

Hypergraph expanders of all uniformities from Cayley graphs

David Conlon* Jonathan Tidor† Yufei Zhao‡

Abstract

Hypergraph expanders are hypergraphs with surprising, non-intuitive expansion properties. In a recent paper, the first author gave a simple construction, which can be randomized, of 3-uniform hypergraph expanders with polylogarithmic degree. We generalize this construction, giving a simple construction of r -uniform hypergraph expanders for all $r \geq 3$.

1 Introduction

An expander is a sparse graph with the property that every small vertex subset expands, that is, is adjacent to many vertices outside of the set. The study of expanders has occupied a central place in both mathematics and computer science for the last forty years, finding numerous applications across both areas. We refer the reader to the detailed surveys [10] and [14] for further information.

That expanders exist was first shown by Pinsker [25], who observed that random regular graphs are almost surely expanders. Since his result, many different methods have been developed for constructing expanders, some completely explicit and some retaining elements of randomness: using representation theory and Kazhdan's property (T), as done by Margulis [20] when he found the first explicit construction; by taking finite quotients of an infinite tree, a procedure used by Margulis [21] and by Lubotzky, Phillips and Sarnak [16] to produce Ramanujan graphs, graphs with optimal spectral properties; the zig-zag product approach of Reingold, Vadhan and Wigderson [26]; through random lifts [2], an idea which has recently led to constructions of bipartite Ramanujan graphs [19] of all degrees; and by taking random Cayley graphs [1, 3]. Since the last of these will be important to us in what follows, let us say a little more.

*Department of Mathematics, California Institute of Technology, Pasadena, CA 91125. Email: dconlon@caltech.edu. Research supported by a Royal Society University Research Fellowship and ERC Starting Grant 676632.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139. Email: jtidor@mit.edu. Researched supported by an MIT Presidential Fellowship.

‡Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139. Email: yufeiz@mit.edu. Research supported by NSF Awards DMS-1764176 and DMS-1362326 and the MIT Solomon Buchsbaum Fund.

The basic result on the expansion of random Cayley graphs is due to Alon and Roichman [1]. Their result is best explained in terms of spectral properties. Writing $\mathbf{A}_G$ for the adjacency matrix of a graph G , we let $\lambda(G)$ be the maximum absolute value of a non-trivial eigenvalue of $\mathbf{A}_G$. We then say that a d -regular graph G is an ϵ -*expander* if $\lambda(G) \leq (1-\epsilon)d$. By Cheeger's inequality, a standard result in the area, the spectral condition implies that every vertex subset U of G with $|U| \leq |V(G)|/2$ has at least $\frac{\epsilon}{2}|U|$ neighbours outside of U , so G conforms with our intuitive idea of what an expander should be. The Alon–Roichman theorem now says that for any $0 < \epsilon < 1$ there exists C such that if a finite group has n elements then the Cayley graph generated by $C \log n$ random elements of this group is almost surely an ϵ -expander. This theorem is easily seen to be best possible when the group is $\mathbb{F}_2^t$, the case which will be of most relevance to us here.

In recent years, a theory of hypergraph, or high-dimensional, expanders has begun to emerge. The literature in this area is already quite large, but a good place to start exploring might be with the recent ICM survey by Lubotzky [15]. Many different definitions have been proposed for hypergraph expansion, each with their own strengths and weaknesses. Intriguingly, one of the few points of consensus among these definitions is that random hypergraphs are not good models for hypergraph expansion.

What then is a good model? The prototype for all subsequent constructions (and, in effect, the foundation on which the entire area is built) are the Ramanujan complexes of Lubotzky, Samuels and Vishne [17], Li [13] and Sarveniazi [27]. In the same way that Ramanujan graphs are built as finite quotients of infinite trees, inheriting many properties of these trees, Ramanujan complexes are finite quotients of Bruhat–Tits buildings and again inherit properties from these buildings.

To say more about their properties, let us fix some terminology. An r -*uniform hypergraph* is a pair $H = (V, E)$ where $E \subseteq \binom{V}{r}$. The elements of V are called vertices and the elements of E are called r -edges. A k -edge of H is a k -element subset of an r -edge. A hypergraph is D -*regular* if every $(r-1)$ -edge is contained in exactly D of the r -edges. Then one property of Ramanujan complexes is that they are regular.

This is already a non-trivial property to obtain, especially if the number of edges in the hypergraph is linear in the number of vertices, as it is for Ramanujan complexes. But one might hope for more. The property that has drawn the most attention in the literature is Gromov's notion of topological expansion [8] and its weaker relative, geometric expansion. We say that an r -uniform hypergraph H , seen as a simplicial complex, is an ϵ -*topological expander* if for any continuous map φ from the complex into $\mathbb{R}^{r-1}$ there exists a point p such that $\varphi^{-1}(p)$ intersects an ϵ -fraction of the edges of H . Geometric expansion is defined similarly but only needs to hold for affine maps ϕ defined by first mapping the vertices of H and then extending the map to the convex hull using linearity. That Ramanujan complexes are geometric expanders was shown in [7], while topological expansion (of a hypergraph derived from Ramanujan complexes) was shown in [6, 11].

The property we will be particularly concerned with here is very different, but also very natural, being a generalization of the notion of graph expansion defined earlier. Given an r -uniform hypergraph H , define its *walk graph* $G_{\text{walk}}(H)$ to be the graph whose vertices are the $(r-1)$ -edges of H with an edge between two $(r-1)$ -edges if they are contained in a common r -edge. We say that H is an ϵ -*expander* if $G_{\text{walk}}(H)$ is an ϵ -expander. In particular, this implies that the random walk defined by starting at any $(r-1)$ -edge and then repeatedly moving to an adjacent $(r-1)$ -edge chosen uniformly at random converges rapidly to the uniform distribution. For Ramanujan complexes, this latter property was verified by Kaufman and Mass [12].

The main result of this paper is a simple construction of such expanders for all uniformities $r \geq 3$, building significantly on earlier work of the first author [4] which applied in the 3-uniform case. The construction in [4] is surprisingly simple: given an expanding Cayley graph over $\mathbb{F}_2^t$ with generating set S , define H to be the hypergraph with vertex set $\mathbb{F}_2^t$ and 3-edges $(x + s_1, x + s_2, x + s_3)$ for $x \in \mathbb{F}_2^t$ and $s_1, s_2, s_3 \in S$ distinct. Then H is a hypergraph expander. For higher uniformities, the construction is not quite so simple, though it also not much more complicated.

1.1 The construction

We work with the following parameters: $r \geq 3$, the uniformity of our hypergraph; t , the dimension of our base vector space; $S \subset \mathbb{F}_2^t$, the generating set of our underlying Cayley graph; and $\epsilon > 0$, the degree of expansion of this Cayley graph. We will typically make the following three assumptions.

Assumption 1. For $r \geq 3$, t and $S \subset \mathbb{F}_2^t$, the set S satisfies $|S| \geq 2^{2r}$.

Assumption 2. For $r \geq 3$, t and $S \subset \mathbb{F}_2^t$, the sum $s_1 + \dots + s_l$ is different for every choice of $0 \leq l \leq 2^r$ and every choice of distinct $\{s_1, \dots, s_l\} \subset S$.

Assumption 3. For t , $\epsilon > 0$ and $S \subset \mathbb{F}_2^t$, the graph $\text{Cay}(\mathbb{F}_2^t, S)$ is an ϵ -expander.

In practice, for $r \geq 3$ and $\epsilon > 0$ fixed, we will let C be a sufficiently large constant in terms of r and ϵ and take t sufficiently large in terms of C , r and ϵ . If we then define S to be a uniform random set $S \subset \mathbb{F}_2^t$ of size $|S| = Ct$, the assumptions are all satisfied with high probability. Indeed, Assumption 1 is obvious, Assumption 2 follows from a first moment calculation and Assumption 3 follows from the Alon–Roichman theorem [1].

Let $\mathcal{P}$ be a collection of subsets of $[r]$ defined as follows:

$$\mathcal{P} = \begin{cases} \{I \subset [r] : 1 \leq |I| < r/2\} & \text{if } r \text{ is odd,} \\ \{I \subset [r] : 1 \leq |I| < r/2 \text{ or } |I| = r/2 \text{ and } 1 \in I\} & \text{if } r \text{ is even.} \end{cases} \quad (1.1)$$

Observe that $\{\emptyset\} \cup \mathcal{P}$ is a downward-closed family of subsets of $[r]$ that contains exactly one of I and I^c for each $I \subseteq [r]$.

Let $\mathcal{T} \subset S^{\mathcal{P}}$ be the set of $|\mathcal{P}|$ -tuples

$$\mathcal{T} = \{(s_I)_{I \in \mathcal{P}} : s_I \in S \text{ distinct}\}. \quad (1.2)$$

For $x \in \mathbb{F}_2^t$ and $\mathbf{s} \in \mathcal{T}$, define $e(x, \mathbf{s})$ to be the ordered r -tuple

$$e(x, \mathbf{s}) = \left(x + \sum_{\substack{I \in \mathcal{P} \\ i \in I}} s_I \right)_{1 \leq i \leq r}.$$

In the above equation and throughout this paper we always use the convention that the bold-face letters $\mathbf{s}, \mathbf{t}, \mathbf{u}$ refer to elements of $\mathcal{T}$. The coordinates of these tuples are then named by subscripted normal-print letters, e.g., $\mathbf{s} = (s_I)_{I \in \mathcal{P}}$.

For $r \geq 3$, t and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, define $\mathbf{H}_{r,t,S}$ to be the r -uniform hypergraph whose vertex set is $\mathbb{F}_2^t$ and with $\{v_1, \dots, v_r\}$ an r -edge if and only if there exists some $x \in \mathbb{F}_2^t$ and $\mathbf{s} \in \mathcal{T}$ such that $e(x, \mathbf{s}) = (v_1, \dots, v_r)$. Note that Assumption 2 implies that every r -tuple $e(x, \mathbf{s})$ for $\mathbf{s} \in \mathcal{T}$ has distinct coordinates.

The 3-uniform hypergraph $\mathbf{H}_{3,t,S}$ has 3-edges $\{x + s_1, x + s_2, x + s_3\}$ for $x \in \mathbb{F}_2^t$ and $s_1, s_2, s_3 \in S$ distinct. This is exactly the hypergraph considered in [4]. The $r = 4$ case will be spelled out explicitly for the sake of illustration in the next subsection.

Recall that for an r -uniform hypergraph H , we use $G_{\text{walk}}(H)$ to denote the graph whose vertices are the $(r-1)$ -edges of H where two $(r-1)$ -edges are connected if and only if they are contained in a common r -edge of H . The main result of this paper is that if $\text{Cay}(\mathbb{F}_2^t, S)$ is an ordinary expander graph then $\mathbf{H}_{r,t,S}$ is a hypergraph expander in the sense that $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ is an expander graph.

Theorem 1.1. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3, there exists a constant $c = c(r) > 0$ such that the walk graph $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ is a $c\epsilon$ -expander.*

We also prove a generalization of this theorem to all lower-order random walks. For an r -uniform hypergraph H and for $1 \leq k \leq r-1$, define the k -th order walk graph $G_{\text{walk}}^{(k)}(H)$ to be the graph whose vertices are the k -edges of H where two k -edges are connected if and only if they are contained in a common $(k+1)$ -edge of H .

Theorem 1.2. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3, there exists a constant $c = c(r) > 0$ such that for every $1 \leq k \leq r-1$, the k -th order walk graph $G_{\text{walk}}^{(k)}(\mathbf{H}_{r,t,S})$ is a $c\epsilon$ -expander.*

We also prove a discrepancy result about $\mathbf{H}_{r,t,S}$ similar to the conclusion of a high-dimensional expander mixing lemma. For an r -uniform hypergraph $H = (V, E)$ and $V_1, \dots, V_r \subseteq V$, define $e_H(V_1, \dots, V_r)$ to be the number of r -tuples $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$ such that $\{v_1, \dots, v_r\} \in E$.

Theorem 1.3. For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3 and $V_1, \dots, V_r \subseteq V$,

$$\left| e_{H_{r,t,S}}(V_1, \dots, V_r) - \frac{|S|(|S|-1) \cdots (|S|-(2^{r-1}-2))}{(2^t)^{r-1}} |V_1| \cdots |V_r| \right| \leq (1-\epsilon) 2r |S|^{2^{r-1}-1} (|V_1| \cdots |V_r|)^{1/r}.$$

When ϵ is sufficiently close to 1 in terms of r , an argument in [23, 24] allows us to deduce that $H_{r,t,S}$ is a geometric expander. The idea is simple. By a result of Pach [22], for any set of n points in $\mathbb{R}^{r-1}$, there are r sets $A_1, \dots, A_r$, each of order at least cn , and a point $p \in \mathbb{R}^{r-1}$ such that p is contained in the convex hull of $\{a_1, \dots, a_r\}$ for all $a_1, \dots, a_r$ with $a_i \in A_i$ for $i = 1, \dots, r$. But Theorem 1.3 implies that a positive fraction of the edges in $H_{r,t,S}$ have one vertex in each of $A_1, \dots, A_r$. Since all of these contain p , the desired conclusion follows.

1.2 Proof techniques (the $r = 4$ case)

We illustrate the main features of the construction by exploring the $r = 4$ case. Fix $S \subset \mathbb{F}_2^t$ such that the sum $s_1 + \dots + s_\ell$ is different for every choice of $0 \leq \ell \leq 16$ and $s_1, \dots, s_\ell \in S$ distinct. Define

$$\mathcal{T} = \{(s_1, s_2, s_3, s_4, s_{12}, s_{13}, s_{14}) \in S^7 : s_1, s_2, s_3, s_4, s_{12}, s_{13}, s_{14} \text{ distinct}\}.$$

Then for $x \in \mathbb{F}_2^t$ and $\mathbf{s} \in \mathcal{T}$, define the ordered 4-tuple

$$e(x, \mathbf{s}) = \begin{pmatrix} x+s_1 & & +s_{12}+s_{13}+s_{14}, \\ x & +s_2 & +s_{12} & , \\ x & +s_3 & +s_{13} & , \\ x & +s_4 & +s_{14} \end{pmatrix}.$$

The 4-uniform hypergraph $H_{4,t,S}$ has vertices $\mathbb{F}_2^t$ and 4-edges $\{v_1, v_2, v_3, v_4\}$ where $e(x, \mathbf{s}) = (v_1, v_2, v_3, v_4)$ for some $x \in \mathbb{F}_2^t$ and $\mathbf{s} \in \mathcal{T}$.

At first sight it might seem that there is a distinguished vertex in each 4-edge, but this is an artifact of our presentation. For example, consider the ordered 4-tuple

$$\begin{pmatrix} x & +s_2 & +s_{12} & , \\ x+s_1 & & +s_{12}+s_{13}+s_{14}, \\ x & +s_3 & +s_{13} & , \\ x & +s_4 & +s_{14} \end{pmatrix}. \quad (1.3)$$

If we define $\mathbf{t}$ by $t_1 = s_2$, $t_2 = s_1$, $t_3 = s_3$, $t_4 = s_4$, $t_{12} = s_{12}$, $t_{13} = s_{14}$, $t_{14} = s_{13}$, and set $y = x + s_{13} + s_{14}$, then one can easily check that

$$e(y, \mathbf{t}) = \begin{pmatrix} x + s_{13} + s_{14} + s_2 & & +s_{12}+s_{14}+s_{13}, \\ x + s_{13} + s_{14} & +s_1 & +s_{12} & , \\ x + s_{13} + s_{14} & +s_3 & +s_{14} & , \\ x + s_{13} + s_{14} & +s_4 & +s_{13} \end{pmatrix}$$

is exactly the same as (1.3).

Define the graph $G_{4,t,S}$ to have vertex set $\mathbb{F}_2^t \times \mathcal{T}$ with an edge between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ if $e(x, \mathbf{s})$ and $e(y, \mathbf{t})$ agree in all but one coordinate. This graph is dual to $G_{\text{walk}}(\mathbf{H}_{4,t,S})$ in a sense that will be made precise later on. A theme that recurs throughout this paper is that we will typically prove that $G_{4,t,S}$ has some desired property and then deduce that the same property holds for $G_{\text{walk}}(\mathbf{H}_{4,t,S})$.

In this case, we wish to prove that $G_{\text{walk}}(\mathbf{H}_{4,t,S})$ is an expander, so we might want to first prove that $G_{4,t,S}$ is an expander. Let us start with the easier problem of showing that $G_{4,t,S}$ is connected.

Start at the vertex $(x, \mathbf{s}) = (x, (s_1, s_2, s_3, s_4, s_{12}, s_{13}, s_{14}))$. This vertex is adjacent to $(x, \mathbf{s}^{(1)}) = (x, (t_{12}, s_2, s_3, s_4, s_{12}, s_{13}, s_{14}))$ since

$$e(x, \mathbf{s}) = \begin{pmatrix} x+s_1 & & +s_{12}+s_{13}+s_{14}, \\ x & +s_2 & +s_{12} \\ x & & +s_3 & +s_{13} \\ x & & +s_4 & +s_{14} \end{pmatrix}, \quad (1.4)$$

$$e(x, \mathbf{s}^{(1)}) = \begin{pmatrix} x+t_{12} & & +s_{12}+s_{13}+s_{14}, \\ x & +s_2 & +s_{12} \\ x & & +s_3 & +s_{13} \\ x & & +s_4 & +s_{14} \end{pmatrix}$$

agree in all but the first coordinate.

Vertex $(x, \mathbf{s}^{(1)})$ is adjacent to $(x, \mathbf{s}^{(2)}) = (x, (s_{12}, s_2, s_3, s_4, t_{12}, s_{13}, s_{14}))$ since

$$e(x, \mathbf{s}^{(1)}) = \begin{pmatrix} x+t_{12} & & +s_{12}+s_{13}+s_{14}, \\ x & +s_2 & +s_{12} \\ x & & +s_3 & +s_{13} \\ x & & +s_4 & +s_{14} \end{pmatrix}, \quad (1.5)$$

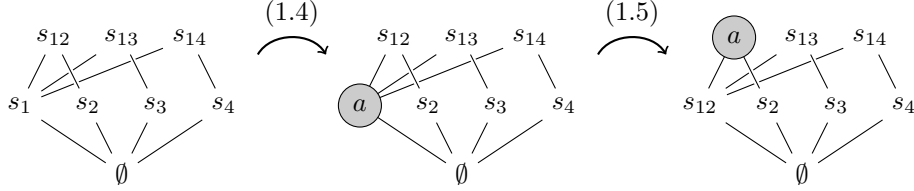
$$e(x, \mathbf{s}^{(2)}) = \begin{pmatrix} x+s_{12} & & +t_{12}+s_{13}+s_{14}, \\ x & +s_2 & +t_{12} \\ x & & +s_3 & +s_{13} \\ x & & +s_4 & +s_{14} \end{pmatrix}$$

agree in all but the second coordinate.

Repeating these two steps two more times produces a path of length four from $(x, \mathbf{s}^{(2)})$ to $(x, \mathbf{s}^{(6)}) = (x, (s_{14}, s_2, s_3, s_4, t_{12}, t_{13}, t_{14}))$, where t_{12}, t_{13}, t_{14} are arbitrary elements of S . Repeating a step similar to (1.4) four times produces a path of length four from $(x, \mathbf{s}^{(6)})$ to $(x, \mathbf{s}^{(10)}) = (x, (t_1, t_2, t_3, t_4, t_{12}, t_{13}, t_{14}))$, where $t_1, t_2, \dots, t_{14}$ are arbitrary.

We call this procedure ‘Boolean bubbling’ because it has the following interpretation. The coordinates of $\mathbf{s}$ are indexed by the bottom half of the Boolean lattice of rank four. Now the above procedure allows us to insert a new value

at the bottom of the Boolean lattice (1.4) and then ‘bubble’ it up the lattice (1.5). Repeating this procedure allows us to change the value of $\mathbf{s}$ arbitrarily.¹



Let us pause to talk about two details of this procedure that we have not mentioned yet. First, we have to make sure that each new value we ‘bubble’ up the lattice does not disturb previous values we have inserted. This is an easy problem to deal with: as long as we bubble up values to the top level before dealing with the lower level, future steps will not disturb previous ones. Second, recall that the coordinates of $\mathbf{s}$ must be distinct at every step of the process. This might be a problem if we wish to find a path from $(x, \mathbf{s})$ to $(x, \mathbf{t})$ where, for example, $s_4 = t_{12}$. The easiest way to get around this problem is to pick a new $\mathbf{u}$ whose coordinates are disjoint from both $\mathbf{s}$ and $\mathbf{t}$ and then use the Boolean bubbling procedure to construct a path from $(x, \mathbf{s})$ to $(x, \mathbf{u})$ and then to $(x, \mathbf{t})$.

Boolean bubbling allows us to construct a path in $\mathbf{G}_{4,t,S}$ from a vertex $(x, \mathbf{s})$ to any vertex of the form $(x, \mathbf{t})$. To walk to an arbitrary $(y, \mathbf{t})$, we only need one further ingredient.

Given a vertex $(x, \mathbf{t}) = (x, (t_1, t_2, t_3, t_4, t_{12}, t_{13}, t_{14}))$, this vertex is adjacent to $(x', \mathbf{t}') = (x + t_3 + t_{14}, (t_1, t_2, t_{14}, t_4, t_{12}, t_{13}, t_3))$ since

$$e(x, \mathbf{t}) = \begin{pmatrix} x+t_1 & +t_{12}+t_{13}+t_{14}, \\ x & +t_2 & +t_{12} \\ x & +t_3 & +t_{13} \\ x & +t_4 & +t_{14} \end{pmatrix}, \quad (1.6)$$

$$e(x', \mathbf{t}') = \begin{pmatrix} x + t_3 + t_{14} + t_1 & +t_{12}+t_{13}+t_3, \\ x + t_3 + t_{14} & +t_2 & +t_{12} \\ x + t_3 + t_{14} & +t_{14} & +t_{13} \\ x + t_3 + t_{14} & +t_4 & +t_3 \end{pmatrix}$$

agree in all but the second coordinate.

Now we claim that for any x, y which are adjacent in $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$, there is a path from $(x, \mathbf{s})$ to $(y, \mathbf{t})$ in $\mathbf{G}_{4,t,S}$. This follows using Boolean bubbling and (1.6). Write $y = x + a_1 + a_2$ for a_1, a_2 distinct elements of S . By Boolean bubbling, there is a path from $(x, \mathbf{s})$ to some $(x, \mathbf{u})$ where $u_3 = a_1$

¹Note that (1.4) and (1.5) are essentially the same operation. The latter switches two coordinates, s_I and $s_{I \sqcup \{j\}}$, while the former can be thought of as switching $s_\emptyset$ and s_j . Since $s_\emptyset$ does not affect the value of $e(x, \mathbf{s})$, its value can be changed arbitrarily. We take advantage of this to insert new values.

and $u_{14} = a_2$. Then $(x, \mathbf{u})$ is adjacent to $(y, \mathbf{u}')$ for some $\mathbf{u}'$ by (1.6). Finally, another application of Boolean bubbling gives a path from $(y, \mathbf{u}')$ to $(y, \mathbf{t})$. Since Assumption 3 easily implies that $\text{Cay}(\mathbb{F}_2^t, (S+S) \setminus \{0\})$ is also an expander, and in particular connected, this argument shows that $\mathbf{G}_{4,t,S}$ is connected as well.

Here is a convenient rephrasing of this argument which shows how we may extend it to prove that $\mathbf{G}_{4,t,S}$ is an expander. Define G_{Cay} to be the graph with vertex set $\mathbb{F}_2^t \times \mathcal{T}$ with an edge between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ if x, y are adjacent in $\text{Cay}(\mathbb{F}_2^t, (S+S) \setminus \{0\})$. We write $\mathbf{G}_{4,t,S}^M$ for the multigraph which has one edge between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ for each walk of length M between these vertices in $\mathbf{G}_{4,t,S}$. The argument above shows that for some M there is a copy of G_{Cay} contained in $\mathbf{G}_{4,t,S}^M$. Since G_{Cay} is connected, this proves that $\mathbf{G}_{4,t,S}$ is connected.

To prove that $\mathbf{G}_{4,t,S}$ is an expander we need a slightly stronger fact. Write $c \cdot G_{\text{Cay}}$ for the multigraph that has c edges for each edge of G_{Cay} . We need to prove that for some M and c the graphs $\mathbf{G}_{4,t,S}^M$ and $c \cdot G_{\text{Cay}}$ are approximately the same in some appropriate sense. Establishing this fact requires showing that for each $(x, \mathbf{s})$ and $(y, \mathbf{t})$ adjacent in G_{Cay} there are many paths between them in $\mathbf{G}_{4,t,S}$. This can be done with a small modification of (1.5) and (1.6).

Suppose we are in the process of Boolean bubbling, having inserted t_{12} at the bottom of the Boolean lattice. Let $(x, \mathbf{s}) = (x, (t_{12}, s_2, s_3, s_4, s_{12}, s_{13}, s_{14}))$ be the current state. We would normally proceed as in (1.5) to bubble t_{12} up the lattice. Instead, what we do is simultaneously bubble t_{12} up and insert an arbitrary element a in place of s_2 . Explicitly, if we set $(x, \mathbf{s}') = (x, (s_{12}, a, s_3, s_4, t_{12}, s_{13}, s_{14}))$, then $(x, \mathbf{s})$ and $(x, \mathbf{s}')$ are adjacent since

$$\begin{aligned}
e(x, \mathbf{s}) &= \begin{pmatrix} x+t_{12} & & +s_{12}+s_{13}+s_{14}, \\ x & +s_2 & +s_{12} \\ x & & +s_3 & +s_{13} \\ x & & +s_4 & +s_{14} \end{pmatrix}, \\
e(x, \mathbf{s}') &= \begin{pmatrix} x+s_{12} & & +t_{12}+s_{13}+s_{14}, \\ x & +a & +t_{12} \\ x & & +s_3 & +s_{13} \\ x & & +s_4 & +s_{14} \end{pmatrix},
\end{aligned} \tag{1.5'}$$

still agree in all but the second coordinate. Replacing every move of type (1.5) with one of type (1.5') in the Boolean bubbling procedure gives not one path of length M from $(x, \mathbf{s})$ to $(x, \mathbf{t})$, but instead on the order of $|S|^{M-7}$ such paths. Here the exponent $M-7$ comes from the fact that our modified Boolean bubbling procedure gives on the order of $|S|^M$ paths of length M starting at $(x, \mathbf{s})$ and ending at a vertex of the form $(x, \mathbf{t})$, while the number of choices for $\mathbf{t}$ is on the order of $|S|^7$. The full details of this procedure in the general r -uniform construction are given in Section 2.2.

Similarly, we upgrade the moves of type (1.6) by inserting an arbitrary element a at the bottom of the lattice. In particular, suppose we are given a vertex $(x, \mathbf{t}) = (x, (t_1, t_2, t_3, t_4, t_{12}, t_{13}, t_{14}))$. This vertex is adjacent to $(x', \mathbf{t}') =$

$(x + t_3 + t_{14}, (t_1, a, t_{14}, t_4, t_{12}, t_{13}, t_3))$ since

$$e(x, \mathbf{t}) = \begin{pmatrix} x+t_1 & & +t_{12}+t_{13}+t_{14}, \\ x & +t_2 & +t_{12} \\ x & & +t_3 & +t_{13} \\ x & & +t_4 & +t_{14} \end{pmatrix}, \quad (1.6')$$

$$e(x', \mathbf{t}') = \begin{pmatrix} x + t_3 + t_{14} + t_1 & & +t_{12}+t_{13}+t_3, \\ x + t_3 + t_{14} & +a & +t_{12} \\ x + t_3 + t_{14} & & +t_{14} & +t_{13} \\ x + t_3 + t_{14} & & +t_4 & +t_3 \end{pmatrix}$$

still agree in all but the second coordinate. This is the only additional ingredient that we need to deduce that $\mathbf{G}_{4,t,S}$ is an expander from the fact that G_{Cay} is an expander.

To complete the proof, all that remains is to deduce the corresponding result for $G_{\text{walk}}(\mathbf{H}_{4,t,S})$. This will be immediate once we write down exactly in what sense $\mathbf{G}_{4,t,S}$ and $G_{\text{walk}}(\mathbf{H}_{4,t,S})$ are dual to each other. We refer the reader to Section 2.3 for the details.

2 Proof of main theorem

2.1 Degree properties

We start by proving a couple of basic properties of our hypergraph $\mathbf{H}_{r,t,S}$. We find the number of representations each r -edge has of the form $e(x, \mathbf{s})$ and compute the number of r -edges of $\mathbf{H}_{r,t,S}$ containing a given $(r-1)$ -edge.

Lemma 2.1. *Recall $\mathcal{T}$ from (1.2). For r, t, S satisfying Assumption 2, the ordered r -tuples $e(x, \mathbf{s})$ for $x \in \mathbb{F}_2^t$ and $\mathbf{s} \in \mathcal{T}$ are all distinct.*

Proof. Suppose $e(x, \mathbf{s}) = e(x', \mathbf{s}')$ for $x, x' \in \mathbb{F}_2^t$ and $\mathbf{s}, \mathbf{s}' \in \mathcal{T}$. Then this implies that

$$e_i(x, \mathbf{s}) + e_j(x, \mathbf{s}) = e_i(x', \mathbf{s}') + e_j(x', \mathbf{s}')$$

for each $1 \leq i < j \leq r$. We rewrite this equation as

$$\sum_{\substack{I \in \mathcal{P} \\ |\{i,j\} \cap I| = 1}} s_I = \sum_{\substack{I \in \mathcal{P} \\ |\{i,j\} \cap I| = 1}} s'_I.$$

By Assumption 2 on S , this implies that

$$\{s_I : I \in \mathcal{P} \text{ with } |\{i,j\} \cap I| = 1\} = \{s'_I : I \in \mathcal{P} \text{ with } |\{i,j\} \cap I| = 1\} \quad (2.1)$$

for each $1 \leq i < j \leq r$.

We claim that the above equation implies that $\mathbf{s} = \mathbf{s}'$. To see this, fix $I \subseteq [r]$ and consider the pairs $(i, j) \in [r]^2$ such that $i \neq j$ and $|\{i, j\} \cap I| = 1$. This

condition is satisfied if and only if $(i, j) \in (I \times I^c) \cup (I^c \times I)$. Note that the sets $(I \times I^c) \cup (I^c \times I)$ are distinct for all $I \in \mathcal{P}$ and are non-empty. This is true simply because $\emptyset, [r] \notin \mathcal{P}$ and if $I \in \mathcal{P}$ then one has $I^c \notin \mathcal{P}$.

Thus for $I \in \mathcal{P}$ the element s_I is present on the left-hand side of (2.1) if and only if $(i, j) \in (I \times I^c) \cup (I^c \times I)$ and no other s_J has the same property. Furthermore, s'_I is present on the right-hand side of (2.1) if and only if $(i, j) \in (I \times I^c) \cup (I^c \times I)$ and no other s'_J has the same property. Since the left- and right-hand sides of (2.1) are equal for all $1 \leq i < j \leq r$, this implies that $s_I = s'_I$ for all $I \in \mathcal{P}$.

To complete the proof, note that

$$x = e_1(x, \mathbf{s}) + \sum_{\substack{I \in \mathcal{P} \\ 1 \in I}} s_I = e_1(x', \mathbf{s}') + \sum_{\substack{I \in \mathcal{P} \\ 1 \in I}} s'_I = x',$$

as required. ■

This implies that given an r -edge $e = \{v_1, \dots, v_r\}$ of $H_{r,t,S}$ there are at most $r!$ possible $e(x, \mathbf{s})$ that it corresponds to, one for each ordering of its vertices. We claim that this is an equality. This is immediate for r odd. Suppose $e(x, \mathbf{s}) = (v_1, \dots, v_r)$. For a permutation $\pi \in \mathfrak{S}_r$ define $\mathbf{s}^\pi$ by $s_I^\pi = s_{\pi(I)}$ where $\pi(I) = \{\pi(i) : i \in I\}$. Then $e(x, \mathbf{s}^\pi) = (v_{\pi(1)}, \dots, v_{\pi(r)})$. This argument does not work for r even since for some $I \in \mathcal{P}$, there are π such that $\pi(I) \notin \mathcal{P}$. However, the same result follows easily from the next lemma.

Lemma 2.2. *Fix r even and $S \subset \mathbb{F}_2^t$. Recall $\mathcal{T}$ from (1.2). For each $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$, there exists $(y, \mathbf{t}) \in \mathbb{F}_2^t \times \mathcal{T}$ such that $e(y, \mathbf{t})$ is the same as $e(x, \mathbf{s})$ but with the first two coordinates swapped.*

Proof. Let π be the transposition that swaps 1 and 2. Write $\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \mathcal{P}_3$, where $\mathcal{P}_1 = \{I \subset [r] : 0 < |I| < r/2\}$, $\mathcal{P}_2 = \{I \subset [r] : |I| = r/2 \text{ and } 1, 2 \in I\}$ and $\mathcal{P}_3 = \{I \subset [r] : |I| = r/2, 1 \in I \text{ and } 2 \notin I\}$. Note that for $I \in \mathcal{P}_1$ one has $\pi(I) \in \mathcal{P}_1$, while for $I \in \mathcal{P}_2$ one has $\pi(I) = I \in \mathcal{P}_2$. Finally, note that for $I \in \mathcal{P}_3$ one has $\pi(I^c) \in \mathcal{P}_3$.

Now let

$$t_I = \begin{cases} s_{\pi(I)} & \text{if } I \in \mathcal{P}_1 \cup \mathcal{P}_2, \\ s_{\pi(I^c)} & \text{if } I \in \mathcal{P}_3 \end{cases}$$

and

$$y = x + \sum_{I \in \mathcal{P}_3} s_I.$$

We claim that $e(y, \mathbf{t})$ has the desired property. We refer the reader to (1.3) for the $r = 4$ case. For the general case, note that for $I \in \mathcal{P}_1 \cup \mathcal{P}_2$ one has $i \in I$ if and only if $\pi(i) \in \pi(I)$. Moreover, for $I \in \mathcal{P}_3$ one has $i \in I$ if and only if

$\pi(i) \notin \pi(I^c)$. Therefore,

$$\begin{aligned}
e_i(y, \mathbf{t}) &= x + \sum_{I \in \mathcal{P}_3} s_I + \sum_{\substack{I \in \mathcal{P}_1 \cup \mathcal{P}_2 \\ i \in I}} s_{\pi(I)} + \sum_{\substack{I \in \mathcal{P}_3 \\ i \in I}} s_{\pi(I^c)} \\
&= x + \sum_{I \in \mathcal{P}_3} s_I + \sum_{\substack{I \in \mathcal{P}_1 \cup \mathcal{P}_2 \\ \pi(i) \in I}} s_I + \sum_{\substack{I \in \mathcal{P}_3 \\ \pi(i) \notin I}} s_I \\
&= x + \sum_{\substack{I \in \mathcal{P}_3 \\ \pi(i) \in I}} s_I + \sum_{\substack{I \in \mathcal{P}_1 \cup \mathcal{P}_2 \\ \pi(i) \in I}} s_I \\
&= e_{\pi(i)}(x, \mathbf{s}). \quad \blacksquare
\end{aligned}$$

Next we use a similar technique to the proof of Lemma 2.1 to compute the degree of each $(r-1)$ -edge in the hypergraph $\mathbf{H}_{r,t,S}$. Recall that the degree of an $(r-1)$ -edge is the number of r -edges that it is contained in. Note that in light of Lemma 2.2 and the discussion preceding it, the degree of the $(r-1)$ -edge formed by the set of all coordinates of $e(x, \mathbf{s})$ but the k -th is exactly the number of $(x', \mathbf{s}') \in \mathbb{F}_2^t \times \mathcal{T}$ such that $e(x, \mathbf{s})$ and $e(x', \mathbf{s}')$ agree in all but the k -th coordinate.

Lemma 2.3. *Recall $\mathcal{T}$ from (1.2). For r, t, S satisfying Assumption 2, $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$ and $1 \leq k \leq r$, there are exactly*

$$(|S| - (2^{r-1} - 2)) 2^{2^{r-2}-1}$$

pairs $(x', \mathbf{s}') \in \mathbb{F}_2^t \times \mathcal{T}$ such that $e(x, \mathbf{s})$ and $e(x', \mathbf{s}')$ agree in all but (possibly) the k -th coordinate.

Proof. First we construct the desired number of pairs $(x', \mathbf{s}')$. To begin, note that for each $I \in \mathcal{P}$ exactly one of the following three statements holds:

1. there exists a unique $J \in \mathcal{P}$ such that either $J = I \sqcup \{k\}$ or $I = J \sqcup \{k\}$;²
2. there exists $J \in \mathcal{P}$ such that $I \sqcup J = [r] \setminus \{k\}$;
3. $I = \{k\}$.

To see this, note that if $k \in I$, then we can let $J = I \setminus \{k\}$ except in the case $I = \{k\}$. On the other hand, if $k \notin I$, we can let $J = I \sqcup \{k\}$ unless $|I| = \lfloor r/2 \rfloor$ or r is even, $k \neq 1$ and $|I| = r/2 - 1$ with $1 \notin I$. But in both these cases, we can take J to be the complement of I in $[r] \setminus \{k\}$.

For $I \in \mathcal{P}$ satisfying condition 1 and $J \in \mathcal{P}$ the unique set satisfying either $J = I \sqcup \{k\}$ or $I = J \sqcup \{k\}$, define $(x', \mathbf{s}')$ by

$$x' = x$$

²We use the notation $J = I \sqcup \{k\}$ to imply that the sets I and $\{k\}$ are disjoint and that J is their union.

and

$$s'_K = \begin{cases} s_J & \text{if } K = I, \\ s_I & \text{if } K = J, \\ s_K & \text{otherwise.} \end{cases}$$

Note that $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for $i \neq k$.

For $I \in \mathcal{P}$ satisfying condition 2 and $J \in \mathcal{P}$ given by $I \sqcup J = [r] \setminus \{k\}$, define $(x', \mathbf{s}')$ by

$$x' = x + s_I + s_J$$

and

$$s'_K = \begin{cases} s_J & \text{if } K = I, \\ s_I & \text{if } K = J, \\ s_K & \text{otherwise.} \end{cases}$$

Note that $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for $i \neq k$. For example, suppose $i \in I$. Then

$$\begin{aligned} e_i(x', \mathbf{s}') &= x + s_I + s_J + \sum_{\substack{K \in \mathcal{P} \\ i \in K}} s'_K \\ &= x + s_I + s_J + s'_I + \sum_{\substack{K \in \mathcal{P} \\ i \in K \text{ and } K \neq I}} s'_K \\ &= x + s_I + \sum_{\substack{K \in \mathcal{P} \\ i \in K \text{ and } K \neq I}} s_K \\ &= e_i(x, \mathbf{s}). \end{aligned}$$

A similar computation covers the complementary case $i \in J$.

For $I = \{k\}$ and any $a \in S$ such that $a \neq s_K$ for all $K \in \mathcal{P} \setminus \{\{k\}\}$, define $(x', \mathbf{s}')$ by

$$x' = x$$

and

$$s'_K = \begin{cases} a & \text{if } K = I, \\ s_K & \text{otherwise.} \end{cases}$$

Note that $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for $i \neq k$.

Note that $\mathcal{P} \setminus \{\{k\}\}$ is partitioned into $2^{r-2} - 1$ disjoint pairs $\{I, J\}$ where either $J = I \sqcup \{k\}$, $I = J \sqcup \{k\}$ or $I \sqcup J = [r] \setminus \{k\}$. Using a combination of the three operations described above, we can find $(|S| - (2^{r-1} - 2)) 2^{2^{r-2}-1}$ pairs $(x', \mathbf{s}') \in \mathbb{F}_2^t \times \mathcal{T}$ satisfying $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for $i \neq k$: we choose whether or not to swap each of the $2^{r-2} - 1$ pairs $\{I, J\}$ and there are $|S| - (2^{r-1} - 2)$ choices for the value of $s'_{\{k\}}$.

Now we prove that these are the only $(x', \mathbf{s}')$ with the desired property. Suppose $(x', \mathbf{s}') \in \mathbb{F}_2^t \times \mathcal{T}$ satisfies $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for $i \neq k$. As in the proof of Lemma 2.1, this implies

$$\{s_I : I \in \mathcal{P} \text{ with } |\{i, j\} \cap I| = 1\} = \{s'_I : I \in \mathcal{P} \text{ with } |\{i, j\} \cap I| = 1\} \quad (2.2)$$

for $1 \leq i < j \leq r$ with $i \neq k$ and $j \neq k$.
 For $I \in \mathcal{P}$, define $A_I \subset ([r] \setminus \{k\})^2$ by

$$A_I = (I \setminus \{k\}) \times (I^c \setminus \{k\}) \cup (I^c \setminus \{k\}) \times (I \setminus \{k\}).$$

Note that s_I appears on the left-hand side of (2.2) if and only if $(i, j) \in A_I$. Note that $A_I \neq \emptyset$ for $I \neq \{k\}$. Now partition $\mathcal{P} \setminus \{\{k\}\}$ into equivalence classes under the relation $I \sim J$ if $A_I = A_J$. The same argument as in the proof of Lemma 2.1 then implies that for any equivalence class $\Pi \subset \mathcal{P} \setminus \{\{k\}\}$, we have

$$\{s_I\}_{I \in \Pi} = \{s'_I\}_{I \in \Pi}.$$

Now it is easy to see that these equivalence classes are exactly the pairs $\{I, J\}$ described earlier with $J = I \sqcup \{k\}$, $I = J \sqcup \{k\}$ or $I \sqcup J = [r] \setminus \{k\}$. Therefore, if $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for $i \neq k$, then $\mathbf{s}'$ must be one of the $(|S| - (2^{r-1} - 2))2^{2^{r-2}-1}$ elements that we have already found. Finally, note that for fixed i and $x, \mathbf{s}, \mathbf{s}'$, there is a unique x' satisfying the equation $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$. Therefore, there are exactly the desired number of pairs $(x', \mathbf{s}')$. ■

2.2 Expansion properties

Given $r \geq 3$, t and $S \subset \mathbb{F}_2^t$, write $\mathbf{G}_{r,t,S}$ for the graph with vertex set $\mathbb{F}_2^t \times \mathcal{T}$ and with an edge between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ if and only if the ordered r -tuples $e(x, \mathbf{s})$ and $e(y, \mathbf{t})$ differ in exactly one coordinate.

For a graph G , we use the notation G^M to refer to the multigraph with the same vertex set as G whose edges between u and v correspond to the M -step walks in G from u to v . We will prove that $\mathbf{G}_{r,t,S}$ is an expander under Assumptions 1, 2, 3 by first showing that there is some constant $M = M(r)$ such that $\mathbf{G}_{r,t,S}^M$ contains an expander as a dense subgraph. The specific expander we will find is a multigraph on the vertex set $\mathbb{F}_2^t \times \mathcal{T}$ with c edges between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ whenever xy is an edge in $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$. In the next section, we will show that this result implies that $\mathbf{G}_{r,t,S}$ is an expander.

We prove the claimed combinatorial properties of $\mathbf{G}_{r,t,S}$ in three stages. First we show how to change the value of s_I for some specific I without changing x or too many of the other s_J 's. Then we iterate this to walk from $(x, \mathbf{s})$ to $(x, \mathbf{t})$ for arbitrary $\mathbf{t}$. Finally, we combine this walk with our second type of move to walk from $(x, \mathbf{s})$ to $(y, \mathbf{t})$ for y an arbitrary neighbor of x in the Cayley graph $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ and $\mathbf{t}$ arbitrary.

Lemma 2.4. *Given r, t, S satisfying Assumption 1, there is a constant $c = c(r) > 0$ such that for each $I \in \mathcal{P}$, each $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$ and each $a \in S$ distinct from all coordinates of $\mathbf{s}$, there are at least $c|S|^{|I|-1}$ walks of length $|I|$ in $\mathbf{G}_{r,t,S}$ which begin at $(x, \mathbf{s})$ and end at an element of the form $(x, \mathbf{s}')$, where $s'_I = a$ and $s_J = s'_J$ for all $J \in \mathcal{P}$ with $J \not\subseteq I$.*

Proof. Write $I = \{k_1, \dots, k_l\}$. Then define $I_i = \{k_1, \dots, k_i\}$. To walk from $(x, \mathbf{s})$ to a vertex of the form $(x, \mathbf{s}')$ in $\mathbf{G}_{r,t,S}$ we insert a at I_1 and then bubble it

up to $I_2, I_3, \dots, I_l = I$. This finds one such path; to find many paths we insert arbitrary values $a_2, \dots, a_l$ at the bottom of the lattice on all steps but the first.

Fix $a_2, \dots, a_l$. Write $\mathbf{s}^0 = \mathbf{s}$ and define $\mathbf{s}^1$ by

$$s_I^1 = \begin{cases} a & \text{if } I = I_1 = \{k_1\}, \\ s_I & \text{otherwise.} \end{cases} \quad (2.3)$$

Inductively define $\mathbf{s}^2, \dots, \mathbf{s}^l$ by

$$s_I^i = \begin{cases} s_{I_{i-1}}^{i-1} = a & \text{if } I = I_i, \\ s_{I_i}^{i-1} & \text{if } I = I_{i-1}, \\ a_i & \text{if } I = \{k_i\}, \\ s_I^{i-1} & \text{otherwise.} \end{cases} \quad (2.4)$$

We claim that $e(x, \mathbf{s}^{i-1})$ and $e(x, \mathbf{s}^i)$ agree everywhere except for the k_i -th coordinate. For $i = 1$, the $r = 4$ case is given in (1.4) and for $2 \leq i \leq l$, the $r = 4$ case is given in (1.5'). Consider

$$e_j(x, \mathbf{s}^1) = x + \sum_{\substack{I \in \mathcal{P} \\ j \in I}} s_I^1.$$

For $j \neq k_1$, the fact that $j \in I$ implies that $I \neq \{k_1\}$, so $s_I^1 = s_I^0$. Thus, $e(x, \mathbf{s}^0)$ and $e(x, \mathbf{s}^1)$ agree in all but the k_1 -th coordinate. Similarly, consider

$$e_j(x, \mathbf{s}^i) = x + \sum_{\substack{I \in \mathcal{P} \\ j \in I}} s_I^i.$$

For $j \neq k_1, \dots, k_i$, the fact that $j \in I$ implies that $s_I^i = s_I^{i-1}$. Furthermore, for $j = k_1, \dots, k_{i-1}$, the sum can be written as

$$\begin{aligned} e_j(x, \mathbf{s}^i) &= x + s_{I_{i-1}}^i + s_{I_i}^i + \sum_{\substack{I \in \mathcal{P} \\ j \in I \text{ and } I \neq I_{i-1}, I_i}} s_I^i \\ &= x + s_{I_i}^{i-1} + s_{I_{i-1}}^{i-1} + \sum_{\substack{I \in \mathcal{P} \\ j \in I \text{ and } I \neq I_{i-1}, I_i}} s_I^{i-1} \\ &= e_j(x, \mathbf{s}^{i-1}). \end{aligned}$$

Thus, $e(x, \mathbf{s}^{i-1})$ and $e(x, \mathbf{s}^i)$ agree in all but the k_i -th coordinate.

Finally, note that as long as $a_2, \dots, a_l$ are distinct from one another and from a and all the coordinates of $\mathbf{s}$, then the tuples $\mathbf{s}^1, \dots, \mathbf{s}^l$ all lie in $\mathcal{T}$. Therefore, we have found a walk of length $l = |I|$ from $e(x, \mathbf{s})$ to $e(x, \mathbf{s}')$ for each valid sequence $a_2, \dots, a_{|I|}$, of which there are $(|S| - 2^{r-1})(|S| - (2^{r-1} + 1)) \cdots (|S| - (2^{r-1} + |I| - 2)) \geq c|S|^{|I|-1}$ by Assumption 1. $\blacksquare$

Remark 2.5. Suppose we are in the setup of Lemma 2.4, but we are also given a set X not containing a or any of the coordinates of $\mathbf{s}$. Suppose we wish to count walks of length $|I|$ from $(x, \mathbf{s})$ to an element of the form $(x, \mathbf{s}')$ such that $s'_I = a$ and $s_J = s'_J$ for all $J \in \mathcal{P}$ with $J \not\subseteq I$ as before, but now with the additional restriction that no coordinate of $\mathbf{s}'$ can lie in X . The argument above still covers this case, giving one walk for each choice of $a_2, \dots, a_{|I|}$ which are distinct from one another and from a , all coordinates of $\mathbf{s}$ and all elements of X . Thus, there are at least $(|S| - (2^{r-1} + |X|))(|S| - (2^{r-1} + |X| + 1)) \cdots (|S| - (2^{r-1} + |X| + |I| - 2))$ walks. For X small enough, say $|X| \leq 2^r$, Assumption 1 still guarantees at least $c|S|^{|I|-1}$ such walks.

Lemma 2.6. *For r, t, S satisfying Assumption 1, there exist constants $c = c(r) > 0$ and $M' = M'(r)$ such that for $x \in \mathbb{F}_2^t$ and $\mathbf{s}, \mathbf{t} \in \mathcal{T}$, there are at least $c|S|^{M' - |\mathcal{P}|}$ walks of length M' from $(x, \mathbf{s})$ to $(x, \mathbf{t})$ in $\mathbf{G}_{r,t,S}$.*

Proof. Fix any $\mathbf{u} \in \mathcal{T}$ whose set of coordinates are disjoint from those of $\mathbf{s}$ and $\mathbf{t}$. Fix an ordering $I_1, I_2, \dots, I_{|\mathcal{P}|}$ of $\mathcal{P}$ so that $I_i \not\subseteq I_j$ for $i < j$. We use the notation

$$u \overset{\ell}{\rightsquigarrow} v$$

to refer to an ℓ -step walk in $\mathbf{G}_{r,t,S}$ from u to v .

Our goal is to walk from $(x, \mathbf{s})$ to $(x, \mathbf{u})$ by bubbling u_{I_1} up to I_1 , then bubbling u_{I_2} up to I_2 , and so on. Lemma 2.4 lets us perform each of these individual steps as long as u_{I_i} , the element we wish to bubble up, is distinct from all coordinates of the current vertex. Moreover, our choice of ordering means that each successive bubbling does not disturb the previous ones. Thus, it suffices to find many walks of the form

$$(x, \mathbf{s}) \overset{|I_1|}{\rightsquigarrow} (x, \mathbf{u}^1) \overset{|I_2|}{\rightsquigarrow} (x, \mathbf{u}^2) \rightsquigarrow \cdots \rightsquigarrow (x, \mathbf{u}),$$

where $\mathbf{u}^i \in \mathcal{T}$ satisfies $u_{I_j}^i = u_{I_j}$ for all $1 \leq j \leq i$ and no coordinate of $\mathbf{u}^i$ is equal to u_{I_j} for $j > i$.

For each $1 \leq i \leq |\mathcal{P}|$, note that if u_{I_i} is distinct from all coordinates of $\mathbf{u}^{i-1}$, Lemma 2.4 guarantees that the number of walks of the form

$$(x, \mathbf{u}^{i-1}) \overset{|I_i|}{\rightsquigarrow} (x, \mathbf{u}^i)$$

is $\Omega(|S|^{|I_i|-1})$.³ Furthermore, Remark 2.5 implies that we can still find this many walks with no coordinate $\mathbf{u}^i$ equal to u_{I_j} for $j > i$.

Multiplying these together, we see that there are $\Omega(|S|^{M'' - |\mathcal{P}|})$ walks of length M'' from $(x, \mathbf{s})$ to $(x, \mathbf{u})$, where $M'' = \sum_{I \in \mathcal{P}} |I|$. Setting $M' = 2M''$, we conclude that there are $\Omega(|S|^{M' - 2|\mathcal{P}|})$ walks of the form

$$(x, \mathbf{s}) \overset{M''}{\rightsquigarrow} (x, \mathbf{u}) \overset{M''}{\rightsquigarrow} (x, \mathbf{t})$$

³We use the notation $\Omega(f)$ to denote a quantity of size at least cf where $c = c(r) > 0$ is a positive constant that only depends on the uniformity r .

for each $\mathbf{u}$ whose coordinates are distinct from the coordinates of both $\mathbf{s}$ and $\mathbf{t}$. Since there are at least $\Omega(|S|^{|\mathcal{P}|})$ choices for $\mathbf{u}$, summing over them gives the desired result. $\blacksquare$

Lemma 2.7. *For r, t, S satisfying Assumption 1, there exist constants $c = c(r) > 0$ and $M = M(r)$ such that, given x adjacent to y in $\text{Cay}(\mathbb{F}_2^t, (S+S) \setminus \{0\})$ and $\mathbf{s}, \mathbf{t} \in \mathcal{T}$, there are at least $c|S|^{M-|\mathcal{P}|-2}$ walks of length M from $(x, \mathbf{s})$ to $(y, \mathbf{t})$ in $\mathbf{G}_{r,t,S}$.*

Proof. Define

$$I_1 = \left\{3, 4, \dots, \left\lceil \frac{r}{2} \right\rceil + 1\right\}$$

and

$$I_2 = \left\{1, \left\lceil \frac{r}{2} \right\rceil + 2, \left\lceil \frac{r}{2} \right\rceil + 3, \dots, r\right\}.$$

The relevant properties of these two sets are that $I_1, I_2 \in \mathcal{P}$, they are disjoint and $I_1 \cup I_2 = [r] \setminus \{2\}$.

Since x is adjacent to y in $\text{Cay}(\mathbb{F}_2^t, (S+S) \setminus \{0\})$, we can write $y = x + a_1 + a_2$ for $a_1, a_2 \in S$. Suppose now that $\mathbf{u} \in \mathcal{T}$ is such that $u_{I_1} = a_1$ and $u_{I_2} = a_2$ and pick $b \in S$ distinct from all coordinates of $\mathbf{u}$. Define $\mathbf{u}'$ by

$$u'_I = \begin{cases} u_{I_2} = a_2 & \text{if } I = I_1, \\ u_{I_1} = a_1 & \text{if } I = I_2, \\ b & \text{if } I = \{2\}, \\ u_I & \text{otherwise.} \end{cases} \quad (2.5)$$

We claim that there are $\Omega(|S|^{2M'-2|\mathcal{P}|})$ walks of the form

$$(x, \mathbf{s}) \xrightarrow{M'} (x, \mathbf{u}) \xrightarrow{1} (y, \mathbf{u}') \xrightarrow{M'} (y, \mathbf{t}).$$

By Lemma 2.6, there are $\Omega(|S|^{M'-|\mathcal{P}|})$ walks of length M' from $(x, \mathbf{s})$ to $(x, \mathbf{u})$ and also from $(y, \mathbf{u}')$ to $(y, \mathbf{t})$. All that remains to be checked is that $(x, \mathbf{u})$ and $(y, \mathbf{u}')$ are adjacent in $\mathbf{G}_{r,t,S}$. We refer the reader to (1.6') for the $r = 4$ example. Now

$$e_i(y, \mathbf{u}') = x + a_1 + a_2 + \sum_{\substack{I \in \mathcal{P} \\ i \in I}} u'_I.$$

For $i \in I_1$, this can be written as

$$\begin{aligned} e_i(y, \mathbf{u}') &= x + a_1 + a_2 + u'_{I_1} + \sum_{\substack{I \in \mathcal{P} \\ i \in I \text{ and } I \neq I_1}} u'_I \\ &= x + a_1 + \sum_{\substack{I \in \mathcal{P} \\ i \in I \text{ and } I \neq I_1}} u_I \\ &= e_i(x, \mathbf{u}). \end{aligned}$$

A similar computation shows that $e_i(x, \mathbf{u}) = e_i(y, \mathbf{u}')$ for $i \in I_2$. Therefore, $e(x, \mathbf{u})$ and $e(y, \mathbf{u}')$ agree in all but the second coordinate.

Finally, note that there are $\Omega(|S|^{|\mathcal{P}|-1})$ choices for the pair $(\mathbf{u}, \mathbf{u}')$ given the constraints that $u_{I_1} = a_1$, $u_{I_2} = a_2$ and $\mathbf{u}$ and $\mathbf{u}'$ are related by (2.5). Setting $M = 2M' + 1$ and summing over all choices of $\mathbf{u}, \mathbf{u}'$ with the above properties, we see that there are $\Omega(|S|^{M-|\mathcal{P}|-2})$ walks of length M from $(x, \mathbf{s})$ to $(y, \mathbf{t})$. ■

2.3 Spectral properties

Lemma 2.8. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3, there exists a constant $c = c(r) > 0$ such that $\mathbf{G}_{r,t,S}$ is a $c\epsilon$ -expander.*

Proof. This follows from Lemma 2.7 in three steps.

First we claim that Assumption 3 implies that $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ is an ϵ -expander. Note that

$$\mathbf{A}_{\text{Cay}(\mathbb{F}_2^t, (S+S) \setminus \{0\})} = \frac{1}{2} \left(\mathbf{A}_{\text{Cay}(\mathbb{F}_2^t, S)}^2 - |S| \mathbf{I} \right).$$

To see this equation, note that $\text{Cay}(\mathbb{F}_2^t, S)^2$ has an edge from x to $x + s_1 + s_2$ for each pair $s_1, s_2 \in S$. This differs from $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ in that the pairs with $s_1 = s_2$ contribute $|S|$ loops at each vertex and every other edge is double-counted since $x + s_1 + s_2 = x + s_2 + s_1$.

Now, by Assumption 3 and the above equation, it follows that the non-trivial eigenvalues of $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ lie in $[-\frac{1}{2}|S|, \frac{1}{2}((1 - \epsilon)|S|^2 - |S|)]$. This interval is contained within $[-\frac{1}{2}(1 - \epsilon)(|S|^2 - |S|), \frac{1}{2}(1 - \epsilon)(|S|^2 - |S|)]$, proving that $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ is an ϵ -expander.

Define the multigraph G_{Cay} to have vertex set $\mathbb{F}_2^t \times \mathcal{T}$ and $c|S|^{M-|\mathcal{P}|-2}$ edges between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ whenever xy is an edge in $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$. Here c, M are the same as in the statement of Lemma 2.7. Note that G_{Cay} is regular of degree $d_1 \geq c'|S|^M$ for some $c' = c'(r) > 0$. We claim that G_{Cay} is an ϵ -expander. To see this, note that G_{Cay} is the tensor product of $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ with another graph, namely, the complete graph on $|\mathcal{T}|$ vertices with self-loops where every edge has multiplicity $c|S|^{M-|\mathcal{P}|-2}$. Then the fact that $\text{Cay}(\mathbb{F}_2^t, (S + S) \setminus \{0\})$ is an ϵ -expander implies the same is true of G_{Cay} .

Second, note that Lemma 2.3 implies that $\mathbf{G}_{r,t,S}^M$ is regular of degree $d_2 \leq c''|S|^M$ for some $c'' = c''(r)$. By Lemma 2.7, we know that G_{Cay} is a subgraph of $\mathbf{G}_{r,t,S}^M$. Thus,

$$\mathbf{A}_{\mathbf{G}_{r,t,S}^M} = \mathbf{A}_{G_{\text{Cay}}} + \mathbf{A}_{\mathbf{G}_{r,t,S}^M \setminus G_{\text{Cay}}}.$$

Here $\mathbf{G}_{r,t,S}^M \setminus G_{\text{Cay}}$ is a $(d_2 - d_1)$ -regular graph, which implies that

$$\lambda(\mathbf{G}_{r,t,S}^M) \leq \lambda(G_{\text{Cay}}) + (d_2 - d_1) \leq (1 - \epsilon) d_1 + (d_2 - d_1) = \left(1 - \frac{d_1}{d_2} \epsilon\right) d_2.$$

Since $d_1 \geq c'|S|^M$ and $d_2 \leq c''|S|^M$, we conclude that $\mathbf{G}_{r,t,S}^M$ is a $\frac{c'}{c''}\epsilon$ -expander.

Third, the equation

$$\mathbf{A}_{\mathbf{G}_{r,t,S}^M} = \mathbf{A}_{\mathbf{G}_{r,t,S}}^M$$

implies that

$$\lambda(\mathbf{G}_{r,t,S}) \leq \left(1 - \frac{c'}{c''M}\epsilon\right)d,$$

where d is the degree of $\mathbf{G}_{r,t,S}$. Since c', c'', M are all constants that depend only on r , this proves the desired result. $\blacksquare$

Finally, we use this result to prove the main theorem.

Theorem 1.1. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3, there exists a constant $c = c(r) > 0$ such that the walk graph $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ is a $c\epsilon$ -expander.*

Proof. In this proof we consider three graphs: $\mathbf{G}_{r,t,S}$, $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ and a third graph that we will call $G'_{\text{walk}}(\mathbf{H}_{r,t,S})$. The graph $\mathbf{G}_{r,t,S}$ was defined in the previous section; its vertices are pairs $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$. The graph $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ is the walk graph whose vertices are $(r-1)$ -edges of $\mathbf{H}_{r,t,S}$ with an edge ef if there is some common r -edge that contains both e and f . We define the graph $G'_{\text{walk}}(\mathbf{H}_{r,t,S})$ dually to have vertices the r -edges of $\mathbf{H}_{r,t,S}$ with an edge ef if there is some common $(r-1)$ -edge contained in both e, f . By Lemma 2.3, we know that there is some D such that every $(r-1)$ -edge of $\mathbf{H}_{r,t,S}$ is contained in exactly D of the r -edges. Then $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ is $(r-1)D$ -regular and $G'_{\text{walk}}(\mathbf{H}_{r,t,S})$ is $r(D-1)$ -regular.

Define a map $\pi: V(\mathbf{G}_{r,t,S}) \rightarrow V(G'_{\text{walk}}(\mathbf{H}_{r,t,S}))$ that sends $(x, \mathbf{s})$ to the unordered underlying set of $e(x, \mathbf{s})$. Then this map defines a graph homomorphism $\pi: \mathbf{G}_{r,t,S} \rightarrow G'_{\text{walk}}(\mathbf{H}_{r,t,S})$. By Lemma 2.2, these two graphs have the same degree and π is r -to-1 on both the vertex- and edge-sets. We claim that this implies that $\lambda(G'_{\text{walk}}(\mathbf{H}_{r,t,S})) \leq \lambda(\mathbf{G}_{r,t,S})$. For any eigenvector $\vec{v} \in \mathbb{R}^{V(G'_{\text{walk}}(\mathbf{H}_{r,t,S}))}$, we can lift $\vec{v}$ to a vector $\pi^{-1}\vec{v} \in \mathbb{R}^{V(\mathbf{G}_{r,t,S})}$ that is constant on the fibers of π . The above properties of π easily imply that $\pi^{-1}\vec{v}$ is an eigenvector of $\mathbf{G}_{r,t,S}$ with the same eigenvalue as $\vec{v}$. Thus, by Lemma 2.8, we conclude that $\lambda(G'_{\text{walk}}(\mathbf{H}_{r,t,S})) \leq (1 - c\epsilon)r(D-1)$ for some $c = c(r)$ satisfying $0 < c < 1/2$. (If the value of c which comes out of Lemma 2.8 is not less than $1/2$, we can reduce it to $1/2$ without changing the validity of this eigenvalue bound.)

Let $\mathbf{B}$ be the incidence matrix whose rows are indexed by the $(r-1)$ -edges of $\mathbf{H}_{r,t,S}$ and whose columns are indexed by the r -edges of $\mathbf{H}_{r,t,S}$. Note that the adjacency matrices of $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ and $G'_{\text{walk}}(\mathbf{H}_{r,t,S})$ satisfy

$$\mathbf{A}_{G'_{\text{walk}}(\mathbf{H}_{r,t,S})} + r\mathbf{I} = \mathbf{B}^T\mathbf{B}$$

and

$$\mathbf{A}_{G_{\text{walk}}(\mathbf{H}_{r,t,S})} + D\mathbf{I} = \mathbf{B}\mathbf{B}^T.$$

Therefore, the non-trivial eigenvalues of $\mathbf{B}^T\mathbf{B}$ lie in the interval $[0, (1 - c\epsilon)r(D-1) + r]$. (The lower bound follows since $\mathbf{B}^T\mathbf{B}$ is positive semi-definite.) Furthermore, $\mathbf{B}\mathbf{B}^T$ has the same eigenvalues as $\mathbf{B}^T\mathbf{B}$ with some 0's removed. Therefore,

the non-trivial eigenvalues of $G_{\text{walk}}(\mathbf{H}_{r,t,S})$ lie in the interval $[-D, (1 - c\epsilon)r(D - 1) + r - D]$. Since $r \geq 3$, $c < 1/2$ and $D \geq r$, this interval is contained in $[-(1 - c\epsilon)(r - 1)D, (1 - c\epsilon)(r - 1)D]$. $\blacksquare$

3 Lower-order random walks

Given an r -uniform hypergraph $H = (V, E)$, recall that a k -edge of H is a k -element subset of V that is contained in an r -edge of H . Two different k -edges are said to be adjacent if there is some $(k + 1)$ -edge of H that they are both contained in. The k -th order random walk on H is a random walk that starts at some k -edge of H and moves to a random adjacent k -edge of H repeatedly.

In the previous section, we showed that the $(r - 1)$ -st order random walk on $\mathbf{H}_{r,t,S}$ mixes rapidly. In this section, we extend this to prove that the k -th order random walk mixes rapidly for all $1 \leq k \leq r - 1$. We do this by proving that the k -th order random walk on $\mathbf{H}_{r,t,S}$ behaves almost exactly the same as the k -th order random walk on $\mathbf{H}_{k+1,t,2^{r-k-1}S'}$ where $2^{r-k-1}S'$ is a modification of the 2^{r-k-1} -fold sumset of S defined below.

In this section, we use $\mathcal{P}_r$ to refer to $\mathcal{P}$ of (1.1) and $\mathcal{T}_r(S)$ to refer to $\mathcal{T}$ of (1.2) when confusion may arise. Define $lS' = \{s_1 + \dots + s_l : s_1, \dots, s_l \in S \text{ distinct}\}$. For a hypergraph H , we write $G_{\text{walk}}^{(k)}(H)$ for the graph whose vertices are the k -edges of H with an edge between each pair of adjacent k -edges.

Let $\mathbf{G}_{r,t,S}^{(k)}$ be the graph whose vertices are the ordered k -tuples $(e_i(x, \mathbf{s}))_{i=1}^k$ for $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$. Two vertices of $\mathbf{G}_{r,t,S}^{(k)}$ are adjacent if they agree in all but one coordinate. Note that previously we defined the graph $\mathbf{G}_{r,t,S}$ to have vertex set consisting of pairs $(x, \mathbf{s})$. This definition agrees with $\mathbf{G}_{r,t,S}^{(r)}$ since for $(x, \mathbf{s}) \neq (y, \mathbf{t})$ the r -tuples $e(x, \mathbf{s})$ and $e(y, \mathbf{t})$ are distinct. However, for $k < r$, there may be many pairs $(x, \mathbf{s})$ and $(y, \mathbf{t})$ such that $(e_i(x, \mathbf{s}))_{i=1}^k = (e_i(y, \mathbf{t}))_{i=1}^k$.

Our proof proceeds in three stages. First we show that $\mathbf{G}_{r,t,S}^{(r-k)}$ is an induced subgraph of $\mathbf{G}_{r-k,t,2^k S'}$. Second, we use this fact to prove that $\mathbf{G}_{r,t,S}^{(r-k)}$ is an expander in the same manner as in Section 2. Third, we deduce from this that $G_{\text{walk}}^{(r-k-1)}(\mathbf{H}_{r,t,S})$ is an expander in the same manner as in the proof of Theorem 1.1.

Define $\mathcal{T}' \subseteq \mathcal{T}_{r-k}(2^k S')$ by

$$\mathcal{T}' = \left\{ \left(\sum_{a=1}^{2^k} s_{a,I} \right)_{I \in \mathcal{P}_{r-k}} : \text{all the } s_{a,I} \in S \text{ are distinct} \right\}. \quad (3.1)$$

Note that this definition differs from that of $\mathcal{T}_{r-k}(2^k S')$ in that the latter set is defined by the two conditions that for each I the elements $\{s_{a,I}\}_{1 \leq a \leq 2^k}$ are distinct and the elements $\{\sum_a s_{a,I}\}_{I \in \mathcal{P}_{r-k}}$ are distinct. Since S satisfies Assumption 2 the condition that defines $\mathcal{T}_{r-k}(2^k S')$ is a weaker condition than the one that defines $\mathcal{T}'$.

Now define $U \subseteq V(\mathbf{G}_{r-k,t,2^k S'})$ by

$$U = \mathbb{F}_2^t \times \mathcal{T}'. \quad (3.2)$$

We write $\mathbf{G}_{r-k,t,2^k S'}[U]$ for the induced subgraph on vertex set U .

Let us briefly discuss the properties of $\mathbf{G}_{r-k,t,2^k S'}$. Assume that r, t, S, ϵ satisfy Assumptions 1, 2, 3. Obviously, $r-k, t, 2^k S'$ still satisfy Assumption 1. Later on, we will show that $t, 2^k S', \epsilon$ still satisfy Assumption 3 for some constant c . However, $r-k, t, 2^k S'$ fail Assumption 2. To see this, suppose $k = 1$ and $a, b, c, d \in S$. Then $a+b, a+c, b+d, c+d \in 2^k S'$, but $(a+b) + (c+d) = (a+c) + (b+d)$. This means that Lemmas 2.1 and 2.3 may not apply to $\mathbf{G}_{r-k,t,2^k S'}$. However, we will see that these lemmas still apply to $\mathbf{G}_{r-k,t,2^k S'}[U]$.

Lemma 3.1. *Recall $\mathcal{T}'$ from (3.1). For r, t, S satisfying Assumption 2 and $0 \leq k \leq r-2$, the ordered $(r-k)$ -tuples $e(x, \mathbf{s})$ for $x \in \mathbb{F}_2^t$ and $\mathbf{s} \in \mathcal{T}'$ are all distinct.*

Proof. This follows from essentially the same proof as Lemma 2.1.

Suppose $e(x, \mathbf{s}) = e(x', \mathbf{s}')$ for $x, x' \in \mathbb{F}_2^t$ and $\mathbf{s}, \mathbf{s}' \in \mathcal{T}'$, where $s_I = \sum_{a=1}^{2^k} s_{a,I}$ and $s'_I = \sum_{a=1}^{2^k} s'_{a,I}$ with $s_{a,I}, s'_{a,I} \in S$. Then we know that

$$e_i(x, \mathbf{s}) + e_j(x, \mathbf{s}) = e_i(x', \mathbf{s}') + e_j(x', \mathbf{s}')$$

for $1 \leq i < j \leq r-k$. This can be rewritten as

$$\sum_{\substack{I \in \mathcal{P}_{r-k} \\ |\{i,j\} \cap I|=1}} \sum_{a=1}^{2^k} s_{a,I} = \sum_{\substack{I \in \mathcal{P}_{r-k} \\ |\{i,j\} \cap I|=1}} \sum_{a=1}^{2^k} s'_{a,I}.$$

Since the $s_{a,I}$'s are all distinct, the $s'_{a,I}$'s are all distinct and S satisfies Assumption 2, the same argument as in Lemma 2.1 implies that $\{s_{a,I}\}_a = \{s'_{a,I}\}_a$ for each $I \in \mathcal{P}_{r-k}$. Thus $\mathbf{s} = \mathbf{s}'$. Finally, since $e_1(x, \mathbf{s}) = e_1(x', \mathbf{s}')$, this implies that $x = x'$, as desired. $\blacksquare$

Lemma 3.2. *Recall U from (3.2). For r, t, S satisfying Assumption 2 and $0 \leq k \leq r-2$, the graph $\mathbf{G}_{r,t,S}^{(r-k)}$ is isomorphic to $\mathbf{G}_{r-k,t,2^k S'}[U]$.*

Proof. We will define maps $\psi: V(\mathbf{G}_{r,t,S}) \rightarrow V(\mathbf{G}_{r,t,S}^{(r-k)})$ and $\phi: V(\mathbf{G}_{r,t,S}) \rightarrow U \subseteq V(\mathbf{G}_{r-k,t,2^k S'})$ and show that $\phi \circ \psi^{-1}$ is a well-defined isomorphism between $\mathbf{G}_{r,t,S}^{(r-k)}$ and $\mathbf{G}_{r-k,t,2^k S'}[U]$.

Define $\psi: V(\mathbf{G}_{r,t,S}) \rightarrow V(\mathbf{G}_{r,t,S}^{(r-k)})$ by $\psi((x, \mathbf{s})) = (e_i(x, \mathbf{s}))_{i=1}^{r-k}$.

Next start by defining a map $\phi: \mathcal{P}_r \rightarrow \mathcal{P}_{r-k} \cup \{\emptyset\}$. For $J \in \mathcal{P}_r$, note that exactly one of $J \cap [r-k]$ and $[r-k] \setminus J$ lies in $\mathcal{P}_{r-k} \cup \{\emptyset\}$. Let $\phi(J)$ be the one of these two sets that lies in $\mathcal{P}_{r-k} \cup \{\emptyset\}$. Note that for any $I \in \mathcal{P}_{r-k} \cup \{\emptyset\}$, there are exactly 2^k sets $J \in \mathcal{P}_r$ such that $\phi(J) = I$. Then $\phi: V(\mathbf{G}_{r,t,S}) \rightarrow V(\mathbf{G}_{r-k,t,2^k S'})$ is given by $\phi((x, \mathbf{s})) = (y, \mathbf{t})$ where

$$t_I = \sum_{\substack{J \in \mathcal{P}_r \\ \phi(J)=I}} s_J$$

and

$$y = x + \sum_{\substack{J \in \mathcal{P}_r \\ J \cap [r-k] \notin \mathcal{P}_{r-k} \cup \{\emptyset\}}} s_J.$$

This map has two important properties. First, $\phi((x, \mathbf{s})) \in U \subseteq V(\mathbf{G}_{r-k, t, 2^k S'})$ for any $(x, \mathbf{s}) \in V(\mathbf{G}_{r, t, S})$. Second, if $\phi((x, \mathbf{s})) = (y, \mathbf{t})$ then $e_i(x, \mathbf{s}) = e_i(y, \mathbf{t})$ for $1 \leq i \leq k-r$. To see this property, write $\mathcal{P}_r = \mathcal{P}_r^{(1)} \sqcup \mathcal{P}_r^{(2)}$, where $\mathcal{P}_r^{(1)}$ consists of those J such that $J \cap [r-k] \in \mathcal{P}_{r-k} \cup \{\emptyset\}$ and $\mathcal{P}_r^{(2)}$ consists of the other J . Then note that

$$\begin{aligned} e_i(y, \mathbf{t}) &= x + \sum_{J \in \mathcal{P}_r^{(2)}} s_J + \sum_{\substack{I \in \mathcal{P}_{r-k} \\ i \in I}} \sum_{\substack{J \in \mathcal{P}_r \\ \phi(J)=I}} s_J \\ &= x + \sum_{J \in \mathcal{P}_r^{(2)}} s_J + \sum_{\substack{J \in \mathcal{P}_r^{(1)} \\ i \in J}} s_J + \sum_{\substack{J \in \mathcal{P}_r^{(2)} \\ i \notin J}} s_J \\ &= e_i(x, \mathbf{s}). \end{aligned}$$

By Lemma 3.1, this implies that $\phi((x, \mathbf{s})) = \phi((x', \mathbf{s}'))$ if and only if $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for all $1 \leq i \leq r-k$. This is the same condition as the one for $\psi((x, \mathbf{s})) = \psi((x', \mathbf{s}'))$. Therefore, we see that $\phi \circ \psi^{-1}: V(\mathbf{G}_{r, t, S}^{(r-k)}) \rightarrow U$ is a well-defined map that sends $(v_1, \dots, v_{r-k})$ to $(y, \mathbf{t}) \in U$ such that $(e_i(y, \mathbf{t}))_{i=1}^{r-k} = (v_1, \dots, v_{r-k})$. Therefore, $\phi \circ \psi^{-1}$ is a graph isomorphism, as desired. ■

Lemma 3.3. *Recall U from (3.2). For r, t, S satisfying Assumption 2 and $0 \leq k \leq r-2$, there exist constants $c = c(r)$ and $D = D(r, k, |S|)$ such that the induced subgraph $\mathbf{G}_{r-k, t, 2^k S'}[U]$ is D -regular with*

$$D \leq c|2^k S'| \leq c|S|^{2^k}.$$

Proof. This follows from a similar argument to Lemma 2.3.

Fix $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}'$ and $1 \leq l \leq r-k$. We wish to count pairs $(x', \mathbf{s}') \in \mathbb{F}_2^t \times \mathcal{T}'$ such that $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for all $i \neq l$. Say $\mathbf{s}, \mathbf{s}'$ are defined by $s_I = \sum_{a=1}^{2^k} s_{a, I}$ and $s'_I = \sum_{a=1}^{2^k} s'_{a, I}$. We know that

$$\sum_{\substack{I \in \mathcal{P}_{r-k} \\ |\{i, j\} \cap I|=1}} \sum_{a=1}^{2^k} s_{a, I} = \sum_{\substack{I \in \mathcal{P}_{r-k} \\ |\{i, j\} \cap I|=1}} \sum_{a=1}^{2^k} s'_{a, I}$$

for $1 \leq i < j \leq r-k$ with $i \neq l$ and $j \neq l$.

As in the proof of Lemma 2.3, we partition $\mathcal{P}_{r-k} \setminus \{\{l\}\}$ into pairs $\{I, J\}$ where either $J = I \sqcup \{l\}$, $I = J \sqcup \{l\}$ or $I \sqcup J = [r-k] \setminus \{l\}$. Using essentially the same argument as in the proof of Lemma 2.3, we conclude that for each such pair $\{I, J\}$,

$$\{s_{a, I}\}_a \cup \{s_{a, J}\}_a = \{s'_{a, I}\}_a \cup \{s'_{a, J}\}_a.$$

Therefore, for each $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}'$ and $1 \leq l \leq r - k$ the number of pairs $(x', \mathbf{s}') \in \mathbb{F}_2^t \times \mathcal{T}'$ satisfying $e_i(x, \mathbf{s}) = e_i(x', \mathbf{s}')$ for all $i \neq l$ is exactly equal to

$$\binom{|S| - (2^{r-1} - 2^{k+1})}{2^k} \binom{2^{k+1}}{2^k}^{2^{r-k-2}-1}.$$

The first term is the number of choices for $s'_{\{l\}} = \sum_{a=1}^{2^k} s'_{a, \{l\}}$ where the $s'_{a, \{l\}}$ are distinct from the $s_{a, I}$ for $I \neq \{l\}$. The second term is the number of ways to choose $s'_I = \sum_{a=1}^{2^k} s'_{a, I}$ and $s'_J = \sum_{a=1}^{2^k} s'_{a, J}$ given the set $\{s'_{a, I}\}_a \cup \{s'_{a, J}\}_a$ for each of the $2^{r-k-2} - 1$ pairs $\{I, J\}$ that $\mathcal{P}_{r-k} \setminus \{\{l\}\}$ is partitioned into. ■

Lemma 3.4. *For r, t, ϵ, S satisfying Assumptions 1, 3 and $0 \leq k \leq r - 2$, there exists a constant $c > 0$ such that $\text{Cay}(\mathbb{F}_2^t, 2^{k+1}S')$ is a $c\epsilon$ -expander.*

Proof. For brevity, write $m = 2^{k+1}$ and $|S| = d$. Let $\mathbf{A}$ denote the adjacency matrix of $\text{Cay}(\mathbb{F}_2^t, S)$ and let its eigenvalues be $d = \lambda_1 \geq |\lambda_2| \geq \dots \geq |\lambda_N|$.

Observe that the edges in $\text{Cay}(\mathbb{F}_2^t, S)^m$ incident to vertex x are in bijection with m -tuples $(s_1, \dots, s_m) \in S^m$, i.e., for each tuple there is an edge $(x, x + s_1 + \dots + s_m)$. Similarly, note that the edges in $m! \cdot \text{Cay}(\mathbb{F}_2^t, mS')$ are in bijection with the tuples $(s_1, \dots, s_m)$ where all m coordinates are distinct.

For a partition P of $[m]$, say that a tuple $(s_1, \dots, s_m)$ is of type P if whenever i, j are in the same part of P then $s_i = s_j$. Write $A_P(S)$ for the multiset with element $s_1 + \dots + s_m$ for each $(s_1, \dots, s_m)$ of type P .

By Möbius inversion, we know that

$$\mathbf{A}_{m! \cdot \text{Cay}(\mathbb{F}_2^t, mS')} = \sum_P \mu(P) \mathbf{A}_{\text{Cay}(\mathbb{F}_2^t, A_P(S))},$$

where $\mu(P)$ is the Möbius function $(-1)^{m-|P|} \prod_{p \in P} (|p| - 1)!$. For a partition P , let $e(P)$ be the number of even-sized parts and $o(P)$ be the number of odd-sized parts. It is not hard to see that

$$\mathbf{A}_{\text{Cay}(\mathbb{F}_2^t, A_P(S))} = d^{e(P)} \mathbf{A}^{o(P)}.$$

Therefore, we conclude that

$$\mathbf{A}_{\text{Cay}(\mathbb{F}_2^t, mS')} = \frac{1}{m!} \sum_P \mu(P) d^{e(P)} \mathbf{A}^{o(P)}.$$

Write

$$f(x) = \frac{1}{m!} \sum_P \mu(P) d^{|P|} x^{o(P)}.$$

Note that since $m = 2^{k+1}$ is even, any partition of $[m]$ has an even number of odd-sized parts. Thus, $f(x)$ is an even function. Therefore, the second largest

eigenvalue of $\mathcal{A}_{\text{Cay}(\mathbb{F}_2^t, mS')}$ is bounded in absolute value by $\sup_{0 \leq x \leq 1-\epsilon} |f(x)|$. We can upper bound $f(x)$ by

$$\begin{aligned} f(x) &\leq (1 - \epsilon/2) \frac{1}{m!} \sum_P \mu(P) d^{|P|} - (1 - \epsilon/2 - x^m) \frac{d^m}{m!} \\ &\quad + \frac{1}{m!} \sum_{P \neq (1, \dots, 1)} |\mu(P)| d^{|P|} \left| 1 - \epsilon/2 - x^{o(P)} \right|. \end{aligned} \quad (3.3)$$

The first term is exactly $(1 - \epsilon/2) \binom{d}{m}$. To bound the last term, we use the following inequality.

Claim 3.5.

$$\sum_{\substack{P \vdash [m] \\ |P| = m-a}} |\mu(P)| \leq m^{2a}.$$

Proof. We use $|\mu(P)| = \prod_{p \in P} (|p| - 1)! \leq \prod_{p \in P} |p|!$. Now

$$\sum_{\substack{P \vdash [m] \\ |P| = m-a}} \prod_{p \in P} |p|!$$

has a combinatorial interpretation as the number of partitions of $[m]$ into $m - a$ parts where the elements of each part are ordered. Such a partition can be represented schematically as a directed graph on m labeled vertices where each connected component is a directed path. The condition that the partition has $m - a$ parts is equivalent to the constraint that the graph has a edges. But there are $m(m - 1)$ choices for a single edge, so $(m(m - 1))^a \leq m^{2a}$ is a simple upper bound on the number of m -vertex a -edge directed graphs. ■

We are interested in the case where $m = 2^{k+1} \leq 2^{r-1}$ which, with Assumption 1, implies that $d \geq 4m^2$. Together with Claim 3.5 this implies that

$$\sum_{P \neq (1, \dots, 1)} |\mu(P)| d^{|P|} = \sum_{a > 1} \sum_{|P| = m-a} |\mu(P)| d^{m-a} \leq d^m \sum_{a > 1} \left(\frac{m^2}{d} \right)^a \leq 2m^2 d^{m-1}.$$

Using this with (3.3) gives

$$\begin{aligned} f(x) &\leq (1 - \epsilon/2) \binom{d}{m} - (1 - \epsilon/2 - x^m) \frac{d^m}{m!} \\ &\quad + \max\{\epsilon/2, 1 - \epsilon/2 - x^m\} \frac{2m^2 d^{m-1}}{m!} \end{aligned}$$

for $0 \leq x \leq 1 - \epsilon$, since $|1 - \epsilon/2 - x^{o(P)}|$ is extremized when $o(P) = 0, m$. But

$$\max\{\epsilon/2, 1 - \epsilon/2 - x^m\} \frac{2m^2 d^{m-1}}{m!} \leq (1 - \epsilon/2 - x^m) \frac{d^m}{m!},$$

since $1 - \epsilon/2 - x^m \geq 1 - \epsilon/2 - (1 - \epsilon) = \epsilon/2$ and $\frac{2m^2 d^{m-1}}{m!} \leq \frac{d^m}{m!}$. Therefore,

$$f(x) \leq (1 - \epsilon/2) \binom{d}{m}.$$

For the lower bound, we use the analogous inequality

$$\begin{aligned} f(x) &\geq (1 - \epsilon/2) \binom{d}{m} - (1 - \epsilon/2 - x^m) \frac{d^m}{m!} \\ &\quad - \max\{\epsilon/2, 1 - \epsilon/2 - x^m\} \frac{2m^2 d^{m-1}}{m!}. \end{aligned}$$

Since

$$\max\{\epsilon/2, 1 - \epsilon/2 - x^m\} \frac{2m^2 d^{m-1}}{m!} \leq \frac{1}{2} (1 - \epsilon/2 - x^m) \frac{d^m}{m!},$$

we conclude

$$f(x) \geq (1 - \epsilon/2) \binom{d}{m} - \frac{3}{2} (1 - \epsilon) \frac{d^m}{m!}.$$

Since $d \geq 4m^2$, we have

$$\binom{d}{m} \geq \left(\frac{d-m}{d} \right)^m \frac{d^m}{m!} \geq \left(1 - \frac{1}{4m} \right)^m \frac{d^m}{m!} \geq e^{-2/7} d^m / m!.$$

Combining these inequalities yields

$$f(x) \geq \left(1 - \frac{3}{2} e^{2/7} \right) (1 - \epsilon/2) \binom{d}{m} \approx -0.996 (1 - \epsilon/2) \binom{d}{m},$$

as required. ■

Note that Lemma 2.7 only requires Assumption 1. Therefore, if we define G_{Cay} to be the multigraph with vertex set $\mathbb{F}_2^t \times \mathcal{T}_{r-k}(2^k S')$ and $c|2^k S'|^{M-|\mathcal{P}_{r-k}|-2}$ edges between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ if xy is an edge in $\text{Cay}(\mathbb{F}_2^t, (2^k S' + 2^k S') \setminus \{0\})$, then we know that $\mathbf{G}_{r-k,t,2^k S'}^M$ contains G_{Cay} as a subgraph. Furthermore, it (essentially) follows from Lemma 3.4 that G_{Cay} is an expander. However, this does not imply that $\mathbf{G}_{r-k,t,2^k S'}$ is an expander since we do not have a degree bound.

Lemma 3.3 gives us the required degree bound on $\mathbf{G}_{r-k,t,2^k S'}[U]$, so to prove that this subgraph is an expander all we need to do is prove that $G_{\text{Cay}}[U]$ is a subgraph of $\mathbf{G}_{r-k,t,2^k S'}[U]^M$. This follows by the same technique as in the proof of Lemma 2.7 with some small modifications.

Lemma 3.6. *Given r, t, S satisfying Assumption 1 and $0 \leq k \leq r-2$, there exists a constant $c = c(r) > 0$ such that for each $I \in \mathcal{P}_{r-k}$, each $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}'$ and each $a \in 2^k S'$ disjoint from all coordinates of $\mathbf{s}$, there are at least $c|2^k S'|^{|I|-1}$ walks of length $|I|$ in $\mathbf{G}_{r-k,t,2^k S'}[U]$ which begin at $(x, \mathbf{s})$ and end at an element of the form $(x, \mathbf{s}')$ where $s'_I = a$ and $s'_J = s'_J$ for all $J \in \mathcal{P}_{r-k}$ with $J \not\subseteq I$.*

Proof. Following the proof of Lemma 2.4, write $|I| = l$ and note that for each sequence $a_2, \dots, a_l \in 2^k S'$ we can define $\mathbf{s} = \mathbf{s}^0, \mathbf{s}^1, \dots, \mathbf{s}^l$ by (2.3) and (2.4) such that each pair $e(x, \mathbf{s}^i)$ and $e(x, \mathbf{s}^{i+1})$ agree in all but one coordinate.

These are exactly the desired paths as long as $\mathbf{s}^1, \dots, \mathbf{s}^l$ lie in $\mathcal{T}'$. To satisfy this condition we simply need to pick $a_2, \dots, a_l$ disjoint from each other and from a and the coordinates of $\mathbf{s}$. The number of ways to do this

$$\binom{|S| - 2^{r-1}}{2^k} \binom{|S| - 2^{r-1} - 2^k}{2^k} \dots \binom{|S| - 2^{r-1} - (l-2)2^k}{2^k} \geq c|2^k S'|^{l-1}.$$

The inequality follows by Assumption 1. $\blacksquare$

Lemma 3.7. *Given r, t, S satisfying Assumption 1 and $0 \leq k \leq r-2$, there exist constants $c = c(r) > 0$ and $M' = M'(r-k)$ such that for $x \in \mathbb{F}_2^t$ and $\mathbf{s}, \mathbf{t} \in \mathcal{T}'$, there are at least $c|2^k S'|^{M' - |\mathcal{P}_{r-k}|}$ walks of length M' from $(x, \mathbf{s})$ to $(x, \mathbf{t})$ in $\mathbf{G}_{r-k, t, 2^k S'}[U]$.*

Proof. Following the proof of Lemma 2.6, fix $\mathbf{u} \in \mathcal{T}'$ all of whose coordinates are disjoint from all the coordinates of both $\mathbf{s}$ and $\mathbf{t}$. By exactly the same argument, we can repeatedly use Lemma 3.6 to find $c|2^k S'|^{M' - 2|\mathcal{P}_{r-k}|}$ walks of length M' from $(x, \mathbf{s})$ to $(x, \mathbf{u})$ to $(x, \mathbf{t})$. Finally, note that there are

$$\begin{aligned} & \binom{|S| - (2^r - 2^{k+1})}{2^k} \binom{|S| - (2^r - 2^{k+1}) - 2^k}{2^k} \\ & \dots \binom{|S| - (2^r - 2^{k+1}) - (|\mathcal{P}_{r-k}| - 1)2^k}{2^k} \geq c'|2^k S'|^{|\mathcal{P}_{r-k}|} \end{aligned}$$

choices for $\mathbf{u}$, completing the proof. $\blacksquare$

Lemma 3.8. *Given r, t, S satisfying Assumption 1 and $0 \leq k \leq r-2$, there exist constants $c = c(r) > 0$ and $M = M(r-k)$ such that, given x adjacent to y in $\text{Cay}(\mathbb{F}_2^t, 2^{k+1} S')$ and $\mathbf{s}, \mathbf{t} \in \mathcal{T}'$, there are at least $c|2^k S'|^{M - |\mathcal{P}_{r-k}| - 2}$ walks of length M from $(x, \mathbf{s})$ to $(y, \mathbf{t})$ in $\mathbf{G}_{r-k, t, 2^k S'}[U]$.*

Proof. Following the proof of Lemma 2.7, let I_1, I_2 be disjoint elements of $\mathcal{P}_{r-k}$ such that $I_1 \sqcup I_2 = [r-k] \setminus \{2\}$. Now since x is adjacent to y in $\text{Cay}(\mathbb{F}_2^t, 2^{k+1} S')$, there exist $a_1, a_2 \in 2^k S'$ such that $y = x + a_1 + a_2$. For any $\mathbf{u} \in \mathcal{T}'$ satisfying $u_{I_1} = a_1$ and $u_{I_2} = a_2$ and any b disjoint from all coordinates of $\mathbf{u}$, define $\mathbf{u}'$ as in (2.5). Then, by the exact same argument, we can use Lemma 3.7 twice to find $c|2^k S'|^{2M' - 2|\mathcal{P}_{r-k}|}$ walks of length $M = 2M' + 1$ from $(x, \mathbf{s})$ to $(x, \mathbf{u})$ to $(y, \mathbf{u}')$ to $(y, \mathbf{t})$. Finally, note that there are

$$\binom{|S| - 2 \cdot 2^k}{2^k} \binom{|S| - 3 \cdot 2^k}{2^k} \dots \binom{|S| - |\mathcal{P}_{r-k}| \cdot 2^k}{2^k} \geq c'|2^k S'|^{|\mathcal{P}_{r-k}| - 1}$$

choices for $(\mathbf{u}, \mathbf{u}')$, completing the proof. $\blacksquare$

Lemma 3.9. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3 and $0 \leq k \leq r-2$, there exists a constant $c = c(r) > 0$ such that $\mathbf{G}_{r, t, S}^{(r-k)}$ is a $c\epsilon$ -expander.*

Proof. Following the proof of Lemma 2.8, define G_{Cay} to be the multigraph with vertex set U and $c|2^k S'|^{M-|\mathcal{P}_{r-k}|-2}$ edges between $(x, \mathbf{s})$ and $(y, \mathbf{t})$ if xy is an edge in $\text{Cay}(\mathbb{F}_2^t, 2^{k+1} S')$. Here c, M are the same constants as in the proof of Lemma 3.8. Note that G_{Cay} is regular of degree $d_1 \geq c'|2^k S'|^M$ for some $c' = c'(r) > 0$. Furthermore, by Lemma 3.4, we know that G_{Cay} is a $c^* \epsilon$ -expander.

By Lemma 3.3, we know that $G_{r-k,t,2^k S'}[U]^M$ is regular of degree $d_2 \leq c''|2^k S'|^M$ for some $c'' = c(r)$ and, by Lemma 3.8, we know that G_{Cay} is a subgraph of $G_{r-k,t,2^k S'}[U]^M$. By the exact same proof as Lemma 2.8, this implies that $G_{r-k,t,2^k S'}[U]$ is a $c_0 \epsilon$ -expander for some c_0 depending on c', c'', c^* and M . Since this graph is isomorphic to $G_{r-k,t,S}^{(r-k)}$, this proves the desired result. $\blacksquare$

Theorem 1.2. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3, there exists a constant $c = c(r) > 0$ such that for every $1 \leq k \leq r-1$, the k -th order walk graph $G_{\text{walk}}^{(k)}(\mathbf{H}_{r,t,S})$ is a $c\epsilon$ -expander.*

Proof. First assume that $k > 1$. This case is essentially the same as the proof of Theorem 1.1 from Lemma 2.8.

By Lemma 3.3, we know that there is some D_k such that every k -edge of $\mathbf{H}_{r,t,S}$ is contained in exactly D_k of the $(k+1)$ -edges.

As before, we consider three graphs: $G_{r,t,S}^{(k+1)}$, $G_{\text{walk}}^{\prime(k+1)}(\mathbf{H}_{r,t,S})$, $G_{\text{walk}}^{(k)}(\mathbf{H}_{r,t,S})$. The vertices of $G_{r,t,S}^{(k+1)}$ are ordered $(k+1)$ -tuples $(v_1, \dots, v_{k+1})$ such that there exists $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$ with $v_i = e_i(x, \mathbf{s})$. We define $G_{\text{walk}}^{\prime(k+1)}(\mathbf{H}_{r,t,S})$ to be the graph whose vertices are the unordered $(k+1)$ -edges of $\mathbf{H}_{r,t,S}$ where two $(k+1)$ -edges are adjacent if they contain a common k -edge. We know that the first two of these graphs are $(k+1)(D_k-1)$ -regular, while the third graph is kD_k -regular.

Now define a graph homomorphism $\pi: G_{r,t,S}^{(k+1)} \rightarrow G_{\text{walk}}^{\prime(k+1)}(\mathbf{H}_{r,t,S})$ that sends a vertex $(v_1, \dots, v_{k+1})$ to the vertex $\{v_1, \dots, v_{k+1}\}$. These two graphs are regular of the same degree and, by Lemma 2.2 and the discussion preceding it, the preimage of each vertex has size $(k+1)!$. As in the proof of Theorem 1.1, it follows that $\lambda(G_{\text{walk}}^{\prime(k+1)}(\mathbf{H}_{r,t,S})) \leq \lambda(G_{r,t,S}^{(k+1)}) \leq (1 - c\epsilon)(k+1)(D_k-1)$.

Next define $\mathbf{B}$ to be the incidence matrix whose rows are indexed by the k -edges of $\mathbf{H}_{r,t,S}$ and whose columns are indexed by the $(k+1)$ -edges of $\mathbf{H}_{r,t,S}$. We know that

$$\mathbf{A}_{G_{\text{walk}}^{\prime(k+1)}(\mathbf{H}_{r,t,S})} + (k+1)\mathbf{I} = \mathbf{B}^T \mathbf{B}$$

and

$$\mathbf{A}_{G_{\text{walk}}^{(k)}(\mathbf{H}_{r,t,S})} + D_k \mathbf{I} = \mathbf{B} \mathbf{B}^T.$$

From the first equation, we conclude that the non-trivial eigenvalues of $\mathbf{B}^T \mathbf{B}$ lie in the interval $[0, (1 - c\epsilon)(k+1)(D_k-1) + k+1]$, so the non-trivial eigenvalues of $G_{\text{walk}}^{(k)}(\mathbf{H}_{r,t,S})$ lie in the interval $[-D_k, (1 - c\epsilon)(k+1)(D_k-1) + k+1 - D_k]$. Since $k \geq 2$, $D_k \geq k+1$ and choosing c such that $c\epsilon < 1/2$, this interval is contained in $[-(1 - c\epsilon)kD_k, (1 - c\epsilon)kD_k]$.

Now assume that $k = 1$. Note that a first order random walk on $H_{r,t,S}$ is exactly the same as an ordinary random walk on the graph $\text{Cay}(\mathbb{F}_2^t, 2^{r-2}S')$. By Lemma 3.4, we know that $\text{Cay}(\mathbb{F}_2^t, 2^{r-2}S')$ is a $c\epsilon$ -expander, completing the proof. $\blacksquare$

4 A discrepancy result

For $H = (V, E)$ an r -uniform hypergraph and $V_1, \dots, V_r \subseteq V$, let $e_H(V_1, \dots, V_r)$ be the number of tuples $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$ such that $\{v_1, \dots, v_r\} \in E$. In this self-contained section, we prove that our hypergraph $H_{r,t,S}$ satisfies the pseudorandomness condition that $e_{H_{r,t,S}}(V_1, \dots, V_r)$ is always close to its expected value. For simplicity, we write $e(V_1, \dots, V_r)$ for $e_{H_{r,t,S}}(V_1, \dots, V_r)$ throughout this section.

A general result of this type was proven by Parzanchevski [23], saying that the desired pseudorandomness condition holds under essentially the hypothesis on the spectra of all the adjacency matrices that we proved in Theorem 3.9. However, Parzanchevski's result only applies when $\frac{|\lambda_2|}{\lambda_1} < \epsilon_0(r)$ where $\epsilon_0(r)$ is a small constant depending on the uniformity, while we have only proved that $\frac{|\lambda_2|}{\lambda_1} < 1 - c(r)\epsilon$ where ϵ is the expansion parameter of the original Cayley graph and $c(r)$ is a very small constant depending on the uniformity. We will therefore use a different proof method.

In this section, $\mathcal{Q}$ will denote an arbitrary multiset supported in $2^{[r]}$. For such a $\mathcal{Q}$ and $(x, \mathbf{s}) \in \mathbb{F}_2^t \times S^{\mathcal{Q}}$, define

$$e_{\mathcal{Q},i}(X, \mathbf{s}) = x + \sum_{\substack{I \in \mathcal{Q} \\ i \in I}} s_I.$$

Write $e_{\mathcal{Q}}(x, \mathbf{s})$ for the r -tuple $(e_{\mathcal{Q},1}(x, \mathbf{s}), \dots, e_{\mathcal{Q},r}(x, \mathbf{s}))$.

We wish to compute $e(V_1, \dots, V_r)$, the number of pairs $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}$ such that $e(x, \mathbf{s}) \in V_1 \times \dots \times V_r$. We will start by computing $f_{\mathcal{Q}}(V_1, \dots, V_r)$, defined to be the number of pairs $(x, \mathbf{s}) \in \mathbb{F}_2^t \times S^{\mathcal{Q}}$ such that $e_{\mathcal{Q}}(x, \mathbf{s}) \in V_1 \times \dots \times V_r$. We will then use Möbius inversion to turn this into a formula for the desired quantity.

Lemma 4.1. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 2, 3, $\mathcal{Q}$ a multiset supported in $2^{[r]}$ and $V_1, \dots, V_r \subseteq \mathbb{F}_2^t$,*

$$\left| f_{\mathcal{Q}}(V_1, \dots, V_r) - \frac{|S|^{\|\mathcal{Q}\|}}{(2^t)^{r-1}} |V_1| \cdots |V_r| \right| \leq (1 - \epsilon)(r - 1) |S|^{\|\mathcal{Q}\|} \sqrt{|V_1| |V_r|}.$$

Proof. For $1 \leq k \leq r$, define $\mathcal{Q}^{(k)} \subseteq \mathcal{Q}$ by

$$\mathcal{Q}^{(k)} = \{I \in \mathcal{Q} : 0 < |I \cap \{1, \dots, k\}| < k\}.$$

Furthermore, write $e_{\mathcal{Q}}^{(k)}(x, \mathbf{s})$ for the k -tuple $(e_{\mathcal{Q}^{(k)},1}(x, \mathbf{s}), \dots, e_{\mathcal{Q}^{(k)},k}(x, \mathbf{s}))$. Let $F^k(V_1, \dots, V_k)$ be the set of pairs $(y, \mathbf{s}) \in \mathbb{F}_2^t \times S^{\mathcal{Q}^{(k)}}$ such that $e_{\mathcal{Q}}^{(k)}(y, \mathbf{s}) \in V_1 \times \dots \times V_k$. We write $f^k(V_1, \dots, V_k) = |F^k(V_1, \dots, V_k)|$.

Define $L^{(k)} = \{I \in \mathcal{Q} : I \cap \{1, \dots, k+1\} = \{k+1\}\}$ and $U^{(k)} = \{I \in \mathcal{Q} : I \cap \{1, \dots, k+1\} = \{1, \dots, k\}\}$. Then

$$\mathcal{Q}^{(k+1)} = \mathcal{Q}^{(k)} \sqcup \mathcal{L}^{(k)} \sqcup \mathcal{U}^{(k)}.$$

By definition, $F^{k+1}(V_1, \dots, V_{k+1})$ can therefore be written as the set of pairs $(x, (\mathbf{s}, \mathbf{t}, \mathbf{u})) \in \mathbb{F}_2^t \times \left(S^{\mathcal{Q}^{(k)}} \times S^{\mathcal{L}^{(k)}} \times S^{\mathcal{U}^{(k)}} \right)$ such that $e_{\mathcal{Q}}^{(k+1)}(x, (\mathbf{s}, \mathbf{t}, \mathbf{u})) \in V_1 \times \dots \times V_{k+1}$. Recall that

$$e_{\mathcal{Q}^{(k+1)}, i}(x, (\mathbf{s}, \mathbf{t}, \mathbf{u})) = x + \sum_{\substack{I \in \mathcal{Q}^{(k)} \\ i \in I}} s_I + \sum_{\substack{I \in \mathcal{L}^{(k)} \\ i \in I}} t_I + \sum_{\substack{I \in \mathcal{U}^{(k)} \\ i \in I}} u_I.$$

Therefore, we can rewrite the condition that $e_{\mathcal{Q}}^{(k+1)}(x, (\mathbf{s}, \mathbf{t}, \mathbf{u})) \in V_1 \times \dots \times V_{k+1}$ as

$$(y, \mathbf{s}) \in F^k(V_1, \dots, V_k) \quad \text{for} \quad y = x + \sum_{I \in \mathcal{U}^{(k)}} u_I$$

and

$$x + \sum_{\substack{I \in \mathcal{Q}^{(k)} \\ k+1 \in I}} s_I + \sum_{I \in \mathcal{L}^{(k)}} t_I \in V_{k+1}.$$

Therefore, $f^{k+1}(V_1, \dots, V_{k+1})$ is the number of edges between the multisets W_k and V_{k+1} in the multigraph $\text{Cay}(\mathbb{F}_2^t, S)^m$ where $m = |\mathcal{L}^{(k)} \sqcup \mathcal{U}^{(k)}|$ and W_k is the multiset with element

$$y + \sum_{\substack{I \in \mathcal{Q}^{(k)} \\ k+1 \in I}} s_I$$

for each $(y, \mathbf{s}) \in F^k(V_1, \dots, V_k)$.

Write w_x for the multiplicity of x in W_k , noting that $w_x \leq |S|^{|\mathcal{Q}^{(k)}|}$ since this is the number of choices for $\mathbf{s}$. Hence, by the ordinary expander-mixing lemma (for multisets), we have

$$\begin{aligned} & \left| f^{k+1}(V_1, \dots, V_{k+1}) - \frac{|S|^{|\mathcal{Q}^{(k+1)}| - |\mathcal{Q}^{(k)}|}}{2^t} f^k(V_1, \dots, V_k) |V_{k+1}| \right| \\ & \leq (1 - \epsilon) |S|^{|\mathcal{Q}^{(k+1)}| - |\mathcal{Q}^{(k)}|} \sqrt{\left(\sum_{x \in W_k} w_x^2 \right) |V_{k+1}|} \\ & \leq (1 - \epsilon) |S|^{|\mathcal{Q}^{(k+1)}| - |\mathcal{Q}^{(k)}|} \sqrt{|S|^{|\mathcal{Q}^{(k)}|} f^k(V_1, \dots, V_k) |V_{k+1}|} \\ & \leq (1 - \epsilon) |S|^{|\mathcal{Q}^{(k+1)}|} \sqrt{|V_1| |V_{k+1}|}. \end{aligned}$$

The last line follows from the easy bound $f^k(V_1, \dots, V_k) \leq |S|^{|\mathcal{Q}^{(k)}|} |V_1|$. To see this inequality, note that $|S|^{|\mathcal{Q}^{(k)}|}$ is the number of choices for $\mathbf{s}$. Once $\mathbf{s}$ is chosen, for each $x \in V_1$, there is a unique $y \in \mathbb{F}_2^t$ such that $e_{\mathcal{Q}^{(k)}, 1}(y, \mathbf{s}) = x$.

Since $f^1(V_1) = |V_1|$, we can telescope this bound to conclude that

$$\begin{aligned} & \left| f^r(V_1, \dots, V_r) - \frac{|S|^{|Q^{(r)}|}}{(2^t)^{r-1}} |V_1| \cdots |V_r| \right| \\ & \leq (1 - \epsilon) |S|^{|Q^{(r)}|} \sum_{i=2}^r \sqrt{|V_1| |V_i|} \prod_{j=i+1}^r \frac{|V_j|}{2^t} \\ & \leq (1 - \epsilon)(r - 1) |S|^{|Q^{(r)}|} \sqrt{|V_1| |V_r|}. \end{aligned}$$

To complete the proof, note that $f_Q(V_1, \dots, V_r) = |S|^{|Q| - |Q^{(r)}|} f^r(V_1, \dots, V_r)$, since $Q \setminus Q^{(r)}$ is a multiset supported on $\{\emptyset, [r]\}$ and these terms do not contribute materially to $e_Q(x, \mathbf{s})$. $\blacksquare$

Proposition 4.2. *For $r \geq 3$, $t, \epsilon > 0$ and $S \subset \mathbb{F}_2^t$ satisfying Assumptions 1, 2, 3 and $V_1, \dots, V_r \subseteq \mathbb{F}_2^t$,*

$$\begin{aligned} & \left| e_{H_{r,t,S}}(V_1, \dots, V_r) - \frac{|S|(|S| - 1) \cdots (|S| - (2^{r-1} - 2))}{(2^t)^{r-1}} |V_1| \cdots |V_r| \right| \\ & \leq (1 - \epsilon) 2r |S|^{2^{r-1}-1} \sqrt{|V_1| |V_r|}. \end{aligned}$$

Proof. For Λ a partition of $\mathcal{P}$, let $\mathcal{T}[\Lambda] \subseteq S^{\mathcal{P}}$ be the set of tuples $(s_I)_{I \in \mathcal{P}}$ such that if I, J are in the same part of Λ then $s_I = s_J$. We define a multiset $\mathcal{P}[\Lambda]$ supported in $2^{[r]}$ of size $|\Lambda|$ by letting the symmetric difference $I_1 \triangle \cdots \triangle I_k$ (that is, the set of elements that appear in an odd number of the I_i 's) be an element of $\mathcal{P}[\Lambda]$ for each $\{I_1, \dots, I_k\}$ a part of Λ .

For $\mathbf{s} \in T[\Lambda]$, define $\mathbf{t} \in S^{\mathcal{P}[\Lambda]}$ by letting $t_{I_1 \triangle \cdots \triangle I_k}$ be equal to $s_{I_1} = \cdots = s_{I_k}$ whenever $\{I_1, \dots, I_k\}$ is a part of Λ . This gives a bijection between $T[\Lambda]$ and $S^{\mathcal{P}[\Lambda]}$. Moreover, this map has the property that $e(x, \mathbf{s}) = e_{\mathcal{P}[\Lambda]}(x, \mathbf{t})$ whenever $\mathbf{s} \in T[\Lambda]$. Thus, the number of pairs $(x, \mathbf{s}) \in \mathbb{F}_2^t \times \mathcal{T}[\Lambda]$ such that $e(x, \mathbf{s}) \in V_1 \times \cdots \times V_r$ is exactly equal to $f_{\mathcal{P}[\Lambda]}(V_1, \dots, V_r)$.

Finally, we want to count $e(V_1, \dots, V_r)$, the number of $(x, \mathbf{s}) \in \mathbb{F}_2^t \times S^{|\mathcal{P}|}$ with $e(x, \mathbf{s}) \in V_1 \times \cdots \times V_r$ and the additional constraint that the s_I 's are distinct. By Möbius inversion, it follows that

$$e(V_1, \dots, V_r) = \sum_{\Lambda} \mu(\Lambda) f_{\mathcal{P}[\Lambda]}(V_1, \dots, V_r),$$

where

$$\mu(\Lambda) = (-1)^{|\mathcal{P}| - |\Lambda|} \prod_{L \in \Lambda} (|L| - 1)!.$$

Combining this formula with the result of Lemma 4.1 gives that

$$\begin{aligned} & \left| e(V_1, \dots, V_r) - \frac{|V_1| \cdots |V_r|}{(2^t)^{r-1}} \sum_{\Lambda} \mu(\Lambda) |S|^{|\Lambda|} \right| \\ & \leq (1 - \epsilon) r \sqrt{|V_1| |V_r|} \sum_{\Lambda} |\mu(\Lambda)| |S|^{|\Lambda|}. \end{aligned}$$

The first sum is exactly equal to $|S|(|S| - 1) \cdots (|S| - (2^{r-1} - 2))$, while we can bound the second sum by $2|S|^{2^{r-1}-1}$ in the same manner as in the proof of Lemma 3.4. $\blacksquare$

Averaging the above result over all renumberings of $V_1, \dots, V_r$, we obtain Theorem 1.3.

5 Concluding remarks

We conclude with two questions. The first is essentially a reiteration of a question in [4]: are the hypergraphs $H_{r,t,S}$ topological expanders for some suitable choice of parameters? In [4], the first author speculated that they might even satisfy a certain combinatorial notion of expansion, known as cosystolic expansion, from which topological expansion follows [5]. Unfortunately, as pointed out to us by Gundert and Luria [9], this is not the case. Nevertheless, the possibility that our hypergraphs are topological expanders remains a tantalizing one.

The second question concerns generalizations of our construction. In a fairly precise sense, abelian groups are the worst possible groups over which to define Cayley graphs if one is trying to produce expanders. A much better choice would be to work over finite simple groups of Lie type with bounded rank, where it is known [3] that two random elements almost surely suffice to generate an expander. However, our constructions depend in an absolutely critical way on commutativity, so much so that they might fittingly be called abelian complexes. On the other hand, r -uniform Ramanujan complexes can be explicitly rendered [18] in terms of Cayley graphs over $\mathrm{PGL}_r(\mathbb{F}_q)$, so it is reasonable to believe that there is some more general mechanism which works over these groups. We consider the problem of distilling out this mechanism to be one of the central problems in the area.

Acknowledgements. This paper was partially written while the first author was visiting the California Institute of Technology as a Moore Distinguished Scholar and he is extremely grateful for their kind support.

References

- [1] N. Alon and Y. Roichman, Random Cayley graphs and expanders, *Random Structures Algorithms* **5** (1994), 271–284.
- [2] Y. Bilu and N. Linial, Lifts, discrepancy and nearly optimal spectral gap, *Combinatorica* **26** (2006), 495–519.
- [3] E. Breuillard, B. Green, R. Guralnick and T. Tao, Expansion in finite simple groups of Lie type, *J. Eur. Math. Soc.* **17** (2015), 1367–1434.

- [4] D. Conlon, Hypergraph expanders from Cayley graphs, *Israel J. Math.* **233** (2019), 49–65.
- [5] D. Dotterrer, T. Kaufman and U. Wagner, On expansion and topological overlap, in *32nd Int. Symposium on Computational Geometry* (2016), 35:1–35:10.
- [6] S. Evra and T. Kaufman, Bounded degree cosystolic expanders of every dimension, in *Proc. of the 48th Annual ACM SIGACT Symposium on Theory of Computing* (2016), 36–48.
- [7] J. Fox, M. Gromov, V. Lafforgue, A. Naor and J. Pach, Overlap properties of geometric expanders, *J. Reine Angew. Math.* **671** (2012), 49–83.
- [8] M. Gromov, Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry, *Geom. Funct. Anal.* **20** (2010), 416–526.
- [9] A. Gundert and Z. Luria, personal communication.
- [10] S. Hoory, N. Linial and A. Wigderson, Expander graphs and their applications, *Bull. Amer. Math. Soc.* **43** (2006), 439–561.
- [11] T. Kaufman, D. Kazhdan and A. Lubotzky, Isoperimetric inequalities for Ramanujan complexes and topological expanders, *Geom. Funct. Anal.* **26** (2016), 250–287.
- [12] T. Kaufman and D. Mass, High dimensional random walks and colorful expansion, arXiv:1604.02947.
- [13] W.-C. W. Li, Ramanujan hypergraphs, *Geom. Funct. Anal.* **14** (2004), 380–399.
- [14] A. Lubotzky, Expander graphs in pure and applied mathematics, *Bull. Amer. Math. Soc.* **49** (2012), 113–162.
- [15] A. Lubotzky, High dimensional expanders, in *Proc. of the Int. Congress of Mathematicians 2018* (2019).
- [16] A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan graphs, *Combinatorica* **8** (1988), 261–277.
- [17] A. Lubotzky, B. Samuels and U. Vishne, Ramanujan complexes of type $\tilde{A}_d$, *Israel J. Math.* **149** (2005), 267–299.
- [18] A. Lubotzky, B. Samuels and U. Vishne, Explicit constructions of Ramanujan complexes of type $\tilde{A}_d$, *European J. Combin.* **26** (2005), 965–993.
- [19] A. W. Marcus, D. A. Spielman and N. Srivastava, Interlacing families I: Bipartite Ramanujan graphs of all degrees, *Ann. of Math.* **182** (2015), 307–325.

- [20] G. A. Margulis, Explicit constructions of expanders, *Problemy Peredači Informacii*, **9** (1973), 71–80.
- [21] G. A. Margulis, Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators, *Problems Inform. Transmission* **24** (1988), 39–46.
- [22] J. Pach, A Tverberg-type result on multicolored simplices, *Comput. Geom.* **10** (1998), 71–76.
- [23] O. Parzanchevski, Mixing in high-dimensional expanders, *Combin. Probab. Comput.* **26** (2017), 746–761.
- [24] O. Parzanchevski, R. Rosenthal and R. Tessler, Isoperimetric inequalities in simplicial complexes, *Combinatorica* **36** (2016), 195–227.
- [25] M. S. Pinsker, On the complexity of a concentrator, in 7th International Teletraffic Conference, 318/1–318/4, 1973.
- [26] O. Reingold, S. Vadhan and A. Wigderson, Entropy waves, the zig-zag graph product, and new constant-degree expanders, *Ann. of Math.* **155** (2002), 157–187.
- [27] A. Sarveniazi, Explicit construction of a Ramanujan $(n_1, n_2, \dots, n_{d-1})$ -regular hypergraph, *Duke Math. J.* **139** (2007), 141–171.