How few edges preserve the diameter of a hypercube?

Tyler Hughes [co-author]

AI-assisted research draft for author review.

Abstract

How many edges can an n-dimensional cube lose while every pair of vertices still have a route of at most n steps? Write $N = 2^n$ and let $f(n)$ be the minimum number of surviving edges. We count the edges between successive distance spheres from a root. The count gives an exact identity relating excess degree, extra path length, and the depths at which shortest paths meet. Averaging over all roots gives $f(n) \geq N + \lceil (N-2n)/(n-1) \rceil$. Applying the same count at opposite corners bounds the contribution from each root separately. A bound on the width of every distance sphere improves the asymptotic lower coefficient from 1 to 3/2 on the scale $N/(n-1)$. In small dimensions we obtain $f(4) = 20$, $39 \leq f(5) \leq 40$, $76 \leq f(6) \leq 81$, and $294 \leq f(8) \leq 324$. The six-dimensional upper endpoint comes from a published construction. Lean verifies the general results and integer certificates through dimension 32. Two evaluations of the same certificate improve the previous formal bounds of 75 and 292 in dimensions six and eight. A separate choice of three radii excludes 149 edges in dimension seven.

MSC 2020. 05C12, 05C35. **Keywords.** Hypercube, diameter, spanning subgraph, graph distance.

1 Introduction and main results

Opposite corners of a cube differ in every coordinate. If a road between two other corners disappears, a detour can be harmless. The requirement is that no detour push a pair farther apart than those opposite corners.

Consider the three paths in Figure 1. Each runs from 000 to 111 in three steps. Together they contain every vertex of Q_3. For points on different paths at positions i, j from 000, the routes through the two ends have lengths $i + j$ and $6 - i - j$. At least one has length at most three. Thus an edge between 100 and 101 can disappear even though the distance between its endpoints triples. Any lower bound on the number of surviving edges must allow this sort of detour.

Write $d(w, v)$ for distance in a diameter-preserving spanning subgraph G and $h(w, v)$ for Hamming distance. A vertex has degree excess $x_v = \deg_G(v) - 2$. The total extra distance from w is $\delta(w) = \sum_v (d(w, v) - h(w, v))$. Count one parent for each nonroot vertex in the distance layers at w. Each additional parent edge ending at depth j contributes $n - j$ to $M(w)$. An extra parent near the root thus costs more than one near the diameter.

Theorem 1 (One root, all roots, opposite roots)**.** *Let G be a diameter-n spanning subgraph of Q_n with $n \geq 2$. Set $N = 2^n$, $|E(G)| = N + t$ and $\mu = (n-1)t - N + 2n$. For every root w, with $\bar{w}$ its bitwise complement,*

$$\sum_v x_v d(w, v) + 2M(w) = (2n-1)t + 2n - N, \tag{1}$$

$$\sum_w \big(x_w \delta(w) + 2M(w)\big) = N\mu, \tag{2}$$

$$M(w) + M(\bar{w}) + \tfrac{1}{2} \sum_v x_v \big(d(w, v) + d(\bar{w}, v) - n\big) = \mu. \tag{3}$$

Every summand on the left in (2) and (3) is nonnegative. In particular $M(w) \leq \mu$ for each root.

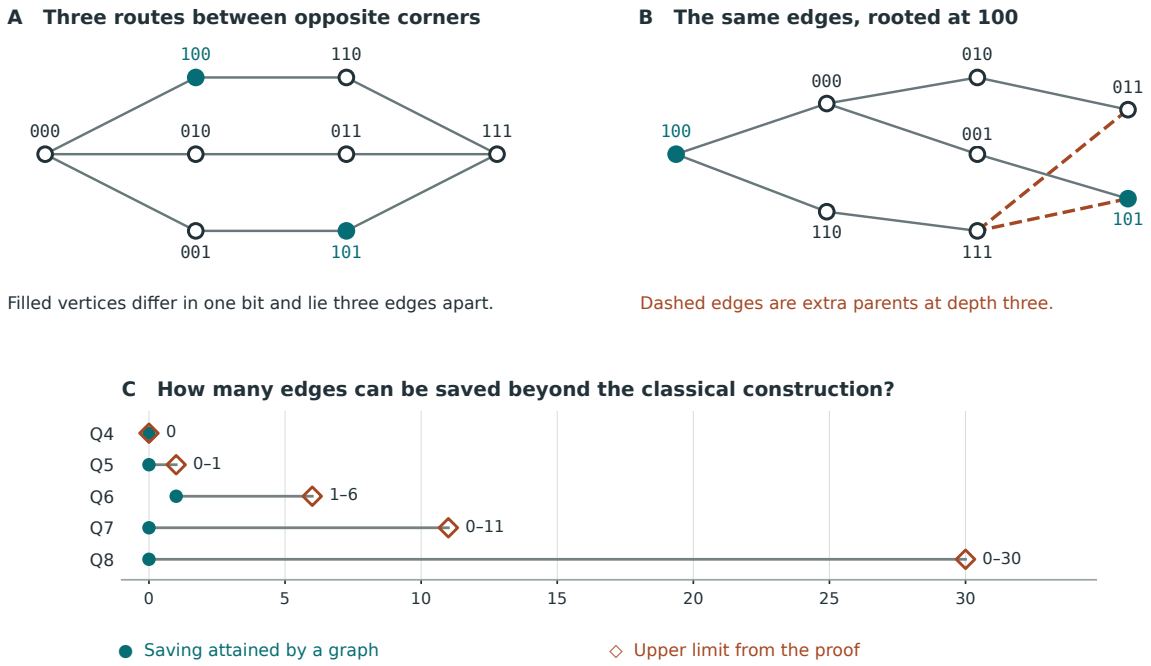


Figure 1. A nine-edge cube and its shortest paths. In B, each vertex except the root needs one incoming edge. The two remaining edges enter the last layer and have weight zero in the edge count below. Applying the count to all roots gives a general lower bound, and pairing opposite roots sharpens it. In C each interval runs from the saving achieved by a construction to the maximum saving permitted by a Lean-checked certificate. Thus one edge can be saved in Q_6 and at most six can. The work began with a two-program census of Q_4, which found 8,320 labelled optima. The exact maximum saving in Q_5 remains open.

Theorem 2 (General and asymptotic bounds)**.** *The minimum size $f(n)$ of a diameter-n spanning subgraph of Q_n satisfies*

$$N + \left\lceil \frac{N - 2n}{n - 1} \right\rceil \le f(n) \le g(n) := N + \binom{n}{\lfloor n/2 \rfloor} - 2, \qquad \liminf_{n \to \infty} \frac{(n-1)(f(n) - N)}{N} \ge \frac{3}{2}.$$

For $n \ge 4$ the unrounded lower bound is strict. The upper bound is the classical symmetric-chain bound [3].

Theorem 3 (Finite dimensions)**.** *The first three minima are $f(2) = 4$, $f(3) = 9$, and $f(4) = 20$. Then*

$$39 \le f(5) \le 40, \quad 76 \le f(6) \le 81, \quad 150 \le f(7) \le 161, \quad 294 \le f(8) \le 324.$$

Table 1 gives lower bounds for $9 \le n \le 32$, including $f(9) \ge 580$ and $f(32) \ge 4{,}454{,}492{,}243$.

An 81-edge witness in ten lines. Charton, Ellenberg, Wagner and Williamson give a diameter-six subgraph one edge below $g(6)$ [4]. The six masks below encode its edges. Bit u in mask i permits flipping bit $5-i$ at vertex u. The assertions check symmetry, size and all distances. The opposite corners certify that the diameter is exactly six.

```
masks = (0xd55d90b5d55d90b5, 0xea80ea805d735d73, 0xaaaa4d4d02025b5b,
         0x88000022aa8822bb, 0xff00005aa0a5a55f, 0xccff3f0000000033)
adj = [{u ^ (32 >> i) for i, m in enumerate(masks) if m >> u & 1}
       for u in range(64)]
assert all(u in adj[v] for u, row in enumerate(adj) for v in row)
assert sum(map(len, adj)) == 2 * 81
for u in range(64):
    seen = {u}
    for _ in range(6): seen |= {w for v in seen for w in adj[v]}
    assert len(seen) == 64
```

The proof starts by orienting each edge away from one root. The resulting count limits both the sizes of distance spheres and how much delay they can accumulate. The first restriction gives the asymptotic bound. The second, applied before averaging over roots, gives the finite improvements.

2 Definitions and distance layers

The vertices of Q_n are binary words of length n. Two words are adjacent when they differ in exactly one position. Throughout, $n \geq 2$. A graph G is *valid* if it contains all cube vertices, uses only cube edges, and is connected with diameter n. For any spanning subgraph one always has $d(u,v) \geq h(u,v)$ when a path exists, since each edge changes just one coordinate. Thus a connected spanning subgraph of Q_n is valid whenever its diameter is at most n. Adding a missing cube edge to a valid graph preserves validity. To rule out all graphs up to a given size, it therefore suffices to rule out that size.

Every vertex of a valid graph has degree at least two. If u had just one neighbor v, its path to $\bar{v}$ would first go to v, then take at least n more steps. Here $\bar{v} \neq u$ because $n \geq 2$. Thus $0 \leq x_u \leq n-2$. The even and odd Hamming-weight classes have equal size, and each edge has one end in each class. Consequently

$$\sum_{u \text{ even}} x_u = \sum_{u \text{ odd}} x_u = t, \qquad \sum_{u} x_u = 2t. \tag{4}$$

Fix a root w and write $S_j = \{v : d(w,v) = j\}$, $s_j = |S_j|$, $X_j = \sum_{v \in S_j} x_v$ and $Y_j = \sum_{i=0}^{j} X_i$. Distances to the two ends of an edge differ by exactly one, since Q_n is bipartite. Orient each edge towards its more distant endpoint. Every vertex except w receives an edge. If M_j counts the edges entering S_j beyond the first at each vertex, then

$$M_1 = 0, \qquad \sum_{j=1}^{n} M_j = t+1, \qquad M(w) = \sum_{j=1}^{n} (n-j) M_j.$$

No choice of parent affects these numbers.

3 Proofs

3.1 The edge count

Proof of Theorem 1. Sum the distances of both endpoints of every edge from w. An edge entering S_j contributes $2j-1$. Choose one incoming edge for each nonroot vertex. These $N-1$ edges

contribute $2\sum_v d(w,v)-(N-1)$. The other $t+1$ edges contribute $(2n-1)(t+1)-2M(w)$, since the difference between $2n-1$ and $2j-1$ is $2(n-j)$. Counting by endpoints gives $\sum_v \deg(v)d(w,v)$. Subtracting $2\sum_v d(w,v)$ proves (1).

Each coordinate differs from that of a fixed vertex v at exactly $N/2$ roots, so $\sum_w d(w,v) = nN/2+\delta(v)$. Sum (1) over w and use $\sum_v x_v = 2t$ to obtain (2). Instead add (1) at w and $\bar{w}$, subtract $n\sum_v x_v$, and divide by two to obtain (3). The terms there are nonnegative because $h(w,v)+h(\bar{w},v)=n$. □

For the example in Figure 1, only the two end vertices have positive excess. All distances from either end are Hamming distances. The two extra parents enter the last layer from every root, as (3) requires when $t=1$ and $\mu=0$.

Lemma 4 (Sphere widths and distance deficits)**.** *For $j \geq 1$ one has $s_j \leq 2+Y_{j-1} \leq 2+2t$. Put $B_k = \sum_{i=0}^{k}\binom{n}{i}$, $C_k = \sum_{i=0}^{k} s_i$ and $g_k = B_k - C_k$. Then $g_k \geq 0$ and $\delta(w) = \sum_{k=0}^{n-1} g_k$.*

Proof. The adjacent levels supply $s_j + M_j$ incoming and $s_{j+1}+M_{j+1}$ outgoing edges at S_j. Their sum is $2s_j + X_j$. Thus

$$s_{j+1} - s_j = X_j - M_j - M_{j+1} \quad (1 \leq j < n). \tag{5}$$

Start at $s_1 = 2+x_w$ and sum this inequality after dropping the nonnegative merge counts. For the second assertion, the graph ball lies inside the Hamming ball. A vertex at distances h and d contributes one to $B_k - C_k$ for precisely the $d-h$ integers $h \leq k < d$. □

Equation (2) gives $\mu \geq 0$, the finite lower bound of Theorem 2. If $\mu = 0$ and $n \geq 4$, then $t > 0$, so choose a root with $x_w > 0$. It has $\delta(w) = M(w) = 0$. The first equality makes $s_j = \binom{n}{j}$ at every radius. The second forces $M_j = 0$ for $j < n$, so (5) makes $s_1, \ldots, s_{n-1}$ nondecreasing. But $s_2 = \binom{n}{2} > n = s_{n-1}$, a contradiction.

3.2 Why sparse cubes have long detours

For $s = 2+2t$ let $U_0 = 1$ and $U_k = \min(B_k, U_{k-1}+s)$. This is a bound on the largest possible graph ball at each radius with the given sphere-width limit. By induction $C_k \leq U_k$. Lemma 4 therefore gives the same lower bound on extra distance at every root:

$$\delta(w) \geq \Delta := \sum_{k=0}^{n-1}(B_k - U_k), \qquad 2t\Delta \leq N\mu. \tag{6}$$

The second inequality follows by summing the first with weights x_w. At $n = 10$ and $t = 112$, for example, $s = 226$ and $\Delta = 36$. The inequality would require $8{,}064 \leq 4{,}096$.

Proof of the asymptotic bound in Theorem 2. Take an increasing sequence of dimensions along which $c_n = (n-1)(f(n)-N)/N$ tends to a finite limit c. If the lower limit is infinite there is nothing to prove. The finite bound gives $c \geq 1$. Use an optimal graph in each dimension, so $s = (2c+o(1))N/n$.

Set $a = \lfloor n/2 - n^{2/3} \rfloor$ and $b = \lceil n/2 + n^{2/3} \rceil$. The normalized Hamming ball B_k/N is the distribution function of a binomial variable with mean $n/2$ and variance $n/4$. Chebyshev's inequality gives $B_a = o(N)$ and $B_b = N - o(N)$. The graph gains at most $(b-a)s = o(N)$ vertices across this window, whence $U_b = o(N)$.

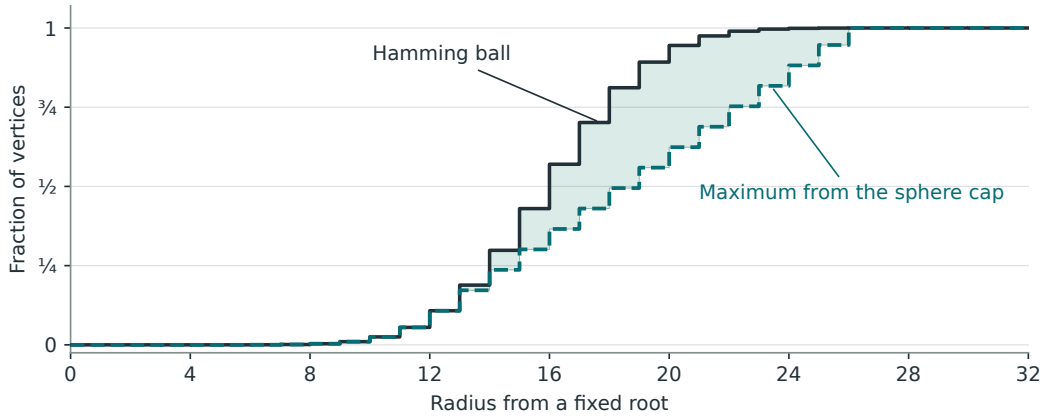


Figure 2. The width obstruction at $n = 32$, using the smallest integer t allowed by $\mu \geq 0$. The Hamming ball grows rapidly at the middle layers. A graph ball can gain at most $2 + 2t$ vertices in one step. Each shaded strip counts vertices that must travel farther. Its width is one radius, so the total shaded area is Δ/N.

For $0 \leq r < n - b$, we have $B_{b+r} - U_{b+r} \geq [B_b - U_b - rs]_+$, where $[z]_+ = \max(z, 0)$. Divide the sum of these terms by nN. The Riemann sums converge to a triangle of height one and base $1/(2c)$:

$$\liminf \frac{\Delta}{nN} \;\geq\; \int_0^{1/2} [1 - 2cy]_+ \, dy = \frac{1}{4c}.$$

Indeed, $(n - b)/n \to 1/2$ and the error in the normalized integrand is uniform for $r/n \leq 1/2$. Divide (6) by N^2. Since $\mu/N \to c - 1$ and $nt/N \to c$, we get $c - 1 \geq 2c/(4c) = 1/2$. □

3.3 Two monotone routes and the classical construction

Fix a corner r and order the vertices by Hamming distance from r. If each interior vertex has both a neighbor below it and a neighbor above it, every vertex has monotone routes to r and $\bar{r}$ (provided the poles have incident edges). Two vertices at heights i, j are then joined in at most $\min(i + j, 2n - i - j) \leq n$ steps. The three paths in Figure 1 realize exactly this construction.

To achieve this at minimum cost across two consecutive Hamming levels, match every vertex of the smaller level to a distinct vertex of the larger. Such a matching exists by Hall's theorem: the level graph is biregular and its smaller side has at least the degree of the larger side, so every subset on that side has at least as many neighbors as vertices. Add one edge for each unmatched vertex on the larger side. Now both sides are covered using as many edges as there are vertices on the larger side. Doing this between every pair of levels gives exactly

$$\sum_{i=0}^{n-1} \max\left\{ \binom{n}{i}, \binom{n}{i+1} \right\} = N + \binom{n}{\lfloor n/2 \rfloor} - 2 = g(n).$$

The equality follows from unimodality. This proves the upper bound in Theorem 2. A symmetric-chain decomposition supplies another choice of the matchings [5].

There is a converse count for this kind of routing. Let β_r count interior vertices missing all neighbors on one side of their Hamming level. Put $m = \lceil n/2 \rceil$. Request a downward edge from each vertex in levels 1 through m, and an upward edge from each vertex in levels m through $n - 1$. There are $g(n)$ requests. Each fulfilled request uses a distinct edge. A middle vertex cannot miss

both requests in a valid graph. Therefore

$$g(n) \le |E(G)| + \beta_r. \tag{7}$$

This explains what an improvement on $g(n)$ must cost: even with the best choice of corner, some vertices lose a monotone step. The 81-edge graph in Q_6 has at least one such vertex for every r.

3.4 What the antipodal cap forces in dimension eight

The root identity has a convenient equivalent form. With $R = \{m : 0 \le m \le n-2,\ m \equiv n \bmod 2\}$,

$$\sum_{m\in R} Y_m = \sum_v x_v \left\lfloor \frac{n - d(w,v)}{2} \right\rfloor = N/2 - n + M(w). \tag{8}$$

Indeed, the floor counts the radii in R containing v. When the floor is doubled, its parity correction has total weight t by (4). Substitution of (1) gives the last equality.

Suppose now that a valid Q_8 graph has 292 edges. Then $t = 36$ and $\mu = 12$. Fix w and abbreviate $x = x_w$. Its first ball deficit is $g_1 = 6 - x$. By (8) and Lemma 4,

$$\begin{aligned} 120 + M(w) &= x + Y_2 + Y_4 + Y_6 \\ &\ge x + s_3 + s_4 + s_5 - 6 = x + 176 + g_2 - g_5. \end{aligned}$$

The last equality uses $\binom{8}{3} + \binom{8}{4} + \binom{8}{5} = 182$ and $C_5 - C_2 = B_5 - B_2 + g_2 - g_5$. Since $M(w) \le \mu$, this forces $g_5 \ge 56 + x + g_2 - \mu$ at *every* root. Consequently $\delta(w) \ge g_1 + g_5 \ge 62 - \mu = 50$. Inserting this into the global count gives

$$3{,}600 = 72 \cdot 50 \le \sum_w x_w \delta(w) \le N\mu = 256 \cdot 12 = 3{,}072.$$

The contradiction rules out 292 edges and all smaller sizes. This gives $f(8) \ge 293$ by hand. Section 4 keeps the two parity classes separate in the same argument and excludes 293 edges as well.

4 Finite certificates and computation

To strengthen the bound at particular dimensions, compare the balls in one parity class at a time. Let P_k count the vertices at graph distance at most k and of the same parity as k from a fixed root. Let H_k be its Hamming counterpart and write $b_k = H_k - P_k$. All $b_k \ge 0$, with $b_0 = b_{n-1} = b_n = 0$ and $b_1 = n - 2 - x_w$. The two parities give $g_k = b_k + b_{k-1}$ and

$$\delta(w) = 2\sum_{k=1}^{n-2} b_k, \qquad s_k = \binom{n}{k} - b_k + b_{k-2} \quad (k \ge 2). \tag{9}$$

Expanding $P_k - P_{k-1}$ into alternating sphere sizes and using $s_{j+1} - s_j \le X_j$ gives $P_k - P_{k-1} \le 1 + \sum_{j<k,\ j\equiv k-1\ (2)} X_j \le t + 1$. The initial terms are $P_1 - P_0 = 1 + x_w$ and $P_2 - P_1 = 1 + s_2 - s_1$. On the cube, Pascal's identity gives $H_k - H_{k-1} = \binom{n-1}{k}$. We get a computable bound $b_k \ge p_k$ by taking

$$p_1 = n - 2 - x_w, \qquad p_k = \left[p_{k-1} + \binom{n-1}{k} - (t+1) \right]_+ \quad (2 \le k \le n-2). \tag{10}$$

Table 1. Lower bounds on $f(n)$ from integer certificates. Each entry is $2^n + t + 1$, where a certificate excludes a graph with $2^n + t$ edges. The verifier records P for each dimension.

n	lower bound	n	lower bound	n	lower bound
9	580	17	140,128	25	35,138,325
10	1,147	18	279,210	26	70,152,107
11	2,271	19	556,449	27	140,081,100
12	4,503	20	1,110,320	28	279,786,071
13	8,941	21	2,214,490	29	558,766,740
14	17,753	22	4,418,155	30	1,116,183,007
15	35,327	23	8,816,922	31	2,229,577,178
16	70,357	24	17,599,972	32	4,454,492,243

The next step uses $M(w) \le \mu$ alongside (10). The details describe how a finite pair of integers can exclude an edge count.

For each positive m in R, choose $2 \le \kappa(m) \le m + 1$. Let u_k count the times k is chosen, taking $u_k = 0$ at unused indices. Set $x = x_w$ and

$$K = x\mathbf{1}_{\{n \text{ even}\}} + \sum_{m \in R,\ m>0} \left(\binom{n}{\kappa(m)} - 2 \right) + u_3(n - 2 - x) - (N/2 - n).$$

Using $Y_m \ge s_{\kappa(m)} - 2$ in (8), and substituting (9), gives $M(w) \ge K + \sum_{k=2}^{n-2}(u_{k+2} - u_k)b_k$. For $\alpha \ge 0$ and $A_k = 2x + \alpha(u_{k+2} - u_k) \ge 0$, it follows that

$$x\delta(w) + 2M(w) \ge 2x(n - 2 - x) + \alpha K - [\alpha - 2]_+\mu + \sum_{k=2}^{n-2} A_k p_k. \tag{11}$$

Indeed, keep $\alpha M(w)$ from the merge term. The remaining $(2 - \alpha)M(w)$ costs at most $[\alpha - 2]_+\mu$.

To check a certificate (t, P), verify for each $x = 1, \dots, n-2$ that some choices of κ, α make the right side of (11) at least $xP/(n-2)$. The case $x = 0$ follows from nonnegativity. Summing over roots would now force $2tP \le (n-2)N\mu$. A certificate demands the strict reverse inequality and hence rules out this size. Adding edges rules out all smaller sizes. For Table 1, only $\alpha \in \{0, 2, 2x\}$ and two choices of κ are needed. One uses $m/2 + 2$ in even dimensions and $(m-1)/2 + 3$ in odd dimensions, with $\kappa(1) = 2$. The other takes $m + 1$ until it reaches $\lceil n/2 \rceil$, then takes that central value and increases by one per remaining radius. The verifier tests these choices with integers.

The same certificate also works below the range in Table 1. The triples (n, t, P) are $(4, 3, 6)$, $(5, 6, 22)$, $(6, 11, 54)$, $(7, 19, 140)$, and $(8, 37, 554)$. In dimension six, taking the largest right side of (11) among these two assignments and three choices of α gives $28, 38, 48, 54$ at $x = 1, 2, 3, 4$, respectively. Each is at least $54x/4$, while $2tP = 1{,}188 > 768 = (n-2)N\mu$. In dimension eight the corresponding values are $134, 226, 314, 398, 478, 554$. Each is at least $554x/6$, while $2tP = 40{,}996 > 29{,}184 = (n-2)N\mu$. This gives $f(6) \ge 76$ and $f(8) \ge 294$ by a direct evaluation of the same inequality.

The other three triples give $f(4) \ge 20$, $f(5) \ge 39$ and $f(7) \ge 148$. Appendix A gives direct combinatorial proofs of the first two. The general lower bound and the construction coincide at $f(2) = 4$ and $f(3) = 9$.

Three radii in dimension seven. A different choice of radii improves the lower bound to 150. Suppose a valid graph has 149 edges. Here $t = 21$ and $\mu = 12$. For $m = 1, 3, 5$ take $\kappa(m) = 2, 3, 4$,

respectively. At a root of excess x, equations (8) and (9) give

$$57 + M = Y_1 + Y_3 + Y_5 \;\geq\; s_2 + s_3 + s_4 - 6 = 90 - x - b_3 - b_4.$$

In this equality the two terms in b_2 cancel and $b_1 = 5 - x$. Keeping b_1, b_3, b_4 in the distance sum gives $\delta \geq 2(b_1 + b_3 + b_4) \geq 76 - 4x - 2M$. For $1 \leq x \leq 5$, applying $M \leq 12$ to the coefficient $2 - 2x$ yields

$$x\delta + 2M \geq 52x - 4x^2 + 24 \geq \frac{184x}{5}.$$

The last inequality follows from $(5 - x)(20x + 24) \geq 0$, and the resulting bound also holds at $x = 0$. Summing over all roots would give $7{,}680 = 5N\mu \geq 184 \cdot 2t = 7{,}728$, a contradiction. This completes the finite bounds in Theorem 3.

The numerical picture in Q_4 comes from two independent exact enumerations, by edge deletion and by symmetry classes. At twenty edges they give 8,320 labelled graphs in 30 classes under coordinate permutations and bit flips. Of these, 15 have stabilizer order one, 13 have order two and two have order twelve. The orbit sizes reconstruct the labelled total as $15{\cdot}384{+}13{\cdot}192{+}2{\cdot}32 = 8{,}320$. A third enumeration supplied with this article checks the same totals by listing edge subsets at the even vertices. The proof in Appendix A is independent of the enumeration.

5 Formal verification

The `Hypercube` Lean 4 library formalizes the potential and antipodal laws, the strict finite bound, the symmetric-chain construction, the blocking bound, the small-dimensional results through $f(8) \geq 292$, and the asymptotic coefficient $3/2$. Its asymptotic theorem is stated with explicit rational quantifiers. The library evaluates the 24 integer pairs in Table 1 and checks the 81-edge graph. The accompanying `SmallCertificates.lean` applies its existing certificate theorem to the five triples above. In particular, Lean proves $76 \leq f(6)$ and $294 \leq f(8)$ by kernel evaluation. The separate `Seven.lean` verifies $150 \leq f(7)$ using the three radii in the preceding proof. The certificate implication alone would require a Boolean hypothesis. Each of the stated finite bounds has that hypothesis discharged in the library.

At source commit `ddd08ca9` the recorded audit covers 237 public lemmas and theorems. It reports zero proof placeholders and zero uses of `native_decide`. Its only axioms are propositional extensionality, classical choice and quotient soundness. The six additional public declarations use the same axioms. The extremal graph enumerations are external. The edge-endpoint derivation (1), the argument by concentration, and the intermediate hand proof of $f(8) \geq 293$ give shorter routes to results with kernel-checked conclusions.

6 Prior work and open questions

Graham and Harary studied diameter-preserving edge deletion [6]. Bouabdallah, Delorme and Djelloul gave the classical upper bound $g(n)$ and the lower bound $N - 1 + \lceil (N-1)/(2n-1) \rceil$ [3]. Their lower second term has coefficient $1/2$ on our scale. Wang, Ho, Ferrero and Sung also record these bounds [8]. Nenadov, Sawhney, Sudakov and Wagner restated the edge-minimization question and proved a lower bound of the same order [7]. Our second-term lower bound is of order N/n while the classical construction gives an upper bound of order $N/\sqrt{n}$. The order of $f(n) - N$, as well as its leading constant, remains open.

Bass and Sudborough tabulate removal counts leaving $18, 35, 69, 137, 272$ edges in dimensions four through eight, respectively [1, 2]. These five numbers coincide with the Bouabdallah–Delorme–Djelloul lower bound on surviving edges. Theorem 2 shows that none is attained.

Charton and coauthors' 81-edge graph shows directly that beating the monotone construction is possible [4]. Whether a valid graph on Q_5 can have 39 edges remains a concrete next case. A second question is whether one can construct valid graphs whose excess has order N/n.

The articles of Wang et al., Nenadov et al., de Bruijn et al. and the later Bass–Sudborough paper were available in full. For the two foundational diameter papers we used abstracts or author summaries and the account of Wang et al. For the earlier Bass–Sudborough paper we checked the original table in the publisher's preview. For the Q_6 construction we checked the relevant section and appendix.

7 Data availability and reproducibility

The supplement contains `recompute.py`, a script for both figures, the Lean certificates, and the claim and reference ledgers. Running `python3 recompute.py` from the paper directory tests the printed witness verbatim, checks the integer certificates against the formal source, and recomputes the dimension-eight contradiction. The graph data in Figure 1 and the integer sums in Figure 2 determine the plotted marks exactly. The formal source and its audit are in `lean/Hypercube`. The README gives the full commands and the computation receipt.

AI tools assisted with drafting, proof exposition, and the accompanying verification code and figures. An independent computational referee supplied the choice of radii in dimension seven and the incidence-row census in the supplement. The named prior constructions and lower bounds retain their cited authorship.

A Direct obstructions in the fourth and fifth cubes

The following proofs of $f(4) \geq 20$ and $f(5) \geq 39$ require no certificate evaluation. Both start from the local excess count (8).

Nineteen edges in Q_4. Suppose $|E| = 19$. Each parity class has excess three. Write $B = \{x > 0\}$, $D = \{x = 0\}$, and let $c(w)$ count vertices of D at graph distance one or two from w. The local inequality $2x_w + X_1 + X_2 \geq 4$ gives

$$4|D| \leq \sum_{w \in B} x_w c(w).$$

Write $b_i = |B \cap S_i(w)|$ and $d_w = 2 + x_w$. Since a graph-distance-two vertex lies at Hamming distance two, (5) gives $c(w) \leq d_w - b_1 + \min(6, d_w + X_1) - b_2$. The branch degrees on each side are either $(3,3,3)$ or $(4,3)$.

If both are $(3,3,3)$, each $c(w) \leq 6$. The required inequality would be $40 \leq 36$. If only one side has a degree-four vertex p, that side contributes at most $2 \cdot 8 + 6$. Let k of the three opposite branch vertices be adjacent to p. A neighbor of p sees the other $k-1$ at distance two, so the opposite contribution is at most $18 + k - k(k-1) \leq 19$. This gives $44 \leq 41$.

For the last case write p, q and p', q' for the vertices of degrees four and three on the two sides respectively. Let a, b, c indicate the edges pp', qp', pq' and let e, e' indicate that the same-side branch pair is at distance two. The preceding bound gives $c(p)+e \leq 8+a(1-c)$ and $c(q)+e \leq 6+b$, and the corresponding primed inequalities interchange b, c. Hence the needed sum is at most

$$44 + 2a(1-b) + 2a(1-c) + b + c - 3e - 3e'.$$

It can reach $4|D| = 48$ only if $a = 1$ and $b = c = e = e' = 0$. But then q sees at most one unit of excess at distance one and none at distance two. Its local inequality would read $4 \leq 3$. This proves $f(4) = 20$.

Thirty-eight edges in Q_5. Here $t = 6$, $\mu = 2$, and the two parity classes each have excess six. Since $M(w) \le 2$ and an extra parent at depth two has weight three, $M_2(w) = 0$ for every w. In particular the graph contains no square. Let $F(w) = \sum_{d(w,u)\ge 4} x_u$ and $Z(w) = \sum_{h(w,u)\le 1} x_u$. From (8), $F(w) + M(w) = Y_1(w) + 1$, while $Z(\bar{w}) \le F(w)$.

Use ordered pairs to define $A = \sum x_u^2$, $T_j = \sum_u x_u X_j(u)$, and $H = \sum_{h(u,v)=4} x_u x_v$. The first three layers in (5) give

$$|\{v : d(u,v) \ge 4\}| = 25 - 3x_u - 2X_1(u) - X_2(u) + M_3(u).$$

Sum the identity for F and interchange u, w on the left. The contribution from Y_1 is $\sum_u (1 + \deg u)x_u = 36 + A$. Rearranging gives

$$4A + 2T_1 + T_2 = 232 + \sum_u x_u M_3(u) + \sum_w M(w) \ge 232.$$

Multiply $Z(\bar{w}) \le Y_1(w) + 1$ by $x_{\bar{w}}$ and sum over w. The neighbor sum on the left includes T_1 and $\sum_w x_w x_{\bar{w}} \le A$. Each neighbor of w is at Hamming distance four from $\bar{w}$, so the remaining term on the right is at most H. It follows that $T_1 \le H + 12$. The pairs defining H and T_2 are disjoint same-parity distinct pairs. Thus $T_2 + H \le 72 - A$ and $232 \le 2A + 168$. It follows that $A \ge 32$.

The positive excesses in one parity class partition six into parts of size at most three. Only $(3, 3)$ has square sum above fourteen, so one class has exactly two degree-five vertices. These must be at Hamming distance four: at distance two their common neighbors make a square. Their neighbor sets are therefore disjoint, giving $T_1 \le 36$. If a is the square sum in the other class, then $A = 18 + a$ and $T_2 \le 36 - a$. Hence $232 \le 180 + 3a$, which forces $a = 18$. Both classes now have the same $(3, 3)$ profile and $T_2 = 0$. The left side of the inequality is at most $4 \cdot 36 + 2 \cdot 36 = 216$, the final contradiction.

References

[1] D. W. Bass and I. H. Sudborough, Removing edges from hypercubes to obtain vertex-symmetric networks with small diameter, *Telecommunication Systems* **13** (2000), 135–146.

[2] D. W. Bass and I. H. Sudborough, Hamilton decompositions and $(n/2)$-factorizations of hypercubes, *J. Graph Algorithms Appl.* **7** (2003), 79–98.

[3] A. Bouabdallah, C. Delorme, and S. Djelloul, Edge deletion preserving the diameter of the hypercube, *Discrete Appl. Math.* **63** (1995), 91–95.

[4] F. Charton, J. S. Ellenberg, A. Z. Wagner, and G. Williamson, PatternBoost: Constructions in mathematics with a little help from AI, arXiv:2411.00566 (2024).

[5] N. G. de Bruijn, C. van Ebbenhorst Tengbergen, and D. Kruyswijk, On the set of divisors of a number, *Nieuw Arch. Wiskd.* (2) **23** (1951), 191–193.

[6] N. Graham and F. Harary, Changing and unchanging the diameter of a hypercube, *Discrete Appl. Math.* **37/38** (1992), 265–274.

[7] R. Nenadov, M. Sawhney, B. Sudakov, and A. Z. Wagner, Bounded degree spanners of the hypercube, *Electron. J. Combin.* **27** (2020), paper P3.3.

[8] J.-J. Wang, T.-Y. Ho, D. Ferrero, and T.-Y. Sung, Diameter variability of cycles and tori, *Inform. Sci.* **178** (2008), 2960–2967.