Eigenvalues of subgraphs of the cube

Béla Bollobás [*†‡] Jonathan Lee [§] Shoham Letzter [*]

March 3, 2022

Abstract

We consider the problem of maximising the largest eigenvalue of subgraphs of the hypercube Q_d of a given order. We believe that in most cases, Hamming balls are maximisers, and our results support this belief. We show that the Hamming balls of radius $o(d)$ have largest eigenvalue that is within $1 + o(1)$ of the maximum value. We also prove that Hamming balls with fixed radius maximise the largest eigenvalue exactly, rather than asymptotically, when d is sufficiently large. Our proofs rely on the method of compressions.

1 Introduction

In the last few decades much research has been done on spectra of graphs, i.e. the eigenvalues of the adjacency matrices of graphs; see Finck and Grohmann [10], Hoffman [16, 17], Nosal [25], Cvetković, Doob and Sachs [7], Neumaier [20], Brigham and Dutton [3, 4], Brualdi and Hoffman [5], Stanley [30], Shearer [29], Powers [26], Favaron, Mahéo and Saclé [8, 9], Hong [18], Liu, Shen and Wang [19], Nikiforov [22, 23, 24, 21], and Cvetković, Rowlinson and Simić [6] for a small selection of relevant publications. Perhaps the most basic property of the spectrum of a graph is its radius, i.e. the maximal eigenvalue: this has received especially much attention. Here we shall mention a small handful of these results.

In what follows, $A(G)$ denotes the adjacency matrix of a graph G and $\lambda_1(G)$ denotes the largest eigenvalue of $A(G)$. As usual, we write $e(G)$ for the number of edges, $\Delta(G)$ for the maximal degree and $\overline{d}(G)$ for the average degree. Trivially, $\overline{d}(G) \le \lambda_1(G) \le \Delta(G)$;

[*]Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, CB3 0WB Cambridge, UK; e-mail: {B.Bollobas,S.Letzter}@dpmms.cam.ac.uk .

[†]Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA and London Institute for Mathematical Sciences, 35a South St., Mayfair, London W1K 2XF, UK.

[‡]Research supported in part by NSF grant DMS-1301614 and EU MULTIPLEX grant 317532.

[§]Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, United Kingdom. e-mail: jonathan.lee@maths.ox.ac.uk

in particular, if G is d-regular then $\lambda_1(G) = d$. In 1985, Brualdi and Hoffman [5] gave an upper bound on $\lambda_1(G)$ in terms of $e(G)$: if $e(G) \le \binom{k}{2}$ for some integer $k \ge 1$ then $\lambda_1(G) \le k-1$, with equality iff G consists of a k-clique and isolated vertices. Extending this result, Stanley [30] showed that if $e(G) = m$ then $\lambda_1(G) \le \frac{1}{2}\left(-1+\sqrt{8m+1}\right)$, with equality only as before. In 1993, Favaron, Mahéo and Saclé [9] published an upper bound on $\lambda_1(G)$ in terms of the local structure of G: writing $s(G)$ for the maximum of the sum of degrees of vertices adjacent to some vertex, we have $\lambda_1(G) \le \sqrt{s(G)}$. Furthermore, if G is connected then equality holds iff G is regular or bipartite semi-regular (i.e. vertices in the same class have equal degrees). In particular, if G is a triangle-free graph with m edges then $s(G) \le m$, so $\lambda_1(G) \le \sqrt{m}$. This inequality was first proved by Nosal [25] in 1970. The star $K_{1,m}$ shows that this inequality is best possible.

Our main aim in this paper is to study the maximal eigenvalue of induced subgraphs of the cube Q_d on 2^d vertices, rather than general graphs restricted by their parameters like order and size. To be precise, our aim is to give a partial answer to the following question posed by Fink [11] and in a weaker form by Friedman and Tillich [12].

Question 1. *Given m, $1 \le m \le 2^d$, what is the maximum of the maximal eigenvalue of $Q_d[U]$, where $|U| = m$?*

This problem can be viewed as a variant of the 'classical' isoperimetric problem in the cube. Indeed, since Q_d is d-regular, the problem of bounding the maximal eigenvalue of the subgraph $Q_d[U]$ of Q_d induced by a set $U \subset V(Q_d) = \{0,1\}^d$ is closely related the the size of the edge boundary of U, the set of edges joining a vertex in U to one not in U. If the maximal eigenvalue of $Q_d[U]$ is λ_1, then $e\left(Q_d[U]\right) \le \lambda_1|U|/2$, so the size of the edge boundary of U is at least $(d-\lambda_1)|U|$. Thus, if $\lambda_1 \le \lambda(m)$ whenever $|U| = m$, then for every set of m vertices of the cube Q_d the edge boundary has size at least $(d-\lambda(m))m$.

The study of eigenvalues as a form of isoperimetric inequality is not new: in 1985, Alon and Milman [1] showed that there is a close relation between the second smallest eigenvalue of the Laplacian of a graph and some expansion properties of the graph. The nature of our problem is very different from this. A vaguely related problem has been studied by Reeves, Farr, Blundell, Gallagher and Fink [27].

Before we state our results, we give some precise definitions. Our ground graph is taken to be Q_d, the *d-dimensional hypercube*, where the vertices are labelled by the $0,1$ strings of length d, so that $V(Q_d) = \{0,1\}^d$. Two vertices are connected by an edge if they differ in exactly one coordinate. We shall often use the obvious correspondence between binary strings of length d and subsets of $[d]$ in which a subset corresponds to its characteristic function. A *subcube* of Q_d of dimension i is the graph induced by a subset of the vertices obtained by fixing the values of all but i coordinates. The *Hamming ball* H_d^i is the subgraph of Q_d induced by the vertices with at most i ones in their strings. We note that the subgraphs minimising the sizes of the vertex and edge boundaries among all subgraphs of Q_d with a given order are well known. In particular, Harper (see [13] and [14]) showed in 1966 that the Hamming balls minimise the size of the vertex boundary among subgraphs of the same order. In 1976, Hart [15] proved a similar result, showing that subcubes minimise the size of the edge boundary among subgraphs of the hypercube of the same order.

As the problem of maximising λ_1 is a form of an isoperimetric problem, it seems natural to believe that either Hamming balls or subcubes should be maximisers of λ_1. Despite the connection between λ_1 and the edge boundary, we believe that in many cases, the task of maximising λ_1 is related to minimising the vertex boundary. More precisely, we believe that for most radii sufficiently smaller than $d/2$, Hamming balls maximise λ_1.

We prove several results in this direction. Our first result, which is relatively easy, gives a precise answer when the number of vertices is at most the dimension of the hypercube.

Theorem 2. *Let G be an induced subgraph of Q_d with $n \leq d$ vertices. Then for $n \geq 103$, $\lambda_1(G) \leq \sqrt{n-1}$ with equality if and only if G is a star.*

We note that the conclusion of Theorem 2 does not hold for all n. Indeed, for $n = 4$, the largest eigenvalue of Q_2 (or C_4) is 2, which is larger than $\sqrt{3}$, the largest eigenvalue of the star $K_{1,3}$.

In order to obtain more general results we evaluate the largest eigenvalue of the Hamming ball H_d^i for radii tending to infinity with the dimension of the cube.

Theorem 3. *If $d, i \to \infty$ and $i \leq \frac{d}{2}$ then*

$$\lambda_1(H_d^i) \; = \; 2\sqrt{i(d+1-i)}\left(1 + O\left(i^{-\frac{1}{2}} \log^{\frac{1}{2}} i\right)\right).$$

Our first main result is a generalisation of Theorem 2. We prove that for a wide range of radii, the Hamming balls have largest eigenvalues which are asymptotically largest among all subgraph of the cube of the same order. We note that Samorodnitsky [28] obtained an equivalent result for a wider range of radii (namely for radii i satisfying $i \to \infty$; our proof works also if i is bounded). His proof methods are very different from ours.

Theorem 4. *Let $i = i(d) = o(d)$ and let G be a subgraph of Q_d with $n = O\left(\left|[d]^{(\leq i)}\right|\right)$ vertices. Then $\lambda_1(G) \leq (1 + o(1))\, \lambda_1(H_d^i)$.*

Finally, our second main result gives an exact answer when the radius is fixed.

Theorem 5. *For every i there is $d_0 = d_0(i)$ such that for $d \geq d_0$ the Hamming ball H_d^i maximises the largest eigenvalue among subgraphs of Q_d with the same number of vertices.*

In the next section, Section 2, we state and prove results about compressions which will be used in the proofs of the above theorems. We prove Theorem 2 in Section 3. In Section 4 we prove Theorem 3 as well as other bounds on the largest eigenvalue of certain subgraphs of the cube. We prove our first main result, Theorem 4, in Section 5 and our second main result, Theorem 5, is proved in Section 6. We conclude with some remarks and open questions in Section 7.

2 Compressions

In this sections we prove the results that we shall need about compressions. We start by introducing notation. Let $v \in \mathbb{R}^{V(G)} \subseteq \mathbb{R}^{V(Q_d)}$. Then $\langle A(G)v, v\rangle = \langle A(Q_d)v, v\rangle$, since the support of v is contained in $V(G)$. Hence

$$\max_{|G|=n} \lambda_1(G) = \max_{|G|=n,\ ||v||=1} \langle A(G)v, v\rangle = \max_{||v||=1,\ \mathrm{supp}(v)=n} \langle A(Q_d)v, v\rangle.$$

We consider a notion of *compressions* acting on vectors in $\mathbb{R}^{V(Q_d)}$. Let $U, V \subseteq [d]$ be such that $U \cap V = \emptyset$ and let $v \in \mathbb{R}^{A(Q_d)}$. We define $C_{U,V}(v) \in \mathbb{R}^V(Q_d)$ as follows, where $S \subseteq [d]$.

$$(C_{U,V}(v))_S = \begin{cases} \max(v_S, v_{S\triangle(U\cup V)}) & V \subseteq S \text{ and } U \cap S = \emptyset \\ \min(v_S, v_{S\triangle(U\cup V)}) & U \subseteq S \text{ and } V \cap S = \emptyset \\ v_S & \text{otherwise} \end{cases}$$

Note that $C_{U,V}$ applies a $U - V$ compression to the support of v, leaving the multiset of entries of v unchanged. In particular, it preserves the size of the support of v and its norm.

The *binary order* on Q_d is defined as follows: $S < T$ if and only if $\max S\triangle T \in T$ for $S, T \in V(Q_d)$. We define the binary i-compression $C_i(v)$ to rearrange the values $(v_S)_{i\in S}$ to be decreasing in the binary order restricted to the subcube $\{S : i \in S\}$, and rearrange the values $(v_S)_{i\notin S}$ to be decreasing in the binary order restricted to $\{S : i \notin S\}$. We define C_i^+ and C_i^- to be the restrictions of v to sets containing i or not containing i respectively. Note that C_i^+ and C_i^- commute with C_i.

We may naturally apply these maps to the indicator function of a set F to obtain another indicator function, coinciding with the usual definitions of these maps on sets. We suppress explicit usage of the indicator function where this can be done without confusion.

Given $i \in [d]$, we abuse notation by denoting the singleton $\{i\}$ by i where this is not likely to cause confusion. Furthermore, if $S \subseteq [d]$ we denote $S \cup \{i\}$ by $S + i$ and similarly we denote $S\setminus\{i\}$ by $S - i$. The following two results show that by applying a $C_{i,\emptyset}$ compression or a $C_{i,j}$ compression to a vector v, we do not decrease the inner product $\langle A(Q_d)v, v\rangle$.

Lemma 6. *Let $i \in [d]$ and $v \in \mathbb{R}^{V(Q_d)}$ and denote $A = A(Q_d)$ and $\bar{v} = C_{i,\emptyset}(v)$. Then $\langle Av, v\rangle \leq \langle A\bar{v}, \bar{v}\rangle$.*

Proof. Consider an edge $ST \in E(Q_d)$ with $S \subset T$. If $T \setminus S = \{i\}$, then v_S and v_T are either swapped or not, and in either case the contribution of ST to the inner product is unchanged. All other edges have either $i \in S \cap T$ or $i \neq S \cup T$. These edges come in pairs $(S, S+j)$, $(S+i, S+i+j)$. By the rearrangement inequality and the definition of $C_{i,\emptyset}$, the contribution of this pair of edges to the inner product is larger in $C_{i,\emptyset}(v)$ than in v. $\square$

Lemma 7. *Let $i, j \in [d]$ be distinct, let $v \in \mathbb{R}^{V(Q_d)}$, and denote $A = A(Q_d)$ and $\bar{v} = C_{i,j}(v)$. Then $\langle Av, v\rangle \leq \langle A\bar{v}, \bar{v}\rangle$.*

Proof. Consider an edge $ST \in E(Q_d)$, $S \subseteq T$. The function $C_{i,j}$ is a composition of conditional swaps, and each vertex of Q_d is involved in at most one of these swaps. If neither S nor T are involved in a swap, then the contribution of the edge ST to the inner product is unchanged.

If both S and T are involved in a swap, then if $i \in S$ we have v_S potentially being swapped with v_{S-i+j} and v_T potentially being swapped with v_{T-i+j}; if $i \notin S$ then $j \in S$, so v_S and v_T are potentially swapped with v_{S-j+i} and v_{T-j+i} respectively. Hence edges ST where both vertices are potentially swapped come in pairs (S, T), $(S - i + j, T - i + j)$. By the rearrangement inequality, the contribution of each of these pairs to the inner product is increased by $C_{i,j}$.

If only S is involved in a swap, then exactly one of i and j, whilst both are in T. Hence such edges come in pairs $(T - i, T)$ and $(T - j, T)$, and the contribution of such pairs to the inner product is unchanged by $C_{i,j}$. Similarly, the edges where only T is involved in a swap come in pairs $(S, S + i)$ and $(S, S + j)$, and the contribution of such pairs to the inner product is unchanged by $C_{i,j}$. $\square$

We say that a vector $v \in \mathbb{R}^{V(Q_d)}$ is *compressed* if $C_{U,\emptyset}(v) = v$ for every $U \subseteq [d]$ and $C_{i,j}(v) = v$ for $1 \leq i < j \leq d$. It follows from Lemma 6 and 7 that in order to find the maximum of $\lambda_1(G)$ over subgraphs of the cube of order n, it suffices to consider induced graphs G whose vertex set is compressed. Furthermore, this maximum equals the maximum of $\langle Av, v\rangle$ over compressed vectors v with support of size n.

2.1 Counting copies of subcubes

The aim of this subsection is to provide an upper bound on the number of copies of a subcube in a subgraph G of the cube in terms of $|G|$.

Given a set $U \subseteq V(Q_d)$ and $d' \leq d$ we denote the number of copies of $Q_{d'}$ in $Q_d[U]$ by $\#(Q_{d'} \subseteq U)$. The following result, which was proved by Bollobás and Radcliffe [2], shows the number of copies of $Q_{d'}$ is maximised by initial segments of the binary order. We present a proof here for the sake of completeness.

Lemma 8. *Let $U, I \subseteq V(Q_d)$ with $|U| = |I|$ and I is an initial segment in binary order. Then for any $d' \le d$,*

$$\#(Q_{d'} \subseteq U) \;\le\; \#(Q_{d'} \subseteq I).$$

Proof. We prove the lemma by induction on d'. $d' = 0$ is trivial, as $|U| = |I|$ Suppose that $d' > 0$. We proceed by induction on $d \ge d'$. For $d = d'$ we have that both $\#(Q_{d'} \subseteq U)$ and $\#(Q_{d'} \subseteq I)$ are 0 if $|U| = |I| < 2^d$ and both are 1 otherwise.

Suppose that $d > d'$ and $C_i U = U' \ne U$. For any H a copy of $Q_{d'}$ in $Q_d[U]$, we have that one of the following three options holds: $C_i^+(H) = H$, $C_i^-(H) = H$, or $C_i^-(H) = C_{i,\emptyset}(C_i^+(H))$. Hence by induction the following holds.

$$\begin{aligned}
\#(Q_{d'} \subseteq U) &\le \#(Q_{d'} \subseteq C_i^+ U) + \#(Q_{d'} \subseteq C_i^- U) + \\
&\qquad \min\Big(\#(Q_{d'-1} \subseteq C_i^+ U), \#(Q_{d'-1} \subseteq C_i^- U)\Big) \\
&\le \#(Q_{d'} \subseteq C_i^+ U') + \#(Q_{d'} \subseteq C_i^- U') + \\
&\qquad \min\Big(\#(Q_{d'-1} \subseteq C_i^+ U'), \#(Q_{d'-1} \subseteq C_i^- U')\Big) \\
&= \#(Q_{d'} \subseteq U')
\end{aligned}$$

The first inequality follows from the fact that $C_i^- U'$ and $C_{i,\emptyset}(C_i^+ U')$ are nested.

Define a finite sequence $\{U_i : i = 0, \dots, K\}$ by taking $U_0 = U$ and $U_{k+1} = C_i U_k$ for the least i such that $C_i U_k \ne U_k$ if such an i exists. It is easy to verify that this sequence cannot be infinite. Denote $W = U_K$. Then $\#(Q_{d'} \subseteq U) \le \#(Q_{d'} \subseteq W)$ and $C_i W = W$ for every $i \in [d]$. If $W = I$ the proof is complete, thus we may assume that $W \ne I$.

Since $W \ne I$, W is not initial, so there exists $S < T$ with $S \notin W$ and $T \in W$. Since $C_i^+ W$ and $C_i^- W$ are both initial in the binary order, we have that $i \in S \triangle T$ for every $i \in [d]$. In other words, $S = T^c$, and there is at most one such pair (S, T), so T is the successor of S in binary order and is the maximal element of W. Hence $T = \{d\}$ and $S = [d-1]$. But then T is in at most one $Q_{d'}$ in W, whilst S is in $\binom{d-1}{d'} \ge 1$ copies of $Q_{d'}$ in $W - T + S$. Hence $I = W - T + S$ has at least as many $Q_{d'}$ subgraphs as W, completing the proof. $\square$

The following upper bound on the number of copies of a subcube follows easily.

Lemma 9. *Let U be a subset of $V(Q_d)$ of size n. Then*

$$\#(Q_{d'} \subseteq U) \le \frac{n}{2^{d'}} \binom{\log_2 n + 1}{d'}$$

Proof. By lemma 8, we may that assume U is initial in binary order, so U is contained in a cube of dimension $\lceil \log_2 n \rceil$. Hence each vertex is in at most $\binom{\log_2 n+1}{d}$ copies of $Q_{d'}$ and each copy of $Q_{d'}$ is counted $2^{d'}$ times. $\square$

In fact, one can prove a smooth version of the above upper bound.

Lemma 10. *Let U be a subset of $V(Q_\infty)$ of size n. Then*

$$\#(Q_d \subseteq U) \le \frac{n}{2^d}\binom{\log_2 n}{d}$$

Proof. Let $T_{d,n} = \#(Q_d \subseteq I_n)$, where I_n is initial in binary order in Q_∞ with $|I| = n$. We prove that $T_{d,n} \le \frac{n}{2^d}\binom{\log n}{d}$ by induction on d. It is clear for $d = 0$ so we assume $d > 0$.

We proceed by induction on the number of non zero digits in the binary representation of n. If n is a power of 2, I_n is a cube of dimension $\log n$ and we have $T_{d,n} = \frac{n}{2^d}\binom{\log n}{d}$.

Now suppose that n has $l > 1$ non zero digits in the binary representation. Write $n = 2^{k_1} + \ldots + 2^{k_l}$ where $k_1 > \ldots > k_l$ and let $r = 2^{k_1}$ and $m = n - r$. Then by the definition of binary order and by induction we have

$$T_{n,d} = T_{r,d} + T_{m,d} + T_{m,d-1} \le \frac{r}{2^d}\binom{\log r}{d} + \frac{m}{2^d}\binom{\log m}{d} + \frac{m}{2^{d-1}}\binom{\log m}{d-1}$$

It remains to prove the following inequality.

$$\frac{r}{2^d}\binom{\log r}{d} + \frac{m}{2^d}\binom{\log m}{d} + \frac{m}{2^{d-1}}\binom{\log m}{d-1} \le \frac{n}{2^d}\binom{\log n}{d} \tag{1}$$

If $m < 2^{d-1}$ the second and third summands are zero, and it is easy to check that the required inequality holds. We assume that $m \ge 2^{d-1}$. Writing $r = (1+\alpha)m$ and rearranging Inequality (1), we need to show that the following expression is non-negative.

$$\begin{aligned}&\frac{(2+\alpha)m}{2^d}\binom{\log((2+\alpha)m)}{d} - \frac{(1+\alpha)m}{2^d}\binom{\log((1+\alpha)m)}{d}\\ &- \frac{m}{2^d}\binom{\log m}{d} - \frac{m}{2^{d-1}}\binom{\log m}{d-1}\end{aligned}$$

Writing $\beta = \log m$, we need to show that the following expression is non-negative for $\alpha > 0$ and $\beta \ge d-1$.

$$\begin{aligned}f_\beta(\alpha) &= (2+\alpha)\binom{\log(2+\alpha)+\beta}{d} - (1+\alpha)\binom{\log(1+\alpha)+\beta}{d}\\ &\quad - \binom{\beta}{d} - 2\binom{\beta}{d-1}\end{aligned}$$

Substituting $\alpha = 0$ we obtain

$$f_\beta(0) = 2\binom{1+\beta}{d} - \binom{\beta}{d} - \binom{\beta}{d} - 2\binom{\beta}{d-1} = 0$$

The derivative $f'_\beta(\alpha)$ at $\alpha > 0$ is

$$\frac{1}{d!\ln 2}\sum_{i=0}^{d-1}\left(\prod_{0\le j\le d-1, j\ne i}(\log(2+\alpha)+\beta-j) - \prod_{0\le j\le d-1, j\ne i}(\log(1+\alpha)+\beta-j)\right)$$

$$+2\binom{\log(2+\alpha)+\beta}{d} - \binom{\log(1+\alpha)+\beta}{d}$$

We conclude that $f_\beta(\alpha) \ge 0$ for $\alpha > 0$ and $\beta \ge d-1$, as required. □

3 The star is best for $n = d$

In this section we prove Theorem 2, showing that the star maximises the largest eigenvalue among all subgraphs of the cube Q_d with at most d vertices.

Theorem 2. *Let G be an induced subgraph of Q_d with $n \le d$ vertices. Then for $n \ge 103$, $\lambda_1(G) \le \sqrt{n-1}$ with equality if and only if G is a star.*

Note that this result is not entirely obvious. Indeed, a natural line of attack is to use the inequality $\lambda_1(G) \le \sqrt{s(G)}$ of Favaron, Mahéo and Saclé [9] that we mentioned in the introduction, where $s(G)$ is the maximum of the sum of degrees of vertices adjacent to some vertex. Taking a vertex u, its k neighbours, and $\binom{k}{2}$ additional vertices, each joined to two of the k neighbours of u, we get a subgraph G of Q_n with $n = 1 + k + \binom{k}{2} = (k^2+k+2)/2$ vertices and $e(G) = s(G) = k^2$. Hence, $\lambda_1(G) \le \sqrt{s(G)} = k$, which is about $\sqrt{2}$ times as large as $\sqrt{n-1}$, the bound we wish to prove. The problem is, of course, that the inequality we have applied is far from sharp in this case.

We shall use the following bound, relating the problem of maximising the largest eigenvalue to the task of maximising a trace of a matrix.

Lemma 11. *Let G be a bipartite graph with bipartition $\{X, Y\}$. Then*

$$\begin{aligned}(\lambda_1(G))^{2k} &\le \frac{1}{2}\operatorname{tr}\left(A(G)^{2k}\right)\\ &= \#(\textit{closed walks of length } 2k \textit{ starting from a vertex in } X)\\ &= \#(\textit{closed walks of length } 2k \textit{ starting from a vertex in } Y).\end{aligned}$$

In particular,

$$(\lambda_1(G))^4 \le \#(\textit{edges in } G) + 2\#(\textit{paths of length } 2 \textit{ in } G) + 4\#(C_4 \textit{ in } G).$$

Proof. Immediate from the fact that $\left(A^k\right)_{i,j}$ is the number of walks of length k from vertex i to vertex j. □

We shall also make use of the following bound on the number of edges and 4-cycles in a $K_{2,3}$-free bipartite graph.

Claim 12. *Let G be a bipartite graph with bipartition $\{X, Y\}$ and assume that G is $K_{2,3}$-free. Set $k = |X|$, $l = |Y|$. Then*

- $\#(C_4 \text{ in } G) \leq \binom{l}{2}$.
- $|E(G)| \leq \#(2\text{-paths with both ends in } Y) + k \leq 2\binom{l}{2} + k$.

Proof. The first part follows directly from the fact that G is $K_{2,3}$-free, so every pair of vertices in Y is contained in at most one 4-cycle. The first inequality in the second part follows from the observation that for every vertex $v \in X$, we have $d(v) \leq \#(\text{2-paths in } G \text{ with } v \text{ as the middle vertex}) + 1$. The second inequality again follows from the assumption that G is $K_{2,3}$-free. □

We now proceed to the proof of Theorem 2.

Proof of theorem 2. Let G be a subgraph of Q_d with $n \leq d$ vertices and assume $\lambda_1(G) \geq \sqrt{n-1}$. Denote by $\{X, Y\}$ the bipartition of the vertices of G where $k = |X| \geq |Y| = l$. For a vertex $v \in V(G)$ denote by $d(v)$ the degree of v.

By Lemma 11 and the fact that G does not have $K_{2,3}$-free, we obtain the following.

$$\begin{aligned}(n-1)^2 &\leq (\lambda_1(G))^4 \\ &\leq 2\sum_{v \in V} \binom{d(v)}{2} + |E(G)| + 4\#(C_4 \text{ in } G) \\ &\leq 2\left(\binom{k}{2} + \binom{l}{2} + 2\#(C_4 \text{ in } G) \right) + |E(G)| + 4\#(C_4 \text{ in } G) \\ &= 2l^2 - 2nl + n^2 - n + |E(G)| + 8\#(C_4 \text{ in } G).\end{aligned}$$

Hence,

$$0 \;\leq\; 2l^2 - 2nl + n - 1 + |E(G)| + 8\#(C_4 \text{ in } G). \tag{2}$$

We replace $|E(G)|$ and $\#(C_4 \text{ in } G)$ by the upper bounds from Lemma 10 to obtain the following inequality.

$$\begin{aligned}0 \;&\leq\; 2l^2 - 2nl + n - 1 + \frac{1}{2}n\log n + 2n\binom{\log n}{2} \\ &= 2l^2 - 2nl + n(\log^2 n - \frac{1}{2}\log n + 1) - 1.\end{aligned}$$

Since $l \leq n/2$, we deduce the following upper bound on l.

$$\begin{aligned} l \ &\leq \frac{1}{4}\left(2n - \sqrt{4n^2 - 8n(\log^2 n - \frac{1}{2}\log n + 1) + 8}\right) \\ &= \frac{1}{2}\left(n - \sqrt{n^2 - 2n\log^2 n + n\log n - 2n + 2}\right). \end{aligned} \tag{3}$$

By Claim 12 and Inequality (2),

$$\begin{aligned} 0 \ &\leq 2l^2 - 2nl + n - 1 + 10\binom{l}{2} + n - l \\ &= (l-1)(7l + 1 - 2n). \end{aligned}$$

If $l \geq 2$ it follows that $l \geq \frac{1}{7}(2n-1)$. Combining this lower bound on l with the upper bound (3), we get the following inequality.

$$\frac{1}{7}(2n-1) \ \leq \ l \ \leq \ \frac{1}{2}\left(n - \sqrt{n^2 - 2n\log^2 n + n\log n - 2n + 2}\right). \tag{4}$$

This is a contradiction if $n \geq 103$. Thus if $n \geq 103$ we must have $l = 1$, implying that G is a star. $\square$

4 The largest eigenvalue of the Hamming ball

In this section we estimate the largest eigenvalue of the Hamming ball H_d^i for several ranges of i and d. We start by proving Theorem 3, where we estimate the eigenvalue of the Hamming ball when the radius goes to infinity.

Theorem 3. *If $d, i \to \infty$ and $i \leq \frac{d}{2}$ then*

$$\lambda_1(H_d^i) \ = \ 2\sqrt{i(d+1-i)}\left(1 + O\left(i^{-\frac{1}{2}}\log^{\frac{1}{2}} i\right)\right).$$

The upper bound from Theorem 3 follows trivially from the following claim. We shall use this claim in subsequent sections, therefore we state it here.

Claim 13. *Let G be a subgraph of Q_∞. Assume that G has maximum degree at most d and that $V(G) \subseteq [d]^{\leq t}$, where $t \leq d/2$. Then $\lambda_1(G) \leq 2\sqrt{td}$.*

Proof. For $j \geq 0$ let $V_j = \{S \in V(G) : |S| = j\}$. Let $G_j = G[V_j \cup V_{j+1}]$. The graph G_j is a bipartite graph whose vertices from one side have degree at most $j+1$ and the vertices from the other side have degree at most d. Thus $\lambda_1(G_j) \leq \sqrt{(j+1)d}$.

Let $v = (v_S)_{S \in V(G)}$ be an eigenvector with norm 1 and eigenvalue $\lambda_1(G)$. Define $\alpha_j^2 = \sum_{S \in V_j} v_s^2$. Note that $E(G) = \bigcup_{0 \le j < t-1} E(G_j)$, hence

$$\begin{aligned}
\lambda_1(G) = \langle A(G)v, v\rangle \; &\le \sum_{j=0}^{t-1} \langle A(G_j)v, v\rangle \\
&\le \sum_{j=0}^{t-1} \left(\alpha_j^2 + \alpha_{j+1}^2\right) \lambda_1(G_j) \\
&\le \sum_{j=0}^{t-1} \left(\alpha_j^2 + \alpha_{j+1}^2\right) \sqrt{(j+1)d} \\
&\le 2\sqrt{td}.
\end{aligned}$$

It follows that $\lambda_1(G) \le 2\sqrt{td}$, completing the proof of Claim 13. $\square$

We now turn to the proof of Theorem 3

Proof of Theorem 3. Denote $\lambda = \lambda_1(H_d^i)$ and $A = A(H_d^i)$. We first note that since every $S \in H_d^i$ satisfies $|S| \le i$ and the maximum degree in H_d^i is $d-i+1$, Claim 13 implies that $\lambda \le 2\sqrt{i(d+1-i)}$.

We now obtain a lower bound on λ. Define the vector $v \in \mathbb{R}^{V(H_d^i)}$ by $v_S = \mathbb{1}_{[d]^i}(S)\binom{d}{i}^{-\frac{1}{2}}$. Note that $||v|| = 1$. For every $k < i$ we have

$$\lambda^{2k} \ge \langle A^{2k}v, v\rangle \ge \binom{d}{i} \frac{1}{k+1} \binom{2k}{k} \left((i-k)(d+1-(i-k))\right)^k \binom{d}{i}^{-1},$$

as $r(d+1-r)$ is the number of choices for which edges to use to move from a set of r to $r-1$ and $r-1$ to r respectively, and $r(d+1-r)$ is an increasing function for $r \le i \le \frac{d}{2}$. Hence

$$\lambda^{2k} \ge k^{-\frac{3}{2}} 2^{2k} ((i-k)(d+1-(i-k)))^k (1+o(1)).$$

Thus

$$\begin{aligned}
\lambda \; &\ge k^{-\frac{3}{2k}} 2\sqrt{(i-k)(d+1-(i-k))} \left(1 + o\left(k^{-1}\right)\right) \\
&= 2\sqrt{(i-k)(d+1-(i-k))} \left(1 + O\left(k^{-1}\log k\right)\right).
\end{aligned}$$

Taking $k = \sqrt{i \log i} \to \infty$, we get $\lambda \ge 2\sqrt{i(d+1-i)}\left(1 + O\left(i^{-\frac{1}{2}} \log^{\frac{1}{2}} i\right)\right)$, completing the proof of Theorem 3. $\square$

We now consider the case where the radius of the Hamming ball is fixed.

Lemma 14. *There exist constants $\lambda_1 < \lambda_2 < \ldots$ such that $\lambda_1(H_d^i) = \lambda_i \sqrt{d}(1 + O(1/d))$.*

Proof. Let A_i be the $(i+1) \times (i+1)$-matrix defined by

$$A_{j,k} = \begin{cases} 1 & j = k+1 \\ j & j = k-1 \\ 0 & \text{otherwise} \end{cases}$$

Denote $\lambda_i = \lambda_1(A_i)$. Since A_i is a submatrix of A_{i+1}, and by monotonicity of the largest eigenvalue for matrices with non negative entries, we have $\lambda_i < \lambda_{i+1}$ for every i. In order to complete the proof, it suffices to show that $\lambda(H_d^i) = \lambda_i \sqrt{d}(1 + O(1/d))$.

By symmetry, the eigenvector of $A(H_d^i)$ with eigenvalue $\lambda_1(H_d^i)$ is uniform on $[d]^{(j)}$ for every $0 \le j \le i$. Denote by x_j the weight of $[j]$ in the eigenvector. The following holds.

$$\lambda_1(H_d^i)\, x_j = \begin{cases} dx_1 & j = 0 \\ jx_{j-1} + (d-j)x_{j+1} & 0 < j < i \\ ix_{i-1} & j = i \end{cases}$$

Letting $\mu = \lambda_1(H_d^i)/\sqrt{d}$ and $y_j = x_j d^{j/2}$, we obtain

$$\mu y_j = \begin{cases} y_1 & i = 0 \\ jy_{j-1} + (1 + O(1/d))y_{j+1} & 0 < j < i \\ iy_{i-1} & j = i \end{cases}$$

Recalling the definition of A_i, this implies that $\mu y_j = (A_i y)_j + O(1/d)$, where $y = (y_0, \ldots, y_i)^T$. It follows (e.g. by looking at the characteristic polynomials) that $|\mu - \lambda_i| = O(1/d)$. Lemma 14 follows. $\square$

5 Hamming ball is asymptotically best for $i = o(d)$

In this section we prove Theorem 4 showing that for $i = o(d)$ the Hamming ball H_d^i asymptotically maximises the λ_1 among subgraphs of Q_d with the same number of vertices. Since our proof is rather technical, we start with the special case $i = 1$.

5.1 Proof of Theorem 4 for $i = 1$

Let us first state the result for the special case $i = 1$.

Lemma 15. *Let $c > 0$ be fixed and let G be a subgraph of Q_d with $n \le cd$ vertices. Then $\lambda_1(G) \le \sqrt{d} + O\left(d^{1/4}(\log d)^{1/2}\right)$.*

Using our results about compressions, we may assume that $V(G)$ is compressed. This enables us to partition $V(G)$ into stars, in such a way that the edges not covered by the stars have a small contribution to the eigenvalue, thus enabling us to obtain the required estimate of $\lambda_1(G)$.

Proof. By Lemmas 6 and 7 we can assume that $V(G)$ is compressed. Namely for every $1 \leq j < k \leq d$ we have $C_{k,\emptyset}(V(G)) = V(G)$ and $C_{k,j}(V(G)) = V(G)$.

We aim to partition $V(G)$ in such a way that each part induces a star and that the graph spanned by the edges not contained in any of these parts is of small maximal degree. This would imply that $\lambda_1(G)$ is at most the eigenvalue of the star with $d+1$ vertices plus an error term which can be controlled by the maximal degree of the 'leftover' edges.

Let $\epsilon = \epsilon(d) = \sqrt{2c/d}$. Let $\mathcal{A}$ be the set of vertices of degree at least ϵd in G. We call these vertices 'heavy'. To minimise the maximal degree of the leftover graph, we wish to have each heavy vertex as a centre of one of the stars in the partition. It may happen e.g. that $\{1\}, \{2\}$ are heavy and $\{1, 2\}$ is not, in which case $\{1, 2\}$ will have to appear in two stars of the partition. To avoid this from happening, we add vertices to the set of heavy vertices as follows.

Let $\mathcal{B} = \{t \in [d] : \{t\} \in \mathcal{A}\}$. Note that since $V(G)$ is compressed, $\mathcal{A}$ is compressed as well, and thus $\mathcal{B}$ is an interval and $m = \max \mathcal{B} = |\mathcal{B}|$. Finally define $\mathcal{D} = \mathcal{P}([m]) \cap V(G)$. Since $\mathcal{A}$ is down-compressed, $\mathcal{A} \subseteq \mathcal{D}$. We note that the maximum degree of $\mathcal{A}$ is at most ϵd. Indeed, suppose that $v \in \mathcal{A}$ has at least ϵd neighbours in $\mathcal{A}$. Denote this set of neighbours by A. The every vertex in A has at least ϵd neighbours in $V(G)$. Note that by the structure of Q_d, no vertex is a neighbour of more than two vertices of A. It follows that $|V(G)| > \frac{|A|\epsilon d}{2} \geq \frac{(\epsilon d)^2}{2} \geq n$, a contradiction. In particular, $m = \deg_{\mathcal{A}}(\emptyset) \leq \epsilon d$.

For $S \in \mathcal{D}$ define $N^*(S) = \{S\} \cup \big(N(S) \setminus \mathcal{D}\big)$, where $N(S)$ denotes the neighbourhood of S in G. We claim that $N^*(S)_{S \in \mathcal{D}}$ is a collection of disjoint sets. Indeed, by the choice of $\mathcal{D}$, the $\mathcal{D} \subseteq \mathcal{P}([m])$ and any vertex in the neighbourhood of D which is not in $\mathcal{D}$ must be of the form $S \cup \{s\}$ where $s \notin [m]$ and $S \in \mathcal{D}$. Furthermore, clearly, each of these sets induces a star.

Let $v = (v_S)_{S \in V(G)}$ be a vector of norm 1 with positive entries such that $A(G)v = \lambda_1(G)v$. Note that the edges of G are covered by the edges of the graphs $G[\mathcal{D}]$, $G \setminus \mathcal{D}$ and $\{N^*(S)\}_{S \in \mathcal{D}}$. We thus obtain the following upper bound on $\lambda_1(G)$.

$$\begin{aligned}
\lambda_1(G) &= \langle A(G)v, v\rangle \\
&\leq \Big(\sum_{S \in \mathcal{D}} \langle A(G[N^*(S)])v, v\rangle\Big) + \langle A(G[\mathcal{D}])v, v\rangle + \langle A(G \setminus \mathcal{D})v, v\rangle \\
&\leq \Big(\sum_{S \in \mathcal{D}} \lambda_1(G[N^*(S)]) \sum_{T \in N^*(S)} v_T^2\Big) + \Big(\lambda_1(G[\mathcal{D}]) \sum_{S \in \mathcal{D}} v_S^2\Big) + \Big(\lambda_1(G \setminus \mathcal{D}) \sum_{S \notin \mathcal{D}} v_S^2\Big).
\end{aligned}$$

It remains to obtain upper bounds on the largest eigenvalue of the graphs $G[\mathcal{D}]$, $G \setminus \mathcal{D}$

and $\{N^*(S)\}_{S\in\mathcal{D}}$. Recall that by Claim 13, given a subgraph of Q_d, we have $\lambda_1(G) \leq 2\sqrt{\Delta(G)t}$, where $\Delta(G)$ is the maximum degree of G and t is the size of the largest set in $V(G)$. Since $\mathcal{D} \subseteq \mathcal{P}([m])$, the maximum degree of $G[\mathcal{D}]$ is bounded by ϵd. Also, by definition of $\mathcal{A}$, the maximum degree of $G \setminus \mathcal{D}$ is at most ϵd. Since $V(G)$ is compressed, the largest set in $V(G)$ is of size at most $\log n$. It follows from Claim 13 that

$$\lambda_1(G[\mathcal{D}]),\ \lambda_1(G \setminus \mathcal{D}) \leq 2\sqrt{\epsilon d \log n}.$$

Furthermore, $\lambda_1(N^*(S)) \leq \sqrt{d}$ since each set $N^*(S)$ is a star with at most $d+1$ vertices. Thus, by the above inequality and using the disjointness of the sets $N^*(S)$, we obtain

$$\lambda_1(G) \leq \sqrt{d} + O(\sqrt{\epsilon d \log d}) = \sqrt{d} + O(d^{1/4}(\log d)^{1/2}),$$

completing the proof of Lemma 15. $\square$

5.2 Proof of Theorem 4

We now prove Theorem 4 in general.

Theorem 4. *Let $i = i(d) = o(d)$ and let G be a subgraph of Q_d with $n = O\left(\left|[d]^{(\leq i)}\right|\right)$ vertices. Then $\lambda_1(G) \leq (1+o(1))\ \lambda_1(H_d^i)$.*

The idea is similar to the special case of $i = 1$. Again we find a partition of G with sets whose largest eigenvalue can be bounded by the largest eigenvalue of a suitable Hamming ball. Furthermore, we ensure that the leftover edges have a small contribution to the largest eigenvalue of G.

Proof of Theorem 4. By Lemmas 6 and 7 we can assume that G is compressed. Similarly to the proof for $i = 1$, we partition the vertices into sets that induce subsets of the Hamming ball of radius approximately i. We choose the partition in such a way that the edges not covered by one of these subsets span a graph with small maximal degree. In this way we can bound the eigenvalue of the both subgraphs of G to obtain the required bound. In order to define the partition we need some notation.

Let $\epsilon = \epsilon(d) < 1$ and define the following sets recursively.

$\mathcal{A}_0 = V(G)$.

$\mathcal{A}_k = \{S \in \mathcal{A}_{k-1} : S \text{ has at least } \epsilon d \text{ neighbours in } \mathcal{A}_{k-1}\}$.

Let $M = \max\{k : \mathcal{A}_k \neq \emptyset\}$. Note that since G is compressed, the sets $(\mathcal{A}_k)_{0\leq k\leq M}$ are compressed. The sets $\mathcal{A}_k$ measure how 'heavy' a vertex is: for a vertex $v \in V(G)$, the larger $\max\{k : v \in \mathcal{A}_k\}$ is, the heavier v is.

As in the proof of the special case, we want to take the heaviest vertices to be the centres of the Hamming balls defining the partition. Since we now have many levels, we first take Hamming balls centred at the heaviest vertices, then take as centres the heaviest vertices among those that weren't covered in the first round, and so on. This process is somewhat complicated by the fact that we want each vertex to appear in at most one such Hamming ball. To ensure this, we add some of the vertices to sets of heavy vertices using the following definitions.

We define sets $\mathcal{B}_k$, $\mathcal{C}_k$, $\mathcal{D}_k$, $\mathcal{E}_k$ and numbers m_k for $0 \leq k \leq M$ as follows. For $k = 0$,

$$
\begin{aligned}
&\mathcal{B}_0 = \{t \in [d] : \{t\} \in \mathcal{A}_M\} \cup \{1\}\\
&m_0 = \max \mathcal{B}_0\\
&\mathcal{C}_0 = \emptyset\\
&\mathcal{E}_0 = \mathcal{D}_0 = \mathcal{P}([m_0]) \cap V(G)
\end{aligned}
$$

For $0 < k \leq M$ define recursively

$$
\begin{aligned}
&\mathcal{B}_k = \{t > m_{k-1} + 1 : \{m_0 + 1, \ldots, m_{k-1} + 1, t\} \in \mathcal{A}_{M-k}\} \cup \{m_{k-1} + 1\}\\
&m_k = \max \mathcal{B}_k\\
&\mathcal{C}_k = \{S \cup \{t\} : S \in \mathcal{C}_{k-1} \cup \mathcal{D}_{k-1}, t > m_{k-1}\}\\
&\mathcal{D}_k = (\mathcal{P}([m_k]) \cap V(G)) \setminus (\mathcal{E}_{k-1} \cup \mathcal{C}_k)\\
&\mathcal{E}_k = \mathcal{C}_0 \cup \ldots \cup \mathcal{C}_k \cup \mathcal{D}_0 \cup \ldots \cup \mathcal{D}_k
\end{aligned}
$$

Before we proceed with the proof, we try to convey the ideas behind the above definitions. The sets $\mathcal{D}_k$ defined above will be the centres of the Hamming balls and the $\mathcal{C}_k$'s will consist of the other vertices covered by these balls. In each stage we define $\mathcal{C}_k$ to be the set of neighbours of vertices which appeared previously. We define $\mathcal{D}_k$ so as to be the up-closure (relatively to $V(G)$) of the vertices in $\mathcal{A}_{M-k}$ which were not covered previously. To this end, in each stage B_k and m_k are defined so that every $t \in S \in \mathcal{A}_{M-k} \setminus \big(\mathcal{E}_{k-1} \cup \mathcal{C}_k\big)$ satisfies $t \leq m_k$. Thus $\mathcal{D}_k$ contains $\mathcal{A}_{M-k} \setminus \big(\mathcal{E}_{k-1} \cup \mathcal{C}_k\big)$ and is up-closed in $V(G)$.

We now define the partition of $V(G)$ into sets inducing subgraphs of Hamming balls with centres in $\bigcup_{0 \leq k < M} \mathcal{D}_k$. For a vertex $S \in V(G)$ and $t \geq 1$, let $N_t(S)$ denote the set of vertices of $V(G)$ in distance t from S. For every $0 \leq k < M$ and every $S \in \mathcal{D}_k$, let

$$N_S^{(k)} = \{S\} \cup \bigcup_{1 \leq j \leq M-k} \Big(N_j(S) \cap \mathcal{C}_{k+j}\Big).$$

In order to show that the sets $N_k(S)$ satisfy our requirement we need the following proposition. Its proof is delayed to the end of this section.

Proposition 16. *The following assertions hold.*

1. *The sets $N^{(k)}(S)$, where $0 \le k \le M-1$ and $S \in \mathcal{D}_k$, are pairwise disjoint.*
2. *The sets $\mathcal{C}_k \cup \mathcal{D}_k$, where $0 \le k \le M$, form a partition of $V(G)$.*
3. $E(G) = \left(\bigcup_{0\le k\le M-1,\, S\in\mathcal{D}_k} E(G[N^{(k)}(S)])\right) \cup \left(\bigcup_{0\le k\le M} E(G[\mathcal{C}_k \cup \mathcal{D}_k])\right)$.
4. *The maximum degree of $G[\mathcal{C}_k \cup \mathcal{D}_k]$, where $0 \le k \le M$, is at most ϵd.*

Let $v = (v_S)_{S\in V(G)}$ be a vector of positive weights on the vertices of G with norm 1, satisfying $A(G)v = \lambda_1 v$. Define

$$\alpha_k^2 = \sum_{S\in\mathcal{C}_k\cup D_k} v_S^2 \quad \text{for } 0 \le k \le M.$$

$$(\beta_{k,S})^2 = \sum_{T\in N_S^{(k)}} v_T^2 \quad \text{for } 0 \le k < M \text{ and } S \in \mathcal{D}_k.$$

By Parts (1) and (2) above, $\sum\limits_{0\le k<M} \sum\limits_{S\in\mathcal{D}_k} (\beta_{k,S})^2 \le 1$ and $\sum\limits_{k=0}^{M} \alpha_k^2 = 1$. Thus, by Part (3),

$$\begin{aligned}
\lambda_1(G) &= \langle A(G)v, v\rangle \\
&\le \sum_{k=0}^{M} \Big\langle A(G[\mathcal{C}_k \cup \mathcal{D}_k])v,\ v\Big\rangle + \sum_{k=0}^{M}\sum_{S\in\mathcal{D}_k} \Big\langle A(G[N_S^{(k)}])v, v\Big\rangle \\
&\le \sum_{k} \alpha_k^2 \cdot \lambda_1(G[\mathcal{C}_k \cup \mathcal{D}_k]) + \sum_{k,S} (\beta_{k,S})^2 \cdot \lambda_1(G[N_S^{(k)}]) \\
&\le \max_k \lambda_1(G[\mathcal{C}_k \cup \mathcal{D}_k]) + \max_{k,S} \lambda_1(G[N_S^{(k)}])
\end{aligned} \tag{5}$$

By Part (4) of Proposition 16, the maximum degree of $G[\mathcal{C}_k \cup \mathcal{D}_k]$ is at most ϵd. Since $V(G)$ is compressed, the largest set in $V(G)$ has size at most $\log n$. Recall that $n = \Theta\left(\binom{d}{i}\right)$, thus $\log n = (1+o(1))i\log(d/i)$. We conclude the following upper bound by Claim 13.

$$\lambda_1(G[\mathcal{C}_k \cup \mathcal{D}_k]) \;\le\; 2\sqrt{\epsilon d \log n} \;=\; 2(1+o(1))\sqrt{\epsilon d i \log(d/i)}. \tag{6}$$

Let us treat first the case where $i \to \infty$. By Theorem 3, using the monotonicity of the largest eigenvalue of a graph,

$$\lambda_1(G[N_S^{(k)}]) \le \lambda_1(H_d^{M-k}) \le \lambda_1(H_d^M) = 2(1+o(1))\sqrt{M(d-M)}. \tag{7}$$

Substituting Inequalities (6) and (7) into the Inequality (5), it follows that

$$\lambda_1(G) \le 2(1+o(1))\Big(\sqrt{\epsilon d i \log(d/i)} + \sqrt{M(d-M)}\Big). \tag{8}$$

The following claim will imply that we can choose ϵ so as to make the above upper bound arbitrarily close to $\lambda_1(H_d^i)$.

Claim 17. *Let $\alpha > 0$ and set $\epsilon = \frac{\alpha}{\log(d/i)}$. Then $M \leq (1+o(1))i$.*

Proof. For arbitrary $\beta > 0$ we show that $M \leq (1+\beta)i$ for large enough d. Let $N = (1+\beta)i$ and $D = \epsilon d$. We need to show that $\mathcal{A}_N = \emptyset$. Assuming the contrary, let $S \in \mathcal{A}_N$. Then S has at least D neighbours in $\mathcal{A}_{N-1}$, which in turn have at least $\binom{D}{2}$ new neighbours in $\mathcal{A}_{N-2}$ and so on. It follows that $n = |V(G)| \geq 1 + D + \binom{D}{2} + \ldots + \binom{D}{N} = \left|[D]^{(\leq N)}\right|$. Recall that $i = o(d)$ and note that $\frac{N}{D} = \frac{1+\beta}{\alpha} \cdot \frac{\log(d/i)}{d/i} = o(1)$, i.e. $N = o(D)$. Thus,

$$\left|[D]^{(\leq N)}\right| = (1+o(1))\frac{1}{\sqrt{2\pi N}}\left(\frac{eD}{N}\right)^N.$$

On the other hand,

$$n \leq c\left|[d]^{(\leq i)}\right| = (1+o(1))\frac{c}{\sqrt{2\pi i}}\left(\frac{ed}{i}\right)^i.$$

Combining the two inequalities, we obtain the following.

$$\begin{aligned}\frac{c}{\sqrt{i}}\left(\frac{ed}{i}\right)^i &\geq (1+o(1))\frac{1}{\sqrt{N}}\left(\frac{eD}{N}\right)^N \\ &= (1+o(1))\frac{1}{\sqrt{(1+\beta)i}}\left(\frac{e\alpha}{1+\beta} \cdot \frac{d/i}{\log(d/i)}\right)^{(1+\beta)i}.\end{aligned}$$

We obtain the following inequality, where c_1, c_2 are constants depending on α, β, c.

$$c_2 \geq \left(c_1\frac{(d/i)^{\frac{\beta}{1+\beta}}}{\log(d/i)}\right)^{(1+\beta)i}$$

Since $i = o(d)$, we have $\log(d/i) = o\left((d/i)^\gamma\right)$ for every fixed $\gamma > 0$ and we have reached a contradiction. This implies that $M \leq (1+\beta)i$ for large d. $\square$

By Inequality (8) with $\epsilon = \frac{\alpha}{\log(d/i)}$ we have

$$\begin{aligned}\lambda_1(G) &\leq 2(1+o(1))(\sqrt{\alpha i d} + \sqrt{i(d-i)}) \\ &= 2(1+\sqrt{\alpha})(1+o(1))\sqrt{i(d-i)}.\end{aligned}$$

Since α can be taken arbitrarily close to 0, it follows that

$$\lambda_1(G) \leq 2(1+o(1))\sqrt{i(d-i)} = (1+o(1))\lambda_1(H_d^i).$$

This completes the proof of Theorem 4 in case $i = \omega(1)$.

It remains to consider the case where i is constant. Take $\epsilon = 2id^{-1/(i+1)}$. It is easy to check that $\binom{\epsilon d}{i+1} > n$, implying that $M \leq i$ similarly to the proof of Claim 17. It follows from Inequalities (5), (6) and Claim 13 that

$$\lambda_1(G) \leq O\left(\sqrt{d^{1-\frac{1}{i+1}} \log d}\right) + \lambda_1(H_d^i) = (1+o(1))\lambda_1(H_d^i),$$

completing the proof of Theorem 4. □

5.3 Proof of Proposition 16

In order to complete the proof of Theorem 4, it remains to prove Proposition 16.

Proposition 16. *The following assertions hold.*

1. *The sets $N^{(k)}(S)$, where $0 \leq k \leq M-1$ and $S \in \mathcal{D}_k$, are pairwise disjoint.*
2. *The sets $\mathcal{C}_k \cup \mathcal{D}_k$, where $0 \leq k \leq M$, form a partition of $V(G)$.*
3. $E(G) = \left(\bigcup_{0 \leq k \leq M-1,\, S \in \mathcal{D}_k} E(G[N^{(k)}(S)])\right) \cup \left(\bigcup_{0 \leq k \leq M} E(G[\mathcal{C}_k \cup \mathcal{D}_k])\right)$.
4. *The maximum degree of $G[\mathcal{C}_k \cup \mathcal{D}_k]$, where $0 \leq k \leq M$, is at most ϵd.*

Proof. Recall the definition of the sets $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k, \mathcal{E}_k$.

$\mathcal{A}_0 = V(G)$.
$\mathcal{A}_k = \{S \in \mathcal{A}_{k-1} : S$ has at least ϵd neighbours in $\mathcal{A}_{k-1}\}$.

For $k = 0$,

$\mathcal{B}_0 = \{t \in [d] : \{t\} \in \mathcal{A}_M\} \cup \{1\}$
$m_0 = \max \mathcal{B}_0$
$\mathcal{C}_0 = \emptyset$
$\mathcal{E}_0 = \mathcal{D}_0 = \mathcal{P}([m_0]) \cap V(G)$

For $0 < k \leq M$,

$\mathcal{B}_k = \{t > m_{k-1} + 1 : \{m_0 + 1, \ldots, m_{k-1} + 1, t\} \in \mathcal{A}_{M-k}\} \cup \{m_{k-1} + 1\}$
$m_k = \max \mathcal{B}_k$
$\mathcal{C}_k = \{S \cup \{t\} : S \in \mathcal{C}_{k-1} \cup \mathcal{D}_{k-1}, t > m_{k-1}\}$
$\mathcal{D}_k = (\mathcal{P}([m_k]) \cap V(G)) \setminus (\mathcal{E}_{k-1} \cup \mathcal{C}_k)$
$\mathcal{E}_k = \mathcal{C}_0 \cup \ldots \cup \mathcal{C}_k \cup \mathcal{D}_0 \cup \ldots \cup \mathcal{D}_k$

We prove the following assertions.

1. For every $S \in \mathcal{C}_k \cup \mathcal{D}_k$ there is a unique $j \le k$ and a unique $T \in D_j$, such that there exist distinct $t_{j+1}, \dots, t_k$ satisfying $t_l > m_{l-1}$ and $S = T \cup \{t_{j+1}, \dots, t_k\}$.
2. $\mathcal{E}_k$ is compressed.
3. There are no edges of G between $\mathcal{E}_k$ and $\mathcal{D}_{k+1} \cup \mathcal{C}_{k+2} \cup \mathcal{D}_{k+2}$.
4. $(\mathcal{C}_k \cup \mathcal{D}_k) \cap \mathcal{E}_{k-1} = \emptyset$.
5. The maximum degree of $G[\mathcal{C}_k \cup \mathcal{D}_k]$ is at most $m_k \le \epsilon d$.
6. $\mathcal{A}_{M-k} \subseteq \mathcal{E}_k$. In particular, $\mathcal{E}_M = V(G)$.
7. The sets $(N^{(k)}(S))_{0 \le k \le M-1, S \in \mathcal{D}_k}$ are pairwise disjoint.

Note that Proposition 16 follows from these assertions. Indeed, Part (2) follows from Assertions (4) and (6), Part (3) follows from Assertions (3) and (6) and Parts (1) and (4) are among these assertions.

Proof of Assertion (1). We prove Assertion (1) by induction on k. It is trivial for $k = 0$, so we assume $k > 0$. Let $S \in \mathcal{C}_k \cup \mathcal{D}_k$. Assume that $S = T \cup \{t_{j+1}, \dots, t_k\} = R \cup \{r_{l+1}, \dots, r_k\}$, where $T \in \mathcal{D}_j$, $R \in \mathcal{D}_l$ and $t_u, r_u > m_{u-1}$ for all u. We show that we must have $j = l$ and $T = R$.

Note that if $j = l = k$, there is nothing to prove. If $j < k$ it follows from the definitions that $S \in \mathcal{C}_k$, thus $S \notin \mathcal{D}_k$ and so $l < k$. By the definitions, there is $s \in S$ with $s > m_{k-1}$ such that $S \setminus \{s\} \in \mathcal{C}_{k-1} \cup \mathcal{D}_{k-1}$. Since $T \subseteq [m_j]$ and $R \subseteq [m_l]$ it follows that $s \in \{t_{j+1}, \dots, t_k\} \cap \{r_{l+1}, \dots, r_k\}$. Without loss of generality, $s = r_k = t_k$. It follows that $T \cup \{t_{j+1}, \dots, t_{k-1}\} = R \cup \{r_{l+1}, \dots, r_{k-1}\} \in \mathcal{C}_{k-1} \cup \mathcal{D}_{k-1}$. By induction, $j = l$ and $R = T$.

Proof of Assertion (2). Again we prove the assertion by induction on k. For $k = 0$ it follows from the definition of $\mathcal{C}_0, \mathcal{D}_0$ and the assumption that G is compressed. Let $k > 0$, $S \in \mathcal{C}_k \cup \mathcal{D}_k$ and choose $a \in S$, $b < a$ such that $b \notin S$ (if such b exists). Let $T = S \setminus \{a\}$ and $R = S \triangle \{a, b\}$. To prove the assertion we show that $R, T \in \mathcal{E}_k$.

If $S \in \mathcal{D}_k$, the claim follows directly from the definition of $\mathcal{D}_k$ and the fact that G is compressed. Thus we assume $S \in \mathcal{C}_k$, so we can write $S = S_1 \cup \{s\}$ where $s > m_{k-1}$ and $S_1 \in \mathcal{C}_{k-1} \cup \mathcal{D}_{k-1}$. If $a \ne s$, by induction we have $R \setminus \{s\}, T \setminus \{s\} \in \mathcal{E}_{k-1}$, thus $R, T \in \mathcal{E}_k$, so we assume $a = s$. Then clearly $T = S_1 \in \mathcal{C}_{k-1} \cup \mathcal{D}_{k-1}$. It remains to show that $R \in \mathcal{E}_k$.

Let $S = S_2 \cup \{s_{j+1}, \dots, s_k\}$, be a representation of S as in Assertion (1) and assume $s = s_k$. If $b \le m_j$, it follows that $S_2 \cup \{b\} \in \mathcal{E}_j$ and thus $S = (S_2 \cup \{b\}) \cup \{s_{j+1}, \dots, s_{k-1}\} \in \mathcal{E}_{k-1}$.

It remains to consider the case $b > m_j$. Let $s'_{j+1} < \ldots < s'_k$ be such that $\{s'_{j+1}, \ldots, s'_k\} = \{s_{j+1}, \ldots, s_{k-1}, b\}$. If $s'_u > m_{u-1}$ for every $j+1 \leq u \leq k$ then $R \in \mathcal{C}_k$. Otherwise let $l = \max\{u : s'_u \leq m_{u-1}\}$ and $S_3 = S_2 \cup \{s'_{j+1}, \ldots, s'_l\}$. Since $S_3 \subseteq [m_{l-1}]$ it follows that $S_3 \in \mathcal{E}_{l-1}$ and $R = S_3 \cup \{s_{l+1}, \ldots, s_k\} \in \mathcal{E}_{k-1}$.

Proof of Assertion (3). Let $S \in \mathcal{E}_k$ and T be a neighbour of S in G. We show that $T \in \mathcal{E}_k \cup \mathcal{C}_{k+1}$, implying that there are no edges of G between $\mathcal{E}_k$ and $\mathcal{D}_{k+1} \cup \mathcal{C}_{k+2} \cup \mathcal{D}_{k+2}$. If $T \subseteq S$, it follows from Assertion (2) that $T \in \mathcal{E}_k$. So we assume $T = S \cup \{t\}$ and set $s = \max S$. If $t > m_k$, then $T \in \mathcal{C}_{k+1}$. If $s, t \leq m_k$, then $T \subseteq [m_k]$, so $T \in \mathcal{E}_k$. Finally, we consider the case $t < m_k \leq s$. Since $\mathcal{E}_k$ is compressed, it follows that $T \setminus \{s\} = S \triangle \{s, t\} \in \mathcal{E}_k$. This implies $T \in \mathcal{E}_k \cup \mathcal{C}_{k+1}$.

Proof of Assertion (4). From the definitions it follows that $\mathcal{E}_{k-1} \cap \mathcal{D}_k = \emptyset$. Since $\mathcal{D}_0 \cup \ldots \cup \mathcal{D}_{k-1} \subseteq \mathcal{P}([m_{k-1}])$, it follows $\mathcal{C}_k \cap (\mathcal{D}_0 \cup \ldots \cup \mathcal{D}_{k-1}) = \emptyset$. Thus it remains to show that $\mathcal{C}_j \cap \mathcal{C}_k = \emptyset$ for $0 \leq j < k$. We prove this by induction on k. For $k = 0$ there is nothing to prove. Assume $0 \leq j < k$ and $S \in \mathcal{C}_k \cap \mathcal{C}_j$. Write $s = \max S$. By considering the representations of S as in Assertion (1), it is easy to see that $S \setminus \{s\} \in (\mathcal{C}_{j-1} \cup \mathcal{D}_{j-1}) \cap (\mathcal{C}_{k-1} \cup \mathcal{D}_{k-1})$. As explained above this implies that $S \setminus \{s\} \in \mathcal{C}_{j-1} \cap \mathcal{C}_{k-1}$, contradicting the induction hypothesis.

Proof of Assertion (5). Let $S, T \in \mathcal{C}_k \cup \mathcal{D}_k$ and $t \in [d]$ be such that $T = S \cup \{t\}$. If $t > m_k$, it follows from the definitions that $T \in \mathcal{C}_{k+1}$, contradicting $\mathcal{C}_{k+1} \cap (\mathcal{C}_k \cup \mathcal{D}_k) = \emptyset$. Thus $t \leq m_k$, implying that the maximum degree of $G[\mathcal{C}_k \cup \mathcal{D}_k]$ is at most m_k.

We now prove by induction on k that $m_k \leq \epsilon d$. Recall that by the definition of M, the maximum degree of $G[\mathcal{A}_M]$ is at most ϵd. Thus for $k = 0$ we have $m_0 = \max(1, \deg_{G[\mathcal{A}_M]}(\emptyset)) \leq \epsilon d$. Now let $k > 0$ and $S = \{m_0 + 1, \ldots, m_{k-1} + 1\}$. It follows from the definition of $\mathcal{B}_{k-1}$ that $S \notin \mathcal{A}_{M-(k-1)}$, so $\deg_{G[\mathcal{A}_{M-k}]}(S) \leq \epsilon d$. Since $\mathcal{A}_{M-k}$ is compressed, $m_k \leq \max(\deg_{G[\mathcal{A}_{M-k}]}(S), m_{k-1}) \leq \epsilon d$.

Proof of Assertion (6). Let $S \in \mathcal{A}_{M-k}$. Note that if $|S| \leq k$ it can be easily shown by induction that $S \in \mathcal{E}_k$. Thus we assume $|S| \geq k+1$. Define $t_k = \max S$ and for $0 \leq j < k$, denote $t_j = \max(S \setminus \{t_{j+1}, \ldots, t_k\})$. Assume first that $m_j < t_j$ for every $0 \leq j < k$. Since $\mathcal{A}_{M-k}$ is compressed, it follows that $\{m_0 + 1, \ldots, m_{k-1} + 1, t_k\} \in \mathcal{A}_{M-k}$. Thus $t_k \leq m_k$, $S \subseteq [m_k]$ and $S \in \mathcal{E}_k$. Otherwise, let $l \geq 0$ be maximal such that $t_l \leq m_l$. It follows from the definitions that $S \cap [m_l] \in \mathcal{E}_l$ and $S \in \mathcal{E}_k$.

Proof of Assertion (7). We show that for every k, j if $S \in \mathcal{C}_k$, $T \in \mathcal{D}_j$ are such that $S \in N_{k-j}(T)$ then there exist $t_{j+1}, \ldots, t_k$ such that $S = T \cup \{t_{j+1}, \ldots, t_k\}$ and $t_l > m_{l-1}$

for every $j < l \le k$. This proves that the sets $(N^{(k)}(S))_{0\le k\le M-1, S\in\mathcal{D}_k}$ are pairwise disjoint using Assertion (1).

By Assertions (2) and (3) the sets $\mathcal{E}_l$ are down-closed and there are no edges between $\mathcal{E}_l$ and $\mathcal{E}_{l+2}$. Thus, since $S \in N_{k-j}(T)$, S is obtained by adding $k-j$ elements to T, and we can write $S = T \cup \{t_{j+1}, \dots, t_k\}$. Assuming that $t_{j+1} < \dots < t_k$ and that there exists $j+1 \le l \le k$ such that $t_l \le m_{l-1}$, we define r to be the maximal such l. It follows that $T \cup \{t_{j+1}, \dots, t_r\} \in \mathcal{E}_{r-1}$ and thus $S \in \mathcal{E}_{k-1}$, contradicting our assumptions. $\square$

The proof of Proposition 16 completes the proof of our first main result, Theorem 4.

6 Hamming ball is best for fixed i

In this section we prove Theorem 5, whose statement is as follows.

Theorem 5. *For every i there is $d_0 = d_0(i)$ such that for $d \ge d_0$ the Hamming ball H_d^i maximises the largest eigenvalue among subgraphs of Q_d with the same number of vertices.*

Let us start with an outline of the proof. We are given a graph G that maximises the largest eigenvalue among subgraphs of Q_d with $|H_d^i|$ vertices. As usual, we assume the graph and the eigenvector v with eigenvalue $\lambda_1(G)$ are compressed. Using the proof of Theorem 4, we conclude that by removing the vertices of level $i+1$ and higher, the largest eigenvalue does not decrease by much. We infer that G has to contain almost all vertices levels i or less. By assuming that G maximises λ_1, given an eigenvector, we know that moving weight from a vertex of level $i+1$ or higher into level i can only decrease the inner product $\langle A(G)v, v\rangle$, enabling us to obtain a lower bound on the weight of a vertex at the highest non empty level. Finally, using the relations between the weights of vertices and their neighbourhoods, and the fact that there are few vertices in level $i+1$ or higher, we reach a contradiction to the assumption that v is compressed, by concluding that there is a vertex of weight higher than the weight of the empty set.

We now proceed to the proof of the theorem.

Proof of Theorem 5. Let G be a subgraph of Q_d with $|H_d^i|$ vertices and assume $\lambda \triangleq \lambda_1(G)$ is maximal among subgraphs of Q_d with the same number of vertices. Let $v = (v_S)_{S\in V(G)}$ be a positive vector of norm 1 giving $\lambda = \langle A(G)v, v\rangle$. By Lemmas 6 and 7, we can assume that $V(G)$ and v are compressed.

We first show that under the above assumptions, the graph obtained from G by removing vertices of level $i+1$ or more still has a large maximal eigenvalue.

Claim 18. *Let $U = V(G) \cap [d]^{(\le i)}$. There exists $\eta = \eta(i) > 0$ such that $\lambda_1(Q_d[U]) \ge \lambda_1(H_d^i) - O(d^{1/2-\eta})$.*

Proof. We use the proof of Theorem 4. Consider Inequality (5) which states the following.

$$\lambda_1(G) \le \max_k \lambda_1(G[\mathcal{C}_k \cup \mathcal{D}_k]) + \max_{k,S} \lambda_1(G[N_S^{(k)}]).$$

As explained in the proof of Claim 17, if $\epsilon = 2id^{-1/(i+1)}$ then $M \le i$, implying that the sets $N_S^{(k)}$ are subsets of Hamming balls of radius at most i. It follows that $\lambda_1(N_S^{(k)}) \le \lambda_1(Q_d[U])$, because G is compressed. Furthermore, for our choice of ϵ, we have

$$\lambda_1(G[\mathcal{C}_k \cup \mathcal{D}_k]) = O\left(\sqrt{d^{1-\frac{1}{i+1}} \log d}\right).$$

Thus for any $\eta < 1/2(i+1)$ we have

$$\lambda_1(G) \le O(d^{1/2-\eta}) + \lambda_1(Q_d[U]).$$

The proof of Claim 18 follows from the assumption that $\lambda_1(G) \ge \lambda_1(H_d^i)$. $\square$

We conclude that $\left|V(G) \setminus [d]^{(\le i)}\right|$ is small.

Claim 19. *There exists $\theta = \theta(i) > 0$ such that $\left|V(G) \setminus [d]^{(\le i)}\right| = O(d^{i-\theta})$.*

Proof. Define

$$\begin{aligned} A &= \left\{a \in [d] \ : \text{there exists } S \in V(G) \cap [d]^{(i)} \text{ such that } a = \min S\right\} \\ B &= \left\{S \in [d]^{(i)} \ : S \cap A \neq \emptyset\right\}. \end{aligned}$$

Since G is compressed, it follows that $A^{(i)} \subseteq V(G) \cap [d]^{(i)} \subseteq B$. Write $|A| = (1-\beta)d$ and let H be the subgraph of Q_d induced by $[d]^{(<i)} \cup B$. Note that $V(G) \cap [d]^{(\le i)} \subseteq V(H)$. It follows from Claim 18 that $\lambda_1(H) \ge \lambda_1(H_d^i) - O(d^{1/2-\eta})$. We shall conclude that $\beta = O(d^{-\theta})$ for some $\theta = \theta(i) > 0$. This implies that $\left|V \cap [d]^{(i)}\right| \ge \binom{(1-\beta)d}{i} = \binom{d}{i} - O(d^{i-\theta})$, as required.

Note that by symmetry, the eigenvector of H with eigenvalue $\lambda_1(H)$ is uniform on vertices from the same level and with the same number of elements in $[(1-\beta)d]$. Let $x_{j,k}$ be the weight of a vertex from $[d]^{(j)}$ with k elements in $[(1-\beta)d]$ in the eigenvector. Let $y_{j,k} = x_{j,k} d^{i/2}$ and denote $\mu = \lambda_1(H) d^{-1/2}$. Consider the following equation.

$$\begin{aligned} \mu\, y_{j,k} &= (j-k)\, y_{j-1,k} + k\, y_{j-1,k-1} \\ &\quad + \mathbb{1}_{j<i} \cdot \Big((1-\beta+O(1/d))\, y_{j+1,k+1} + (\beta - O(1/d))\, y_{j+1,k} \Big). \end{aligned}$$

This system of equations, taken for $0 \le j \le i$ and $0 \le k \le j$ describes H_d^i, so the corresponding maximal eigenvalue is $\lambda_1(H_d^i)/\sqrt{d}$.

We obtain a similar system, by dropping the equation for $j = i$ and $k = 0$ (and taking $x_{i,0} = 0$). Denote the corresponding eigenvalue by μ_β. By monotonicity of the eigenvalue, μ_β is decreasing with β.

Now consider the systems obtained from the above two systems by omitting the $O(1/d)$ terms. The eigenvalue from the first system is the constant λ_i from Lemma 14. Let ν_β be the eigenvalue of the second system with the $O(1/d)$ terms omitted. Then $|\mu_\beta - \nu_\beta| = O(1/d)$. Furthermore, by monotonicity of the maximal eigenvalue, $\nu_\beta < \lambda_i$ for $0 < \beta < 1$.

Note that we can conclude that $\beta = o(1)$. Suppose to the contrary that $\beta > c$ where $c > 0$ is a constant. Then $\lambda_i - \nu_c = \Omega(1)$ and $\mu_\beta \leq \nu_c + O(1/d)$, implying that $\mu_\beta = \lambda_i - \Omega(1)$, a contradiction.

We now prove a sharper upper bound on β. Denote by $P_\beta(x)$ the characteristic polynomial obtained by the second system with the $O(1/d)$ terms omitted. Then $P_\beta(\lambda_i)$ is a polynomial in β which has a root at $\beta = 0$ but is not identically 0. Thus $P_\beta(\lambda_i) = \Omega(\beta^t)$ where t is the smallest power of β in $P_\beta(\lambda_i)$ with a non zero coefficient. The degree of this polynomial is at most $(1+2+\ldots+i)+i \leq 2i^2$, thus $P_\beta(\lambda_i) = \Omega(\beta^{2i^2})$. Note that the derivative of $P_\beta(x)$ satisfies $\frac{\partial P_\beta(x)}{\partial x} = O(1)$ for $|x| \leq \lambda_i$. It follows that $P_\beta(\lambda_i) - P_\beta(\nu_\beta) = O(\lambda_i - \nu_\beta)$. Since $P_\beta(\nu_\beta) = 0$ and $P_\beta(\lambda_i) = \Omega(\beta^{2i^2})$, this implies that $\lambda_i - \nu_\beta = \Omega(\beta^{2i^2})$.

If $\beta^{2i^2} = O(1/d)$, we are done. Otherwise, this implies that $\frac{\lambda_1(H_d^i) - \lambda_1(H)}{\sqrt{d}} = \Omega(\beta^{2i^2})$. Recall that we have $\frac{\lambda_1(H_d^i) - \lambda_1(H)}{\sqrt{d}} = O(d^{-\eta})$. Thus $\beta = O(d^{-\eta/2i^2})$, completing the proof. $\square$

Let $l = \max\{j : V(G) \cap [d]^{(j)} \neq \emptyset\}$. Assuming that $V(G) \neq [d]^{(\leq i)}$, we have $l > i$. Note that since G is down-compressed, we have $l = O(\log d)$.

Claim 20. *If $l > i$ then $v_{[l]} = \Omega(v_\emptyset \lambda^{-i})$.*

Proof. We first show by induction on $|S|$ that for every $S \in V(G)$ we have $v_S \geq v_\emptyset \lambda^{-|S|}$. It is clearly true for $S = \emptyset$. Now let $S \neq \emptyset$ and let $a \in S$, $T = S \setminus \{a\}$. Then λv_S is the sum of weights of the neighbours of S, and in particular $\lambda v_S \geq v_T \geq v_\emptyset \lambda^{-|T|}$.

Since we assume $l > i$, the set $[d]^{(\leq i)} \setminus V(G)$ is non empty. Pick a minimal element S in it. By moving the weight $v_{[l]}$ from $[l]$ to S, the value of $\langle A(G)v, v\rangle$ decreases by $v_{[l]} \cdot (2 \sum_{j\in[l]} v_{[l]\setminus\{j\}} - 2 \sum_{j\in S} v_{S\setminus\{j\}})$. This amount is non negative, because G has the largest maximal eigenvalue among subgraphs of Q_d with the same number of vertices. Hence,

$$\lambda v_{[l]} = \sum_{j\in[l]} v_{[l]\setminus\{j\}} \geq \sum_{j\in S} v_{S\setminus\{j\}} \geq v_\emptyset \lambda^{-(|S|-1)} \geq v_\emptyset \lambda^{-(i-1)}.$$

and Claim 20 follows. $\square$

In the following claim we conclude that $v_{[l-i]} = \Omega(v_\emptyset / l^i) = \Omega(v_\emptyset / \log^i d)$.

Claim 21. $v_{[l-i]} = \Omega\big(v_{[l]}(\frac{\lambda}{l})^i\big) = \Omega(v_\emptyset/l^i)$.

Proof. We shall show that $v_{[l-j-1]} = \Omega(\lambda v_{[l-j]}/l)$ for every $0 \le j < i$. Claim 21 then follows from Claim 20.

Assume that this assertion does not hold and denote $t = \min\{j : v_{[l-j-1]} = o(\lambda v_{[l-j]}/l)\}$. Define for $j \in [t]$

$$\begin{aligned}
\mathcal{A}_j &= \big\{S \in V(G) \cap [d]^{(l-t+j)} : [l-t] \subseteq S\big\} \\
\mathcal{B}_j &= \big\{T \in \big(V(G) \cap [d]^{(l-t+j-1)}\big) \setminus \mathcal{A}_{j-1} : \text{there exists } S \in \mathcal{A}_j \text{ s.t. } T \subseteq S\big\} \\
W_j &= \sum_{S \in \mathcal{A}_j} v_S \qquad\qquad U_j = \sum_{S \in \mathcal{B}_j} v_S.
\end{aligned}$$

We obtain the following inequalities, using the fact that every vertex in $\mathcal{A}_j$ has at most $d-j$ up-neighbours in $\mathcal{A}_{j+1}$.

$$\lambda W_j \le \begin{cases} W_1 + U_0 & j = 0 \\ (d-j+1)W_{j-1} + (j+1)W_{j+1} + U_j & 0 < j < t \\ (d-t+1)W_{t-1} + U_t & j = t \end{cases}$$

Define $k = \min\{j : W_{j+1} = o(\lambda W_j)\}$. Clearly $k \le t$ since $W_{t+1} = 0$ (by the definition of l). Thus $W_j = \Omega(\lambda^j W_0)$ for $0 \le j \le k$.

Note that $U_0 \le l v_{[l-t-1]} = o(\lambda v_{[l-t]}) = o(\lambda W_0)$, by our assumption. Also, $\lambda U_j \ge (j+1)U_{j+1}$ for $0 \le j \le t$, thus $U_j = O(\lambda^j U_0) = o(\lambda^{j+1} W_0) = o(\lambda W_j)$ for $0 \le j \le k$. Hence, the above inequalities can be rewritten as follows.

$$\lambda W_j \le \begin{cases} W_1 + o(\lambda W_0) & j = 0 \\ (d-j+1)W_{j-1} + (j+1)W_{j+1} + o(\lambda W_j) & 0 < j < k \\ (d-k+1)W_{k-1} + o(\lambda W_k) & j = k \end{cases}$$

Denote $W = (W_0, \dots, W_k)^T$, and let A be the matrix with the above coefficients with the $o(\lambda W_j)$ terms dropped. The above inequalities translate to $\lambda W_j \le (AW)_j + o(\lambda W_j)$. We obtain the following chain of inequalities, where $X_j = o(\lambda W_j)$.

$$\lambda_1(A) \ge \frac{\langle AW, W\rangle}{\langle W, W\rangle} \ge \frac{\langle \lambda W, W\rangle}{\langle W, W\rangle} - \frac{\langle X, W\rangle}{\langle W, W\rangle} = \lambda - o(\lambda) \ge \lambda_1(H_d^i) - o(\sqrt{d}).$$

However, $\lambda_1(A) = \lambda_1(H_d^k)$, thus we obtained $\lambda_1(H_d^i) - \lambda_1(H_d^k) = o(\sqrt{d})$ for some $0 \le k < i$, which is a contradiction to Lemma 14. $\square$

Claim 21 implies that $v_{[l-i]} = \omega(v_{[l-i-1]}\lambda/l)$ (otherwise, $v_{[l-i]} > v_\emptyset$, a contradiction to the assumption that v is compressed). Similarly to the proof of the claim, denote by X_j the sum of weights of vertices in the $(l-i+j)^{\text{th}}$ level containing $[l-i]$. By the same arguments we obtain a contradiction unless $X_j = \Omega(\lambda X_{j-1})$ for $1 \le j \le i$. Thus $X_i = \Omega(X_0 \lambda^i) = \Omega(v_\emptyset \lambda^i / \log^i d)$. Recall that by Claim 19, there are at most $O(d^{i-\theta})$ vertices in $V \cap [d]^{(l)}$. Hence, using the fact that v is compressed, $v_{[l]} = \Omega(d^{-i/2+\theta/2})$. By Claim 20, this implies that $v_{[l-i]} = \Omega(v_{[l]} d^{\theta/2} / \log^i d) = \omega(v_\emptyset)$, contradicting the assumption that v is compressed. Hence we cannot have $l > i$, so $G = H_d^i$, as required. This proves that H_d^i maximises λ_1 among subgraphs of the cube with the same number of vertices, thus completing the proof of our second main result, Theorem 5. $\square$

7 Conclusion

The question of characterising the subgraphs of the cube maximising λ_1 is far from being completely answered. For radii tending to infinity with the dimension of the cube, our results as well as Samorodnitsky's results [28] only show that the Hamming balls have largest eigenvalues which are asymptotically largest among subgraphs of the same order. We believe that, similarly to Theorem 2, the Hamming balls maximise the maximal eigenvalues exactly rather than just asymptotically, for large d and a large range of radii.

Question 22. *Is it true that if $d/2 - i$ is sufficiently large, then H_d^i maximises λ_1 among subgraphs of Q_d with the same number of vertices?*

We point out that for radii that are very close to $d/2$ the Hamming ball does not achieve the largest maximal eigenvalue, as can be seen by the following example.

Example 23. *Assume that d is even and consider the Hamming ball of radius $d/2 - 1$, $H = H_d^{d/2-1}$. We show that $\lambda_1(H) = d - 2$. Put $\lambda = d - 2$ and let x be the vector with weight $x_i = 1 - 2i/d$ on the vertices of level i. The following can be easily verified.*

$$\lambda x_i = \begin{cases} d x_1 & i = 0 \\ i x_{i-1} + (d-i) x_{i+1} & 0 < i < d/2 - 1 \\ i x_{i-1} & i = d/2 - 1 \end{cases}$$

Thus we have $A(H)x = (d-2)x$. Since all the weights x_i are positive, this implies that $\lambda_1(H) = d - 2$. Note that $|H| = 2^{d-1}(1 - \Theta(1/\sqrt{d})) > 2^{d-2}$. Thus, since the largest eigenvalue of the subcube of dimension $d - 2$ is $d - 2$, we can achieve a larger maximal eigenvalue with a (connected) subgraph with $|H|$ vertices containing the subcube of dimension $d - 2$.

Similarly, it can be shown that if $i = d/2 - \sqrt{d}$ is an integer, then $\lambda_1(H_d^i) = d - 4$. Since $|H_d^i| > 2^{d-4}$, also in this case the largest eigenvalue of the Hamming ball is not maximal among subgraphs of the cube of the same order.

As seen by this example, it may be interesting to consider subgraphs whose largest eigenvalue is very close to d. For instance, determining the range of radii for which the Hamming balls maximise the maximal eigenvalue, especially for large radii, seems like a challenging problem. The following weaker problem also seems hard. Is it true that for every fixed $c > 0$, if a subgraph H of Q_d has $\lambda_1(H) \geq d - c$, then $|H| = \Omega(2^d)$?

References

[1] N. Alon and V.D. Milman, λ_1, *isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theory Ser. B **38** (1985), 73–88.

[2] B. Bollobás and A.J. Radcliffe, *Isoperimetric inequalities for faces of the cube and the grid*, European J. Combin. **11** (1990), 323–333.

[3] R.C. Brigham and R.D. Dutton, *Bounds on graph spectra*, J. Combin. Theory Ser. B **37** (1984), 228–234.

[4] ______, *Bounds on graph spectra and girth*, Ars Combin. **20** (1985), 91–100.

[5] R.A. Brualdi and A.J. Hoffman, *On the spectral radius of* $(0,1)$*-matrices*, Linear Algebra Appl. **65** (1985), 133–146.

[6] D. Cvetković, P. Rowlinson, and S. Simić, *Graphs with least eigenvalue* -2*: ten years on*, Linear Algebra Appl. **484** (2015), 504–539.

[7] D.M. Cvetković, M. Doob, and H. Sachs, *Spectra of graphs*, third ed., Johann Ambrosius Barth, Heidelberg, 1995, Theory and applications.

[8] O. Favaron, M. Mahéo, and J.-F. Saclé, *Some results on conjectures of Graffiti. I*, Ars Combin. **29** (1990), no. C, 90–106, Twelfth British Combinatorial Conference (Norwich, 1989).

[9] ______, *Some eigenvalue properties in graphs (conjectures of Graffiti. II)*, Discrete Math. **111** (1993), no. 1-3, 197–220, Graph Theory and Combinatorics (Marseille-Luminy, 1990).

[10] H.J. Finck and G. Grohmann, *Vollständiges Produkt, chromatische Zahl und charakteristisches Polynom regulärer Graphen. I (in German)*, Wiss. Z. Techn. Hochsch. Ilmenau **11** (1965), 1–3.

[11] T.M.A. Fink, *Private communication*, (2012).

[12] J. Friedman and J-P. Tillich, *Generalized Alon-Boppana theorems and error-correcting codes*, SIAM J. Discrete Math. **19** (2005), no. 3, 700–718.

[13] L.H. Harper, *Optimal assignments of numbers to vertices*, J. Soc. Indust. Appl. Math. **12** (1964), 131–135.

[14] ———, *Optimal numberings and isoperimetric problems on graphs*, J. Combin. Theory **1** (1966), 385–393.

[15] S. Hart, *A note on the edges of the n-cube*, Discrete Math. **14** (1976), no. 2, 157–163.

[16] A.J. Hoffman, *Estimation of eigenvalues of a matrix and the theory of linear inequalities*, Mathematics of the Decision Sciences, Part I (Seminar, Stanford, Calif., 1967), Amer. Math. Soc., Providence, R.I., 1968, pp. 295–300.

[17] ———, *On eigenvalues and colorings of graphs*, Graph Theory and its Applications (Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., 1969), Academic Press, New York, 1970, pp. 79–91.

[18] Y. Hong, *Bounds of eigenvalues of graphs*, Discrete Math. **123** (1993), 65–74.

[19] B. Liu, J. Shen, and X. Wang, *On the largest eigenvalue of non-regular graphs*, J. Combin. Theory Ser. B **97** (2007), 1010–1018.

[20] A. Neumaier, *The second largest eigenvalue of a tree*, Linear Algebra Appl. **46** (1982), 9–25.

[21] V. Nikiforov, *Extrema of graph eigenvalues*, preprint, arXiv:1502.00359, pp. 32.

[22] ———, *Walks and the spectral radius of graphs*, Linear Algebra Appl. **418** (2006), 257–268.

[23] ———, *Bounds of graph eigenvalues. i.*, Linear Algebra Appl. **420** (2007), 667–671.

[24] ———, *Bounds on graph eigenvalues. ii.*, Linear Algebra Appl. **427** (2007), 183–189.

[25] E. Nosal, *Eigenvalues of Graphs*, Master's thesis, University of Calgary, 1970.

[26] D.L. Powers, *Bounds on graph eigenvalues*, Linear Algebra Appl. **117** (1989), 1–6.

[27] T. Reeves, R.S. Farr, J. Blundell, A. Gallagher, and T.M.A. Fink, *Eigenvalues of neutral networks: interpolating between hypercubes*, preprint, arXiv:1504.03065, 8 pp.

[28] A. Samorodnitsky, *A modified logarithmic Sobolev inequality for the Hamming cube and some applications*, preprint, arXiv:0807.1679, 18 pp.

[29] J.B. Shearer, *On the distribution of the maximum eigenvalue of graphs*, Linear Algebra Appl. **114/115** (1989), 17–20.

[30] R.P. Stanley, *A bound on the spectral radius of graphs with e edges*, Linear Algebra Appl. **87** (1987), 267–269.