

Edge multiset dimension of subdivided complete graphs

Pedro M. M. de Castro 

Centro de Informática, Universidade Federal de Pernambuco
Recife, Pernambuco, Brazil
pmmc@cin.ufpe.br

October 26, 2026

Abstract

An edge of a graph can be identified by its distances to a suitable set of vertices. If the reports omit the names of these vertices, only a multiset of distances remains, and finding a suitable set becomes more difficult. We determine the asymptotic cost of this loss of information for subdivided complete graphs. Let K_n be the complete graph on $n \geq 3$ vertices, and let $S_L(K_n)$ replace each edge with a path of L edges, where $L \geq 2$ is an integer. Let edim_m be the minimum number of vertices whose distance multisets distinguish all edges. For every fixed $L \geq 3$, we prove

$$\text{edim}_m(S_L(K_n)) = \Theta_L\left(n^{1+1/\lfloor L/2 \rfloor}\right) \quad (n \rightarrow \infty).$$

Retaining the names requires only $\Theta(n)$ vertices. Without them, the cost is quadratic at length three, has order $n^{3/2}$ at lengths four and five, and has smaller exponents at longer fixed lengths. We also settle existence at the two shortest lengths for every $n \geq 3$: no selection works at length two, whereas at length three at most $\binom{n}{2}$ vertices suffice. The proofs combine a count of integer vectors with realizable vertex labels. Sidon sets distinguish endpoint pairs at odd lengths, and exact finite certificates complete the length-three construction.

Keywords: Edge multiset dimension; graph subdivision; Sidon set; degree sequence; incidence matrix.

Mathematics Subject Classification (2020): Primary 05C12; Secondary 05C07, 11B13.

1 Introduction

Distances from a few observation vertices can identify an unknown edge of a network. Each observation vertex, called a *landmark*, reports its distance to the nearer endpoint of the edge. Kelenc, Tratnik and Yero [12] introduced the edge metric dimension to measure the smallest number of landmarks needed for this task. Their starting point was that a set distinguishing all vertices can still leave two edges indistinguishable. Identifying connections therefore requires its own choice of observation vertices.

The usual record associates each distance with the landmark that supplied it. If this association is lost, the values and their repetitions form a multiset. Ikhlaq et al. [11] introduced edge multiset dimension to study identification from these records. The graph and the landmark set remain known. Their results include families where a small selection suffices and families where no selection works. This raises two questions: when does losing the names make identification impossible, and when it remains possible, how many additional landmarks are required?

We study these questions on subdivisions of complete graphs. For integers $n \geq 3$ and $L \geq 2$, let K_n be the complete graph on n vertices, and let $S_L(K_n)$ replace every edge of K_n with a path of L edges. The replacement paths have disjoint internal vertices. We call the inherited vertices *original* and the others *internal*. Lengthening the paths adds observation positions while preserving the symmetry among the originals, so this family separates the freedom to place landmarks from the difficulty of distinguishing equivalent parts of the graph.

Subdivisions already serve as examples for comparing distance identification parameters. Knor, Škrekovski and Yero [13, Proposition 2.5] determined the named-distance minimum for $S_2(K_n)$, which grows linearly with n , in their study of vertex and edge identification in 2-connected graphs. For multiset records, the survey of Farhan et al. [9, Section 6.1] gives computed values for small subdivisions and asks whether $S_2(K_n)$ has no resolving set for every $n \geq 3$ [9, Problem 24]. These results leave a broader issue: how does the replacement length affect both existence and the cost of identification without names?

Consider $S_3(K_5)$, with 25 vertices and 30 edges (Fig. 1). Choose the five originals as landmarks. A middle edge is one step from the two originals at the ends of its path and four steps from the other three. The two reports equal to one identify the path when their names are retained. Once the names are removed, all ten middle edges give $\{1, 1, 4, 4, 4\}$. The record still identifies the middle position, leaving ten possibilities for the path.

Choosing more vertices does not by itself settle this ambiguity. If every vertex is selected, a permutation of the five originals still exchanges paths while preserving the selection. A successful choice must break this symmetry and distinguish all 30 edges. Ten suitably placed landmarks achieve this, as the construction in Section 6.2 will show.¹

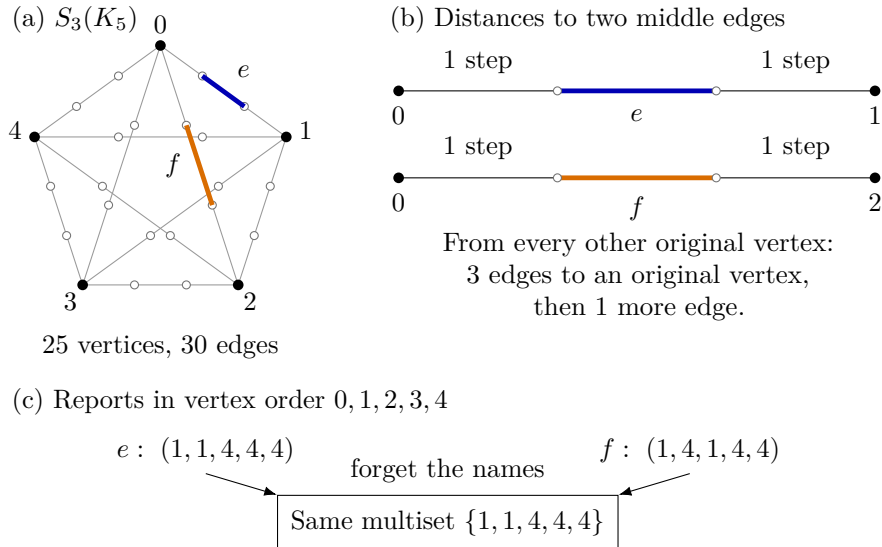


Figure 1: Losing the names of the landmarks in $S_3(K_5)$. In (a), the five filled vertices are selected; the highlighted middle edges are different. In (b), each is one step from the original vertices at the ends of its replacement path. Every other original vertex is four steps away. In (c), the two ordered records become the same multiset. Crossings without a vertex do not join paths.

¹An interactive demonstration of distance reports and landmark choices is available at <https://www.cin.ufpe.br/~pmmc/demos/edge-multiset/>.

We now make the measurement precise. Throughout, graphs are finite, simple, connected and unweighted. For a graph G with vertex set $V(G)$ and edge set $E(G)$, let $d(u, v)$ denote the number of edges in a shortest path between vertices $u, v \in V(G)$. For an edge $e = xy \in E(G)$ and a vertex $s \in V(G)$, put

$$d(e, s) = \min\{d(x, s), d(y, s)\}.$$

Fig. 2 illustrates the replacement length and this distance on an expanded path.

A landmark set $W \subseteq V(G)$ is *edge multiset resolving* if the multisets $\{d(e, s) : s \in W\}$ differ for every two distinct edges $e, f \in E(G)$. Each selected vertex contributes one entry; equal distances retain their repetitions. The least size of such a set is the *edge multiset dimension* $\text{edim}_m(G)$, with value ∞ if no such set exists. Retaining the names of the landmarks instead gives the *edge metric dimension* $\text{edim}(G)$. In both definitions, each vertex may be selected at most once.

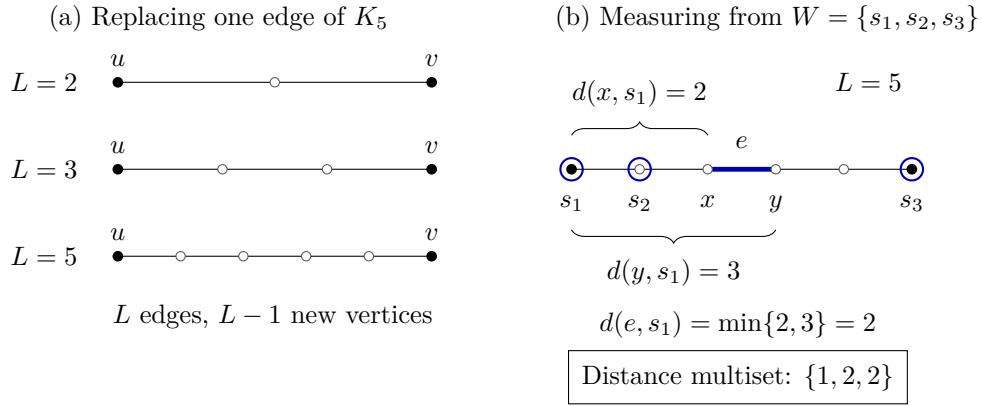


Figure 2: Replacement length and distance to an edge. In (a), the same original edge uv is replaced by paths of two, three and five edges. This operation is applied to every edge of K_5 , keeping its five original vertices. In (b), the circled vertices form $W = \{s_1, s_2, s_3\}$, and $e = xy$ is the highlighted edge on a length-five path. The distances from e to s_1, s_2, s_3 are two, one and two, respectively, giving $\{1, 2, 2\}$. Filled vertices are original.

Theorem 1.1. *For every $n \geq 3$,*

$$\text{edim}_m(S_2(K_n)) = \infty, \quad \text{edim}_m(S_3(K_n)) \leq \binom{n}{2}.$$

For every fixed integer $L \geq 3$,

$$\text{edim}_m(S_L(K_n)) = \Theta_L(n^{1+1/\lfloor L/2 \rfloor}) \quad \text{as } n \rightarrow \infty,$$

where the constants and the sufficient order may depend on L .

The lower bound is proved in Section 3, the matching constructions for $L \geq 4$ in Sections 4 and 5, and the two short-length statements in Section 6, including its finite certificates. The first statement answers the length-two question above. Length three then marks a change in existence at every order: a suitable selection always exists, and its minimum size grows quadratically in n .

At lengths four and five the minimum has order $n^{3/2}$, at lengths six and seven it has order $n^{4/3}$, and longer fixed lengths give smaller exponents. With the landmark names retained, only $\Theta(n)$ vertices are needed at every length (Proposition 2.2). Thus the theorem gives matching bounds for the cost of losing correspondence throughout this family, beyond the previously

computed finite cases. It also identifies how additional positions along the paths reduce that cost.

The lower bound comes from comparing edges next to the same original vertex. After accounting for the choices on each edge's own path, the changing part depends on landmark counts in the first half of the other endpoint's paths. Many different count vectors require many selected vertices. For an upper bound, we choose these vectors and realize them by selecting internal vertices. The middle edge of an odd-length path sees the sum of its two endpoint vectors, so these sums must identify the pair. Sidon sets supply this property. Length three requires a separate matrix construction, completed at the remaining finite orders by exact certificates. Section 7 gives finite existence ranges and quantifies restrictions on the available landmark positions. The corresponding refinements use exact lattice counts, label constructions and probabilistic bounds, developed in Appendices C–F.

2 What the distance multiset records

We first express edge distances through landmark positions on the replacement paths. This will isolate the counts that distinguish edges and turn equality of multisets into equations for those counts.

Throughout the article, $n \geq 3$ and $L \geq 2$ are integers. Index the original vertices by $0, \dots, n-1$. Sums over original vertices and the rows and columns of binary matrices encoding landmark selections use these indices. For distinct originals i, j and an integer $0 \leq t \leq L$, let $x_{ij}(t)$ be the vertex in position t on the path from i to j . Thus $x_{ij}(0) = i$, $x_{ij}(L) = j$, and $x_{ij}(t) = x_{ji}(L-t)$. We call ij the *support* of each edge on this path.

Orient an edge so that its nearer original endpoint is i , and write

$$e = x_{ij}(a)x_{ij}(a+1), \quad 0 \leq a \leq \lfloor (L-1)/2 \rfloor, \quad b = L-1-a.$$

Here a and b are the distances from e to i and j , respectively, and $a \leq b$. We call a the *edge depth* and put $\delta = b - a$; either orientation is permitted when $a = b$. An internal vertex $x_{ij}(t)$ has *vertex depth* $\min(t, L-t)$.

The distances depend on how the landmark's support meets ij . On ij itself, the endpoints of e are in positions a and $a+1$, as shown in Fig. 3. If $t \leq a$, the nearer endpoint is in position a , at distance $a-t$; if $t \geq a+1$, it is in position $a+1$, at distance $t-(a+1)$. These two cases cover every vertex because t is an integer. In particular, the distance is zero at either endpoint of e .

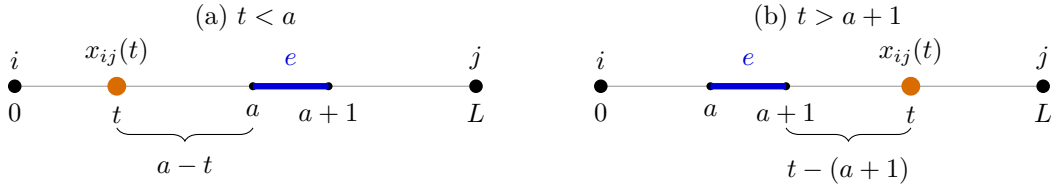


Figure 3: Distance on the edge's own support. Positions are measured from i . In (a), the nearest endpoint of e is in position a ; in (b), it is in position $a+1$. Each brace measures the distance to that endpoint. Moving the orange landmark to either endpoint gives distance zero. The paths are schematic.

For the other supports, compare routes through their original endpoints, using distance L between distinct originals. Fig. 4 shows the two configurations where both endpoint routes can matter.

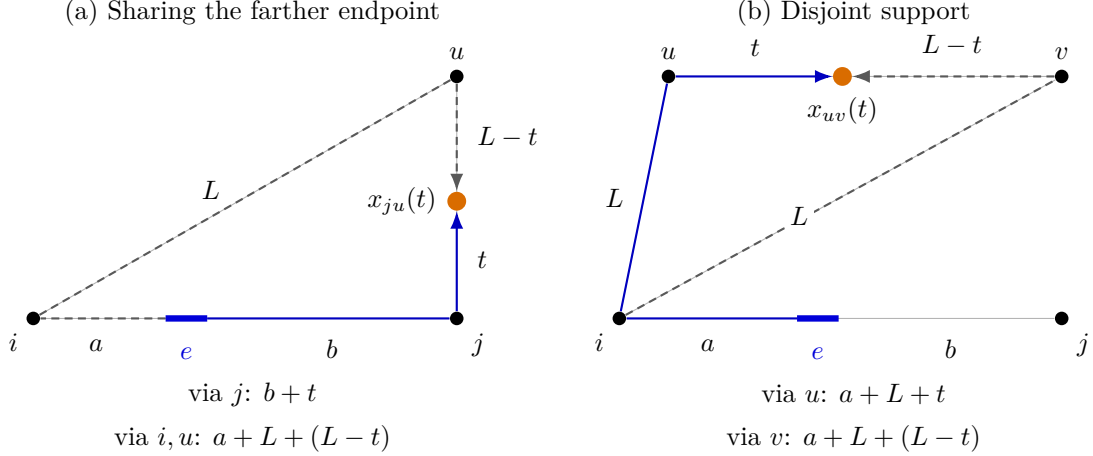


Figure 4: Routes to a landmark on another support. In (a), entering ju through j or u gives lengths $b + t$ and $a + L + (L - t)$. In (b), entering the disjoint support through u or v gives $a + L + t$ or $a + L + (L - t)$. Orange vertices are landmarks. Each complete support has length L ; intermediate positions are schematic.

Table 1 collects these distances. In its first row, leaving the support of e cannot shorten the segment drawn in Fig. 3; in the remaining rows, the shorter endpoint route gives the stated value.

Table 1: Distances from e , with $a \leq b$, $a + b = L - 1$, $1 \leq t < L$, and distinct $u, v \notin \{i, j\}$.

Landmark	Distance from e
$x_{ij}(t)$, on the support	$\begin{cases} a - t, & t \leq a, \\ t - (a + 1), & t \geq a + 1 \end{cases}$
$x_{iu}(t)$, sharing i	$a + t$
$x_{ju}(t)$, sharing j	$\min\{b + t, 2L + a - t\}$
$x_{uv}(t)$, on a disjoint support	$L + a + \min(t, L - t)$
Originals i, j, u	$a, b, L + a$, respectively

Fix a landmark set $W \subseteq V(S_L(K_n))$. To record which internal vertices are selected, define

$$\eta_{ij}(t) = \begin{cases} 1, & x_{ij}(t) \in W, \\ 0, & x_{ij}(t) \notin W, \end{cases} \quad i \neq j, \quad 1 \leq t < L.$$

The two orientations describe the same vertex, so $\eta_{ij}(t) = \eta_{ji}(L - t)$. Summing these indicators over paths leaving i gives the *endpoint profile*

$$D_i(t) = \sum_{j \neq i} \eta_{ij}(t), \quad 1 \leq t < L.$$

Thus $D_i(t)$ is the number of landmarks in position t on paths leaving i ; it summarizes their individual choices in one integer.

We will count landmarks at each distance from an edge. For $e \in E(S_L(K_n))$ and a nonnegative integer h , put

$$H_e^W(h) = |\{s \in W : d(e, s) = h\}|, \quad H_e(z) = \sum_{h \geq 0} H_e^W(h) z^h.$$

Here z is an indeterminate and $H_e \in \mathbb{Z}[z]$; the fixed set W is suppressed in H_e . Equality of these histograms is exactly equality of the distance multisets. The polynomial collects the integer counts: its exponent is a distance and its coefficient is a number of landmarks.

For a landmark on a support disjoint from ij , only its vertex depth matters. Use this distance as a common baseline. A landmark on a support sharing i or j changes one count from that baseline; the following polynomials record the change:

$$\begin{aligned} A_t(z) &= z^{a+t} - z^{L+a+\min(t, L-t)}, \\ B_t(z) &= z^{\min(b+t, 2L+a-t)} - z^{L+a+\min(t, L-t)}, \quad 1 \leq t \leq L. \end{aligned} \quad (1)$$

Although suppressed in the notation, these polynomials depend on a, b, L . We also regard each polynomial as a column vector indexed by distance, so its coefficient at z^h is the change at distance h .

Let T_s count all internal landmarks of depth s , where $1 \leq s \leq k = \lfloor L/2 \rfloor$. For the original vertices, put $\alpha_i = \mathbf{1}_{\{i \in W\}}$ and $A_0 = \sum_i \alpha_i$. We use $\mathbf{1}_{\{\mathcal{P}\}} = 1$ when the condition $\mathcal{P}$ holds and 0 otherwise. The next identity begins with all landmarks at the disjoint-support baseline and corrects the two families of paths through the endpoints of e . Landmarks on ij belong to both families, so one final correction restores their actual distances.

Lemma 2.1. *For a fixed landmark set W , let $\eta_{ij}(t)$ be its internal selection indicators as defined above. For $e = x_{ij}(a)x_{ij}(a+1)$, its histogram is*

$$\begin{aligned} H_e(z) &= \sum_{s=1}^k T_s z^{L+a+s} + \sum_{t=1}^{L-1} (D_i(t)A_t(z) + D_j(t)B_t(z)) + E_e(z) \\ &\quad + \alpha_i z^a + \alpha_j z^b + (A_0 - \alpha_i - \alpha_j) z^{L+a}, \end{aligned} \quad (2)$$

where the correction from the support of e is

$$\begin{aligned} E_e(z) &= \sum_{t=1}^{L-1} \eta_{ij}(t) [z^{\min(|t-a|, |t-(a+1)|)} - z^{a+t} \\ &\quad - z^{\min(b+L-t, L+a+t)} + z^{L+a+\min(t, L-t)}]. \end{aligned} \quad (3)$$

With $\|\cdot\|_\infty$ denoting the largest absolute coefficient, any two edges e, f of the same graph, using the same set W , satisfy

$$\|E_e\|_\infty \leq 2, \quad \|E_e - E_f\|_\infty \leq 4. \quad (4)$$

Proof. Initially count each internal landmark as though its support were disjoint from ij . The columns A_t, B_t correct the two families of paths, using Table 1. A landmark on ij appears in both corrections; the summand in (3) restores its actual distance. Original landmarks give the last line of (2).

For the coefficient bound, each of the two positive exponents in (3) has at most two preimages as t varies. Their ranges are disjoint: the first is at most b , whereas the second is at least $L+a+1 > b$. Their combined contribution to any coefficient is therefore at most two. The exponent $a+t$ is injective and at most $L+a-1$. The other negative exponent is

$$L+a+\min(\delta-t, t).$$

Its increasing branch starts at $L+a+1$, above the first negative range, and its decreasing branch is injective. Thus the negative contribution also has magnitude at most two at every coefficient. Subtracting these contributions proves the first bound; the second follows by the triangle inequality. $\square$

We will recover the edge depth before comparing its endpoint profiles. All distances in Table 1 are at most $L+a+k$. If a depth- k landmark lies on a disjoint support, it attains this value, so

$$\max\{t : H_e^W(t) > 0\} = L+a+k. \quad (5)$$

The constructions in Sections 4 and 5 ensure such a landmark for every support. Equation (5) then determines a , and hence b and δ .

Proposition 2.2. *For $n \geq 3$ and $L \geq 2$,*

$$\left\lceil \frac{n-1}{2} \right\rceil \leq \text{edim}(S_L(K_n)) \leq n.$$

Proof. The original vertices resolve every edge: the two smallest named distances identify its support endpoints, and either endpoint distance specifies its position. This proves the upper bound. For the lower bound, associate to each landmark its original vertex, or the two endpoints of its support if it is internal. Fewer than $(n-1)/2$ landmarks leave at least two originals unassociated. Interchanging these originals induces a nontrivial automorphism that fixes every landmark and moves an edge of their replacement path. Those two edges have equal named distance records, a contradiction. $\square$

3 Counting possible distance records

To bound the number of landmarks from below, compare the $n-1$ edges next to a fixed original vertex i , one on each path ij . The selection around i is common to all comparisons; the changing part comes from the first $k = \lfloor L/2 \rfloor$ positions measured from j . The remaining positions give the same distances as a path sharing neither endpoint with ij . Choices on each edge's own path allow only bounded repetitions of a count vector, so many different vectors force a large total count. Fig. 5 shows the cancellation for $L = 6$, when the first three positions remain.

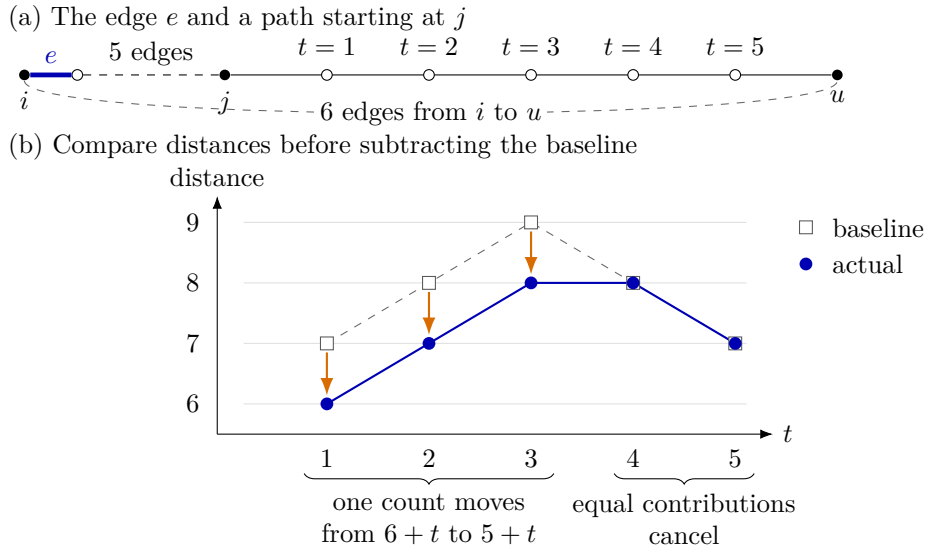


Figure 5: The three counts that matter when $L = 6$. In (a), e is the edge next to i , and t measures position from j . In (b), squares give the distances for the comparison path and circles give the actual distances. At $t = 1, 2, 3$, the orange arrows move one count from distance $6 + t$ to $5 + t$. At $t = 4, 5$, the values agree. Thus the changing part depends only on $D_j(1), D_j(2), D_j(3)$.

The relevant counting fact is elementary. Call the coordinate sum of a nonnegative integer vector its *cost*. For a positive integer d and a real $r \geq 0$, the number of vectors in $\mathbb{Z}_{\geq 0}^d$ with coordinate sum at most r is $\binom{\lfloor r \rfloor + d}{d}$, the Ehrhart count of the standard simplex [3]. It is at most $(r+1)^d$, because each coordinate is at most $\lfloor r \rfloor$. For example, in two coordinates, allowing each vector twice gives two occurrences of cost zero, four of cost one and six of cost

two. To choose fifteen occurrences, three more of cost three are needed. Their total cost is 25 (Fig. 6).

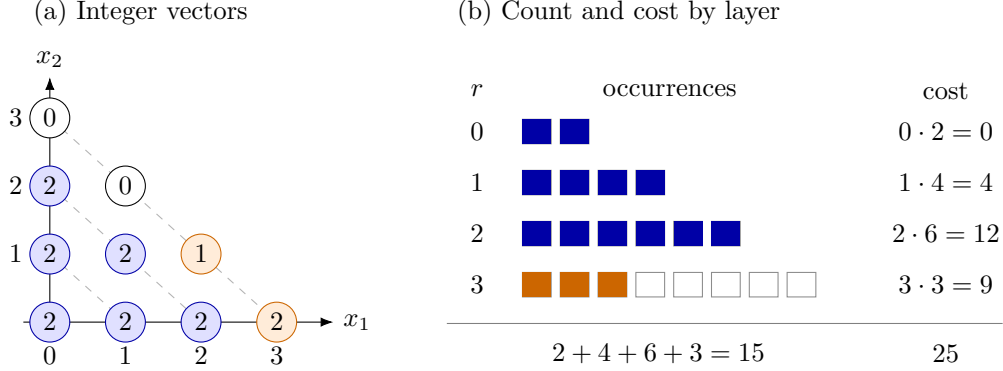


Figure 6: Fifteen occurrences of integer vectors, with at most two per vector. In (a), the number inside a point is its number of uses. In (b), squares count the allowed occurrences in the layer $r = x_1 + x_2$; filled squares are used. The last column gives each layer's contribution to the total cost 25.

Theorem 3.1. *Let $L \geq 2$, put $k = \lfloor L/2 \rfloor$, $Q = 2^{2k+1}$, and suppose $n > Q + 1$. Every edge multiset resolving set in $S_L(K_n)$ contains at least*

$$\frac{n}{2} \left(\left(\frac{n}{2Q} \right)^{1/k} - 1 \right)$$

internal vertices. In particular, for each fixed $L \geq 3$, $\text{edim}_m(S_L(K_n)) = \Omega_L(n^{1+1/k})$ whenever a resolving set exists.

Proof. Let B be the number of selected internal vertices. For the edge e_{ij} next to i , the far-endpoint columns in (1) are

$$B_t(z) = \begin{cases} z^{L-1+t} - z^{L+t}, & 1 \leq t \leq k, \\ 0, & k < t < L. \end{cases}$$

Thus (2) becomes

$$H_{e_{ij}}(z) = C_i(z) + \alpha_j(z^{L-1} - z^L) + \sum_{t=1}^k D_j(t)(z^{L-1+t} - z^{L+t}) + E_{e_{ij}}(z),$$

where C_i is independent of j , and the support correction $E_{e_{ij}}$ in (3) depends only on selections along ij .

For odd $L = 2k + 1$, assign to j the vector $p_j = (D_j(1), \dots, D_j(k))$. Once i, p_j are fixed, only α_j and the $2k$ choices on ij can change the histogram. There are at most Q possibilities. For even $L = 2k$, use $p_j = (D_j(1), \dots, D_j(k-1), \lfloor D_j(k)/2 \rfloor)$. The parity of $D_j(k)$, together with α_j and the $2k - 1$ choices on ij , again gives at most Q possibilities.

Consequently no vector occurs more than Q times among the p_j . Indeed, for a larger proper class of equal vectors, choose i outside that class; two of the corresponding edges would have equal histograms. If all n vectors agree, any i gives $n - 1 > Q$ edges and the same contradiction.

The cost $\|p_j\|_1$ of a profile is the sum of its coordinates. At odd length, each internal landmark contributes once to $\sum_j \|p_j\|_1$, so the sum is B . At even length, a middle landmark is counted at both ends before the last coordinate is divided by two. Rounding down can only decrease the sum. Hence $\sum_j \|p_j\|_1 \leq B$ in both cases.

At least half the n profiles consequently have cost at most $2B/n$. There are at most $(1 + 2B/n)^k$ possible vectors of this cost, and each occurs at most Q times. Counting these profiles gives

$$\frac{n}{2} \leq Q \left(1 + \frac{2B}{n}\right)^k.$$

Solving for B gives the required number of internal landmarks. $\square$

Corollary 3.2. *For $n \geq 10$, every resolving set in $S_3(K_n)$ contains at least $n^2/16 - n/2$ internal vertices.*

Proof. In the preceding proof $k = 1$, $Q = 8$, and $\sum_j p_j = B$. The minimum sum of $n = 8q + r$ nonnegative integers, each used at most eight times, fills $0, \dots, q-1$ eight times and then uses q another r times. Therefore

$$B \geq 4q(q-1) + rq = \frac{n^2}{16} - \frac{n}{2} + \frac{r(8-r)}{16}, \quad 0 \leq r < 8.$$

$\square$

The same filling argument gives an exact finite bound in any dimension. Lemma C.1 computes the minimum total cost by filling successive layers of integer vectors, and Corollary C.3 applies this count to the endpoint profiles above. These refinements are proved in Appendix C.

4 Symmetric endpoint profiles

The lower bound suggests assigning a different short integer label to each original vertex and realizing its coordinates as landmark counts. We first do this with symmetric profiles. At each original vertex i , we require equal landmark counts in positions s and $L-s$, measured along all paths leaving that same vertex. For a fixed depth s , these counts will form the degree sequence of a simple graph. Selecting the corresponding path positions turns these graphs into the landmark set.

The main issue is recovery. An off-center edge sees the two endpoint labels differently, and the equations below recover them separately. A central edge at an odd length sees only their sum. Thus arbitrary distinct labels suffice at even lengths, while symmetric profiles at odd lengths require distinct pair sums. This distinction will explain why the odd-length construction of Section 5 uses asymmetric profiles.

Put $k = \lfloor L/2 \rfloor$. We seek a selection with endpoint counts

$$D_i(s) = D_i(L-s) = d_{i,s} \quad (1 \leq s < L/2).$$

If $L = 2k$, the midpoint in position k has count $D_i(k) = d_{i,k}$. For each pair of positions $s, L-s$ with $1 \leq s < L/2$, combine their contributions from (1) into $\Phi_s = A_s + A_{L-s}$ and $\Psi_s = B_s + B_{L-s}$. Explicitly,

$$\begin{aligned} \Phi_s &= z^{a+s} + z^{a+L-s} - 2z^{L+a+s}, \\ \Psi_s &= z^{b+s} + z^{\min(b+L-s, L+a+s)} - 2z^{L+a+s}. \end{aligned}$$

When $L = 2k$, position k is a single midpoint, so its contribution occurs once. Thus $\Phi_k = A_k$ and $\Psi_k = B_k$, namely

$$\Phi_k = z^{a+k} - z^{L+a+k}, \quad \Psi_k = z^{b+k} - z^{L+a+k}.$$

The histogram for a selection without original vertices is

$$H_e(z) = \sum_{s=1}^k T_s z^{L+a+s} + \sum_{s=1}^k (d_{i,s} \Phi_s + d_{j,s} \Psi_s) + E_e(z).$$

Lemma 4.1 (Recovery from symmetric profiles). *For $u = (u_1, \dots, u_k)$ and $v = (v_1, \dots, v_k)$ in $\mathbb{R}^k$, consider the real linear map $\mathcal{S}_a : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}[z]$ given by*

$$\mathcal{S}_a(u, v) = \sum_{s=1}^k (u_s \Phi_s + v_s \Psi_s).$$

If $\delta = b - a > 0$, this map is injective. If $\delta = 0$, which occurs only for odd L , its kernel is $\{(u, v) : u + v = 0\}$.

Proof. Suppose $\mathcal{S}_a(u, v) = 0$. Reading the coefficients at $a + 1, \dots, a + k$ gives

$$u_s = 0 \quad (s \leq \min(k, \delta)), \quad u_s + v_{s-\delta} = 0 \quad (\delta < s \leq k).$$

Every other monomial of the columns has exponent above this range. For $\delta = 0$, the two column families agree; their distinct first exponents show that their common map is injective. Thus the kernel is exactly $u + v = 0$.

If $\delta \geq k$, the low coefficients give $u = 0$. Each Ψ_s has smallest exponent $b + s$ with positive coefficient. These exponents strictly increase, so elimination in increasing order gives $v = 0$.

It remains to consider $0 < \delta < k$. For a paired coordinate s , the coefficient at $L + a + s$ gives

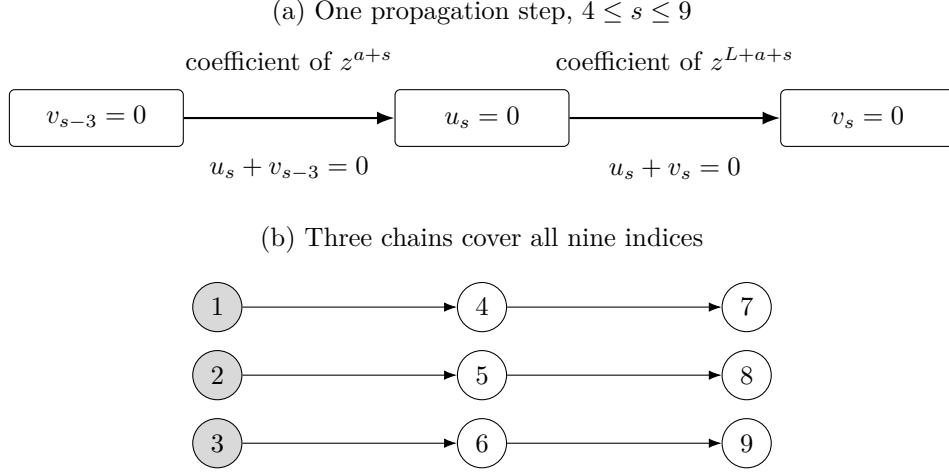
$$\begin{aligned} u_s + v_s &= 0, & 2s > \delta, \\ 2u_s + v_s &= 0, & 2s = \delta, \\ 2u_s + v_s - v_{\delta-s} &= 0, & 2s < \delta. \end{aligned}$$

Indeed, the negative terms contribute $-2(u_s + v_s)$. A positive second term of Ψ_t reaches this exponent precisely when $\min(\delta - t, t) = s$. For $2s < \delta$, the two indices are $s, \delta - s$; for $2s = \delta$, there is the single index s ; for $2s > \delta$, there is none. The relevant indices in the first two cases are below $\delta < k$, so they are always paired coordinates. The equality case can occur only when L is odd, since then δ is even. For even L , the middle coordinate instead gives

$$u_k + v_k = 0.$$

No positive second term reaches $L + a + k$.

The low coefficients first give $u_1 = \dots = u_\delta = 0$. For $\delta/2 < s \leq \delta$, the high equations give $v_s = 0$. They also give $v_s = 0$ when $2s = \delta$. For $s < \delta/2$, the coordinate $v_{\delta-s}$ is already zero, so $v_s = 0$ follows. Finally, for $s > \delta$, the earlier coordinate $v_{s-\delta} = 0$ gives $u_s = 0$ through the low equation. The high equation, or the middle equation if applicable, then gives $v_s = 0$. Induction proves injectivity. Fig. 7 displays these dependencies for one even length. $\square$



Index s means $u_s = v_s = 0$. Shading marks the initial zeros.

Figure 7: Recovery in Lemma 4.1, for $L = 18$, $k = 9$, $a = 7$, $b = 10$, and $\delta = 3$. In (a), the two indicated coefficients successively force $u_s = 0$ and $v_s = 0$ once $v_{s-3} = 0$. The final step $s = 9$ uses the single middle-position columns. In (b), the shaded indices are the initial zeros, and the same step reaches all remaining indices.

Proposition 4.2 (Realizing symmetric codes). *Let $L \geq 4$, $k = \lfloor L/2 \rfloor$, and let Q be a positive integer with $n \geq 14Q$. Suppose that n labels*

$$c_i \in \{0, \dots, Q-1\}^k$$

are distinct when L is even. When L is odd, suppose $c_i + c_j = c_u + c_v$ implies $\{i, j\} = \{u, v\}$ for all pairs with $i \neq j$ and $u \neq v$. Then $S_L(K_n)$ has an edge multiset resolving set with symmetric profiles and size less than

$$6(L-1)nQ.$$

Proof. Choose the target degrees

$$d_{i,s} = 6(Q + c_{i,s}).$$

For each coordinate these degrees have even sum and lie between $m_0 = 6Q$ and $M_0 = 12Q - 6 < n$. Lemma B.2 applies, since

$$\frac{(M_0 + m_0 + 1)^2}{4m_0} = \frac{(18Q - 5)^2}{24Q} < 13.5Q < n.$$

Thus there is a simple graph G_s with these degrees. For each edge of G_s , select the two positions $s, L-s$ if $s < L/2$. If $L = 2k$ and $s = k$, select its single middle vertex. The selections use distinct vertices for different coordinates and realize the required profiles.

For a unified count, let $w_s = 2$ at a paired coordinate and $w_k = 1$ at a middle coordinate. The number of depth- s landmarks is

$$T_s = \frac{w_s}{2} \sum_i d_{i,s} \geq 3w_s nQ.$$

Among these, at most

$$w_s(d_{i,s} + d_{j,s}) < 24w_s Q$$

have supports meeting a fixed pair i, j . Since $n \geq 14Q > 8$, there is a depth- k landmark on a support disjoint from every given support. The largest occupied histogram distance is therefore $L + a + k$, and it identifies a .

For two edges of equal depth, subtract their histogram identities. The global terms cancel. Every coefficient of the difference of the linear terms is divisible by six, whereas the support-correction difference has coefficient norm at most four. Equality of the histograms therefore forces equality of the linear terms.

For $\delta > 0$, Lemma 4.1 recovers the two degree vectors in order. The labels are distinct in both parities of L : in the odd case, $c_i = c_j$ with $i \neq j$ would give $c_i + c_h = c_j + c_h$ for a third index h , contrary to the pair-sum hypothesis. Distinct labels therefore determine the ordered endpoints. For $\delta = 0$, it recovers their sum. This pair-sum property determines the unordered endpoints, and their support has one central edge. This distinguishes every edge. Finally,

$$|W| = \sum_s \frac{w_s}{2} \sum_i d_{i,s} < 6nQ \sum_s w_s = 6(L-1)nQ.$$

□

Corollary 4.3. *Let $L = 2k \geq 4$ and $Q = \lceil n^{1/k} \rceil$. If $n \geq 14Q$, then*

$$\text{edim}_m(S_{2k}(K_n)) < 6(2k-1)nQ.$$

For fixed $k \geq 2$, this is $O_k(n^{1+1/k})$.

Proof. There are $Q^k \geq n$ labels in the box. Use any n distinct labels in Proposition 4.2. □

5 Odd lengths: Sidon labels and projected kernels

Let $L = 2k + 1$, with $k \geq 2$. Symmetric profiles require $\Theta_k(n^{1+2/k})$ landmarks when $k \geq 3$ (Corollary D.1). To reach the exponent $1 + 1/k$, we let the counts at opposite positions vary separately. This gives $2k$ coordinates for each endpoint profile. At the central edge, $a = b = k$, so the two endpoint columns in (1) coincide. Their contribution records only the sum of the endpoint profiles. We will choose Sidon labels so that this sum identifies the pair. Away from the center, one endpoint's change can be cancelled by a change at the other. We must also prevent these cancellations. For each edge position, we describe the changes that can be hidden and bound the dimension of the space they form. This bound will allow us to discard the conflicting pairs from a Sidon set while keeping enough labels for all n originals.

5.1 The linear information retained by a noncentral edge

A noncentral edge may hide a change at one endpoint through cancellation at the other. We bound the dimension of these changes by viewing the histogram columns as edges between distance levels and counting the connected components of this auxiliary graph.

Write $\delta = 2r$, where $1 \leq r \leq k$, and remove the common factor z^a from the columns. Let e_j be the coordinate vector for exponent j in a space with coordinates $1, \dots, 3k + 1$. Denote the resulting matrices by $\mathcal{A}$ and $\mathcal{B}_r$. Their columns are

$$\begin{aligned} \mathcal{A}_t &= e_t - e_{L+\min(t, L-t)}, & 1 \leq t \leq 2k, \\ (\mathcal{B}_r)_t &= e_{t+2r} - e_{L+\min(t, L-t)}, & 1 \leq t < L-r, \\ (\mathcal{B}_r)_t &= 0, & L-r \leq t \leq 2k. \end{aligned} \tag{6}$$

Each nonzero column has one entry $+1$ and one entry -1 , so we can draw it as an edge joining their coordinate indices. These indices are distance levels shifted by a , because we removed z^a . The incidence-matrix rank formula converts the component count into the dimension bound

below. Allikvere [2, Lemma 11] uses such graphs in a probabilistic analysis of hypercube histograms. Fix a field $\mathbb{F}$, and set

$$U_r = \{u \in \mathbb{F}^{2k} : \mathcal{A}u + \mathcal{B}_r v = 0 \text{ for some } v \in \mathbb{F}^{2k}\}. \quad (7)$$

A vector in U_r is therefore a change at the nearer endpoint whose effect can be cancelled at the farther endpoint. To recover the nearer endpoint, we need labels whose nonzero differences avoid every U_r .

Lemma 5.1. *Over every field, $\mathcal{A}$ is injective, $\ker \mathcal{B}_r$ consists of vectors supported on their last r coordinates, and*

$$\dim U_r \leq k - 1.$$

Proof. The $\mathcal{A}$ -columns form k disjoint two-edge stars. Star s has leaves $s, L - s$ and center $L + s$, for $1 \leq s \leq k$; exponent L is isolated. Thus $\text{rank } \mathcal{A} = 2k$. Each nonzero $\mathcal{B}_r$ -column has smallest nonzero coordinate $t + 2r$, with coefficient one. These coordinates strictly increase with t , so these columns are independent. Hence $\text{rank } \mathcal{B}_r = 2k - r$ and its kernel is as stated.

To bound $\dim U_r$, we now show that adding the $\mathcal{B}_r$ -columns joins enough of the stars. Figure 8 follows this step for $k = 3, r = 1$: the columns $(\mathcal{B}_1)_5 = e_7 - e_9$ and $(\mathcal{B}_1)_4 = e_6 - e_{10}$ leave two components, and $(\mathcal{B}_1)_3 = e_5 - e_{10}$ connects them.

Contract the $\mathcal{A}$ -stars to vertices $1, \dots, k$, and call the isolated exponent L vertex 0. For every $j > r$, the column $(\mathcal{B}_r)_{L-j}$ joins j to $|j - 2r| < j$. These edges form a forest with roots $0, \dots, r$. The component of an index is determined by its residue modulo $2r$, up to sign: each step preserves that class and strictly decreases the index until it reaches its unique representative in $\{0, \dots, r\}$.

The additional column $(\mathcal{B}_r)_k$ joins k to $|k + 1 - 2r|$. Its endpoints lie in different forest components. Otherwise $k \equiv \pm(k + 1) \pmod{2r}$, which would require the even integer $2r$ to divide either one or $2k + 1$. The contracted graph therefore has at most r components.

Contraction preserves the number of components, so the original incidence graph on $3k + 1$ vertices also has at most r components. An oriented incidence matrix on v vertices with c components has rank $v - c$ over every field, as follows by taking a spanning forest and eliminating its leaves. Consequently,

$$\text{rank}[\mathcal{A} \ \mathcal{B}_r] \geq 3k + 1 - r.$$

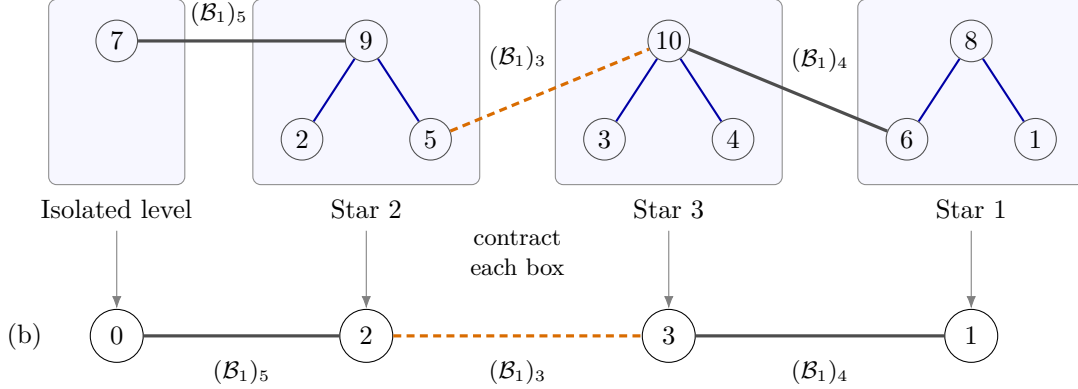
Since $\mathcal{A}$ is injective, it maps U_r bijectively to the intersection of the two column spaces. Thus

$$\dim U_r = \text{rank } \mathcal{A} + \text{rank } \mathcal{B}_r - \text{rank}[\mathcal{A} \ \mathcal{B}_r] \leq k - 1.$$

□

(a) A graph of coefficient levels

For example, $(\mathcal{B}_1)_5 = e_7 - e_9$ gives the edge from 7 to 9.



The solid edges leave two components. The dashed edge joins them.

$$\begin{array}{c} \text{The shown columns connect all ten levels.} \\ \text{rank}[\mathcal{A} \ \mathcal{B}_1] \geq 10 - 1 = 9 \quad \implies \quad \dim U_1 \leq 6 + 5 - 9 = 2. \end{array}$$

Figure 8: How connectivity bounds the ambiguity in Lemma 5.1, for $L = 7$, $k = 3$, $r = 1$, and $a = 2$. In (a), the vertices are coefficient indices after removing z^2 . The six blue edges are the $\mathcal{A}$ -columns; three $\mathcal{B}_1$ -columns connect their boxed components. In (b), each box becomes one vertex, retaining those connections. The solid edges leave two components and the dashed edge joins them. Thus the displayed columns connect all ten levels, giving rank at least nine. Since $\mathcal{A}$ and $\mathcal{B}_1$ have ranks six and five, at most two independent changes at one endpoint can be cancelled by changes at the other.

5.2 Removing the ambiguous differences

We remove labels whose differences lie in the spaces U_r , so that cancellation cannot hide the nearer endpoint's label. For an abelian group, call a subset C *Sidon* if every equality $c_1 + c_2 = c_3 + c_4$ between elements of C gives the same unordered pair, including repetitions. Every nonzero difference then has at most one ordered representation, so counting the forbidden differences bounds how many labels must be discarded.

Lemma 5.2. *Let q be an odd prime power and d, s positive integers. Let $C \subseteq \mathbb{F}_q^d$ be Sidon, and let $U_1, \dots, U_s$ be subspaces of $\mathbb{F}_q^d$. There is a subset $C' \subseteq C$ on which each quotient map modulo U_j is injective, with*

$$|C'| \geq |C| - \frac{1}{2} \sum_{j=1}^s (q^{\dim U_j} - 1).$$

Proof. For each j , a nonzero difference in U_j has at most one ordered representation in C . The two directions of an unordered pair give distinct opposite differences because q is odd. There are therefore at most $(q^{\dim U_j} - 1)/2$ unordered pairs whose difference lies in U_j . Remove one endpoint of each such pair for all j . The remaining set has the stated size and no two of its elements agree in any of the quotients. Fig. 9 illustrates the removal for one quotient. $\square$

Choose a prime $q > k$, and fix a basis of $\mathbb{F}_{q^k}$ over $\mathbb{F}_q$. Since $k \geq 2$, the characteristic is odd. The classical Sidon parabola is

$$C = \{(x, x^2) : x \in \mathbb{F}_{q^k}\} \subseteq \mathbb{F}_q^{2k}. \quad (8)$$

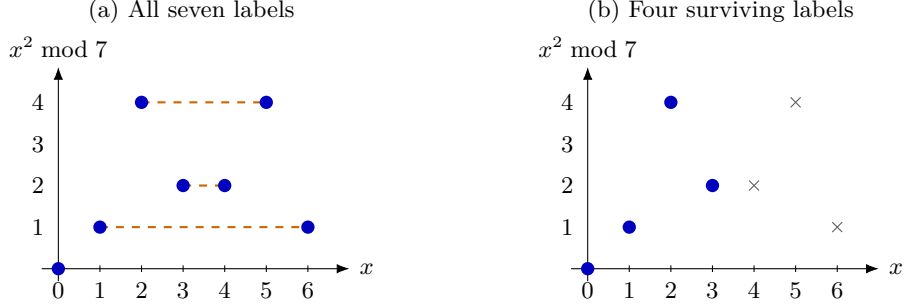


Figure 9: Deleting ambiguous label differences. In (a), the seven points (x, x^2) in $\mathbb{F}_7^2$ form a Sidon set. Projection modulo the horizontal subspace keeps only the second coordinate; dashed lines join the three conflicting pairs. In (b), deleting one point from each pair leaves four different second coordinates. The deletion bound is $(7-1)/2 = 3$. Elements of $\mathbb{F}_7$ are drawn using their representatives in $\{0, \dots, 6\}$.

Its first k coordinates specify x . To recall its Sidon property, equality of two pair sums gives equal sums of the two field elements and of their squares. Since the characteristic is odd, the products also agree. The unordered pairs are the roots of the same monic quadratic; see [5, 15].

Apply Lemma 5.2 to the k projected kernels of Lemma 5.1, now over $\mathbb{F}_q$. It gives a Sidon subset C' with injective quotient maps and

$$|C'| \geq q^k - \frac{k}{2}(q^{k-1} - 1) > \frac{q^k}{2}. \quad (9)$$

Thus $q^k \geq 2n$ supplies n distinct labels $c_i \in C'$. In what follows, each coordinate is represented by its integer in $\{0, \dots, q-1\}$.

5.3 Balancing the labels and recovering the edges

For each pair of positions $s, L-s$, a zero-diagonal binary matrix specifies the selection: a one in entry (i, j) selects $x_{ij}(s)$. Its row i sums to $D_i(s)$, and its column j sums to $D_j(L-s)$. Summing all row margins and summing all column margins count the same selected vertices, so these two totals must agree. The Sidon labels need not have this balance. We will adjust their last k coordinates to obtain it, after separating the integer labels by a fixed scale. That separation will prevent the adjustments and the support correction from hiding a nonzero label difference.

Proposition 5.3. *Let $L = 2k + 1$, $k \geq 2$. Choose a prime $q > k$ with $q^k \geq 2n$, and put*

$$K = 2k + 5, \quad m = K(q - 1) + 1.$$

If $n > 12$ and $2n \geq 9m + 3$, then

$$\text{edim}_m(S_{2k+1}(K_n)) < k(2m + 1)n.$$

For every fixed $k \geq 2$, suitable primes give the bound $O_k(n^{1+1/k})$ for all sufficiently large n .

Proof. Take the labels just constructed, indexed by $0 \leq i < n$. We choose target values for the endpoint counts $D_i(t)$, for $1 \leq t \leq 2k$. First let

$$\mu_t = \left\lfloor \frac{K \sum_i c_{i,t}}{n} \right\rfloor, \quad R_t = K \sum_i c_{i,t} - n\mu_t \in \{0, \dots, n-1\}.$$

Begin with degrees $2m + Kc_{i,t} - \mu_t$. For every $1 \leq s \leq k$, modify coordinate $L - s$ in exactly $|R_s - R_{L-s}|$ entries by $\text{sgn}(R_s - R_{L-s})$. Denote these corrections by $\varepsilon_{i,t}$, setting them to zero in the first k coordinates, and write

$$D_{i,t} = 2m + Kc_{i,t} - \mu_t + \varepsilon_{i,t}. \quad (10)$$

Here $D_{i,t}$ is the target value of $D_i(t)$. Write $\varepsilon_i = (\varepsilon_{i,1}, \dots, \varepsilon_{i,2k}) \in \mathbb{Z}^{2k}$. The number of modified entries is less than n , so the choices are possible. Within a fixed coordinate all corrections belong to $\{0, \sigma_t\}$ for a single sign σ_t . In particular,

$$\begin{aligned} \sum_i D_{i,s} &= \sum_i D_{i,L-s} = 2mn + R_s, & 1 \leq s \leq k, \\ m &\leq D_{i,t} \leq 3m. \end{aligned} \quad (11)$$

The bounds use $0 \leq \mu_t \leq K(q-1)$ and $|\varepsilon_{i,t}| \leq 1$.

For each pair $s, L-s$, realize the margins by Lemma B.1. Their total is at least $2mn$, and their maximum is at most $3m$. The required inequality is

$$(3m+1)3m \leq 2mn,$$

which is exactly $2n \geq 9m+3$. It also implies $3m < n$. Select the internal vertices specified by these k matrices, and no originals. Opposite positions are distinct since L is odd. Equation (11) gives

$$|W| = \sum_{s=1}^k \sum_i D_{i,s} = 2kmn + \sum_{s=1}^k R_s < k(2m+1)n. \quad (12)$$

There are at least nm landmarks of depth k . The supports meeting either endpoint of a given edge contain at most $12m$ such landmarks. Since $n > 12$, a depth- k landmark lies on a disjoint support, and (5) identifies a .

Now compare edges $e = x_{ij}(a)x_{ij}(a+1)$ and $f = x_{uv}(a)x_{uv}(a+1)$ of the same depth. The global terms in (2) cancel, as do the constants $2m - \mu_t$ in their endpoint profiles. For $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^{2k}$, define $\mathcal{T}_a(\mathbf{x}, \mathbf{y}) = \sum_{t=1}^{2k} (x_t A_t + y_t B_t) \in \mathbb{Z}[z]$. Equality of histograms gives

$$\begin{aligned} 0 &= K\mathcal{T}_a(c_i - c_u, c_j - c_v) \\ &\quad + \mathcal{T}_a(\varepsilon_i - \varepsilon_u, \varepsilon_j - \varepsilon_v) + E_e - E_f. \end{aligned} \quad (13)$$

Each correction difference vanishes in the first k coordinates and has every coordinate of absolute value at most one, since its sign is fixed within each remaining coordinate. Each endpoint map has at most k such columns, each with coefficient norm at most one. The correction image therefore has coefficient norm at most $2k$. Together with (4), the second line of (13) has coefficient norm at most $2k+4 < K$. The first line has coefficients divisible by K . Therefore

$$\mathcal{T}_a(c_i - c_u, c_j - c_v) = 0 \quad (14)$$

over the integers, and hence modulo q .

If the edge is central, the two column families agree. After removing z^a , each is $\mathcal{A}$, which is injective. Equation (14) gives $c_i + c_j = c_u + c_v$ over $\mathbb{F}_q$. The Sidon property identifies the unordered support, which determines its central edge.

For a noncentral edge, write $b-a = 2r$. We recover the nearer endpoint first. Equation (14) gives $c_i - c_u \in U_r$, so injectivity modulo U_r gives $i = u$. The equation then becomes $\mathcal{B}_r(c_j - c_v) = 0$. By Lemma 5.1, this kernel is supported on the last $r \leq k$ coordinates. Thus the first k coordinates of c_j, c_v agree. These specify the field element x in (8), and hence the entire label, so $j = v$. The ordered support and the already known a identify the edge. Thus every pair is distinguished.

Finally, Bertrand's postulate supplies a prime between X and $2X$, where

$$X = \max\{k + 1, 3, \lceil (2n)^{1/k} \rceil\}.$$

For fixed k , this gives $q = O_k(n^{1/k})$, while $m = O_k(n^{1/k})$. Since $k \geq 2$, both hypotheses on n eventually hold. Equation (12) proves the asserted asymptotic bound. $\square$

The scale $K = 2k + 5$ used above follows from a bound of $2k$ on the balancing correction and four on the support error. Lemma D.2 reduces the first bound to two by examining the last k columns more closely. The combined error is therefore at most six, allowing the fixed scale $K = 7$; Proposition D.3 gives the resulting finite bound.

6 The two shortest subdivisions

Increasing a replacement path from two to three edges gives each support two independently selectable internal vertices. We show how this changes the answer from impossibility to existence for every order. In the matrix description, the single shared selection at length two imposes symmetry; length three permits different entries in the two directions. The first part proves that symmetry forces a collision. The remaining parts construct the nonsymmetric matrices and complete their finite verification.

6.1 Length two: every selection leaves a collision

At length two, the two directions of a support share the same internal landmark choice. We show that this symmetry forces a collision by encoding a selection as a symmetric binary matrix and comparing its row sums.

Theorem 6.1. *For every integer $n \geq 3$, $\text{edim}_m(S_2(K_n)) = \infty$.*

The proof follows in this subsection, after Lemmas 6.2 and 6.3. The first finds a collision when two adjusted degrees coincide; the second gives an explicit matrix form for the remaining case.

Write $v_0, \dots, v_{n-1}$ for the original vertices, and $w_{ij} = w_{ji}$ for the vertex inserted on $v_i v_j$. Thus $w_{ij} = x_{ij}(1)$, and we abbreviate its selection indicator $\eta_{ij}(1)$ to η_{ij} throughout the length-two argument. Every edge is $e_{ij} = v_i w_{ij}$ for a unique ordered pair $i \neq j$. For an arbitrary landmark set W , put

$$\alpha_i = \mathbf{1}_{\{v_i \in W\}}, \quad \eta_{ij} = \mathbf{1}_{\{w_{ij} \in W\}}, \quad d_i = \sum_{j \neq i} \eta_{ij}, \quad A_0 = \sum_i \alpha_i, \quad B = \sum_{i < j} \eta_{ij}.$$

The counts of selected vertices at distances 0, 1, 2, 3 from e_{ij} are

$$(\alpha_i + \eta_{ij}, \alpha_j + d_i - \eta_{ij}, A_0 - \alpha_i - \alpha_j + d_j - \eta_{ij}, B - d_i - d_j + \eta_{ij}). \quad (15)$$

For example, the vertices at distance one are v_j and the w_{ik} with $k \notin \{i, j\}$; at distance three they are precisely the $w_{k\ell}$ with support disjoint from $\{i, j\}$. Put $t_i = d_i + \alpha_i$ and $r_i = d_i - \alpha_i$. Equality of the histograms in (15) is equivalent to equality of

$$\Gamma(i, j) = (\alpha_i + \eta_{ij}, t_i + \alpha_j, r_j). \quad (16)$$

Indeed, a triple (x, y, z) recovers the histogram as $(x, y - x, A_0 + z - x, B - y - z + x)$.

Lemma 6.2. *If $t_u = t_v$ for distinct u, v , then two distinct edges have equal distance multisets.*

Proof. If $\alpha_u = \alpha_v$, then $d_u = d_v$ and $r_u = r_v$, so $\Gamma(u, v) = \Gamma(v, u)$. Otherwise interchange the indices to have $\alpha_u = 0, \alpha_v = 1$. Then $d_u = d_v + 1$, whence

$$\sum_{j \notin \{u, v\}} (\eta_{uj} - \eta_{vj}) = 1.$$

For some such j , $\eta_{uj} = 1, \eta_{vj} = 0$, and (16) gives $\Gamma(u, j) = \Gamma(v, j)$. $\square$

The remaining case is described by the following matrix form of Aguilar's classification [1, Theorem 2.1]. The diagonal entries represent loops counted once in each degree.

Lemma 6.3. *Let M be a symmetric binary matrix of order n with distinct row sums. After possibly replacing M by its entrywise complement and simultaneously ordering its rows and columns, its entries are*

$$M_{ij} = \mathbf{1}_{\{i+j \geq n\}}, \quad 0 \leq i, j < n.$$

Proof. The n distinct row sums belong to $\{0, \dots, n\}$. A zero row and an all-ones row cannot coexist in a symmetric matrix. Thus the row sums are either $0, \dots, n-1$ or $1, \dots, n$. Complement if necessary, and order them increasingly. Row zero vanishes. Row $n-1$ is one except in column zero. Deleting these two rows and columns leaves row sums $0, \dots, n-3$. Induction, starting with orders one and two, gives the stated formula. $\square$

Proof of Theorem 6.1. Repeated values of t_i are covered by Lemma 6.2. In the remaining case, apply Lemma 6.3 to the symmetric matrix whose diagonal is α_i and whose off-diagonal entries are η_{ij} . Complementation is legitimate: the full vertex set gives every edge the same histogram

$$\left(2, n-1, 2n-4, \binom{n-2}{2}\right),$$

so complementing W preserves every collision.

We may consequently assume

$$t_i = i, \quad \alpha_i = \mathbf{1}_{\{2i \geq n\}}, \quad \eta_{ij} = \mathbf{1}_{\{i+j \geq n\}}, \quad r_i = i - 2\alpha_i.$$

For $n = 3, 4$, substitution into (16) gives $\Gamma(0, 2) = \Gamma(1, 0) = (0, 1, 0)$. For $n \geq 5$, put $h = \lceil n/2 \rceil$. Then $h \geq 3$, $h+1 \leq n-1$, and

$$\Gamma(0, h+1) = \Gamma(1, h-1) = (0, 1, h-1).$$

Indeed, both first coordinates vanish, both second coordinates equal one, and $r_{h+1} = r_{h-1} = h-1$. These are valid distinct edges. Complementation transfers the collision back to the original landmark set when needed. $\square$

Distinct row sums still leave two edges indistinguishable. For $n = 6$, Fig. 10 follows the entries and row sums that produce their common record.

For $n = 2$, the graph is a path and an endpoint resolves its edges. Thus the restriction $n \geq 3$ is sharp.

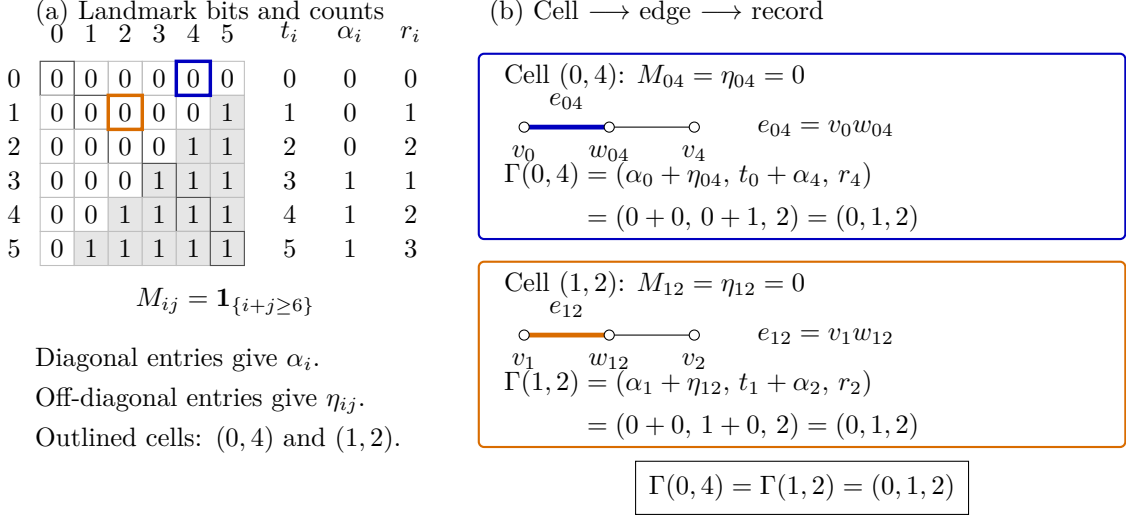


Figure 10: From landmark bits to a collision at length two, for $n = 6$. In (a), the adjacent columns give $t_i = i$, $\alpha_i = M_{ii}$, and $r_i = i - 2\alpha_i$. In (b), the cells (0, 4) and (1, 2) give the internal-vertex bits η_{04} and η_{12} ; the highlighted edges are $e_{04} = v_0 w_{04}$ and $e_{12} = v_1 w_{12}$. Substitution from (a) gives $\Gamma(0, 4) = \Gamma(1, 2) = (0, 1, 2)$. Here $A_0 = 3$ and $B = 6$, so (16) gives the common histogram (0, 1, 5, 3), ordered by distances zero through three.

6.2 Length three: encode the landmarks by a matrix

The two internal vertices on each support can now be selected independently. We encode this freedom in a binary matrix and express the edge histograms through its row sums, column sums and diagonal, so that conditions on these quantities will ensure that all edges are distinguished. Fig. 11 illustrates this encoding with an order-five matrix from the CSV data and its ten landmarks. Write each replacement path as $v_i - x_{ij} - x_{ji} - v_j$, where $x_{ij} = x_{ij}(1)$. In this length-three argument, abbreviate $\eta_{ij}(1)$ to η_{ij} ; now η_{ij} and η_{ji} refer to two different vertices. Set

$$\alpha_i = \mathbf{1}_{\{v_i \in W\}}, \quad \eta_{ij} = \mathbf{1}_{\{x_{ij} \in W\}}, \quad p_i = \sum_{j \neq i} \eta_{ij}, \quad q_i = \sum_{j \neq i} \eta_{ji}, \quad A_0 = \sum_i \alpha_i, \quad B = \sum_i p_i.$$

Here p_i, q_i are integers; the letter p without a subscript will denote a prime. Denote terminal edges by $T_{ij} = v_i x_{ij}$, for ordered pairs $i \neq j$, and middle edges by $E_{ij} = x_{ij} x_{ji}$, for unordered pairs $\{i, j\}$. Put $s_{ij} = \eta_{ij} + \eta_{ji}$. It is convenient to encode W by a full binary matrix M , allowing its diagonal to be nonzero:

$$\alpha_i = M_{ii}, \quad \eta_{ij} = M_{ij} \ (i \neq j), \quad R_i = \sum_j M_{ij} = p_i + \alpha_i, \quad Q_i = \sum_j M_{ji} = q_i + \alpha_i. \quad (17)$$

Every entry equal to one selects exactly one vertex. Hence $|W| = \sum_i R_i$. The diagonal entries correspond to original vertices.

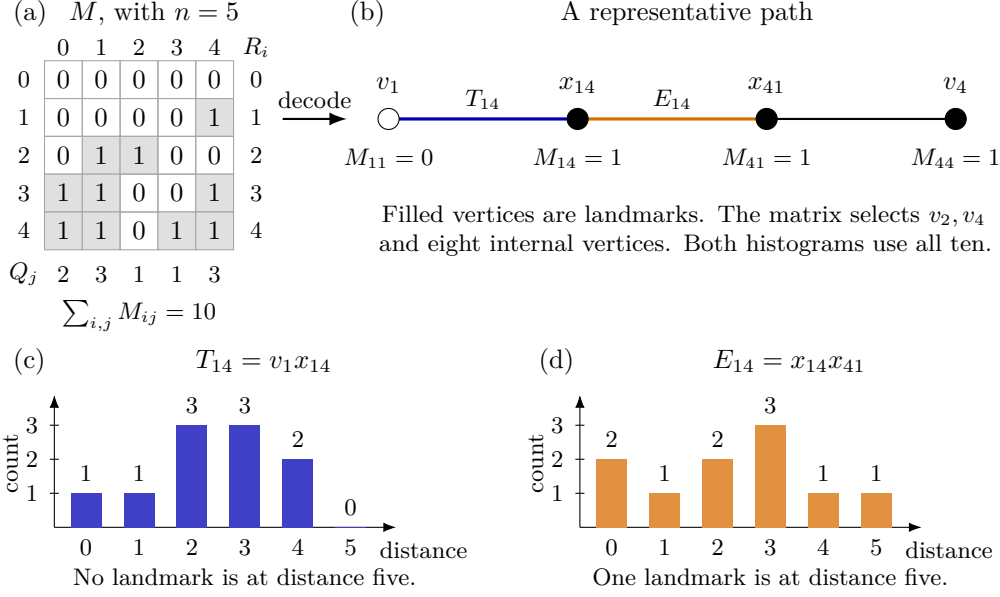


Figure 11: From a matrix to unnamed distances. In (a), a binary matrix of order five selects ten landmarks. Its row and column sums are shown in the margins. Diagonal ones select originals; off-diagonal ones select the corresponding internal vertices. In (b), the path from v_1 to v_4 illustrates four entries; filled vertices are selected. In (c) and (d), the histograms of T_{14} and E_{14} count distances to all ten landmarks encoded by the matrix. Both have total mass ten, and their multiplicities distinguish these two edges; distance five already separates their types in this example. The CSV certificate checks all edge pairs.

The two histograms, ordered by distance from zero to five, are

$$H_T(i, j) = (\alpha_i + \eta_{ij}, p_i - \eta_{ij} + \eta_{ji}, \alpha_j + q_i - \eta_{ji}, A_0 - \alpha_i - \alpha_j + p_j - \eta_{ji}, B - p_i - q_i - p_j + \eta_{ji}, 0), \quad (18)$$

$$H_E(i, j) = (s_{ij}, \alpha_i + \alpha_j, p_i + p_j - s_{ij}, q_i + q_j - s_{ij}, A_0 - \alpha_i - \alpha_j, B - p_i - p_j - q_i - q_j + s_{ij}). \quad (19)$$

For a terminal edge, distance zero selects its two endpoints; distance one selects x_{ji} and the other x_{ik} ; distance two selects v_j and the other x_{ki} ; distance three selects the remaining original vertices and the x_{jk} with $k \neq i, j$. The remaining internal vertices have distance four. For a middle edge, the two originals have distance one, the other internals near these originals have distance two, their partners have distance three, the other originals have distance four, and internals supported outside $\{i, j\}$ have distance five. This proves both formulas.

We seek matrices satisfying

$$R_i = i, \quad 2Q_i - \alpha_i \equiv f(i) + d \pmod{p}, \quad f(i) = \frac{i(i+3)}{2}, \quad 0 \leq i < n, \quad (20)$$

where $p \geq n+5$ is prime and $d \in \{0, \dots, p-1\}$ is a common shift. Throughout this construction, $x \bmod p$ denotes the representative in $\{0, \dots, p-1\}$ of an integer x . The pair $(i, f(i))$ lies on an affine image of the Sidon parabola. The added linear term makes its consecutive differences $i+2$, which will also distinguish terminal edges.

Proposition 6.4. *Let $n \geq 7$, let $p \geq n+5$ be prime, and suppose a binary matrix has row sums $R_i = i$, column sums Q_i , and diagonal α_i satisfying (20) for some d . The corresponding landmark set resolves $S_3(K_n)$ and has size $n(n-1)/2$.*

Proof. Interpret the matrix using (17). Every entry equal to one selects one vertex, so $|W| = \sum_i R_i = n(n-1)/2$.

Middle edges. Equation (19) recovers $\alpha_i + \alpha_j$, $R_i + R_j = i + j$, and $Q_i + Q_j$. It therefore recovers $i + j$ and $f(i) + f(j) + 2d$ modulo p . The latter quantity, together with $i + j$, determines $i^2 + j^2$, then ij , and hence the unordered pair $\{i, j\}$. Here p is odd, and the labels $0, \dots, n - 1$ are distinct modulo p . Thus all middle edges are distinguished.

Terminal edges with a common origin. The sum of the first two entries of (18) is $i + \eta_{ji}$, so it identifies η_{ji} for fixed i . The third entry then identifies α_j , and the fourth, which equals $A_0 + j - \alpha_i - 2\alpha_j - \eta_{ji}$, identifies j .

Terminal edges with different origins. Suppose $H_T(i, j) = H_T(u, v)$ with $i < u$. The sum of the first two entries forces $u = i + 1$, $\eta_{ji} = 1$, and $\eta_{vu} = 0$. Set $\Delta_1 = \alpha_u - \alpha_i$ and $\Delta_2 = \alpha_j - \alpha_v$. The other equalities give

$$Q_u - Q_i = \Delta_1 + \Delta_2 - 1, \quad j - v = 1 - \Delta_1 + 2\Delta_2.$$

Writing $Z_i = 2Q_i - \alpha_i$, a collision consequently requires $Z_u - Z_i = \Delta_1 + 2\Delta_2 - 2 \in \{-5, -4, \dots, 1\}$. The value -5 forces $\Delta_1 = \Delta_2 = -1$ and $j = v$, contradicting $\Delta_2 = -1$. Hence the difference must lie in $\{-4, -3, -2, -1, 0, 1\}$. On the other hand, (20) implies

$$Z_{i+1} - Z_i \equiv f(i + 1) - f(i) = i + 2 \pmod{p}.$$

The residue $i + 2$ belongs to $\{2, \dots, n\}$, which is disjoint from $\{p - 4, p - 3, p - 2, p - 1, 0, 1\}$, because $p \geq n + 5$. This excludes every terminal-edge collision.

Separating the two types. For any middle support $\{i, j\}$, one of the three largest row indices k is outside this support. Its internal outdegree is $p_k = k - \alpha_k \geq n - 4 \geq 3$. At most two of its selected vertices $x_{k\ell}$ have $\ell \in \{i, j\}$. At least one is therefore supported entirely outside $\{i, j\}$, and is at distance five from the middle edge. Every terminal histogram has distance-five coordinate zero. This proves that W resolves all edges. $\square$

6.3 Producing the matrix margins

The remaining task is to make the column sums add up to $N = 0 + 1 + \dots + (n - 1) = n(n - 1)/2$, while keeping enough room for a chosen diagonal. Each residue produced by $f(i) + d$ offers two choices. Write $p = 2m + 1$. For a shift d , put

$$r_i = (f(i) + d) \bmod p \in \{0, \dots, 2m\}, \quad L_i = \lceil r_i/2 \rceil, \quad \alpha_i^0 = r_i \bmod 2. \quad (21)$$

The low choice $(Q_i, \alpha_i) = (L_i, \alpha_i^0)$ satisfies $2Q_i - \alpha_i = r_i$. The high choice is

$$Q_i = L_i + m + 1 - \alpha_i^0, \quad \alpha_i = 1 - \alpha_i^0, \quad (22)$$

and satisfies $2Q_i - \alpha_i = r_i + p$. Thus the column increases by m at odd residues and by $m + 1$ at even residues.

Start with all low choices. If their sum is S , the deficit is $D = N - S$. Raising X odd-residue columns and Y even-residue columns gives the exact balance precisely when

$$D = mX + (m + 1)Y.$$

We choose the smallest eligible low values in each class, ordered by (L_i, i) , and leave index zero unraised. This moves the columns towards the middle of their possible range. The resulting concentration provides room in the matrix inequalities, even after the diagonal is fixed. The relevant error is controlled by the quadratic character $\chi : \mathbb{Z} \rightarrow \{-1, 0, 1\}$ modulo p : it is 0 on multiples of p , 1 on nonzero squares modulo p , and -1 otherwise. An interval bound $C_\chi \geq 0$ means that

$$\left| \sum_{j=u+1}^{u+\ell} \chi(j) \right| \leq C_\chi \quad (u \in \mathbb{Z}, \quad \ell \in \{0, \dots, p\}).$$

The character is periodic, so this includes intervals that wrap modulo p . Appendix A proves the following quantitative statement.

Proposition 6.5. *Let $p = 2m + 1$ be prime, let $n = p - g$, and suppose $g \geq 5$ and $m \geq 150$. If the quadratic character modulo p has absolute sum at most C_χ on every interval of residues and*

$$H := C_\chi + g + 2 \leq m/12, \quad (23)$$

then a binary matrix satisfying (20) exists. The shift d can be chosen even in $[m/2, 3m/2]$, with $\alpha_0 = 0$.

The proof in Appendix A first uses Lemma A.1 to quantify the room needed for the chosen diagonal. Varying the modular shift then makes the column total match the row total, as illustrated in Fig. 15. Lifting the smallest columns preserves enough room in every matrix cut. Fig. 16 illustrates the concentration mechanism; the tail estimates in Appendix A establish the required finite bounds.

Dusart's prime-gap estimate and Lapkova's character-sum estimate imply these hypotheses for every $n \geq 10\,000$ (Corollary A.2). Exact certificates cover all remaining orders in Section 6.4. Corollary 3.2 supplies the lower bound in the theorem below. Together these results give:

Theorem 6.6. *For every $n \geq 3$,*

$$\text{edim}_m(S_3(K_n)) \leq \binom{n}{2}.$$

For $n \geq 10$, the lower bound is $n^2/16 - n/2$. Consequently $\text{edim}_m(S_3(K_n)) = \Theta(n^2)$ as $n \rightarrow \infty$.

6.4 Certifying the remaining orders

The analytic construction leaves the bounded range $3 \leq n < 10\,000$. We complete it with two forms of exact certificate. Small orders are supplied as explicit binary matrices. For the larger orders, it suffices to give margins and diagonal entries and verify the following criterion.

Lemma 6.7. *Let $\alpha_i \in \{0, 1\}$, with $\alpha_0 = 0$, and let the integers Q_i satisfy $\sum_i Q_i = n(n-1)/2$ and $0 \leq Q_i - \alpha_i \leq n-1$. Put $q_i = Q_i - \alpha_i$. For each $1 \leq k \leq n$, form the weights*

$$w_i^{(k)} = i - \alpha_i + \mathbf{1}_{\{q_i \geq k\}},$$

and write $w_{(1)}^{(k)} \geq \dots \geq w_{(n)}^{(k)}$ for their nonincreasing rearrangement. A binary matrix with row sums i , column sums Q_i , and diagonal α_i exists if and only if

$$\sum_{i=0}^{n-1} \min(q_i, k) \geq \sum_{j=1}^k w_{(j)}^{(k)} \quad (1 \leq k \leq n). \quad (24)$$

Proof. Reserve the diagonal entries and apply the forbidden-diagonal flow criterion in Lemma B.1. For a row subset I of size k , its right side is

$$\sum_j \min(q_j, k - \mathbf{1}_{\{j \in I\}}) = \sum_j \min(q_j, k) - \sum_{i \in I} \mathbf{1}_{\{q_i \geq k\}}.$$

Thus its cut condition is $\sum_{i \in I} w_i^{(k)} \leq \sum_j \min(q_j, k)$. The largest left side is the sum of the k largest weights. The empty subset satisfies the condition, and the integer flow criterion proves the equivalence. Fig. 12 shows this conversion for one cut in the order-five example. $\square$

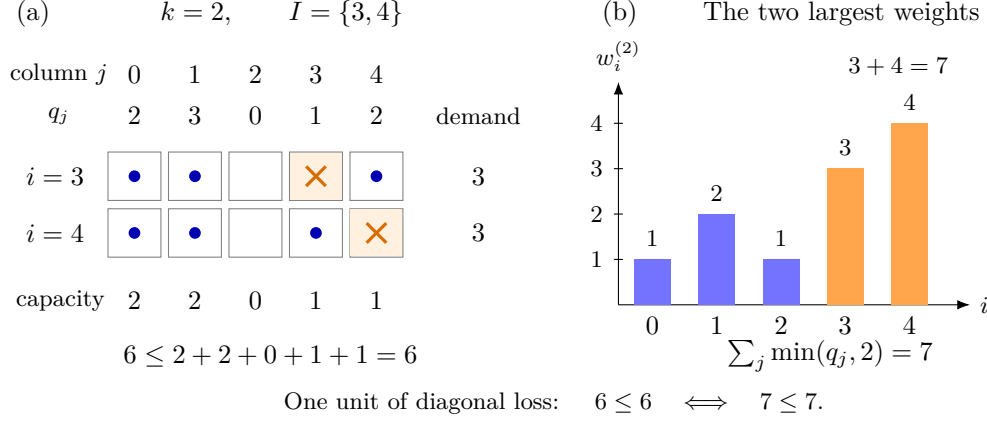


Figure 12: The cut with $k = 2$ and $I = \{3, 4\}$ for the matrix in Fig. 11. In (a), crosses mark forbidden diagonal positions after reserving the diagonal entries. Column capacities are $\min(q_j, 2 - \mathbf{1}_{j \in I})$, giving $(2, 2, 0, 1, 1)$ for total demand six. In (b), weights $(1, 2, 1, 3, 4)$ have two largest values summing to seven, equal to $\sum_j \min(q_j, 2)$. Moving the one-unit diagonal loss to the left converts the same inequality from $6 \leq 6$ to $7 \leq 7$. Every other cut size must also be checked.

Proposition 6.8. *For every integer $3 \leq n < 10\,000$, the graph $S_3(K_n)$ has an edge multiset resolving set of size at most $n(n-1)/2$.*

Proof by exact finite certificates. The accompanying data files² cover each integer in this range once. For $3 \leq n \leq 29$, `small_matrices.csv` gives all entries of one binary matrix of each order. Its fields `n` and `row` identify a row, and fields `c0`, ..., `c28` contain its entries followed by empty padding. Algorithm 1 decodes the ones using (17) and verifies the resolving property from graph distances. All 27 matrices satisfy the size bound and have distinct histograms for their 12 177 graph edges in total.

For $30 \leq n \leq 9999$, the file `degree_certificates.csv` contains one record (n, p, d, X, Y) for each order. Algorithm 2 reconstructs the column sums and diagonal, then checks every inequality in (24). These checks use integer arithmetic. They passed for all 9970 records, including all 49 994 565 inequalities indexed by (n, k) . The diagonal-bit identity gives $2Q_i - \alpha_i \equiv f(i) + d \pmod{p}$. Lemma 6.7 therefore realizes each record, and Proposition 6.4 proves that its matrix encodes a resolving set of size $n(n-1)/2$. Table 2 records the coverage. $\square$

Table 2: Complete coverage of the length-three case. The finite rows contain exact certificates for every integer in the indicated interval.

Orders	Certificate or construction	Number of orders
$3 \leq n \leq 29$	Explicit matrices; graph distances	27
$30 \leq n \leq 9999$	Margins and all integer cut inequalities	9970
$n \geq 10\,000$	Corollary A.2 and Proposition 6.4	All remaining orders

Proposition 6.8 and the analytic construction prove the upper bound in Theorem 6.6 for every $n \geq 3$. Corollary 3.2 supplies its lower bound for $n \geq 10$, completing the theorem.

The record $(30, 37, 6, 1, 6)$ provides a worked reconstruction (Fig. 13). Its low total is 303, its deficit is 132, and lifting one entry with odd residue and six with even residue adds exactly $18 + 6 \cdot 19 = 132$. These are finite certificate parameters; their validity follows from the cuts, independently of the sufficient analytic hypotheses of Proposition 6.5.

²CSV files and reading instructions:
<https://github.com/tashimir/subdivided-complete-edge-multiset-data>.

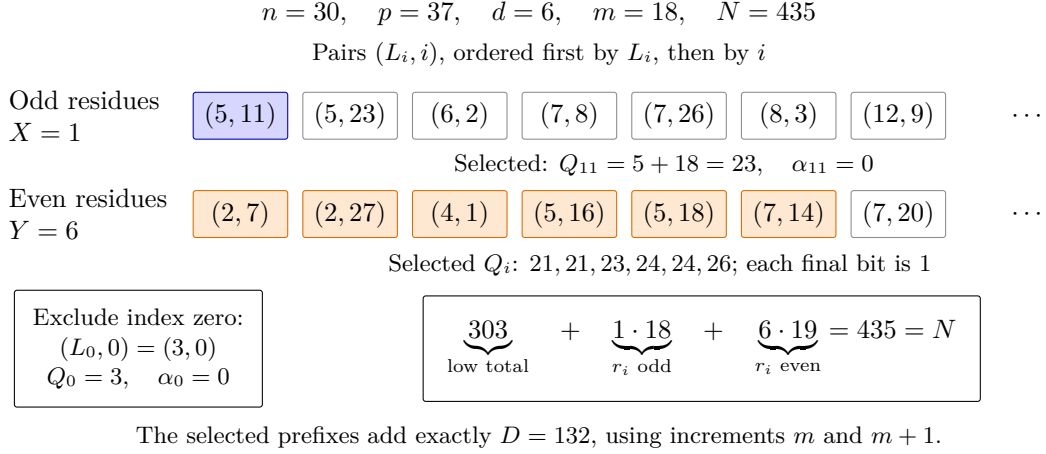


Figure 13: Exact balancing in the certificate $(n, p, d, X, Y) = (30, 37, 6, 1, 6)$. Pairs (L_i, i) are ordered lexicographically within each parity class of the residues r_i . Blue selects the first pair with odd r_i ; orange selects the first six eligible pairs with even r_i . The following pairs make the tie-breaking rule visible. Index zero is excluded from the even-residue class. The total rises from 303 to $303 + 18 + 6 \cdot 19 = 435 = n(n-1)/2$.

The two algorithms below check different mathematical objects. Algorithm 1 builds the graph from a small matrix and compares the actual distance histograms. Algorithm 2 reconstructs the larger certificates and checks the necessary and sufficient inequalities for a matrix to exist. The latter relies on Proposition 6.4 for the resolving property.

The CSV conventions are specified in `data/README.txt`. Read the fields by name and verify that every required order and matrix row occurs exactly once before applying the algorithms. In Algorithm 1, `BREADTHFIRSTDISTANCES( $G, s$ )` returns the function $v \mapsto d_G(s, v)$.

Algorithm 1 Check a small matrix by graph distances

Input: A binary matrix M of order n , with $3 \leq n \leq 29$.

Output: Accept exactly when its selected vertices resolve $S_3(K_n)$ using at most $\binom{n}{2}$ landmarks.

- 1: Create originals $v_0, \dots, v_{n-1}$ in an empty graph G .
 - 2: **for** each pair $0 \leq i < j < n$ **do**
 - 3: Create fresh vertices x_{ij}, x_{ji} .
 - 4: Add the path $v_i - x_{ij} - x_{ji} - v_j$ to G .
 - 5: **end for**
 - 6: $W \leftarrow \{v_i : M_{ii} = 1\} \cup \{x_{ij} : i \neq j, M_{ij} = 1\}$.
 - 7: **if** $|W| > \binom{n}{2}$ **then return** Reject.
 - 8: $(H_e(0), \dots, H_e(5)) \leftarrow (0, 0, 0, 0, 0, 0)$ for each edge e .
 - 9: **for** each $s \in W$ **do**
 - 10: $d \leftarrow \text{BREADTHFIRSTDISTANCES}(G, s)$.
 - 11: **for** each edge $e = uv$ of G **do**
 - 12: $t \leftarrow \min(d(u), d(v))$.
 - 13: $H_e(t) \leftarrow H_e(t) + 1$.
 - 14: **end for**
 - 15: **end for**
 - 16: Sort the six-entry vectors $(H_e)_{e \in E(G)}$ lexicographically.
 - 17: **return** Accept if no consecutive vectors agree; otherwise Reject.
-

The standard breadth-first search in Algorithm 1 returns the distance from s to every vertex.

All edge-to-vertex distances in $S_3(K_n)$ are at most five, so six counters record a complete histogram. Each landmark increments exactly one counter for each edge. Consequently acceptance is equivalent to the definition of an edge multiset resolving set. There are $O(n^2)$ vertices, edges, and possible landmarks; the algorithm takes $O(n^4)$ operations and $O(n^2)$ storage. It checks distances independently of the algebraic formulas used to find the matrices.

Algorithm 2 uses the low values and parity bits from (21). The initial parity bit is denoted by $\sigma_i = \alpha_i^0$ to distinguish it from the final diagonal entry α_i . Ordering by (L_i, i) makes every selection unambiguous. Primality may be checked by trial division through $\lfloor \sqrt{p} \rfloor$.

Algorithm 2 Check a degree certificate

Input: Integers (n, p, d, X, Y) , with $30 \leq n \leq 9999$.

Output: Accept exactly when the reconstruction below satisfies the prime, diagonal, total, and matrix-cut conditions.

```

1: if  $p$  is not prime or  $p < n + 5$  then return Reject.
2: if  $d \notin \{0, \dots, p-1\}$  or  $\min(X, Y) < 0$  then return Reject.
3:  $m \leftarrow (p-1)/2$ ;  $N \leftarrow n(n-1)/2$ .
4: for  $i = 0, \dots, n-1$  do
5:    $r_i \leftarrow (i(i+3)/2 + d) \bmod p$ .
6:    $L_i \leftarrow \lfloor (r_i + 1)/2 \rfloor$ ;  $\sigma_i \leftarrow r_i \bmod 2$ ;  $\epsilon_i \leftarrow 0$ .
7: end for
8:  $O \leftarrow \{i : \sigma_i = 1\}$ , ordered by  $(L_i, i)$ .
9:  $E \leftarrow \{i : \sigma_i = 0, i \neq 0\}$ , ordered by  $(L_i, i)$ .
10: if  $X > |O|$  or  $Y > |E|$  then return Reject.
11: Set  $\epsilon_i \leftarrow 1$  at the first  $X$  indices of  $O$  and the first  $Y$  indices of  $E$ .
12: for  $i = 0, \dots, n-1$  do
13:    $Q_i \leftarrow L_i + \epsilon_i(m+1-\sigma_i)$ .
14:    $\alpha_i \leftarrow \sigma_i + \epsilon_i(1-2\sigma_i)$ ;  $q_i \leftarrow Q_i - \alpha_i$ .
15: end for
16: if  $\alpha_0 \neq 0$ , or  $\sum_i Q_i \neq N$ , or some  $q_i \notin \{0, \dots, n-1\}$  then return Reject.
17: for  $k = 1, \dots, n$  do
18:   Form  $w_i \leftarrow i - \alpha_i + \mathbf{1}_{\{q_i \geq k\}}$  for all  $i$ .
19:   Sort these weights as  $w_{(1)} \geq \dots \geq w_{(n)}$ .
20:   if  $\sum_{j=1}^k w_{(j)} > \sum_i \min(q_i, k)$  then return Reject.
21: end for
22: return Accept.

```

Every reconstructed pair satisfies $2Q_i - \alpha_i = r_i + p\epsilon_i$, so it has the target label value modulo p . The last loop checks all the inequalities of Lemma 6.7; acceptance therefore proves matrix realizability, and Proposition 6.4 proves the resolving property. No approximate arithmetic is used. The displayed cut verification takes $O(n^2 \log n)$ arithmetic operations and $O(n)$ storage per record.

7 Finite existence and restrictions on the selection

The fixed-length theorem leaves two quantitative questions: which finite orders admit a resolving set, and what is lost by restricting the available landmark positions? We collect the existence ranges first, then combine the constructions and counting bounds developed below to compare constrained selections. We again allow all integers $n \geq 3$ and $L \geq 2$, with the conventions of Section 2.

7.1 Existence at finite orders

To locate finite orders where a resolving set exists, we combine the known cycle case with explicit thresholds from the constructions. For $n = 3$, we have $S_L(K_3) = C_{3L}$, where C_m is the cycle on m vertices. Since $\text{edim}_m(C_m) = 3$ for $m \geq 7$ [11, Theorem 12], every $L \geq 3$ works at $n = 3$; the remaining bounds concern larger orders.

Corollary 7.1. *For every $n \geq 210$,*

$$\text{edim}_m(S_4(K_n)) < 18n \lceil \sqrt{n} \rceil.$$

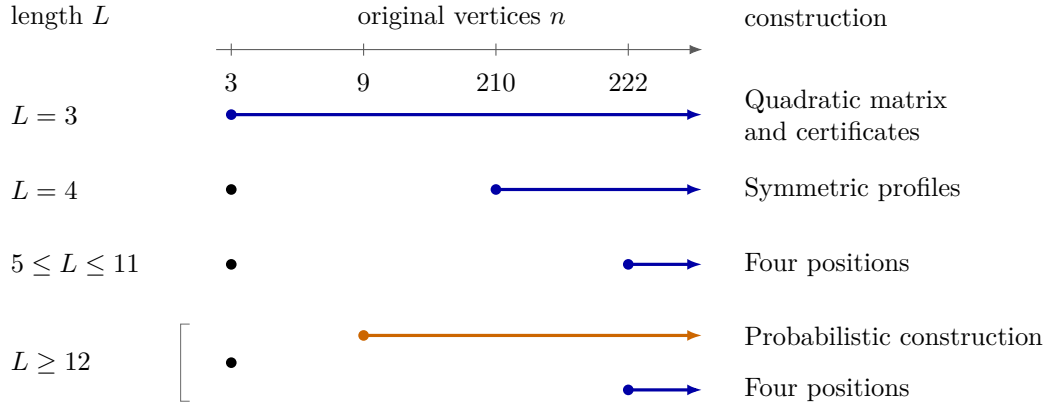
Proof. Apply Corollary 4.3 with $k = 2$. For $210 \leq n \leq 225$, its condition is $n \geq 14 \cdot 15$. For $n \geq 226$, put $Q = \lceil \sqrt{n} \rceil \geq 16$. Then $n \geq (Q - 1)^2 + 1 \geq 14Q$, because $Q^2 - 16Q + 2 \geq 0$. $\square$

Theorem 7.2. *For every $n \geq 222$ and every $L \geq 5$, there is a resolving set using only internal landmarks of depths one and two, of size less than*

$$24n^{3/2} + 18n.$$

The proof is given in Appendix E, starting with four-coordinate Sidon labels (Lemma E.1) and showing that balancing preserves their distinct pair sums (Lemma E.2). Proposition E.3 then recovers each edge from its histogram; the recovery cases are collected in Table 4. Both the constant and the order threshold are independent of L . The finite prime intervals needed for the threshold 222 appear in Table 3.

For smaller orders at long lengths, Proposition F.4 gives a resolving set using at most half of all vertices when $n \geq 9$ and $L \geq 12$. Its proof bounds the probability of equal histograms through a forest of distance levels (Lemma F.3). Fig. 17 in Appendix F shows how this forest separates the constraints on a random selection. Fig. 14 collects these bounds with thresholds uniform over the indicated lengths. The length-dependent conditions in Corollary 4.3 and Proposition 5.3 cover additional pairs (n, L) .



Black dots at $n = 3$: cycles. Arrows include every larger order.

Horizontal scale is categorical. Unmarked ranges receive no assertion here.

Figure 14: Uniform existence ranges. Each line starts at an included order and continues to the right. The points at $n = 3$ use the known cycle case. In the bottom row, the two constructions overlap for $n \geq 222$. The scale groups orders by the four displayed thresholds. Unmarked intervals receive no assertion from this summary; length-dependent conditions can cover further pairs. At length two no resolving set exists for $n \geq 3$.

7.2 The cost of restricting the selection

The uniform construction uses only the first two depths, whatever the path length. For fixed $L \geq 5$, comparing it with the restricted-depth count in Section C shows the cost of this choice: the minimum keeps exponent $3/2$, even when further depths permit a smaller exponent.

Corollary 7.3. *For each fixed $L \geq 5$, the minimum size among resolving sets whose internal landmarks have depth at most two is $\Theta(n^{3/2})$. For $n > 33$, at least $F_{32,2}(n)$ internal landmarks are needed, where $F_{Q,d}$ is the minimum total cost of integer vectors defined in Section C.*

Proof. In Theorem C.2, set $s = 2$, so $Q_s = 32$. It gives the finite lower bound $F_{32,2}(n)$ for $n > 33$, whose order is $n^{3/2}$ by (48) with $d = 2$. Theorem 7.2 gives the matching upper order. $\square$

For comparison, Corollary D.1 gives $\Theta_k(n^{1+2/k})$ for symmetric profiles at length $2k + 1$, when $k \geq 3$. Its lower bound already follows by distinguishing only the central edges, as Proposition F.1 shows. The unrestricted order is $\Theta_k(n^{1+1/k})$. Thus using further depths helps, and allowing opposite profiles to differ is also essential to attain the sharp odd exponent.

The comparison with named distances follows from Theorem 1.1: Proposition 2.2 bounds $\text{edim}(S_L(K_n))$ between two linear functions of n , so, for every fixed $L \geq 3$,

$$\frac{\text{edim}_m(S_L(K_n))}{\text{edim}(S_L(K_n))} = \Theta_L(n^{1/\lfloor L/2 \rfloor}). \quad (25)$$

This is the factor required to replace named reports by a multiset.

The growth exponents leave the optimal leading constants open. Proposition F.2 improves the lower constant at length three to $c_* = 0.080208390067\dots$, giving

$$\text{edim}_m(S_3(K_n)) \geq c_* n^2 - O(n).$$

The upper constant supplied by Theorem 6.6 is $1/2$. Proposition F.5 also gives bounds using sparse independent selections, for $L \geq 14$ and the explicit order threshold stated there. These bounds complement the fixed-length theorem when the length grows with n .

8 Conclusion and further questions

We have determined the asymptotic number of landmarks needed to identify edges of subdivided complete graphs from unnamed distance reports. For each fixed $L \geq 3$, this number has order $n^{1+1/\lfloor L/2 \rfloor}$, whereas named reports require only $\Theta(n)$ landmarks. The two shortest lengths are also settled for every $n \geq 3$: no selection resolves $S_2(K_n)$, while $S_3(K_n)$ has a resolving set of size at most $\binom{n}{2}$. These results extend the earlier finite computations to sharp asymptotic bounds and identify a change in existence between lengths two and three.

The proofs explain this behavior through the information retained by endpoint counts. Counting the possible vectors gives the lower bound, while realizing suitable labels as degrees or matrix margins gives the matching constructions. At odd lengths, Sidon labels identify endpoint pairs, and the projected kernels describe the changes that a noncentral edge can hide. The quantitative results also show that the freedom in these profiles matters: for fixed $L \geq 5$, using only the first two depths has cost $\Theta(n^{3/2})$, and symmetric profiles at length $2k + 1$, with fixed $k \geq 3$, have cost $\Theta_k(n^{1+2/k})$. The constructions uniform in L and the probabilistic bounds complement this picture with explicit finite existence ranges.

Two questions follow naturally. At length three, does $\text{edim}_m(S_3(K_n))/n^2$ converge, and what is its limiting value if it does? The current bounds place such a value between $c_* = 0.080208390067\dots$ and $1/2$, leaving room for sharper counting and more economical matrix

constructions. A second question is to classify the finite pairs (n, L) , with $L \geq 4$, that admit a resolving set. The criteria for realizing endpoint profiles and the exact certificate methods provide concrete ways to investigate the orders below the sufficient thresholds obtained here.

AI-use disclosure. The author directed the research and made substantive use of OpenAI’s ChatGPT and Codex as artificial intelligence (AI) tools for literature retrieval, mathematical exploration, code development, symbolic and numerical checks, and manuscript preparation. The author critically assessed this assistance and determined what to incorporate into the work. To refine the English, grammar, and presentation of results, proofs, and figures, the author made revisions directly or specified changes for the tools to implement, critically assessing successive versions as the exposition developed. AI tools also assisted in checking and maintaining the manuscript’s internal consistency. The author checked all results and takes full responsibility for the mathematical content and the final manuscript.

A Effective realization of the quadratic margins

We prove Proposition 6.5. Throughout this appendix, $p = 2m+1$, $n = p-g$, and $N = n(n-1)/2$. The quadratic character χ and its interval bound C_χ are as defined in Section 6.3, with $H = C_\chi + g + 2 \leq m/12$. Each value of the quadratic modulo p offers two column sums. We must choose between them so that their total matches the row sums and the chosen diagonal still fits. Equal totals alone do not guarantee such a matrix. We first quantify the room required for the diagonal, then show that lifting the smallest columns leaves enough room in every cut.

A.1 Room for the chosen diagonal

We need enough spare capacity in each matrix cut to accommodate the chosen diagonal. For each integer $0 \leq k \leq n$, define the margin gap

$$G_k = \sum_{i=0}^{n-1} \min(Q_i, k) - \frac{k(2n-k-1)}{2}. \quad (26)$$

The subtracted term is the sum of the largest k desired row sums; the next lemma shows how a lower bound on this gap absorbs the diagonal corrections.

Lemma A.1. *Suppose the integers Q_i satisfy $\sum_i Q_i = N$ and $1 \leq Q_i \leq n-2$. Let $\alpha_i \in \{0, 1\}$ with $\alpha_0 = 0$. If*

$$G_k \geq n \quad \text{whenever} \quad \min_i Q_i \leq k \leq \max_i Q_i,$$

then a binary matrix with row sums i , column sums Q_i , and diagonal α_i exists.

Proof. Reserve the diagonal ones, subtract α_i from row and column i , and forbid all diagonal positions in the remaining matrix. The residual demands are nonnegative and have the same total. The integral-flow criterion for matrices with a forbidden diagonal [10, p. 832, equation (2)], applied with rows and columns interchanged, requires the following inequality for every $I \subseteq \{0, \dots, n-1\}$:

$$\sum_{i \in I} (i - \alpha_i) \leq \sum_j \min(Q_j - \alpha_j, k - \mathbf{1}_{\{j \in I\}}) \quad (k = |I|).$$

Moving the diagonal corrections to the left gives the equivalent condition

$$\sum_{i \in I} i + \#\{i \in I : \alpha_i = 0, Q_i \geq k\} + \#\{i \notin I : \alpha_i = 1, Q_i \leq k\} \leq \sum_i \min(Q_i, k). \quad (27)$$

For k in the stated range the two counts sum to at most n , so the hypothesis verifies the cut.

If $k < \min_i Q_i$, the second count vanishes and the first is at most k . Here $G_k = k(k+1)/2 \geq k$, so the cut also holds. If $k > \max_i Q_i$, the first count vanishes and the right side is N . The cut follows from $\sum_{i \notin I} \alpha_i \leq \sum_{i \notin I} i$, since $\alpha_0 = 0$ and $\alpha_i \leq 1 \leq i$ for $i \geq 1$. The empty set satisfies the criterion as well. Integral flow gives the required residual matrix, and restoring the diagonal completes the proof. $\square$

A.2 Two representatives and one-sided counts

The balancing step will lift the smallest low representatives from (21)–(22), so we need bounds on how many are available in each parity class. Completing the square expresses their multiplicities through the quadratic character, whose interval bounds will control the cumulative counts. For counting purposes, extend $f(i) = i(i+3)/2$, $r_i(d) = (f(i) + d) \bmod p$, and $L_i(d) = \lceil r_i(d)/2 \rceil$ to $0 \leq i < p$; the construction retains only $0 \leq i < n$. We omit d when the shift is fixed. Over all p indices, the multiplicity of a residue $r \in \{0, \dots, p-1\}$ is $1 + \chi(8r + 9 - 8d)$, by completing the square. Each multiplicity is at most two. Let g_e and g_o count the even and odd output residues among the removed indices $n, \dots, p-1$, so $g_e + g_o = g$. Let $C_e(t)$ and $C_o(t)$ count retained low values at most t in each parity class. For $0 \leq t \leq m$,

$$\begin{aligned} \lfloor t \rfloor + 1 - C_\chi - g_e &\leq C_e(t) \leq \lfloor t \rfloor + 1 + C_\chi, \\ \lfloor t \rfloor - C_\chi - g_o &\leq C_o(t) \leq \lfloor t \rfloor + C_\chi. \end{aligned} \quad (28)$$

Indeed, the even residues $r = 2\ell$ produce $\chi(16\ell + 9 - 8d)$, and the odd residues $r = 2\ell - 1$ produce $\chi(16\ell + 1 - 8d)$. Since 16 is a square, each sum is a translated character sum on an interval. Deleting inputs can only decrease a count.

The sum of all low values has a useful additional estimate. For integer $0 \leq t < m$, the full count with $L_i \leq t$ is $2t + 1 + E_t$, where

$$E_t = \sum_{r=0}^{2t} \chi(8r + 9 - 8d), \quad |E_t| \leq C_\chi.$$

Factoring the nonzero coefficient 8 reduces this to another interval sum. Consequently,

$$S_{\text{full}} := \sum_{i=0}^{p-1} L_i = mp - \sum_{t=0}^{m-1} (2t + 1 + E_t), \quad |S_{\text{full}} - m(m+1)| \leq mC_\chi.$$

Put $S(d) = \sum_{i=0}^{n-1} L_i(d)$, $D(d) = N - S(d)$, and $\delta_0 = g(g-1)/(2m)$. Write S, D when the shift is fixed. The deleted low values have sum between zero and mg . Expanding $N = (2m+1-g)(2m-g)/2$ therefore gives

$$m - 2g + \delta_0 - C_\chi \leq D/m \leq m - g + \delta_0 + C_\chi. \quad (29)$$

A.3 Balancing the total exactly

The remaining deficit must be a sum of the two available increments, m and $m+1$. We vary the even shift d until $D(d)$ modulo m enters an interval that gives admissible numbers of lifts, using a walk with bounded increments to guarantee such a shift.

For $t \in \{0, \dots, p-1\}$, let $c_t(d) = |\{i : 0 \leq i < n, r_i(d) = t\}| \leq 2$. Increasing d by two raises each low value by one, except that residues $p-2$ and $p-1$ wrap to zero and one. It follows that

$$S(d+2) - S(d) = n - (m+1)c_{p-2}(d) - mc_{p-1}(d),$$

and hence

$$D(d+2) - D(d) \equiv g - 1 + c_{p-2}(d) \pmod{m}. \quad (30)$$

The consecutive even shifts in $[m/2, 3m/2]$ have $J = \lfloor 3m/4 \rfloor - \lfloor m/4 \rfloor \geq m/2 - 2$ steps. Lift the residues $D(d) \bmod m$ to an increasing walk whose increments lie in $[g-1, g+1]$, as permitted by (30). Its total increase is at least $4J \geq 2m - 8 > m$. Also $g+1 < m/3$, by (23). The next translated left endpoint of $[m/3, 2m/3]$ is at most m away, and the first crossing overshoots it by less than $m/3$. Fig. 15 shows why the crossing stays in the target interval. Thus an even shift satisfies

$$Y = D(d) \bmod m \in [m/3, 2m/3], \quad K = \lfloor D(d)/m \rfloor, \quad X = K - Y. \quad (31)$$

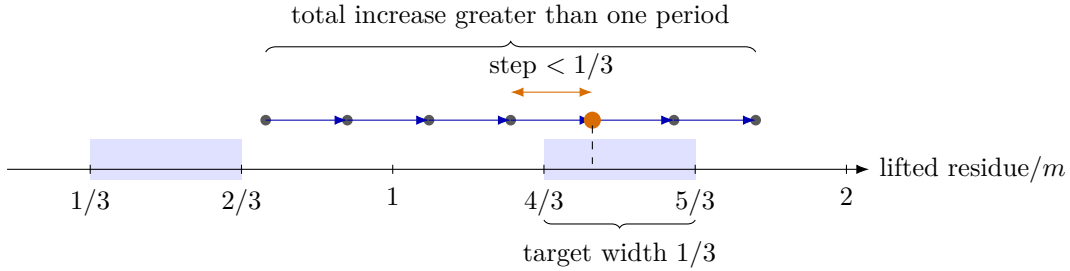


Figure 15: Why a suitable shift exists in (30). Divide the lifted values of $D(d) \bmod m$ by m . Consecutive values advance by less than $1/3$, and the walk travels more than one period. It therefore enters a translate of $[1/3, 2/3]$. The orange point marks the first crossing in a schematic walk; the plotted values illustrate the interval argument rather than a particular degree certificate.

In particular,

$$D = mX + (m+1)Y \quad (32)$$

holds exactly. Since $g < m$, one has $\delta_0 \leq g$. Equation (29) gives

$$m - 2g + \delta_0 - C_\chi - 1 \leq K \leq m - g + \delta_0 + C_\chi. \quad (33)$$

Hence $0 < X \leq 2m/3 + C_\chi$. The odd class has at least $m - C_\chi - g$ indices by (28), which exceeds this upper bound for X . The even class after excluding index zero has at least $m - C_\chi - g > 2m/3$ indices. These inequalities follow from $C_\chi + g + 2 \leq m/12$.

Choose the X smallest low values among indices with odd r_i , and the Y smallest among indices with even r_i and index different from zero. Ties may be resolved arbitrarily. Set $\epsilon_i = 1$ at these indices and 0 elsewhere, and define

$$Q_i = L_i + \epsilon_i(m+1 - \alpha_i^0), \quad \alpha_i = \begin{cases} \alpha_i^0, & \epsilon_i = 0, \\ 1 - \alpha_i^0, & \epsilon_i = 1. \end{cases} \quad (34)$$

Then $2Q_i - \alpha_i = r_i + p\epsilon_i$, and (32) gives $\sum_i Q_i = N$. Moreover, $r_0 = d$ is even, so

$$\alpha_0 = 0, \quad Q_0 = d/2 \in [m/4, 3m/4]. \quad (35)$$

This exclusion respects the zero row sum at index zero.

A.4 Bounds for the column tails

The selection concentrates the columns. In the ideal case of uniform low values and $X = m - Y$, its two classes occupy the intervals $[Y, Y+m]$ and $[m-Y, 2m-Y]$, shown in Fig. 16. We bound the finite errors through the two tails to verify that this concentration leaves enough room for the chosen diagonal. For a real number x , write $x_+ = \max\{x, 0\}$ for its *positive part*.

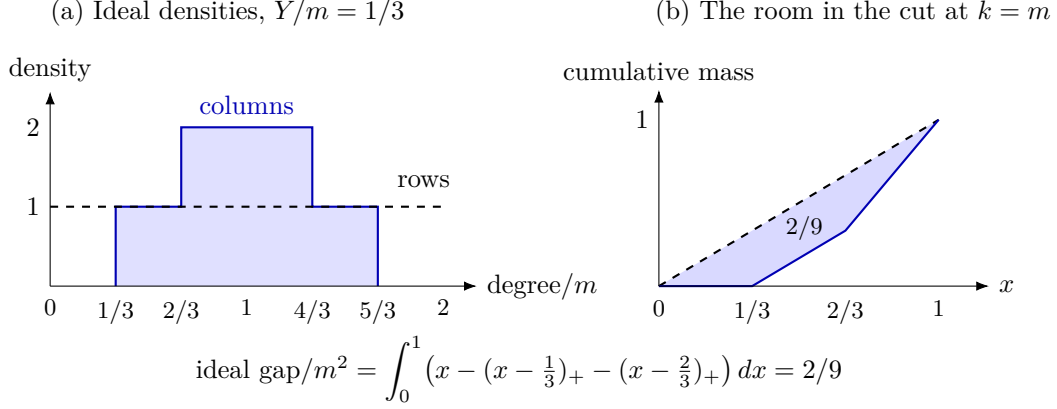


Figure 16: How concentration of the columns creates room for the chosen diagonal. In (a), the ideal column measure is the sum of uniform densities on $[1/3, 4/3]$ and $[2/3, 5/3]$; the dashed row measure is uniform on $[0, 2]$. The horizontal coordinate is a row or column sum divided by m ; both measures have mass two and mean one. In (b), their cumulative masses up to the normalized value $x \leq 1$ enclose an area of $2/9$. Multiplying this area by m^2 gives the ideal gap at the cut $k = m$. The finite construction requires the error estimates below to retain enough of this gap.

Put

$$\ell_1 = Y - 2H, \quad \ell_2 = m - Y - 2H, \quad \rho = m/6. \quad (36)$$

Both ℓ_1 and ℓ_2 are at least ρ . An unselected low value with odd r_i is at least $X - C_\chi$, by the upper bound in (28); an eligible unselected low value with even r_i is at least $Y - C_\chi - 1$. Equations (33) and (36) imply

$$X - C_\chi \geq m - Y - 2g - 2C_\chi - 1 \geq \ell_2, \quad Y - C_\chi - 1 \geq \ell_1.$$

The lifted columns are at least $m + 1$, and (35) handles index zero. Thus $Q_i \geq \rho$ for every i .

Let $\tilde{C}_e$ exclude index zero from C_e . The lower bounds in (28) imply, for $0 \leq t \leq m$,

$$C_o(t) \geq t - 1 - C_\chi - g_o, \quad \tilde{C}_e(t) \geq t - 1 - C_\chi - g_e. \quad (37)$$

Therefore every lifted column with odd r_i is at most U_o , and every lifted column with even r_i is at most U_e , where

$$U_o = m + X + C_\chi + g_o + 1, \quad U_e = m + Y + C_\chi + g_e + 2.$$

For example, the X -th low value in the odd-residue class is at most $X + C_\chi + g_o + 1$ by (37); the same reasoning applies to the even class. Using (33) gives

$$\begin{aligned} n - U_o &\geq Y - \delta_0 - 2C_\chi - g_o \geq \ell_1, \\ n - U_e &= m - Y - g - C_\chi - g_e - 1 \geq \ell_2. \end{aligned} \quad (38)$$

Here $\delta_0 \leq g$ and $g_e, g_o \leq g$. Unlifted columns are at most $m < n - \rho$. We conclude that

$$\rho \leq Q_i \leq n - \rho \quad (0 \leq i < n). \quad (39)$$

Since $m \geq 150$, the columns are integers between one and $n - 2$.

For real t , put $C_Q(t) = \#\{i : Q_i \leq t\}$. For $0 \leq t \leq m$, lifted values do not contribute, and selection of the smallest values gives

$$C_Q(t) = (C_o(t) - X)_+ + (\tilde{C}_e(t) - Y)_+ + \mathbf{1}_{\{Q_0 \leq t\}}.$$

The preceding estimates yield

$$C_Q(t) \leq (t - \ell_1)_+ + (t - \ell_2)_+ + \mathbf{1}_{\{m/4 \leq t\}}. \quad (40)$$

For $k \geq m$, only lifted columns contribute to $\sum_i (Q_i - k)_+$. In a class where the smallest T low values were selected, the number of selected values above a real threshold t is $(T - C(t))_+$. Here $C = C_o$ for odd residues and $C = \tilde{C}_e$ for eligible even residues. Integrating (37) therefore gives

$$\sum_i (Q_i - k)_+ \leq \frac{(U_o - k)_+^2}{2} + \frac{(U_e - k)_+^2}{2}. \quad (41)$$

For the even class at $k = m$, the low-value integration starts at -1 . On $[-1, 0]$, its tail count is at most Y , which is bounded by $Y + C_\chi + g_e + 1 - t$. Thus (41) also holds at that endpoint. Above m , the low-value tail is zero, and extending the positive-part bound only increases the integral.

A.5 Uniform room in the matrix inequalities

The tail estimates must now give $G_k \geq n$ throughout the range required by Lemma A.1. We use the lower and upper tails through the respective identities

$$\begin{aligned} G_k &= \frac{k(k+1)}{2} - \int_0^k C_Q(t) dt, \\ G_k &= \frac{(n-k)(n-k-1)}{2} - \sum_i (Q_i - k)_+ \end{aligned} \quad (42)$$

from (26) and $\sum_i Q_i = N$. These forms reduce the small and large cuts to a common one-variable bound. For real $x \in [\rho, m]$, define

$$F(x) = \frac{x(x-1)}{2} - \frac{(x - \ell_1)_+^2}{2} - \frac{(x - \ell_2)_+^2}{2}. \quad (43)$$

For $k \leq m$, integrating (40) in the first identity of (42) gives

$$G_k \geq F(k) + k - (k - m/4)_+ \geq F(k).$$

For $k \geq m$, use (38) and (41) in the second identity to obtain $G_k \geq F(n - k)$. When $\min_i Q_i \leq k \leq \max_i Q_i$, the relevant argument is in $[\rho, m]$, by (39) and $n - m = m + 1 - g < m$.

The minimum of F on this interval is attained at an endpoint. Its derivative is initially $x - 1/2 > 0$, then $\min(\ell_1, \ell_2) - 1/2 > 0$, and finally $\ell_1 + \ell_2 - x - 1/2$. The last expression decreases, so it can produce a maximum but cannot produce an interior minimum. At the endpoints,

$$\begin{aligned} F(\rho) &= m^2/72 - m/12, \\ F(m) &= Y(m - Y) - 2mH - 4H^2 - m/2 \geq m^2/36 - m/2. \end{aligned} \quad (44)$$

The second estimate uses $m/3 \leq Y \leq 2m/3$ and $H \leq m/12$. Both expressions are at least $2m$ when $m \geq 150$. Thus $G_k \geq 2m \geq n$ for every cut in the occupied column range. Lemma A.1 realizes the margins and diagonal in (34), proving Proposition 6.5.

A.6 An explicit threshold

To verify the hypotheses of Proposition 6.5 for every $n \geq 10\,000$, we combine published bounds on prime gaps and quadratic-character sums. Dusart [7, Theorem 1, p. 55] proves that for every $x \geq 3275$ there is a prime

$$x < p \leq x \left(1 + \frac{1}{2 \log^2 x} \right). \quad (45)$$

For the quadratic character modulo a prime $p > 2$, Lapkova [14, Lemma 3A] bounds every interval sum. Her two cases, $\chi(-1) = 1$ and $\chi(-1) = -1$, have leading constants $2/\pi^2$ and $1/(2\pi)$, respectively. Both imply the common bound

$$C_\chi = (0.203 \log p + 0.947) \sqrt{p} + 1.668. \quad (46)$$

The constants for $\chi(-1) = 1$ are $2/\pi^2, 0.9467, 1.668$; those for $\chi(-1) = -1$ are smaller. All logarithms are natural.

Corollary A.2. *For every integer $n \geq 10\,000$, a prime $p \geq n + 5$ and a binary matrix satisfying (20) exist.*

Proof. Apply (45) to $x = n + 4$, and put $g = p - n$ and $m = (p - 1)/2$. Since $\log(n + 4) > 9$,

$$\frac{g}{n} \leq \frac{4}{10\,000} + \frac{1.0004}{162} < 0.006576, \quad p < 1.01n.$$

The functions $\log(1.01n)/\sqrt{n}$ and $1/\sqrt{n}$ decrease for $n \geq 10\,000$. Using $\log 10\,100 < 10$ and $\sqrt{1.01} < 1.01$ in (46),

$$\frac{C_\chi}{n} < \frac{1.01(2.03 + 0.947)}{100} + \frac{1.668}{10\,000} = 0.0302345.$$

Consequently

$$\frac{C_\chi + g + 2}{n} < 0.0302345 + 0.006576 + 0.0002 = 0.0370105 < \frac{1}{24}.$$

Since $g \geq 5$, one has $m > n/2$, so $C_\chi + g + 2 < m/12$. Also $m \geq 150$. Proposition 6.5 applies. $\square$

B Realizing endpoint profiles

The constructions in Sections 4 and 5 use two standard sufficient conditions for realizing chosen landmark counts. The first produces binary matrices for different endpoint profiles; the second handles symmetric selections, including middle vertices. We recall their short proofs and describe how each realization yields an ordinary landmark set.

Lemma B.1. *Let $r_i, c_i \in \{0, \dots, n - 1\}$, for $0 \leq i < n$, have common sum E , and suppose $r_i \leq M_r$, $c_i \leq M_c$, with nonnegative integer bounds. If*

$$(M_r + 1)M_c \leq E,$$

there is a binary $n \times n$ matrix with zero diagonal, row sums r_i , and column sums c_i .

Proof. This sufficient condition is due to Burstein and Rubin [4]; we give the flow argument. If $E = 0$, use the zero matrix. Otherwise $M_c > 0$. The bipartite flow criterion with diagonal entries forbidden [10] requires, for each set I of s rows,

$$\sum_{i \in I} r_i \leq \sum_{j=0}^{n-1} \min\{c_j, s - \mathbf{1}_{j \in I}\}.$$

For $1 \leq s \leq M_c$, the right side is at least

$$\sum_j \min(c_j, s) - s \geq \frac{sE}{M_c} - s \geq sM_r.$$

For $s \geq M_c + 1$, it equals E . For the empty row set, both sides are zero. All cuts therefore have sufficient capacity, and integral flow supplies the binary matrix. $\square$

Given such a matrix for positions $t, L - t$, where t is an integer with $1 \leq t < L/2$, select $x_{ij}(t)$ when its (i, j) -entry is one. Equivalently, on support ij with $i < j$, its two relevant entries select positions t and $L - t$. The row sums become $D_i(t)$, and the column sums become $D_i(L - t)$.

Lemma B.2. *Let m, M be positive integers. Suppose positive integers $d_0, \dots, d_{n-1}$ have even sum and satisfy*

$$m \leq d_i \leq M < n, \quad 4nm \geq (M + m + 1)^2.$$

Then they are the degrees of a simple graph.

Proof. This is the sufficient condition of Zverovich and Zverovich [16, Theorem 6], also recalled in [6, Theorem 2]. Order the degrees nonincreasingly and use the Erdős–Gallai inequalities. For $s \leq m$, their right side is $s(n - 1)$, which dominates the sum of the first s degrees. For $s > m$, the right side minus that sum is at least

$$s(s - 1) + (n - s)m - sM = s^2 - (M + m + 1)s + nm \geq 0.$$

The last inequality follows from the assumed discriminant bound. Thus all the inequalities hold. $\square$

For a simple graph with degrees d_i , selecting both positions $t, L - t$ on each of its edges gives $D_i(t) = D_i(L - t) = d_i$. If $L = 2k$, selecting just the middle vertex on each graph edge gives $D_i(k) = d_i$. In that case the total number of middle landmarks is $\frac{1}{2} \sum_i d_i$.

C Exact lattice counts and restricted depths

To turn the profile count into a finite lower bound, we fill integer-vector layers in increasing order of their coordinate sum. For positive integers Q, d and a nonnegative integer N , let $F_{Q,d}(N)$ be the minimum total coordinate sum of N vectors in $\mathbb{Z}_{\geq 0}^d$, with each vector allowed at most Q occurrences. This allowance accounts for endpoints with the same vector that can still be distinguished by their local choices. The minimization ignores the cost of those local selections; in particular, $F_{Q,d}(0) = 0$.

Lemma C.1. *Let ℓ be the smallest nonnegative integer satisfying $Q \binom{\ell+d}{d} \geq N$. Then*

$$F_{Q,d}(N) = Qd \binom{\ell+d-1}{d+1} + \ell \left(N - Q \binom{\ell+d-1}{d} \right). \quad (47)$$

For fixed Q, d ,

$$F_{Q,d}(N) = \frac{d}{d+1} \left(\frac{d!}{Q} \right)^{1/d} N^{1+1/d} + O_{Q,d}(N). \quad (48)$$

Here $\binom{u}{v} = 0$ for nonnegative integers $u < v$.

Proof. There are $\binom{r+d}{d}$ nonnegative integer vectors of coordinate sum at most r . This is the Ehrhart polynomial of the standard simplex [3]. A minimum fills each layer below ℓ with Q copies of every vector, and then uses the required number of vectors in layer ℓ . The completed layers cost

$$Q \sum_{r=0}^{\ell-1} \binom{r+d-1}{d-1} = Qd \binom{\ell+d-1}{d+1}.$$

This gives the exact formula. As N tends to infinity,

$$N = \frac{Q}{d!} \ell^d + O_{Q,d}(\ell^{d-1}), \quad F_{Q,d}(N) = \frac{Qd}{(d+1)!} \ell^{d+1} + O_{Q,d}(\ell^d).$$

Eliminating ℓ proves the asymptotic formula. $\square$

Theorem C.2 (The number of active depths). *Let s be an integer with $1 \leq s \leq \lfloor L/2 \rfloor$, put $Q_s = 2^{2s+1}$, and suppose $n > Q_s + 1$. If an edge multiset resolving set in $S_L(K_n)$ has all its internal landmarks at depth at most s , then its number B of internal landmarks satisfies*

$$B \geq F_{Q_s,s}(n).$$

Original vertices may also be selected. In particular, for fixed s ,

$$B \geq a_s n^{1+1/s} - O_s(n), \quad a_s = \frac{s}{s+1} \left(\frac{s!}{2^{2s+1}} \right)^{1/s}.$$

The lower bound is uniform over all integers $L \geq 2s$.

Proof. Fix an original vertex i , and let e_{ij} be the edge adjacent to i on support ij . The far-endpoint columns of its histogram are

$$B_t(z) = \begin{cases} z^{L-1+t} - z^{L+t}, & 1 \leq t \leq \lfloor L/2 \rfloor, \\ 0, & \lfloor L/2 \rfloor < t < L. \end{cases}$$

Consequently its histogram has the form

$$H_{e_{ij}}(z) = C_i(z) + \alpha_j(z^{L-1} - z^L) + \sum_{t=1}^{\lfloor L/2 \rfloor} D_j(t)(z^{L-1+t} - z^{L+t}) + E_{e_{ij}}(z),$$

where C_i is independent of j , and the support correction $E_{e_{ij}}$ from (3) depends only on the landmark bits on support ij .

First suppose $L > 2s$. The only allowed internal positions are $1, \dots, s, L-s, \dots, L-1$. The second group lies above $\lfloor L/2 \rfloor$, so assign to j the profile

$$p_j = (D_j(1), \dots, D_j(s)).$$

For fixed i, p_j , the displayed histogram is determined by α_j and the $2s$ bits on support ij . There are at most $2^{2s+1} = Q_s$ possibilities.

If $L = 2s$, use instead

$$p_j = (D_j(1), \dots, D_j(s-1), \lfloor D_j(s)/2 \rfloor).$$

The additional parity bit of $D_j(s)$ recovers its unrounded value. Together with α_j and the $2s - 1$ support bits, this again gives at most 2^{2s+1} histograms.

No profile class can have more than Q_s members. For a proper class with more than Q_s members, take i outside it and apply the preceding bound to its targets. If the class contains all n original vertices, any origin sees $n - 1 > Q_s$ targets. Either case would give a collision.

For $L > 2s$, every internal landmark contributes once to $\sum_j \|p_j\|_1$, so this sum is B . For $L = 2s$, every nonmiddle landmark still contributes once, while a middle landmark is counted at both endpoints before the division by two. Hence

$$\sum_j \|p_j\|_1 = B - \frac{1}{2} |\{j : D_j(s) \text{ is odd}\}| \leq B.$$

The multiplicity bound gives $B \geq F_{Q_s, s}(n)$, and Lemma C.1 gives the asymptotic statement. $\square$

Corollary C.3. *For $L \geq 2$, put $k = \lfloor L/2 \rfloor$. If $n > 2^{2k+1} + 1$, every resolving set in $S_L(K_n)$ has at least $F_{2^{2k+1}, k}(n)$ internal landmarks. For each fixed L ,*

$$\text{edim}_m(S_L(K_n)) \geq a_k n^{1+1/k} - O_k(n)$$

whenever a resolving set exists. In particular $a_2 = 1/6$.

Proof. Every internal vertex has depth at most k , so Theorem C.2 applies with $s = k$. Its finite bound and asymptotic constant give the assertion; at $k = 2$, the constant is $(2/3)(2/32)^{1/2} = 1/6$. $\square$

D Restrictions and constants in the label constructions

The label constructions leave two quantitative questions at odd lengths: the cost of symmetric profiles and the scale needed for asymmetric ones. We answer the first by matching a Sidon construction with a central-edge count, and improve the second by bounding the correction terms more closely.

D.1 Symmetric profiles at odd lengths

To determine the cost of symmetric profiles at odd lengths, we match a construction with the central-edge lower bound in Proposition F.1. Integer Sidon labels keep the endpoint sums distinct, and writing them in a suitable base turns them into short coordinate vectors.

Corollary D.1. *Let $L = 2k + 1 \geq 5$, suppose $n > 2$, choose a prime P with $n \leq P < 2n$, and put $Q = \lceil (2P^2)^{1/k} \rceil$. If $n \geq 14Q$, there is a resolving set with symmetric profiles and size less than $12knQ$. For every fixed $k \geq 3$, the minimum size of a resolving set with symmetric profiles in $S_{2k+1}(K_n)$ is $\Theta_k(n^{1+2/k})$.*

Proof. For $i = 0, \dots, n - 1$, take the integer labels

$$z_i = 2Pi + (i^2 \bmod P) \in [0, 2P^2).$$

These are the classical integer Sidon labels of Erdős and Turán [8]. For completeness, an equality $z_i + z_j = z_u + z_v$ first gives $i + j = u + v$ over the integers: the difference between two

residue sums has absolute value below $2P$. The residue sums then agree, so reduction modulo P gives $i^2 + j^2 = u^2 + v^2$. Since P is odd, the products agree as well. The two unordered pairs are the roots of the same monic quadratic over $\mathbb{F}_P$, and the indices $0, \dots, n-1$ are distinct modulo P .

Write each z_i in base Q , using k digits. The resulting vectors belong to $\{0, \dots, Q-1\}^k$ because $Q^k \geq 2P^2$. Equality of their vector pair sums would imply equality of the corresponding integer pair sums. Hence they form a Sidon set, and Proposition 4.2 applies.

For fixed $k \geq 3$, Bertrand's postulate provides the stated prime, and $Q \leq 8^{1/k} n^{2/k} + 1$. Thus $n \geq 14Q$ eventually holds. The symmetric central-edge bound with $s = k$ in Proposition F.1 supplies the matching lower order. $\square$

D.2 A smaller scale for odd-length labels

The construction in Proposition 5.3 remains valid with the fixed scale $K = 7$ if its correction terms are bounded independently of k . We obtain this improvement by checking the nonzero columns among the last k positions of each incidence matrix; within each family, their distance coordinates are all distinct.

Lemma D.2. *Let $k \geq 1$ and $0 \leq r \leq k$ be integers, and put $L = 2k + 1$. Use the integer coefficient matrices $\mathcal{A}, \mathcal{B}_r$ from (6), after removing z^a , with $\delta = 2r$. For $r = 0$, put $\mathcal{B}_0 = \mathcal{A}$. If vectors $u, v \in \mathbb{Z}^{2k}$ vanish in coordinates $1, \dots, k$ and satisfy $\|u\|_\infty, \|v\|_\infty \leq 1$, where this norm is the largest absolute coordinate, then*

$$\|\mathcal{A}u + \mathcal{B}_rv\|_\infty \leq 2 \quad (0 \leq r \leq k).$$

Proof. In the last k coordinates the columns of $\mathcal{A}$ are

$$\mathcal{A}_t = e_t - e_{2L-t}, \quad k+1 \leq t \leq 2k.$$

All $2k$ displayed endpoints are distinct: positive exponents lie in $[k+1, 2k]$, and negative ones lie in $[2k+2, 3k+1]$. Consequently $\|\mathcal{A}u\|_\infty \leq 1$.

For $r > 0$, the only nonzero columns of $\mathcal{B}_r$ among these coordinates are

$$(\mathcal{B}_r)_t = e_{t+2r} - e_{2L-t}, \quad k+1 \leq t \leq 2k-r.$$

Their positive exponents are distinct and at most $2k+r$. Their negative exponents are distinct and at least $2k+r+2$. Thus all endpoints within this restricted family are again distinct, and $\|\mathcal{B}_rv\|_\infty \leq 1$. The family is empty if $r = k$. For $r = 0$, the same estimate follows from $\mathcal{B}_0 = \mathcal{A}$. The triangle inequality proves the assertion. $\square$

Proposition D.3. *Let $L = 2k + 1$, $k \geq 2$. Choose a prime $q > k$ with $q^k \geq 2n$, and put $m = 7(q-1) + 1$. If $n > 12$ and $2n \geq 9m + 3$, then*

$$\text{edim}_m(S_{2k+1}(K_n)) < k(2m+1)n.$$

Proof. Use the labels and balanced degrees of Proposition 5.3, with $K = 7$. The margin, size, and depth arguments are unchanged. In (13), each correction difference is supported on the last k coordinates and has norm at most one. Lemma D.2 bounds its image by two. The support error adds at most four, so the total is at most $6 < 7$. The integer divisibility argument therefore gives (14), and the same recovery proves the claim. $\square$

E A bound uniform in the path length

We prove Theorem 7.2 using only positions $1, 2, L-2, L-1$, so the construction uses four endpoint counts for every $L \geq 5$. A four-coordinate Sidon labeling provides the counts; we balance and realize them as landmark selections, then recover each edge from its histogram.

E.1 Four-coordinate labels

The four counts must identify endpoint pairs at central edges and still permit recovery away from the center. We obtain labels with both properties by writing a finite-field parabola in four coordinates, then scale and balance them so that binary matrices realize the counts. Fix integers $n \geq 222$ and $L \geq 5$, and let q be the least prime strictly larger than $\lceil \sqrt{n} \rceil$. Choose a nonsquare $\nu \in \mathbb{F}_q$ and n distinct pairs

$$(u_i, v_i) \in \mathbb{F}_q \times \mathbb{F}_q^\times.$$

Index these pairs by $0 \leq i < n$. There are enough pairs because $q(q-1) \geq n$. With representatives in $\{0, \dots, q-1\}$, put

$$c_i = (u_i, v_i, u_i^2 + \nu v_i^2 \bmod q, 2u_i v_i \bmod q). \quad (49)$$

The first two coordinates identify i , and $v_i \neq 0$.

Lemma E.1. *The vectors c_i have distinct unordered pair sums in $\mathbb{Z}^4$, including pairs with repetition.*

Proof. In $\mathbb{F}_{q^2} = \mathbb{F}_q(\sqrt{\nu})$, put $x_i = u_i + v_i\sqrt{\nu}$. An equality of integer vector sums implies

$$x_i + x_j = x_k + x_\ell, \quad x_i^2 + x_j^2 = x_k^2 + x_\ell^2.$$

Since q is odd, the products also agree. Both pairs are the roots of the same monic quadratic. The pairs of indices therefore agree as multisets. This is the finite-field parabola construction of [5, 15], in coordinates. $\square$

The four coordinates will determine counts at positions $1, 2, L-2, L-1$, in that order. Their margins must have matching totals. For $t \in \{1, 2, 3, 4\}$, define

$$\mu_t = \left\lfloor \frac{3 \sum_i c_{i,t}}{n} \right\rfloor, \quad R_t = 3 \sum_i c_{i,t} - n\mu_t \in \{0, \dots, n-1\},$$

and set $m = 3q-2$, $T = 2m = 6q-4$. The number T is the common base of each target margin. Start with $T + 3c_{i,t} - \mu_t$. In coordinate 3, add $\sigma_3 = \text{sgn}(R_2 - R_3)$ to exactly $|R_2 - R_3|$ entries. In coordinate 4, add $\sigma_4 = \text{sgn}(R_1 - R_4)$ to exactly $|R_1 - R_4|$ entries. The chosen entries may, for example, be the first ones in the indexing of the vertices. Write these corrections as $\varepsilon_{i,t}$, with $\varepsilon_{i,1} = \varepsilon_{i,2} = 0$, and define

$$D_{i,t} = T + 3c_{i,t} - \mu_t + \varepsilon_{i,t}, \quad D_i = (D_{i,1}, \dots, D_{i,4}). \quad (50)$$

Then

$$\sum_i D_{i,1} = \sum_i D_{i,4} = nT + R_1, \quad \sum_i D_{i,2} = \sum_i D_{i,3} = nT + R_2, \quad m \leq D_{i,t} \leq 3m. \quad (51)$$

The last bounds use $0 \leq \mu_t \leq 3(q-1)$ and $|\varepsilon_{i,t}| \leq 1$.

Lemma E.2. *The vectors D_i retain distinct unordered pair sums. In each of their first two coordinates, the possible values differ by multiples of three.*

Proof. In each corrected coordinate all corrections belong to $\{0, \sigma_t\}$ for one fixed sign σ_t . The difference between two sums of two corrections has absolute value at most two. Hence an equality $D_i + D_j = D_k + D_\ell$ forces $c_i + c_j = c_k + c_\ell$ coordinatewise: the other part of the difference is a multiple of three. Apply Lemma E.1. The first two coordinates have no corrections, which proves the second assertion. $\square$

Apply Lemma B.1 to $(D_{\cdot,1}, D_{\cdot,4})$ and $(D_{\cdot,2}, D_{\cdot,3})$. By (51), their totals are at least nT , and their maximum is at most $\Delta = 3m$. The sufficient inequality becomes

$$\Delta(\Delta + 1) \leq nT \iff 2n \geq 27q - 15. \quad (52)$$

It also implies $\Delta < n$: $n \geq (3/2)(\Delta + 1) > \Delta$.

For $n \geq 1024$, Bertrand's postulate gives $q < 2\lceil\sqrt{n}\rceil < 2\sqrt{n} + 2$. The inequality $2n \geq 54\sqrt{n} + 39$ holds at $n = 1024$ and its left side minus its right side is increasing thereafter. This proves (52). Table 3 covers the remaining interval; the left endpoint of each interval suffices.

Table 3: Prime intervals establishing (52) for $222 \leq n < 1024$. For each interval, q is the least prime strictly exceeding $\lceil\sqrt{n}\rceil$.

Range of n	q	$27q - 15$	Range of n	q	$27q - 15$
222–256	17	444	485–784	29	768
257–324	19	498	785–900	31	822
325–484	23	606	901–1023	37	984

Let $M^{(1)}, M^{(2)}$ be any two matrices realizing these pairs of margins. For every $i < j$, select vertices according to

$$\eta_{ij}(1) = M_{ij}^{(1)}, \quad \eta_{ij}(2) = M_{ij}^{(2)}, \quad \eta_{ij}(L-2) = M_{ji}^{(2)}, \quad \eta_{ij}(L-1) = M_{ji}^{(1)}. \quad (53)$$

Let W consist precisely of these selected vertices. The four positions are distinct for $L \geq 5$. Their degree profile is

$$(D_i(1), D_i(2), D_i(L-2), D_i(L-1)) = D_i.$$

Writing $T_1 = \sum_i D_i(1)$, $T_2 = \sum_i D_i(2)$, we obtain

$$|W| = T_1 + T_2 = 2nT + R_1 + R_2 < (12q - 6)n < 24n^{3/2} + 18n. \quad (54)$$

Thus the size estimate will follow once every realization is shown to resolve.

E.2 Recovering every edge from its histogram

To prove that the constructed selection resolves the graph, we recover an edge from its histogram. The last occupied level determines its depth, and the remaining levels recover the endpoints through separate counts, a Sidon pair sum, or the mixed coordinates listed in Table 4.

Table 4: Recovery after the depth is known. The three cases distinguish whether the endpoint counts can be read separately or first appear as a sum.

Case	Information recovered	How it identifies the edge
$\delta \geq 2$	u_i, v_i , then u_j, v_j	Read the nearer endpoint from (56), then the farther one from (57) or (58).
$\delta = 0$	$D_i + D_j$	The Sidon property identifies the unordered pair. Its central edge is unique.
$\delta = 1$	$u_i, v_j, v_i + u_j, 2u_jv_j \bmod q$	Equation (59) gives these four values. Divide the fourth by $2v_j \neq 0$ in $\mathbb{F}_q$ to recover u_j , then subtract from the third to recover v_i . The ordered pair fixes the central position.

Proposition E.3. *Every selection defined by (53), using any matrices with the margins (50), resolves $S_L(K_n)$.*

Proof. We first recover the depth a of an edge. By Table 1, every landmark is at distance at most $L + a + 2$. The bounds in (51) give at least nm landmarks of depth two, while the supports sharing i or j contain at most $4\Delta = 12m$ such landmarks. Since $n > 12$, a depth-two landmark exists on a disjoint support. Consequently,

$$\max\{t : H_e^W(t) > 0\} = L + a + 2. \quad (55)$$

It remains to distinguish edges at the same a .

We repeatedly use the following elementary observation. If $y = 3x + C + \xi$, where C is known and ξ lies in a known interval of three consecutive integers, then y determines the integer x . The intervals for consecutive values of x are disjoint.

Write $\eta_t = \eta_{ij}(t)$, with value zero at positions outside $1, 2, L - 2, L - 1$, and $D_i(t) = 0$ at inactive positions. There are three cases.

Case 1: $\delta = b - a \geq 2$. For $s = 1, 2$, Table 1 gives

$$H_e^W(a + s) = D_i(s) - \eta_s + \eta_{2a+s+1}. \quad (56)$$

The paths from i supply $D_i(s) - \eta_s$; the supporting path can supply the displayed additional bit. Other supports have larger distances. The error is in $[-1, 1]$. Therefore these layers recover u_i, v_i , and hence i .

If $a \geq 1$, the next useful layers are

$$H_e^W(b + s) = D_j(s) + D_i(\delta + s) - \eta_{L-s} - \eta_{\delta+s}, \quad s = 1, 2. \quad (57)$$

These layers precede contributions from disjoint supports. Contributions from the supporting path itself vanish, since $b + s > b$, its largest edge-to-vertex distance on that path. After subtracting the known $D_i(\delta + s)$, the remaining error is in $[-2, 0]$. The two coordinates of j are therefore recovered.

If $a = 0$, the second layer in (57) would receive contributions from disjoint supports. Instead use

$$\begin{aligned} H_e^W(L) &= D_j(1) - \eta_{L-1}, \\ H_e^W(L + 1) &= T_1 - D_i(1) - D_i(L - 1) - D_j(1) + D_j(2) + \eta_{L-1} - \eta_{L-2}. \end{aligned} \quad (58)$$

The first equation recovers u_j , so $D_j(1)$ is known. After moving the other known terms in the second equation, the remaining error is in $[-1, 1]$, recovering v_j . Thus the histogram identifies the ordered endpoints and the position.

Case 2: $\delta = 0$. Here $L = 2h + 1$ and the edge is central. For $s = 1, 2$, put $\gamma_{ij,s} = \eta_{ij}(s) + \eta_{ij}(L - s)$, the number of selected vertices of depth s on its own support. For each active position $t \in \{1, 2, L - 2, L - 1\}$, put

$$X_{ij,t} = D_i(t) + D_j(t) - \eta_{ij}(t) - \eta_{ij}(L - t).$$

This counts landmarks at that position on paths sharing one endpoint, with ij excluded. The histogram identity (60) records $\gamma_{ij,s}$ at distance $h - s$ and $X_{ij,t}$ at distance $h + t$; these levels are distinct. Since

$$D_i(t) + D_j(t) = X_{ij,t} + \gamma_{ij,\min(t, L-t)}$$

at each active position, it determines $D_i + D_j$. Lemma E.2 identifies the unordered support, which contains exactly one central edge.

Case 3: $\delta = 1$. Here $L = 2h$, $h \geq 3$, and $a = h - 1, b = h$. Directly from Table 1,

$$\begin{aligned} H_e^W(h) &= D_i(1) - \eta_1, \\ H_e^W(h+1) &= D_i(2) + D_j(1) - \eta_2 - \eta_{L-1}, \\ H_e^W(h+2) &= D_j(2) - \eta_{L-2}, \\ H_e^W(3h-1) &= D_j(L-1) - \eta_1. \end{aligned} \tag{59}$$

For the last equation, only landmarks at position $L - 1$ on paths leaving j reach that layer; landmarks on disjoint supports begin at $3h$. The first and third equations recover u_i, v_j , and the second recovers the integer $v_i + u_j$. The last recovers the fourth code coordinate $2u_jv_j \bmod q$. Indeed, its total correction $\varepsilon_{j,4} - \eta_1$ lies in $[-1, 1]$ when $\sigma_4 = 1$, in $[-2, 0]$ when $\sigma_4 = -1$, and in $[-1, 0]$ when $\sigma_4 = 0$. Each is a known interval of width at most two.

Since $v_j \neq 0$, division by $2v_j$ in $\mathbb{F}_q$ gives u_j . The integer sum already recovered then gives v_i . The ordered support and edge position are determined. This explains the nonzero second coordinate imposed in (49). All three cases identify the support and position once the depth is known. Together with the size estimate (54), this proves Theorem 7.2. $\square$

F Additional quantitative bounds

We refine the finite lower bounds and extend the known existence ranges for long replacement paths. Central-edge histograms give stronger counts, while independent selections, controlled through distance-level forests, give further existence estimates and allow the selection probability to vary.

F.1 Finite central-edge bounds

Counting only the central edges gives another finite estimate and, at some odd lengths, a larger leading constant. Its symmetric specialization also quantifies the cost of equal profiles at opposite positions.

Proposition F.1. *Let $n \geq 3$, let k, s be integers with $k \geq 1$ and $1 \leq s \leq k$, and put $L = 2k + 1$. Suppose a landmark set distinguishes the central edges and has no internal landmark of depth greater than s . Its number B of internal landmarks satisfies*

$$B \geq \left\lceil \frac{F_{3^{s+1}, 2s}(\binom{n}{2})}{2(n-2)} \right\rceil.$$

If its endpoint profiles are symmetric, $D_i(t) = D_i(L - t)$, then also

$$B \geq \left\lceil \frac{F_{3^{s+1}, s}(\binom{n}{2})}{n-2} \right\rceil.$$

In particular the respective asymptotic orders are $\Omega_s(n^{1+1/s})$ and $\Omega_s(n^{1+2/s})$. The first bound has leading constant

$$c_s = \frac{2s}{2s+1} \left(\frac{(2s)!}{3^{s+1}} \right)^{1/(2s)} 2^{-2-1/(2s)}.$$

Proof. For a support ij , put

$$\begin{aligned} \gamma_{ij,t} &= \eta_{ij}(t) + \eta_{ij}(L - t) & (1 \leq t \leq s), \\ X_{ij,t} &= D_i(t) + D_j(t) - \eta_{ij}(t) - \eta_{ij}(L - t) & (1 \leq t < L). \end{aligned}$$

Each $X_{ij,t}$ counts selected vertices on paths sharing an endpoint, other than ij , and is therefore nonnegative. Only the positions

$$J_s = \{1, \dots, s, L-s, \dots, L-1\}$$

can have a nonzero X -coordinate. Let $A_0 = \sum_i \alpha_i$, and let T_t be the total number of selected internal vertices of depth t . Write $H_{ij} = H_{e_{ij}}^W$ for the histogram of the unique central edge e_{ij} on ij . Its nonzero ranges are

$$\begin{aligned} H_{ij}(k-t) &= \gamma_{ij,t} & (1 \leq t \leq s), \\ H_{ij}(k) &= \alpha_i + \alpha_j, \\ H_{ij}(k+t) &= X_{ij,t} & (t \in J_s), \\ H_{ij}(L+k) &= A_0 - \alpha_i - \alpha_j, \\ H_{ij}(L+k+t) &= T_t - \gamma_{ij,t} - X_{ij,t} - X_{ij,L-t} & (1 \leq t \leq s). \end{aligned} \tag{60}$$

All other entries are zero. These ranges are disjoint. For the fixed landmark set, the central histogram is thus equivalent to the tuple consisting of the $s+1$ tags $\gamma_{ij,1}, \dots, \gamma_{ij,s}, \alpha_i + \alpha_j$, each in $\{0, 1, 2\}$, and the $2s$ coordinates $(X_{ij,t})_{t \in J_s}$. The $\binom{n}{2}$ such tagged vectors must be distinct.

Every selected internal vertex on support uv contributes once for each support containing u but not v , and once for each support containing v but not u . Hence

$$\sum_{i < j} \sum_{t \in J_s} X_{ij,t} = 2(n-2)B.$$

Apply Lemma C.1 with $Q = 3^{s+1}$, $d = 2s$, and $N = \binom{n}{2}$, then divide by $2(n-2)$. Integrality gives the first finite bound, and the asymptotic formula gives the stated constant.

Under symmetry, $X_{ij,t} = X_{ij,L-t}$. The first s coordinates therefore determine the tuple, and their total sum is $(n-2)B$. The same argument with $d = s$ gives the second bound. $\square$

The unrestricted central bound is the specialization $s = k$. For example, at $L = 5$, its leading constant is $c_2 = 2/(5\sqrt{6})$, whereas the star bound gives $a_2 = 1/6$. At $L = 7$,

$$a_3 = \frac{3}{4}(3/64)^{1/3} \approx 0.270421794, \quad c_3 = \frac{6}{7}(80/9)^{1/6} 2^{-13/6} \approx 0.274765879.$$

Thus retaining both finite bounds preserves information beyond the common growth exponent.

F.2 A refined lower bound at length three

The count in Proposition F.1 treats the local landmark choices as a fixed number of tags. We obtain a stronger bound by retaining the number of supports of each local type. The counting ingredient is again the lattice-point count for a simplex [3].

For an integer $t \geq 0$, write $F(t) = F_{3,2}(t)$, the minimum total cost of t distinct elements of $\{0, 1, 2\} \times \mathbb{Z}_{\geq 0}^2$, where the cost of (a, u, v) is $u + v$. Thus $F(0) = 0$. If $t > 0$ and ℓ is the least nonnegative integer with $3\binom{\ell+2}{2} \geq t$, then

$$F(t) = \ell t - \frac{\ell(\ell+1)(\ell+2)}{2}, \quad F(t) = \frac{2}{3}\sqrt{\frac{2}{3}}t^{3/2} + O(t). \tag{61}$$

Indeed, there are $3(j+1)$ points of cost j . Filling the levels below ℓ and then the required part of level ℓ gives the exact formula. The relations $t = 3\ell^2/2 + O(\ell)$ and $F(t) = \ell^3 + O(\ell^2)$ give its asymptotic form.

Put $M = \binom{n}{2}$. For integers $0 \leq b \leq 2M$, define

$$J_M(b) = \min_{\substack{x \in \mathbb{Z} \\ \max(0, b-M) \leq x \leq \lfloor b/2 \rfloor}} (F(M-b+x) + F(b-2x) + F(x)). \tag{62}$$

The interval in this definition is nonempty.

Proposition F.2. *Let $n \geq 3$, and suppose that a set of landmarks distinguishes the central edges of $S_3(K_n)$. If B is the number of its internal landmarks, then*

$$2(n-2)B \geq J_M(B). \quad (63)$$

In particular, with

$$h(n) = \min\{b \in \{0, \dots, M\} : 2(n-2)b \geq J_M(b)\},$$

one has

$$h(n) \leq B \leq n(n-1) - h(n). \quad (64)$$

Moreover,

$$B \geq c_* n^2 - O(n), \quad c_* = 0.080208390067\dots, \quad (65)$$

where c_* is the unique solution in $(0, 1/10)$ of

$$2c_* = \frac{2}{3} \sqrt{\frac{2}{3}} ((1/2 - c_*)^{3/2} + c_*^{3/2}). \quad (66)$$

Proof. Write each replacement path as $v_i - x_{ij} - x_{ji} - v_j$, and let α_i and η_{ij} indicate the selection of v_i and x_{ij} , respectively. Here $x_{ij} = x_{ij}(1)$ and $\eta_{ij} = \eta_{ij}(1)$. Put

$$d_i^+ = \sum_{j \neq i} \eta_{ij}, \quad d_i^- = \sum_{j \neq i} \eta_{ji}, \quad A_0 = \sum_i \alpha_i, \quad B = \sum_i d_i^+.$$

For a support $\{i, j\}$, define

$$s_{ij} = \eta_{ij} + \eta_{ji}, \quad u_{ij} = d_i^+ + d_j^+ - s_{ij}, \quad v_{ij} = d_i^- + d_j^- - s_{ij}.$$

Both u_{ij} and v_{ij} are nonnegative. Its central-edge histogram, ordered by distances zero to five, is

$$(s_{ij}, \alpha_i + \alpha_j, u_{ij}, v_{ij}, A_0 - \alpha_i - \alpha_j, B - u_{ij} - v_{ij} - s_{ij}).$$

Consequently, within a fixed class $s_{ij} = s$, the triples $(\alpha_i + \alpha_j, u_{ij}, v_{ij})$ are distinct.

Let M_s be the number of supports in this class. Then

$$M_0 + M_1 + M_2 = M, \quad M_1 + 2M_2 = B,$$

and

$$\sum_{i < j} (u_{ij} + v_{ij}) = 2(n-2)B.$$

For the last identity, the sums of the outdegrees and indegrees over all pairs each contribute $(n-1)B$, while the two copies of $\sum_{i < j} s_{ij}$ subtract $2B$. Applying the definition of F in each of the three classes gives

$$2(n-2)B \geq F(M_0) + F(M_1) + F(M_2). \quad (67)$$

Taking $x = M_2$ proves (63).

The minimum defining $h(n)$ exists. At $b = M$, the choice $x = 0$ gives $J_M(M) \leq F(M) \leq (n-2)M$: the levels up to $n-2$ contain $3\binom{n}{2} = 3M$ points, each of cost at most $n-2$. Equation (63) therefore gives $B \geq h(n)$ when $B \leq M$, and the same conclusion is immediate when $B > M$. For the full vertex set, every central edge has histogram

$$(2, 2, 2n-4, 2n-4, n-2, (n-2)(n-3)).$$

Subtracting the chosen histogram from this common vector shows that the complement also distinguishes all central edges. It has $2M - B$ internal landmarks. Applying the lower bound to that complement proves (64). This complementation argument concerns the central edges.

It remains to derive the constant. Write $C = (2/3)\sqrt{2/3}$. If $B \geq n^2/10$, the desired asymptotic bound follows from $c_* < 1/10$. Otherwise put $c = B/n^2$ and $d = M_2/n^2$. The error term in (61) is uniform for arguments between zero and M . Divide (67) by n^3 and use $M/n^2 = 1/2 + O(1/n)$ to obtain

$$2c \geq C \min_{0 \leq d \leq c/2} ((1/2 - c + d)^{3/2} + (c - 2d)^{3/2} + d^{3/2}) - O(1/n).$$

The function being minimized is convex. Its right derivative at zero is

$$\frac{3}{2}(\sqrt{1/2 - c} - 2\sqrt{c}) \geq 0 \quad (0 \leq c \leq 1/10).$$

The minimum is therefore attained at $d = 0$. Define

$$g(c) = 2c - C((1/2 - c)^{3/2} + c^{3/2}).$$

On $[0, 1/10]$, one has

$$g(0) < 0, \quad g(1/10) = \frac{5 - \sqrt{15}}{25} > 0, \quad g'(c) = 2 + \frac{3C}{2}(\sqrt{1/2 - c} - \sqrt{c}) > 2.$$

These facts prove the existence and uniqueness of c_* . The inequality $g(c) \geq -O(1/n)$, together with the derivative bound, gives $c \geq c_* - O(1/n)$, proving (65). $\square$

Every edge multiset resolving set distinguishes the central edges. Proposition F.2 thus improves the asymptotic lower constant at length three. The simpler bound $n^2/16 - n/2$ remains an explicit finite consequence of the terminal-edge count.

F.3 A probabilistic bound for long replacement paths

The forest estimate of Allikvere [2, Lemma 11] converts independent landmark choices into a bound on histogram collisions. We recall the estimate and specialize its distance-level graph to complete subdivisions. This gives an existence range at smaller orders than the four-position construction. For every nonnegative integer j , write

$$\beta(j) = 2^{-j} \binom{j}{\lfloor j/2 \rfloor}.$$

Lemma F.3 (Forest estimate). *Select every vertex of a finite connected graph independently with probability $1/2$, and let W be the selected set. Fix two distinct edges e, f . Suppose that a forest F has vertices among the nonnegative distance levels and that, for each edge $\{\lambda, \mu\} \in E(F)$, a group of $m_{\lambda\mu}$ graph vertices has $\{d(e, x), d(f, x)\} = \{\lambda, \mu\}$. Here $m_{\lambda\mu}$ is a nonnegative integer and $\lambda \neq \mu$. Then*

$$\Pr(H_e = H_f) \leq \prod_{\{\lambda, \mu\} \in E(F)} \beta(m_{\lambda\mu}).$$

Proof. The groups are disjoint because their unordered distance pairs differ. Condition on all landmark choices outside them. Choose an orientation of each forest edge. The contribution of its group to $H_e - H_f$ is its oriented incidence column multiplied by a signed sum of independent Bernoulli variables. Replacing each negatively signed variable X by $1 - X$ expresses this sum as a translate of $\text{Bin}(m_{\lambda\mu}, 1/2)$. The greatest probability of any single value is $\beta(m_{\lambda\mu})$.

The incidence columns of a forest are independent. Equivalently, successive elimination of leaves shows that fixing the net flow at each vertex determines every edge flow. Thus $H_e - H_f = 0$ allows at most one value for each group sum. Conditional independence gives the product bound, which also holds after removing the conditioning. This is the forest argument of [2, Lemma 11]. $\square$

Proposition F.4. *For all integers $n \geq 9$ and $L \geq 12$,*

$$\text{edim}_m(S_L(K_n)) \leq \left\lfloor \frac{n + (L-1)\binom{n}{2}}{2} \right\rfloor.$$

Proof. Put $r = \lfloor L/2 \rfloor$ and $N = n + (L-1)\binom{n}{2}$. Select each of the N vertices independently with probability $1/2$, and call the selected set W .

Distance forests. For an edge e , let q_e be its distance from its nearer original endpoint. Its distances from its two supporting originals are $q_e, L-1-q_e$; every other original is at distance $L+q_e$. Fix $e \neq f$, let I be the union of the originals supporting them, and interchange e, f if necessary so that $q_e \leq q_f$.

First suppose some nearer original u of e satisfies $d(f, u) > q_e$. For every $k \notin I$, the internal vertices on the path uk , measured from u , satisfy

$$\begin{aligned} d(e, x_{uk}(t)) &= q_e + t, \\ d(f, x_{uk}(t)) &= \min\{d(f, u) + t, 2L + q_f - t\} > q_e + t, \quad 1 \leq t < L. \end{aligned}$$

The last inequality follows from $d(f, u) > q_e$, $q_f \geq q_e$, and $t < L$. The corresponding $L-1$ unordered pairs of distance levels form a forest. Their smaller endpoints are distinct, so the smallest vertex of a hypothetical cycle would have to be the smaller endpoint of two different edges. For each forest edge, the group has $n - |I|$ vertices. Lemma F.3 bounds this pair's collision probability by $\beta(n - |I|)^{L-1}$.

If this case fails, every nearer original of e is also a nearer original of f , and $q_e = q_f$. Two distinct edges with the same support cannot have this property. Nor can a central edge of an odd path: its two nearer originals would force the supports, and then the central edges, to coincide. The remaining pairs have supports with a common endpoint 01, 02, and both edges are at the same depth $q \in \{0, \dots, r-1\}$ from the common original zero.

For such a pair and $k \notin \{0, 1, 2\}$, the distances on path $1k$ are

$$\min\{L-1-q+t, 2L+q-t\}, \quad \min\{L+q+t, 2L+q-t\}.$$

They differ exactly when $1 \leq t \leq r+q$. Their smaller levels $L-1-q+t$ are distinct, so these $r+q$ pairs also form a forest. The paths $2k$ give the same pairs with the opposite orientations. Each group therefore contains $2(n-3)$ vertices. Its collision probability is at most $\beta(2(n-3))^{r+q}$. Fig. 17 displays this forest for one length and depth.

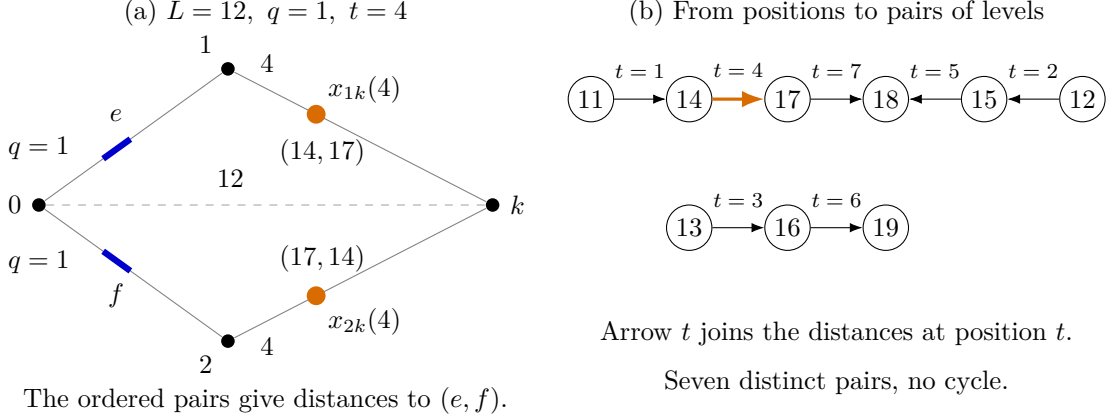


Figure 17: A shared-endpoint comparison for $L = 12$ and $q = 1$. In (a), the edges e and f have the same distance q from the original vertex 0. The vertices at position $t = 4$ on the supports $1k$ and $2k$ have ordered distance pairs $(14, 17)$ and $(17, 14)$. In (b), each position $t = 1, \dots, 7$ gives an arrow from the smaller level ℓ_t to the larger level h_t ; the highlighted arrow corresponds to (a). These seven pairs form a forest. For each pair, varying k outside $\{0, 1, 2\}$ supplies $2(n - 3)$ vertices. The drawn supports in (a) are schematic paths of length 12.

Summing the collision probabilities. We use the elementary bound

$$\beta(j) \leq (3\lceil j/2 \rceil + 1)^{-1/2}. \quad (68)$$

For even indices, induction starts with $\beta(0) = 1$, uses $\beta(2k + 2) = \beta(2k)(2k + 1)/(2k + 2)$, and observes that

$$(2k + 2)^2(3k + 1) - (2k + 1)^2(3k + 4) = k \geq 0.$$

The odd indices follow from $\beta(2k - 1) = \beta(2k)$.

Write

$$\hat{\beta}(j) = (3\lceil j/2 \rceil + 1)^{-1/2}, \quad \theta_n = (3(n - 3) + 1)^{-1/2}.$$

The numbers of unordered pairs of supports that are equal, sharing an endpoint, or disjoint are respectively $\binom{n}{2}$, $n\binom{n-1}{2}$, and $3\binom{n}{4}$. There are $\binom{L}{2}$ pairs of different positions on one support and L^2 position pairs on two supports. Hence the total probability for the pairs in the first case is bounded above by

$$G(n, L) = L^2 \left[\frac{1}{2} \binom{n}{2} \hat{\beta}(n - 2)^{L-1} + n \binom{n-1}{2} \hat{\beta}(n - 3)^{L-1} + 3 \binom{n}{4} \hat{\beta}(n - 4)^{L-1} \right]. \quad (69)$$

The middle term uses all pairs of positions on supports with a common endpoint to overestimate the number in the first case. The exceptional pairs are bounded separately. There is one exceptional position pair at each depth q for each pair of supports with a common endpoint, giving the bound

$$E(n, L) = n \binom{n-1}{2} \frac{\theta_n^r}{1 - \theta_n}. \quad (70)$$

The union bound is therefore $\Pr(\text{some collision}) \leq G(n, L) + E(n, L)$.

Uniform numerical bounds. For $n \geq 9$, the ratio of the expression in (69) at successive lengths is at most

$$\left(\frac{13}{12} \right)^2 \frac{1}{\sqrt{10}} < 1 \quad (L \geq 12).$$

For each parity of n , every summand also decreases with n . Up to a positive constant it has the form

$$\frac{n(n-1)\cdots(n-k+1)}{(n-k+d)^{(L-1)/2}}, \quad k \in \{2, 3, 4\}, \quad d \in \{2/3, 5/3\}.$$

Its logarithmic derivative is at most

$$\frac{k}{n-k+1} - \frac{L-1}{2(n-k+d)} < 0.$$

Indeed, it suffices that $11(n-k+1) - 2k(n-k+d) > 0$. Its minimum for the displayed choices and $n \geq 9$ is $38/3$, at $k=4, d=5/3, n=9$. It follows that

$$\begin{aligned} G(n, L) &\leq \max\{G(9, 12), G(10, 12)\} = G(10, 12) \\ &= \frac{55080}{13^5 \sqrt{13}} + \frac{90720}{10^5 \sqrt{10}} \\ &< \frac{18360}{371293} + \frac{189}{625} < \frac{9}{25}. \end{aligned}$$

For the comparison, $G(9, 12)$ has the same second term and numerator 2592 in its first term.

Since $r \geq 6$, the expression in (70) is at most $n \binom{n-1}{2} \theta_n^6 / (1 - \theta_n)$, which decreases with $n \geq 9$. For example,

$$n \binom{n-1}{2} \theta_n^6 = \frac{1}{54} \prod_{j=0}^2 \frac{n-j}{n-8/3},$$

and all three factors decrease; so does $(1 - \theta_n)^{-1}$. Therefore

$$E(n, L) \leq \frac{252}{19^3(1 - 1/\sqrt{19})} < \frac{336}{6859} < \frac{1}{20}.$$

The probability of any collision is consequently less than $9/25 + 1/20 = 41/100 < 1/2$.

Finally, the symmetry of $\text{Bin}(N, 1/2)$ gives $\Pr(|W| \leq \lfloor N/2 \rfloor) \geq 1/2$. With positive probability the landmark set both satisfies this size bound and has no histogram collision. Such a set proves the proposition. $\square$

F.4 Sparse independent selections

The same distance forests give a finite bound with the selection probability as a parameter. This can be useful when L varies with n , although its fixed-length exponent is weaker than Theorem 1.1.

Proposition F.5. *Let $L \geq 14$ be an integer, and put $r = \lfloor L/2 \rfloor > 6$, $R = L - 1$, and $\alpha_L = 3R/r - 4 > 0$. Define*

$$n_{\text{prob}}(L) = \left\lceil \max \left\{ 8, 8^{r/(r-6)}, L^{2/\alpha_L} \right\} \right\rceil.$$

For every $n \geq n_{\text{prob}}(L)$, writing $N = n + (L-1)\binom{n}{2}$, we have

$$\text{edim}_m(S_L(K_n)) \leq \left\lfloor 8Nn^{-1+6/r} \right\rfloor.$$

Proof. For a signed sum of independent Bernoulli variables with total variance $V > 0$, every atom has probability at most $\sqrt{\pi/(8V)}$. Indeed, replace a negative Bernoulli variable by its complement and absorb the constant. Fourier inversion bounds each atom by

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-2V \sin^2(t/2)} dt \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-2Vt^2/\pi^2} dt = \sqrt{\frac{\pi}{8V}}.$$

Here $\sin(|t|/2) \geq |t|/\pi$ for $|t| \leq \pi$.

Select each vertex independently with probability $0 < p \leq 1/2$. Let W be the selected set and put $w = np$. For $n \geq 8$, each generic forest group contains at least $n/2$ vertices and has variance at least $w/4$. Each exceptional group contains $2(n-3) \geq n$ vertices and has variance at least $w/2$. Their largest atoms are therefore less than $\sqrt{2/w}$ and $1/\sqrt{w}$, respectively. The conditioning and leaf-elimination proof of Lemma F.3 applies with these atom bounds. For $w > 1$, the two classes of pairs have total collision probability less than

$$\frac{L^2 n^4}{8} \left(\frac{2}{w}\right)^{R/2} \quad \text{and} \quad \frac{n^3}{2} \frac{w^{-r/2}}{1 - w^{-1/2}}.$$

The first count bounds all pairs among the $L\binom{n}{2}$ edges. The second sums the exceptional depths as a geometric series.

Choose $p = 4n^{-1+6/r}$. The threshold gives $p \leq 1/2$, $w = 4n^{6/r} \geq 4$, and $n^{\alpha_L} \geq L^2$. Thus the two displayed probabilities are at most $1/8$ and $2^{-r} \leq 1/128$. Also $\alpha_L \geq 2 - 3/r > 0$, since $R \geq 2r - 1$, so the threshold is well defined. By Markov's inequality, $\Pr(|W| > 2Np) \leq 1/2$. The sum of this probability and the collision bound is less than one. Some selection resolves all edges and has size at most $\lfloor 2Np \rfloor$, as required. $\square$

References

- [1] Cesar O. Aguilar, *Anti-regular graphs with loops and their spectrum*, 2018, Preprint, revised 2019, arXiv:1805.08287v2.
- [2] Jaan Allikvere, *The edge multiset dimension of hypercubes*, 2026, Preprint, arXiv:2608.09983v1.
- [3] M. Beck and S. Robins, *Computing the continuous discretely: Integer-point enumeration in polyhedra*, 2nd ed., Undergraduate Texts in Mathematics, Springer, New York, 2015.
- [4] D. Burstein and J. Rubin, *Sufficient conditions for graphicality of bidegree sequences*, SIAM J. Discrete Math. **31** (2017), no. 1, 50–62.
- [5] J. Cilleruelo, *Combinatorial problems in finite fields and Sidon sets*, Combinatorica **32** (2012), no. 5, 497–511.
- [6] Brian Cloteaux, *A sufficient condition for graphic sequences with given largest and smallest entries, length, and sum*, Discrete Mathematics & Theoretical Computer Science **20** (2018), no. 1, 25.
- [7] Pierre Dusart, *Inégalités explicites pour $\psi(X)$, $\theta(X)$, $\pi(X)$ et les nombres premiers*, C. R. Math. Rep. Acad. Sci. Canada **21** (1999), 53–59.
- [8] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory, and on some related problems*, J. London Math. Soc. **16** (1941), no. 4, 212–215.
- [9] M. Farhan, S. Klavžar, D. Kuziak, and I. G. Yero, *Multiset resolvability parameters in graphs: A survey with new results and open problems*, 2026, Preprint, arXiv:2607.10311v1.
- [10] D. R. Fulkerson, *Zero-one matrices with zero trace*, Pacific J. Math. **10** (1960), no. 3, 831–836.
- [11] H. M. Ikhlaiq, R. Ismail, H. M. A. Siddiqui, and M. F. Nadeem, *A new technique to uniquely identify the edges of a graph*, Symmetry **15** (2023), no. 3, 762.
- [12] A. Kelenc, N. Tratnik, and I. G. Yero, *Uniquely identifying the edges of a graph: The edge metric dimension*, Discrete Appl. Math. **251** (2018), 204–220.
- [13] Martin Knor, Riste Škrekovski, and Ismael G. Yero, *A note on the metric and edge metric dimensions of 2-connected graphs*, Discrete Applied Mathematics **319** (2022), 454–460.
- [14] Kostadinka Lapkova, *Correction to: Explicit upper bound for the average number of divisors of irreducible quadratic polynomials*, Monatshefte für Mathematik **186** (2018), 675–678.
- [15] D. Ruiz and C. Trujillo, *Construction of $B_h[g]$ sets in product of groups*, Rev. Colombiana Mat. **50** (2016), no. 2, 165–174.

- [16] I. E. Zverovich and V. E. Zverovich, *Contributions to the theory of graphic sequences*, Discrete Mathematics **105** (1992), no. 1–3, 293–303.