604 of *LNCS*, pages 295–309. Springer, 2012. doi:10.1007/978-3-642-33475-7_21.

[16] Endre Csóka, Balázs Gerencsér, Viktor Harangi, and Bálint Virág. Invariant Gaussian processes and independent sets on regular graphs of large girth. *Random Struct. Algorithms*, 47(2):284–303, 2015. doi:10.1002/RSA.20547.

[17] Radu Curticapean, Holger Dell, and Dániel Marx. Homomorphisms are a good basis for counting small subgraphs. In *STOC 2017*, pages 210–223. ACM, 2017. doi:10.1145/3055399.3055502.

[18] Marek Cygan, Łukasz Kowalik, and Arkadiusz Socała. Improving TSP tours using dynamic programming over tree decompositions. *ACM Trans. Algorithms*, 15(4):54:1–54:19, 2019. doi:10.1145/3341730.

[19] 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.

[20] 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.

[21] Joaquin B. Diaz and Frederic T. Metcalf. Stronger forms of a class of inequalities of G. Pólya–G. Szegő, and L. V. Kantorovich. *Bull. Am. Math. Soc.*, 69(3):415–418, 1963. doi:10.1090/S0002-9904-1963-10953-2.

[22] Josep Díaz, Jordi Petit, and Maria J. Serna. A survey of graph layout problems. *ACM Comput. Surv.*, 34(3):313–356, 2002. doi:10.1145/568522.568523.

[23] 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.

[24] John A. Ellis, Ivan Hal Sudborough, and Jonathan S. Turner. The vertex separation and search number of a graph. *Inf. Comput.*, 113(1):50–79, 1994. doi:10.1006/INCO.1994.1064.

[25] Paul Erdős and John L. Selfridge. On a combinatorial game. *J. Comb. Theory Ser. A*, 14(3):298–301, 1973. doi:10.1016/0097-3165(73)90005-8.

[26] Uriel Feige. Approximating the bandwidth via volume respecting embeddings. *J. Comput. Syst. Sci.*, 60(3):510–539, 2000. doi:10.1006/JCSS.1999.1682.

[27] László Fejes Tóth. Extremum properties of the regular polyhedra. *Can. J. Math.*, 2:22–31, 1950. doi:10.4153/CJM-1950-003-x.

[28] 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.

[29] Fedor V. Fomin, Fabrizio Grandoni, and Dieter Kratsch. Some new techniques in design and analysis of exact (exponential) algorithms. *Bull. EATCS*, 87:47–77, 2005.

[30] 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.

[31] Fedor V. Fomin and Kjartan Høie. Pathwidth of cubic graphs and exact algorithms. *Inf. Process. Lett.*, 97(5):191–196, 2006. doi:10.1016/j.ipl.2005.10.012.

[32] Fedor V. Fomin, Kazuo Iwama, Dieter Kratsch, Petteri Kaski, Mikko Koivisto, Łukasz Kowalik, Yoshio Okamoto, Johan M. M. van Rooij, and Ryan Williams. 08431 open problems – moderately exponential time algorithms. In *Moderately Exponential Time Algorithms, Dagstuhl Seminar 08431*, volume 08431 of *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany, 2008. doi:10.4230/DagSemProc.08431.3.

[33] 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.

[34] Serge Gaspers. Path decompositions of sparse graphs via Gaussian sweeps and faster exponential algorithms. Full version, 2026. URL: https://serggasp.github.io/pathwidth-site/papers/sparse-pathwidth/sparse-pathwidth.pdf.

[35] 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.

[36] Serge Gaspers and Edward J. Lee. Exact algorithms via multivariate subroutines, 2017. Version 1. URL: https://arxiv.org/abs/1704.07982v1.

[37] 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.

[38] Serge Gaspers and Gregory B. Sorkin. A universally fastest algorithm for Max 2-Sat, Max 2-CSP, and everything in between. *J. Comput. Syst. Sci.*, 78(1):305–335, 2012. doi:10.1016/J.JCSS.2011.05.010.

[39] Serge Gaspers and Gregory B. Sorkin. Separate, Measure and Conquer: Faster polynomial-space algorithms for Max 2-CSP and counting dominating sets. *ACM Trans. Algorithms*, 13(4):44:1–44:36, 2017. doi:10.1145/3111499.

[40] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995. doi:10.1145/227683.227684.

[41] Alexander Golovnev and Konstantin Kutzkov. New exact algorithms for the 2-constraint satisfaction problem. *Theor. Comput. Sci.*, 526:18–27, 2014. doi:10.1016/J.TCS.2014.01.010.

[42] Viktor Harangi and Bálint Virág. Independence ratio and random eigenvectors in transitive graphs. *Ann. Probab.*, 43(5):2810–2840, 2015. doi:10.1214/14-AOP952.

[43] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *J. Comput. Syst. Sci.*, 62(2):367–375, 2001. doi:10.1006/JCSS.2000.1727.

[44] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001. doi:10.1006/JCSS.2001.1774.

[45] Fredrik Johansson. Arb: Efficient arbitrary-precision midpoint-radius interval arithmetic. *IEEE Trans. Comput.*, 66(8):1281–1292, 2017. doi:10.1109/TC.2017.2690633.

[46] Harry Kesten. Symmetric random walks on groups. *Trans. Am. Math. Soc.*, 92(2):336–354, 1959. doi:10.1090/S0002-9947-1959-0109367-6.

[47] Nancy G. Kinnersley. The vertex separation number of a graph equals its path-width. *Inf. Process. Lett.*, 42(6):345–350, 1992. doi:10.1016/0020-0190(92)90234-M.

[48] Joachim Kneis, Daniel Mölle, Stefan Richter, and Peter Rossmanith. A bound on the pathwidth of sparse graphs with applications to exact algorithms. *SIAM J. Discrete Math.*, 23(1):407–427, 2009. doi:10.1137/080715482.

[49] Łukasz Kowalik and Konrad Majewski. The asymmetric travelling salesman problem in sparse digraphs, 2020. Version 2. URL: https://arxiv.org/abs/2007.12120v2.

[50] 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.

[51] Lyuben Lichev and Dieter Mitsche. On the minimum bisection of random 3-regular graphs. *Electron. J. Comb.*, 30(2):P2.40, 2023. doi:10.37236/11085.

[52] Yuxi Liu and Mingyu Xiao. Solving Co-Path/Cycle Packing and Co-Path Packing faster than 3^k. *Inf. Comput.*, 309:105406, 2026. doi:10.1016/J.IC.2026.105406.

[53] Anand Louis, Prasad Raghavendra, and Santosh S. Vempala. The complexity of approximating vertex expansion. In *FOCS 2013*, pages 360–369. IEEE Computer Society, 2013. doi:10.1109/FOCS.2013.46.

[54] Russell Lyons. Factors of IID on trees. *Comb. Probab. Comput.*, 26(2):285–300, 2017. doi:10.1017/S096354831600033X.

[55] Fillia Makedon and Ivan Hal Sudborough. On minimizing width in linear layouts. *Discrete Appl. Math.*, 23(3):243–265, 1989. doi:10.1016/0166-218X(89)90016-4.

[56] Burkhard Monien and Robert Preis. Upper bounds on the bisection width of 3- and 4-regular graphs. *J. Discrete Algorithms*, 4(3):475–498, 2006. doi:10.1016/J.JDA.2005.12.009.

[57] Abhijeet Mulgund. Stochastic domination of Gaussian maxima by the regular simplex, 2026. Version 2. URL: https://arxiv.org/abs/2609.28452v2.

[58] 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.

[59] 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.

[60] 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.

[61] Robin L. Plackett. A reduction formula for normal multivariate integrals. *Biometrika*, 41(3-4):351–360, 1954. doi:10.1093/biomet/41.3-4.351.

[62] PVC (Research Infrastructure), UNSW Sydney. Katana. UNSW, Sydney, 2010. doi:10.26190/669X-A286.

[63] Prabhakar Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *J. Comput. Syst. Sci.*, 37(2):130–143, 1988. doi:10.1016/0022-0000(88)90003-7.

[64] Alexander D. Scott and Gregory B. Sorkin. Linear-programming design and analysis of fast algorithms for Max 2-CSP. *Discrete Optim.*, 4(3-4):260–287, 2007. doi:10.1016/J.DISOPT.2007.08.001.

[65] William F. Sheppard. On the application of the theory of error to cases of normal distribution and normal correlation. *Philos. Trans. R. Soc. Lond. Ser. A*, 192:101–167, 1899. doi:10.1098/rsta.1899.0003.

[66] David Slepian. The one-sided barrier problem for Gaussian noise. *Bell Syst. Tech. J.*, 41(2):463–501, 1962. doi:10.1002/j.1538-7305.1962.tb02419.x.

[67] Blair D. Sullivan and Andrew van der Poel. A fast parameterized algorithm for co-path set. In *IPEC 2016*, volume 63 of *LIPIcs*, pages 28:1–28:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPICS.IPEC.2016.28.

[68] 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.

[69] Ryan Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theor. Comput. Sci.*, 348(2-3):357–365, 2005. doi:10.1016/j.tcs.2005.09.023.

[70] Wolfgang Woess. *Random Walks on Infinite Graphs and Groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2000. doi:10.1017/CBO9780511470967.