OVERLAP PROPERTIES OF GEOMETRIC EXPANDERS

JACOB FOX, MIKHAIL GROMOV, VINCENT LAFFORGUE, ASSAF NAOR, AND JÁNOS PACH

Abstract. The *overlap number* of a finite $(d+1)$-uniform hypergraph H is the largest constant $c(H) \in (0,1]$ such that no matter how we map the vertices of H into $\mathbb{R}^d$, there is a point covered by at least a $c(H)$-fraction of the simplices induced by the images of its hyperedges. In [17], motivated by the search for an analogue of the notion of graph expansion for higher dimensional simplicial complexes, it was asked whether or not there exists a sequence $\{H_n\}_{n=1}^{\infty}$ of arbitrarily large $(d+1)$-uniform hypergraphs with bounded degree, for which $\inf_{n\geqslant 1} c(H_n) > 0$. Using both random methods and explicit constructions, we answer this question positively by constructing infinite families of $(d+1)$-uniform hypergraphs with bounded degree such that their overlap numbers are bounded from below by a positive constant $c = c(d)$. We also show that, for every d, the best value of the constant $c = c(d)$ that can be achieved by such a construction is asymptotically equal to the limit of the overlap numbers of the complete $(d+1)$-uniform hypergraphs with n vertices, as $n \to \infty$. For the proof of the latter statement, we establish the following geometric partitioning result of independent interest. For any d and any $\varepsilon > 0$, there exists $K = K(\varepsilon, d) \geqslant d+1$ satisfying the following condition. For any $k \geqslant K$, for any point $q \in \mathbb{R}^d$ and for any finite Borel measure μ on $\mathbb{R}^d$ with respect to which every hyperplane has measure 0, there is a partition $\mathbb{R}^d = A_1 \cup \ldots \cup A_k$ into k measurable parts of equal measure such that all but at most an ε-fraction of the $(d+1)$-tuples $A_{i_1}, \ldots, A_{i_{d+1}}$ have the property that either all simplices with one vertex in each A_{i_j} contain q or none of these simplices contain q.

Contents

*****Dedicated to Endre Szemerédi on his 70th birthday*****.

1. Introduction

Let $G = (V, E)$ be an n-vertex graph. Think of G as a 1-dimensional simplicial complex, i.e., each edge is present in G as an actual interval. Assume that for every subset $S \subseteq V$ of size $\lfloor \frac{n}{2} \rfloor$ the number of edges joining S and $V \setminus S$ is at least $\alpha|E|$, for $\alpha \in (0, 1]$. It follows that for every $f : V \to \mathbb{R}$, if we extend f to be a linear (or even just continuous) function defined also on the edges of G, there must necessarily exist a point $x \in \mathbb{R}$ such that $|f^{-1}(x)| \geqslant \alpha|E|$. Indeed, x can be chosen to be a median of the set $f(V) \subseteq \mathbb{R}$. In other words, no matter how we draw G on the line, its edges will heavily overlap.

As illustrated by this simple example, the above expander-like condition[1] on G implies that all of its embeddings in $\mathbb{R}$ satisfy a geometric overlap condition. This condition naturally extends to higher-dimensional simplicial complexes, and can thus serve as a potential definition of a higher-dimensional analogue of edge expansion[2]. Such investigations of high-dimensional geometric analogues of edge expansion were initiated by Gromov in [17]. The present paper follows this approach.

In 1984, answering a question of Kárteszi, two undergraduates at Eötvös University, Boros and Füredi [8], proved the following theorem.

Theorem 1.1 ([8]). *For every set P of n points in the plane, there is a point (not necessarily in P) that belongs to at least $\left(\frac{2}{9} - o(1)\right)\binom{n}{3}$ closed triangles induced by the elements of P.*

The factor $\frac{2}{9}$ in Theorem 1.1 is asymptotically tight, as shown by Bukh, Matoušek and Nivasch in [10]. A short and elegant "book proof" of Theorem 1.1 was given by Bukh [9]. In Section 2, we present an alternative "topological" argument.

The theorem of Boros and Füredi has been generalized to higher dimensions. Bárány [4] proved that for every $d \in \mathbb{N}$ there exists a constant $c_d > 0$ such that given any set P of n points in $\mathbb{R}^d$, one can always find a point in at least $c_d n^d$ closed simplices whose vertices belong to P. In fact, the following stronger statement due to Pach [27] holds true.

Theorem 1.2. *([27]) Every set P of n points in $\mathbb{R}^d$ has $d+1$ disjoint $\lfloor c'_d n \rfloor$-element subsets, $P_1, \ldots, P_{d+1}$, such that* all *closed simplices with one vertex from each P_i have a point in common. Here $c'_d > 0$ is a constant depending only on the dimension d.*

Recall that a hypergraph $H = (V, E)$ consists of a set V and a set E of non-empty subsets of V. The elements of V are called vertices and the elements of E are called hyperedges. H is *d-uniform* if every hyperedge $e \in E$ contains exactly d vertices. The *degree* of a vertex $v \in V$ in H is the number of hyperedges containing v. To simplify the presentation, we introduce the following terminology.

Definition 1.1. Given a $(d+1)$-uniform hypergraph $H = (V, E)$, its *overlap number* $c(H)$ is the largest constant $c \in (0, 1]$ such that for every embedding $f : V \to \mathbb{R}^d$, there exists a point $p \in \mathbb{R}^d$ which belongs to at least $c|E|$ simplices whose vertex sets are hyperedges of H, i.e., there exists a set of hyperedges $S \subseteq E$ with $|S| \geqslant c|E|$ and $p \in \bigcap_{e \in S} \operatorname{conv}(f(e))$

[1] It isn't quite edge expansion since we do not care about boundaries of small sets.

[2] To be precise, what we are detecting here is only that G contains a large expander, rather than being an expander itself.

(where $\mathrm{conv}(A)$ denotes the convex hull of $A \subseteq \mathbb{R}^d$). An infinite family $\mathscr{H}$ of $(d+1)$-uniform hypergraphs is *highly overlapping* if there exists an absolute constant $c > 0$ such that $c(H) > c$ for every $H \in \mathscr{H}$. An infinite family of *d-dimensional simplicial complexes* is called *highly overlapping* if the family of $(d+1)$-uniform hypergraphs consisting of the vertex sets of their d-dimensional faces (their d-skeletons) is highly overlapping[3].

Using this terminology, the Boros-Füredi theorem states that the family of all finite complete 3-uniform hypergraphs (or 2-skeletons of all complete simplicial complexes) is highly overlapping. Bárány's theorem says that the same is true for the family of complete $(d+1)$-uniform hypergraphs (or d-skeletons of complete simplicial complexes). The fact that the family of all finite complete graphs (1-skeletons of complete simplicial complexes) is highly overlapping (with $c = 1/2$) is trivial, but its higher dimensional generalizations are much more subtle.

It was a simple but very important graph-theoretic discovery by Pinsker [29] and others that there exist arbitrarily large edge *expanders* of bounded degree [18]. As we have seen at the beginning of this paper, expanders with a fixed rate of expansion are necessarily highly overlapping. This fact motivated Gromov's question [17] whether there exist infinite families of higher dimensional simplicial complexes with bounded degree that are highly overlapping. In other words, Gromov's question [17] for 2-dimensional simplicial complexes asks whether a Boros-Füredi type theorem remains true if instead of *all* triangles determined by n points in the plane, we consider only "sparse" systems of triangles. In particular, do there exist arbitrarily large 3-uniform hypergraphs H, in which every vertex belongs to at most a constant number k of triples, and whose overlap numbers are bounded from below by an absolute positive constant?

In Section 3.1, we answer this question in the affirmative, by proving the following result.

Theorem 1.3. *For any $\varepsilon > 0$, there exists a positive integer $k = k(\varepsilon)$ satisfying the following condition. There is an infinite sequence of 3-uniform hypergraphs H_n with n vertices and n tending to infinity, each of degree k, such that, for any embedding of the vertex set $V(H_n)$ in $\mathbb{R}^2$, there is a point belonging to at least a $(\frac{2}{9} - \varepsilon)$-fraction of all closed triangles induced by images of hyperedges of H_n. Here the constant $\frac{2}{9}$ cannot be improved.*

We also generalize Theorem 1.3 to $(d+1)$-uniform hypergraphs with $d \geqslant 2$.

Theorem 1.4. *For every integer $d \geqslant 2$, there exist positive constants c_d and k_d with the following property. There is an infinite sequence of $(d+1)$-uniform hypergraphs H_n with n vertices and n tending to infinity, each of degree k_d, such that, for any embedding of the vertex set $V(H_n)$ in $\mathbb{R}^d$, there is a point in $\mathbb{R}^d$ that belongs to at least a c_d-fraction of all closed simplices induced by images of hyperedges of H_n.*

Among the most natural and powerful methods to construct good expanders is the use of certain Cayley graphs of finitely generated groups (see [20, 23, 14]), via arguments related to Kazhdan's property (T) (see [5]). Such graphs yield explicit constructions of expanders that have extremal spectral properties, namely Ramanujan graphs [20]. Being Cayley graphs of

[3] Gromov [17] calls such simplicial complexes "polyhedra with large cardinalities."

finitely generated groups, these constructions can be viewed as quotients of trees (Cayley graphs of free groups). It is natural to study hypergraph versions of this type of construction, based on quotients buildings (a type of higher dimensional simplicial complexes that extends the notion of a tree [33]). In particular, a notion of Ramanujan complex, which is a simplicial complex with extremal spectral properties analogous to Ramanujan graphs, was introduced and constructed in [3, 11, 19, 22, 21, 31]. Here we show that such constructions can yield highly overlapping bounded degree hypergraph families. Specifically, we show that for every integer $r \geqslant 2$, for a large enough odd prime power q, certain finite quotients of the building of $PGL_r(F)$, where F is a non-archimedian local field with residue field of order q, are highly overlapping r-uniform hypergraphs (with degree and overlap number depending only q, r). Rather than defining the relevant notions in the introduction, we refer to Section 5 for precise definitions and statements. Instead, we state below the following concrete special case of our result, which follows from our argument in Section 5, in combination with a construction of Lubotzky, Samuels and Vishne [21].

Theorem 1.5. *For every odd prime p and every integer $r \geqslant 3$ there exist $k(p,r) \in \mathbb{N}$ and $c(p,r) > 0$ with the following property. For every $m \in \mathbb{N}$, the finite group $G = PGL_r(\mathbb{F}_{p^m})$, where $\mathbb{F}_{p^m}$ is the field of cardinality p^m, has a symmetric generating set $S \subseteq G$ of size bounded above by $k(p,r)$, such that the following holds. Consider the r-regular hypergraph H whose vertex set is G and whose hyperedges are those r-tuples $\{g_1, \ldots, g_r\} \subseteq G$ with $g_i g_j^{-1} \in S$ for all distinct $1 \leqslant i, j \leqslant r$ (i.e., H is the hypergraph consisting of all cliques of size r in the Cayley graph induced by S). Then there exist arbitrarily large integers m for which the hypergraph H has overlap number at least $c(p,r) > 0$.*

By Theorem 1.3, the best value of the constant c_2 in Theorem 1.4 is close to $\frac{2}{9}$, but in higher dimensions $d > 2$, we do not have very good estimates for c_d. Our goal is to show, roughly speaking, that the best constant in Theorem 1.4 is the same as the best constant in the Boros-Füredi-Bárány theorem (Theorem 1.1). To state this formally, it will be convenient to introduce some notation. Let $c(K_n^{d+1})$ be the overlap number of K_n^{d+1}, the complete $(d+1)$-uniform hypergraph on n vertices, and set

$$c(d) = \lim_{n \to \infty} c\left(K_n^{d+1}\right).$$

It is easy to show, via a straightforward point duplication argument, that the limit defining $c(d)$ exists, and the Boros-Füredi-Bárány theorem shows that $c(d) > 0$, for every d.

One might suspect that if H is a $(d+1)$-uniform hypergraph without isolated vertices, then $c(H) \leqslant c(d) + o(1)$, where the $o(1)$ term goes to 0 as the number of vertices of H tends to infinity. This is not the case. Consider, for example, the $(d+1)$-hypergraph H_n^{d+1} on n vertices, whose hyperedges are those sets of size $d+1$ that contain the first d vertices. In any general position embedding of the vertices of H_n^{d+1} in $\mathbb{R}^d$, any segment joining a pair of points sufficiently close and on opposite sides of the face consisting of the first d vertices stabs all the simplices induced by the images of hyperedges of H_n^{d+1}. Hence, $c(H_n^{d+1}) \geqslant 1/2$. However, $c(d)$ decays to 0 at least exponentially in d (see, e.g., [4, 10]). Despite this example, we show that our suspicion is *correct* for bounded degree hypergraphs.

Theorem 1.6. *For any d, $\Delta \in \mathbb{N}$, and $\varepsilon > 0$, there is $n(d, \Delta, \varepsilon) \in \mathbb{N}$ such that every $(d+1)$-uniform hypergraph H on $n \geqslant n(d, \Delta, \varepsilon)$ nonisolated vertices with maximum degree Δ satisfies $c(H) \leqslant c(d) + \varepsilon$.*

In the other direction, we show that there are regular $(d+1)$-uniform hypergraphs H of bounded degree such that $c(H)$ is at least $c(d) - \varepsilon$ for any given $\varepsilon > 0$.

Theorem 1.7. *For each $d \in \mathbb{N}$ and $\varepsilon > 0$, there is $r(d, \varepsilon) \in \mathbb{N}$ such that for every $r \geqslant r(d, \varepsilon)$ and sufficiently large n which is a multiple of $d+1$, there is a $(d+1)$-uniform, r-regular hypergraph H on n vertices with $c(H) \geqslant c(d) - \varepsilon$.*

The previous two theorems essentially show that $c(d)$ is the largest possible overlap number for bounded degree hypergraphs with sufficiently many nonisolated vertices.

The proof of the last theorem is based on a geometric partitioning result of independent interest. A $(d+1)$-tuple of subsets $S_1, \ldots, S_{d+1} \subseteq \mathbb{R}^d$ is said to be *homogeneous* with respect to a point $q \in \mathbb{R}^d$ if either all simplices with one vertex in each of the sets $S_1, \ldots, S_{d+1}$ contain q, or none of these simplices contain q.

Theorem 1.8. *For a positive integer d and $\varepsilon > 0$, there exists another positive integer $K = K(\varepsilon, d) \geqslant d+1$ such that for any $k \geqslant K$ the following statement is true. For any point $q \in \mathbb{R}^d$ and for any finite Borel measure μ on $\mathbb{R}^d$ with respect to which every hyperplane has measure 0, there is a partition $\mathbb{R}^d = A_1 \cup \ldots \cup A_k$ into k measurable parts of equal measure such that all but at most an ε-fraction of the $(d+1)$-tuples $A_{i_1}, \ldots, A_{i_{d+1}}$ are homogenous with respect to q.*

An *equipartition* of a finite set is a partition of the set into subsets whose sizes differ by at most one. A discrete version of Theorem 1.8 is the following.

Corollary 1.9. *Given a positive integer d and $\varepsilon > 0$, there exists another positive integer $K = K(\varepsilon, d) \geqslant d+1$ such that for any $k \geqslant K$ the following statement is true. For any finite set $P \subseteq \mathbb{R}^d$ and for any point $q \in \mathbb{R}^d$, there is an equipartition $P = P_1 \cup \ldots \cup P_k$ such that all but at most an ε-fraction of the $(d+1)$-tuples $P_{i_1}, \ldots, P_{i_{d+1}}$ are homogenous with respect to q.*

Notice that due to Bárány's result [4] that $c(d) > 0$, by taking $\varepsilon \ll c(d)$, Corollary 1.9 immediately implies Theorem 1.2.

The rest of the paper is organized as follows. Section 2 contains a detailed topological proof of the Boros-Füredi theorem (Theorem 1.1), following Gromov's approach in [17]. In the two subsections of Section 3, we present randomized constructions for Theorems 1.3 and 1.4. In the plane, these constructions are nearly optimal; their overlap numbers are close to the value $\frac{2}{9}$. In Section 4, we give a deterministic recipe how to turn certain families of explicitly given expander graphs into families of highly overlapping $(d+1)$-uniform hypergraphs. In Section 5 we give a criterion which ensures that certain finite quotients of the building of $PGL_r(F)$ are highly overlapping r-uniform hypergraphs; this criterion implies in particular Theorem 1.5. In Section 6, we establish a Szemerédi-type theorem for infinite hypergraphs with a measure on their vertex sets (Theorem 6.1). This is used in Section 7 for the proof of the geometric partition result Theorem 1.8. In Section 8, we show how this result can be applied to obtain Theorem 1.7. Section 9 contains the proof of Theorem 1.6.

For the sake of clarity of the presentation, in the rest of this paper, we systematically omit floor and ceiling signs whenever they are not crucial. We shall also assume throughout that all embeddings of hypergraphs into $\mathbb{R}^d$ are such that the vertices are mapped to points in general position. Even though the corresponding statements for degenerate embeddings will then follow from standard limiting arguments, it is convenient to make this assumption in order to not deal explicitly with such degeneracies in each of the proofs.

2. A topological proof of the Boros-Füredi theorem

We will prove a somewhat stronger statement. Given a set P of n points in the plane, a *ray* (closed half-line) is said to be *exposed* if it has nonempty intersection with fewer than $n^2/9$ segments connecting point pairs in P. The set of all segments connecting two elements of P forms a *complete geometric graph* $K(P)$ on the vertex set P, and we refer to these segments as the *edges* of $K(P)$.

Proposition 2.1. *Given a set P of n points in the plane, one can always find a point q not necessarily in P such that no ray emanating from q is exposed.*

Suppose that such a point q does not belong to P. For each $p \in P$, ray emanating from q in the direction opposite to p intersects at least $n^2/9$ edges of $K(P)$. Each such edge, together with p, spans a triangle that contains q. Every triangle is counted at most *three* times, therefore the total number of triangles containing q is at least $n(n^2/9)/3 = n^3/27$. If q belongs to P, the number of (closed) triangles containing q is larger than $n^3/27$.

Thus, it is sufficient to prove Proposition 2.1. Suppose for a contradiction that for each point q of the plane, there is an exposed ray emanating from q. Let D denote a large disk around the origin O, which contains all elements of P, and let S^1 denote the boundary of D. For $\sigma \in \mathbb{R}^2 \setminus \{O\}$, we denote by $\mathrm{ray}(q, \sigma)$ the ray emanating from q in the direction parallel to $\overrightarrow{O\sigma}$.

Notice that for any two exposed rays, $\mathrm{ray}(q, \sigma)$ and $\mathrm{ray}(q, \tau)$, emanating from the same point, one of the two closed regions bounded by them contains fewer than $n/3$ points of P. Otherwise, one of the regions has x points of P with $n/3 \leqslant x \leqslant 2n/3$, and the two boundary rays together would intersect at least $x(n-x) \geqslant (n/3)(2n/3) = 2n^2/9$ edges, which implies at least one of them was not exposed.

Let I denote the set of all pairs $(q, \varrho) \in D \times S^1$, for which $\mathrm{ray}(q, \varrho)$ is exposed or belongs to the closed region bounded by two exposed rays, $\mathrm{ray}(q, \sigma)$ and $\mathrm{ray}(q, \tau)$, that contains fewer than $n/3$ points of P.

Claim 2.2. *The set I has the following properties:*

(a) *I is an* open *subset of $D \times S^1$,*
(b) *$(\varrho, \varrho) \in I$ for all $\varrho \in S^1$,*
(c) *for every $q \in D$, the set $I_q \overset{\text{def}}{=} \{\varrho \in S^1 : \ (q, \varrho) \in I\}$ is a nonempty proper subinterval of S^1.*

Proof. Parts (a) and (b) directly follow from the definition. It is also clear, by our contrapositive assumption, that I_q is a nonempty interval for every $q \in D$.

We have to show only that $I_q \neq S^1$. To see this, let $\mathrm{ray}(q, \varrho)$ be an exposed ray emanating from q, and let $\varrho' \in S^1$ be a direction such that both closed regions bounded by $\mathrm{ray}(q, \varrho)$ and $\mathrm{ray}(q, \varrho')$ contain at least $n/2$ points of P.

We claim that $\varrho' \notin I_q$. Otherwise, we can select two exposed rays, $\mathrm{ray}(q, \sigma)$ and $\mathrm{ray}(q, \tau)$, such that $\mathrm{ray}(q, \varrho')$ belongs to the closed region bounded by them which contains fewer than $n/3$ points. The *three* rays, $\mathrm{ray}(q, \varrho)$, $\mathrm{ray}(q, \sigma)$, and $\mathrm{ray}(q, \tau)$, cut the plane into *three* closed regions, and it is easy to see that each of them must contain fewer than $n/3$ points, which is a contradiction. Indeed, if e.g. the region bounded by $\mathrm{ray}(q, \varrho)$ and $\mathrm{ray}(q, \sigma)$ that does not contain $\mathrm{ray}(q, \tau)$ had at least $n/3$ points, then by the discussion above the closure of its complement had fewer than $n/3$ points, contradicting our assumption that both closed regions bounded $\mathrm{ray}(q, \varrho)$ and $\mathrm{ray}(q, \varrho')$ contain at least $n/2$ points. □

Now we can obtain the desired contradiction, thus completing the proof of Proposition 2.1, by applying to $J \overset{\mathrm{def}}{=} (D \times S^1) \setminus I$ the following version of the Brouwer fixed point theorem.

Lemma 2.3. *Let J be a closed subset of $D \times S^1$ with the property that for every $q \in D$ we have that $J_q \overset{\mathrm{def}}{=} \{\varrho \in S^1 :\ (q, \varrho) \in J\}$ is a nonempty proper (closed) subinterval of S^1. Then $(\varrho, \varrho) \in J$, for some $\varrho \in S^1$.*

To see why Lemma 2.3 holds true, assume for contradiction that $(\varrho, \varrho) \notin J$, for all $\varrho \in S^1$. Write $J_S \overset{\mathrm{def}}{=} J \cap (S^1 \times S^1)$, and let $Proj_1, Proj_2 : J_S \to S^1$ denote the projections onto the first and second coordinates, respectively. The fibers of $Proj_1$ are nonempty proper closed intervals, and therefore $Proj_1$ induces a bijection between $\pi_1(J_S)$ and $\pi_1(S^1) = \mathbb{Z}$. But, the contrapositive assumption implies that $Proj_1$ and $Proj_2$ are homotopic, and therefore $Proj_2$ also induces a bijection between $\pi_1(J_S)$ and $\pi_1(S^1)$. This is a contradiction since $Proj_2$ extends to J, and $\pi_1(J) = 0$ since J is fibered over D with fibers equal to intervals.

Clearly, Lemma 2.3 contradicts part (b) of Claim 2.2. □

3. Sparse constructions using the probabilistic method

In this section, we prove Theorems 1.3 and 1.4 using the probabilistic method. Our planar construction is nearly optimal, but in higher dimensions the overlap numbers of our hypergraphs will be far from maximal. We note that our proofs use a non-uniformly random choice of $(d+1)$-uniform hypergraphs of degree k_d, which is designed especially for our purposes. Nevertheless, the argument in Section 8, which uses Theorem 1.8, shows that assuming the degree r satisfies a large enough lower bound depending on d (which is inferior to the bound on k_d obtained in this section), for a hypergraph H chosen uniformly at random among all $(d+1)$-uniform hypergraphs of degree r, with high probability $c(H)$ will be bounded below by a positive constant depending only on d (which is also inferior to the bound on c_d obtained in this section).

3.1. **Highly overlapping triple systems—Proof of Theorem 1.3.** The outline of the proof of Theorem 1.3 is the following. We first pick t randomly and independently selected partitions of the set $[n] = \{1, 2, \ldots, n\}$ into parts of equal size b. We define H_n to be the

3-uniform hypergraph with vertex set $[n]$, consisting of all triples that lie in the same part in at least one of the t partitions. Finally, we will show that H_n meets the requirements of Theorem 1.3.

We need the following simple technical lemma. A key ingredient that is used in the proof is the Chernoff bound for negatively associated random variables (see, e.g., [15]). It implies that if $A_1, \ldots, A_n$ are n mutually negatively correlated events in an arbitrary probability space such that A_i has probability p_i, then the probability that the number of A_i which occur exceeds the expected number $p_1 + \cdots + p_n$ by at least a is at most $e^{-2a^2/n}$.

Lemma 3.1. *Suppose that $\delta > 0$, and let $b = \delta^{-3}$, $\beta = 2e^{-2\delta^2 b}$, $r = 4\beta^{-2}b$, $t = r\delta^{-1}$. If n is a sufficiently large multiple of b, then there exist t partitions $\mathcal{P}_1, \ldots, \mathcal{P}_t$ of $[n]$, each consisting of n/b parts of size b, with the following two properties:*

1. *any two parts of size b in different partitions have at most* two *elements in common,*
2. *for every subset $S \subseteq [n]$, there are fewer than r partitions $\mathcal{P}_i$ for which at least $\beta n/b$ parts contain at least $\left(\frac{|S|}{n} + \delta\right) b$ elements of S.*

Proof. We verify that t *randomly* selected partitions of $[n]$ into parts of equal size b almost surely have the desired properties. Fix a set $S \subseteq [n]$, and consider a random partition $\mathcal{P}$ of $[n]$ into parts $I_1, \ldots, I_{n/b}$ of size b. For any $1 \leqslant i \leqslant n/b$, let A_i denote the event that $|I_i \cap S| \geqslant \left(\frac{|S|}{n} + \delta\right) b$. For any $1 \leqslant j \leqslant b$, let $A_{i,j}$ denote the event that the jth element of I_i is in S. The events $A_{i,1}, \ldots, A_{i,b}$ are mutually negatively correlated and each of them has probability $|S|/n$. Thus, by Chernoff's bound [15], we have

$$\Pr[A_i] \leqslant e^{-2(\delta b)^2/b} = e^{-2\delta^2 b} = \frac{\beta}{2}.$$

Let X denote the event that at least $\beta n/b$ of the events $A_1, \ldots, A_{n/b}$ occur. Since the events $A_1, \ldots, A_{n/b}$ are also mutually negatively correlated and each has probability at most $\beta/2$, we can again apply the Chernoff bound [15] to obtain

$$\Pr[X] \leqslant e^{-2(\frac{\beta n}{2b})^2/(n/b)} = e^{-\frac{1}{2}\beta^2 n/b}.$$

Take t independent random partitions of $[n]$, $\mathcal{P}_1, \ldots, \mathcal{P}_t$, each consisting of n/b parts of size b. The probability that a given pair of parts of size b have at least 3 elements in common is at most $\binom{b}{3}\left(\frac{b}{n}\right)^3 \leqslant \frac{b^6}{6n^3}$. Since there are $\binom{tn/b}{2}$ such pairs, by linearity of expectation, the probability that there is a pair sharing at least 3 elements is at most $\binom{tn/b}{2}\frac{b^6}{6n^3} < \frac{t^2 b^4}{12n}$. Hence, by our choice of parameters, almost surely condition (1) will be satisfied.

For a fixed $S \subseteq [n]$, the probability that for at least r of the partitions $\mathcal{P}_1, \ldots, \mathcal{P}_t$, at least $\beta n/b$ of the b-element subsets of the partition have at least $\left(\frac{|S|}{n} + \delta\right) b$ elements in S is at most

$$\binom{t}{r}(\Pr[X])^r \leqslant \binom{t}{r} e^{-r\frac{1}{2}\beta^2 n/b} = \binom{t}{r} e^{-2n} \leqslant e^{-n}.$$

The number of subsets S of $[n]$ is 2^n. Hence, by linearity of expectation, the expected number of subsets S with property (2) is $o(1)$. We conclude that there are t such partitions with the desired properties. $\square$

Let $\delta = \varepsilon/50$ and $k = t\binom{b-1}{2}$. Consider the 3-uniform hypergraph H_n with $V(H_n) = [n]$, the hyperedges of which are those triples that lie in the same part in at least one (hence, precisely one) of the partitions $\mathcal{P}_1, \dots, \mathcal{P}_t$ meeting the requirements of Lemma 3.1. Clearly, in H_n, each vertex belongs to $k = t\binom{b-1}{2}$ hyperedges.

The proof of Theorem 1.3 can now be completed by adapting the idea of Bukh [9]. Consider an embedding of the vertices of H_n in the plane. We shall use the following lemma of Ceder [13]:

Lemma 3.2 (Ceder [13]). *Assume that n is divisible by 6. Given any set of n points in the plane, there are three concurrent lines that divide the plane into 6 angular regions, each containing roughly the same number of points. More precisely, there are disjoint $\frac{n}{6}$-element point sets $S_1, \dots, S_6$ such that S_i is contained in the closure of region i.*

We shall assume throughout the n is divisible by 6. Let $S_1, \dots, S_6$ be the sets from Lemma 3.2, and let p denote the intersection point of the *three* lines from Lemma 3.2. By a simple case analysis, Bukh [9] showed that, for every choice of *six* points, one from each S_i, at least 8 of the $\binom{6}{3} = 20$ triangles induced by them contain p.

Let $I \subseteq [n]$ be a b-element set such that $|I \cap S_i| \leqslant (\frac{|S_i|}{n} + \delta)b = (1 + 6\delta)\frac{b}{6}$, for $1 \leqslant i \leqslant 6$. Obviously, we have

$$|I \cap S_i| \geqslant b - 5(1 + 6\delta)\frac{b}{6} \geqslant (1 - 30\delta)\frac{b}{6},$$

for every i. Each of the

$$\prod_{i=1}^{6} |I \cap S_i| \geqslant (1 - 30\delta)^6 \left(\frac{b}{6}\right)^6$$

6-element sets with one vertex from each $I \cap S_i$ induces at least 8 triangles that contain point p. Each of these triangles belongs to at most $(1 + 6\delta)^3(\frac{b}{6})^3$ such 6-element sets. Thus, there are at least

$$8\frac{(1 - 30\delta)^6 \left(\frac{b}{6}\right)^6}{(1 + 6\delta)^3(\frac{b}{6})^3} \geqslant \frac{1}{27}(1 - 200\delta)b^3 > (1 - 200\delta)\frac{2}{9}\binom{b}{3}$$

triangles induced by *three* vertices in I which contain p.

According to part 2 of Lemma 3.1, for every i, $1 \leqslant i \leqslant 6$, fewer than r partitions $\mathcal{P}_j$ have the property that at least $\beta\frac{n}{b}$ of their parts contain at least $(1 + 6\delta)\frac{b}{6}$ elements of S_i. Hence, the total number of b-element parts I in all t partitions, for which $|I \cap S_i| > (1 + 6\delta)\frac{b}{6}$ for some i, $1 \leqslant i \leqslant 6$, is smaller than

$$6r\frac{n}{b} + 6t\beta\frac{n}{b} = 6\delta t\frac{n}{b} + 6\beta t\frac{n}{b} \leqslant 10\delta t\frac{n}{b}.$$

It follows that the fraction of the $t\frac{n}{b}\binom{b}{3}$ hyperedges of H_n that contain point p in this embedding is at least

$$(1 - 10\delta)(1 - 200\delta)\frac{2}{9} \geqslant (1 - 210\delta)\frac{2}{9} \geqslant \frac{2}{9} - \varepsilon,$$

which completes the proof of Theorem 1.3. $\square$

3.2. **Higher dimensions—Proof of Theorem 1.4.** As in the proof of Theorem 1.3, we establish Theorem 1.4 using Lemma 3.1. We may assume that $c'_d = 1/m$ with m an integer, where c'_d is the constant in Theorem 1.2, and let n be a multiple of m. Set $\delta = \frac{1}{2m(m-1)}$ and apply Lemma 3.1. Consider now the $(d+1)$-uniform hypergraph H_n with $V(H_n) = [n]$, the hyperedges of which are those $(d+1)$-element sets that lie in the same part in at least one (hence, precisely one) of the partitions $\mathcal{P}_1, \ldots, \mathcal{P}_t$ meeting the requirements of Lemma 3.1. Clearly, in H_n, each vertex belongs to $k_d = t\binom{b-1}{d}$ hyperedges.

Consider now any embedding of $V(H_n)$ into $\mathbb{R}^d$, and let P denote the image of $V(H_n)$. By Theorem 1.2, one can find disjoint $c'_d n$-element subsets $P_1, P_2, \ldots, P_{d+1} \subseteq P$ and a point q such that picking one element from each subset P_i, their convex hull always contains q. We extend this to a partition $P = P_1 \cup \ldots \cup P_m$ into subsets of size n/m by picking the P_i for $d+1 < i \leqslant m$ of size n/m arbitrarily.

Let $I \subseteq [n]$ be a b-element set such that

$$\forall 1 \leqslant i \leqslant m, \quad |I \cap P_i| \leqslant \left(\frac{|P_i|}{n} + \delta\right) b = \left(1 + \frac{1}{2(m-1)}\right) \frac{b}{m}.$$

Obviously, we have

$$|I \cap P_i| \geqslant b - (m-1)\left(1 + \frac{1}{2(m-1)}\right)\frac{b}{m} = \frac{b}{2m},$$

for every $1 \leqslant i \leqslant m$. Each of the

$$\prod_{i=1}^{d+1} |I \cap P_i| \geqslant \left(\frac{b}{2m}\right)^{d+1}$$

$(d+1)$-element sets with one vertex from each $I \cap P_1, \ldots, I \cap P_{d+1}$ induces a closed simplex containing point q. Hence, the fraction of $(d+1)$-element subsets of I which induce a closed simplex that contains point q is at least

$$\left(\frac{b}{2m}\right)^{d+1} \binom{b}{d+1}^{-1} \geqslant (d+1)! \left(\frac{c'_d}{2}\right)^{d+1}.$$

According to part (2) of Lemma 3.1, for every $1 \leqslant i \leqslant m$, fewer than r partitions $\mathcal{P}_j$ have the property that at least $\beta\frac{n}{b}$ of their parts contain at least $(\frac{|P_i|}{n} + \delta)b$ elements of P_i. Hence, the total number of b-element parts I in all t partitions, for which $|I \cap P_i| > (\frac{|P_i|}{n} + \delta)b$ for some $1 \leqslant i \leqslant m$, is smaller than

$$mr\frac{n}{b} + mt\beta\frac{n}{b} = m\delta t\frac{n}{b} + m\beta t\frac{n}{b} \leqslant \frac{3}{4} \cdot t \cdot \frac{n}{b}.$$

Hence, the fraction of the $t\frac{n}{b}\binom{b}{d+1}$ hyperedges of H_n that contain the point q in this embedding is at least $\frac{1}{4}(d+1)!\left(\frac{c'_d}{2}\right)^{d+1}$. □

4. Deterministic constructions using expander graphs

In the next two subsections, we present deterministic constructions based on expander graphs, to provide alternative proofs of Theorem 1.3 and Theorem 1.4. These proofs yield significantly better bounds on $k(\varepsilon)$ and k_d in Theorem 1.3 and Theorem 1.4, respectively.

As in the previous section, the proof gives a nearly optimal bound in the plane, but not in higher dimension.

4.1. **Highly overlapping triple systems—second proof of Theorem 1.3.** Fix integers $k, n \in \mathbb{N}$, with n divisible by 6, and let $G = (\{1, \dots, n\}, E)$ be a k-regular graph on the vertex set $\{1, \dots, n\}$. Let $k = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ be the eigenvalues of the adjacency matrix of G in decreasing order, and write $\lambda = \max_{i \in \{2,\dots,n\}} |\lambda_i|$. For any $S, T \subseteq \{1, \dots, n\}$ let $E(S,T)$ denote the number of ordered pairs $(i,j) \in S \times T$ such that $ij \in E$. The expander mixing lemma (see Corollary 9.2.5 in [1]) states that

$$\left| E(S,T) - \frac{k|S| \cdot |T|}{n} \right| \leqslant \lambda \sqrt{|S| \cdot |T|}. \tag{1}$$

For every $i \in \{1, \dots, n\}$ let $N_G(i) \stackrel{\mathrm{def}}{=} \{j \in \{1, \dots, n\} : \ ij \in E\}$ denote its neighborhood in G. Define a hypergraph H on the vertex set $\{1, \dots, n\}$ by letting $E(H)$ consist of those triples $\{i, j, \ell\}$ for which there exists $r \in \{1, \dots, n\}$ such that $ir, jr, \ell r \in E$, i.e., $i, j, \ell \in N_G(r)$. Assume from now on that the graph G is quadrilateral-free. This implies that the hyperedges in H corresponding to three vertices $i, j, \ell \in N_G(r)$ cannot arise from neighborhoods of vertices of G other than r itself. Hence the 3-uniform hypergraph corresponding to H is $k\binom{k-1}{2}$-regular and $|E(H)| = \binom{k}{3} n$.

Fix $\varepsilon, \delta \in (0,1)$. Let $\{P_i\}_{i=1}^{6}$ be a partition of $\{1, \dots, n\}$ such that $|P_j| = \frac{n}{6}$ for all $1 \leqslant j \leqslant 6$. Write

$$A_j = \left\{ i \in \{1, \dots, n\} : \ |N_G(i) \cap P_j| < \frac{(1-\delta)k}{6} \right\}.$$

Then, by definition, we have $E(A_j, P_j) < |A_j| \frac{(1-\delta)k}{6}$. An application of (1) yields the inequality:

$$|A_j| \frac{(1-\delta)k}{6} \geqslant \frac{k|A_j| \cdot |P_j|}{n} - \lambda \sqrt{|A_j| \cdot |P_j|} = \frac{k|A_j|}{6} - \lambda \sqrt{\frac{n|A_j|}{6}},$$

which simplifies to

$$|A_j| \leqslant \frac{6\lambda^2 n}{\delta^2 k^2}.$$

Thus, if we define

$$A = \left\{ i \in \{1, \dots, n\} : \ |N_G(i) \cap P_j| \geqslant \frac{(1-\delta)k}{6} \ \forall j \in \{1, \dots, 6\} \right\}, \tag{2}$$

then

$$A \geqslant n - \sum_{j=1}^{6} |A_j| \geqslant n \left(1 - \frac{36\lambda^2}{\delta^2 k^2} \right). \tag{3}$$

We shall assume from now on that $\frac{36\lambda^2}{\delta^2 k^2} < 1$. We also note that for every $i \in A$ and $j \in \{1, \dots, 6\}$ we have

$$|N_G(i) \cap P_j| \leqslant k - \sum_{r \in \{1,\dots,6\} \setminus \{j\}} |N_G(i) \cap P_r| \leqslant k - 5\frac{(1-\delta)k}{6} = \frac{(1+5\delta)k}{6}. \tag{4}$$

Let $x_1, \dots, x_n \in \mathbb{R}^2$ be an embedding of $\{1, \dots, n\}$ in the plane. Let $S_1, \dots, S_6$ be a partition of $\{x_1, \dots, x_n\}$, as in the first proof of Theorem 1.3, which corresponds to the three concurrent lines from Lemma 3.2, whose common intersection point is $p \in \mathbb{R}^2$. We shall use the above reasoning (and notation) for the partition $P_1, \dots, P_6$ of $\{1, \dots, n\}$ given by $P_j = \{i \in \{1, \dots, n\} : \ x_i \in S_i\}$.

Fix $i \in A$, where A is as in (2). For every $(j_1, \dots, j_6) \in \prod_{r=1}^{6}(N_G(i) \cap P_r)$ at least 8 of the 20 triangles induced by the points $\{x_{j_1}, \dots, x_{j_6}\}$ contain p. By the definition of A, there are at least $\left(\frac{(1-\delta)k}{6}\right)^6$ such 6-tuples, while, using (4), each of these triangles that contains p belongs to at most $\left(\frac{(1+5\delta)k}{6}\right)^3$ such 6-tuples. Observe also that by the definition of H, since all of these triangles correspond to neighbors of i, their corresponding triples of indices belong to $E(H)$, and since G is quadrilateral-free, they cannot arise from the above reasoning with i replaced by any other vertex. Thus, the number of triangles that are images of hyperedges of H and contain p is at least

$$8 \cdot \frac{\left(\frac{(1-\delta)k}{6}\right)^6}{\left(\frac{(1+5\delta)k}{6}\right)^3} \cdot |A| \overset{(3)}{\geqslant} \frac{(1-\delta)^6 k^3}{27(1+5\delta)^3} n \left(1 - \frac{36\lambda^2}{\delta^2 k^2}\right) = \left(1 - O\left(\delta + \frac{\lambda^2}{\delta^2 k^2} + \frac{1}{k}\right)\right) \cdot \frac{2}{9}\binom{k}{3} n. \quad (5)$$

For arbitrarily large n, we can choose the graph G so that it is quadrilateral-free and $\lambda \leqslant 2\sqrt{k}$ (e.g., Ramanujan graphs work—see [20, 18]). By choosing $\delta \asymp \varepsilon$ and $k \asymp \frac{1}{\varepsilon^3}$ in (5), we get that p is in at least $\left(\frac{2}{9} - \varepsilon\right)|E(H)|$ of the triangles in that are images of hyperedges of H. Note that the degree of H is $O(k^3) = O\left(\frac{1}{\varepsilon^9}\right)$. This proves Theorem 1.3 with the bound $k(\varepsilon) = O\left(\frac{1}{\varepsilon^9}\right)$. $\square$

4.2. **Higher dimensions—second proof of Theorem 1.4.** Here we shall use a variant of the construction in Section 4.1, to give an alternative proof of Theorem 1.4. We use the notation from Section 4.1, and we assume that $k \geqslant d$. Fix n vectors $x_1, \dots, x_n \in \mathbb{R}^d$. Define a set of d-dimensional simplices H' whose vertices are in $\{x_1, \dots, x_n\}$ by taking the simplex whose vertices are the distinct vectors $\{x_{j_1}, x_{j_2}, \dots, x_{j_{d+1}}\}$ if and only if we have $j_1 j_2, j_2 j_3, \dots j_d j_{d+1} \in E$. In other words, the simplices in H' correspond to non-returning walks of length d in G. Thus, $|H'| \leqslant k^d n$.

Let $P_1, \dots, P_{d+1} \subseteq \{x_1, \dots, x_n\}$ be the disjoint subsets from Theorem 1.2, i.e., $|P_i| \geqslant c'_d n$, and all the closed simplices with one vertex in each of the sets $\{P_1, \dots, P_{d+1}\}$ have a point in common. Set $Q_i = \{j \in \{1, \dots, n\} : \ x_j \in P_i\}$. Define $\widetilde{Q}_{d+1} = Q_{d+1}$ and inductively for $i \in \{2, \dots, d+1\}$,

$$\widetilde{Q}_{i-1} = \left\{ j \in Q_{i-1} : \ \exists \ell \in \widetilde{Q}_i \ j\ell \in E \right\}.$$

Then, by definition, there are no edges between $Q_{i-1} \setminus \widetilde{Q}_{i-1}$ and $\widetilde{Q}_i$. It follows from (1) that

$$\frac{k}{n}\left|Q_{i-1} \setminus \widetilde{Q}_{i-1}\right| \cdot \left|\widetilde{Q}_i\right| \leqslant \lambda \sqrt{\left|Q_{i-1} \setminus \widetilde{Q}_{i-1}\right| \cdot \left|\widetilde{Q}_i\right|}.$$

Thus, we have

$$\frac{\lambda^2 n^2}{k^2} \geqslant \left(|Q_{i-1}| - \left|\widetilde{Q}_{i-1}\right|\right)\left|\widetilde{Q}_i\right| \geqslant \left(c'_d n - \left|\widetilde{Q}_{i-1}\right|\right)\left|\widetilde{Q}_i\right|,$$

or

$$\left|\widetilde{Q}_{i-1}\right| \geqslant c'_d n - \frac{\lambda^2 n^2}{k^2 \left|\widetilde{Q}_i\right|}. \tag{6}$$

Assuming that $\lambda \leqslant \frac{c'_d}{2}k$, inequality (6) implies by induction that for all $i \in \{1, \ldots, d+1\}$ we have $\left|\widetilde{Q}_{i-1}\right| \geqslant \frac{c'_d}{2}n$ (for $i = d+1$ this follows from our assumption, arising from Theorem 1.2, on the cardinality of P_{d+1}). Thus, $\left|\widetilde{Q}_1\right| \geqslant \frac{c'_d}{2}n$, and by construction any point $j \in \widetilde{Q}_1$ can be completed to a walk in G of length d whose ith vertex is in Q_i. Each such walk corresponds to a simplex in H', and by Theorem 1.2, all of these simplices have a common point. Thus, the number of simplices in H' which have a common point is at least $\frac{c'_d}{2}n \geqslant \frac{c'_d}{2k^d}|H'|$. Since there exist arbitrarily large graphs G with $\lambda \leqslant \frac{c'_d}{2}k$ and $k \leqslant k_d$ (e.g., for Ramanujan graphs we can take $k_d \asymp \frac{1}{(c'_d)^2}$), this completes our deterministic proof of Theorem 1.4. □

5. Finite quotients of buildings

Let F be a non-archimedean local field, $\mathcal{O}_F$ its ring of integers, π_F a uniformizer and $q = |\mathcal{O}_F/\pi_F\mathcal{O}_F|$ the cardinality of the residue field of F. For example, we may take $F = \mathbb{Q}_p$, $\mathcal{O}_F = \mathbb{Z}_p$, $\pi_F = p$ and $q = p$. We assume below that q is odd.

Let $r \geqslant 3$ be an integer and $G = PGL_r(F)$. Now r is fixed but F will be chosen such that q is big enough. We recall that $K = PGL_r(\mathcal{O}_F)$ is a maximal compact subgroup of G. We also recall that G/K is the set of vertices of a building, and is also equal to the set of lattices in F^r up to homothety (a lattice in F^r is a free $\mathcal{O}_F$-submodule of rank r and a homothety is the multiplication by an element of the multiplicative group $F^\times$). We refer to [33] for an elementary introduction to the building of $PGL_r(F)$ (we will not use the definition of a building below—all simplicial complexes will be defined explicitly). We have the map

$$\mathrm{type} : G/K \to \mathbb{Z}/r\mathbb{Z}$$

such that if $x \in G/K$ is the homothety class of a lattice $M \subseteq F^r$ and $\det(M) = \pi_F^a \mathcal{O}_F^\times$ with $a \in \mathbb{Z}$, then $\mathrm{type}(x) = a \bmod r\mathbb{Z}$. We denote by $\mathrm{vdet}(\cdot)$ the composition

$$G \xrightarrow{\det} F^\times/(F^\times)^r \xrightarrow{\text{valuation}} \mathbb{Z}/r\mathbb{Z},$$

and for $i \in \mathbb{Z}/r\mathbb{Z}$ we write $G_i = \mathrm{vdet}^{-1}(\{i\})$. Thus, G_0 is a subgroup of index r in G and the G_i are the left and right cosets for G_0 in G. We remark that $K \subseteq G_0$. For $g \in G_i$ and $x \in G/K$ we have $\mathrm{type}(gx) = i + \mathrm{type}(x)$. Moreover, G_i/K is the subset of G/K of vertices of type i.

Let $\Lambda = \{(\lambda_1, ..., \lambda_r) \in \mathbb{Z}^r/\mathbb{Z}(1, ..., 1) : \ \lambda_1 \leqslant ... \leqslant \lambda_r\}$ be the set of dominant coweights of $PGL_r(F)$. For $(\lambda_1, ..., \lambda_r) \in \Lambda$ we write

$$D(\lambda_1, ..., \lambda_r) = \begin{pmatrix} \pi_F^{\lambda_1} & 0 & \ldots & \ldots & 0 \\ 0 & \pi_F^{\lambda_2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \pi_F^{\lambda_3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \pi_F^{\lambda_r} \end{pmatrix} \in G.$$

Then the mapping $\delta \stackrel{\text{def}}{=} (\lambda_1, ..., \lambda_r) \mapsto KD(\lambda_1, ..., \lambda_r)K$ is a bijection from Λ to $K\backslash G/K$. For $x, y \in G/K$ we have $x^{-1}y \in K\backslash G/K$, and we define the relative position of x and y in the building as $\sigma(x, y) = \delta^{-1}(x^{-1}y) \in \Lambda$. In other words, for $(\lambda_1, ..., \lambda_r) \in \Lambda$ we have $\sigma(x, y) = (\lambda_1, ..., \lambda_r)$ if and only if there exists a basis $(e_1, ..., e_r)$ of F^r such that x is the homothety class of $\mathcal{O}_F e_1 + ... + \mathcal{O}_F e_r$ and y is the homothety class of $\pi_F^{\lambda_1}\mathcal{O}_F e_1 + ... + \pi_F^{\lambda_r}\mathcal{O}_F e_r$. The inverse image of $K\backslash G_0/K$ under δ is

$$\Lambda_0 = \left\{ (\lambda_1, ..., \lambda_r) \in \mathbb{Z}^r/\mathbb{Z}(1, ..., 1) : \lambda_1 \leqslant ... \leqslant \lambda_r \ \wedge \ \sum_{i=1}^{r} \lambda_i = 0 \text{ in } \mathbb{Z}/r\mathbb{Z} \right\} \subseteq \Lambda,$$

which is also the set of dominant coweights of $SL_r(F)$.

We recall that for $d \in \{0, ..., r-1\}$ a d-dimensional face of the building is a $(d+1)$-tuple $\{y_0, ..., y_d\} \subseteq G/K$ such that there are lattices $M_0, ..., M_d$ in F^r satisfying

(1) $M_0 \subsetneq M_1 \subsetneq ... \subsetneq M_d \subsetneq \pi_F^{-1}M_0$,
(2) $y_0, ..., y_d$ are the homothety classes of $M_0, ..., M_d$.

The type of a face $\{y_0, ..., y_d\}$ is $\{\text{type}(y_0), ..., \text{type}(y_d)\} \subseteq \mathbb{Z}/r\mathbb{Z}$. For any non-empty subset $I \subseteq \mathbb{Z}/r\mathbb{Z}$ we denote by Y_I the set of faces of type I in the building. More precisely, we denote by Y_I the set of families $(y_i)_{i\in I}$ with $y_i \in G_i/K$ such that $\{y_i : i \in I\}$ is a $(|I| - 1)$-dimensional face of the building.

Let $G^+ \subseteq G$ be the subgroup generated by unipotent elements of G. Then G^+ is also the image of $SL_r(F)$ in $G = PGL_r(F)$, and $G^+ \subseteq G_0$. We have an exact sequence

$$1 \to G^+ \to G \xrightarrow{\det} F^\times/(F^\times)^r \to 1.$$

We will apply Theorem 1.1 of [26] (a quantitative form of Kazhdan's property (T)) to the strongly orthogonal system of the set of roots of $G = PGL_r(F)$ consisting of the single root $e_1 - e_r$ (where $e_i - e_j$ stands for the root corresponding to the character of the maximal torus of diagonal matrices which sends $D(\lambda_1, ..., \lambda_r) \in G$ to $\pi_F^{\lambda_i - \lambda_j}$, for $(\lambda_1, ..., \lambda_r) \in \Lambda$). By this theorem, for any unitary representation (H, π) of G without a nonzero G^+-invariant vector, for any $(\lambda_1, ..., \lambda_r) \in \Lambda$ and for any K-invariant vectors $\xi, \eta \in H$ we have

$$\left|\langle \eta, \pi(D(\lambda_1, ..., \lambda_r))\xi\rangle\right| \leqslant \frac{(\lambda_r - \lambda_1)(q-1) + (q+1)}{q^{\frac{1}{2}(\lambda_r - \lambda_1)}(q+1)} \|\xi\| \cdot \|\eta\|. \tag{7}$$

For any nonzero $(\lambda_1, ..., \lambda_r) \in \Lambda_0$ we have $\lambda_r \geqslant \lambda_1 + 2$ and therefore inequality (7) implies

$$\left|\langle \eta, \pi(D(\lambda_1, ..., \lambda_r))\xi\rangle\right| \leqslant \frac{3}{q} \|\xi\| \cdot \|\eta\| \ \text{ for any nonzero } \ (\lambda_1, ..., \lambda_r) \in \Lambda_0. \tag{8}$$

Note that in (7) we did not use the full strength of Theorem 1.1 of [26] because we used the very poor strongly orthogonal system $\{e_1 - e_r\}$, whereas [26] uses the maximal strongly orthogonal system $\left\{ e_1 - e_r, e_2 - e_{r-1}, ..., e_{\lfloor \frac{r}{2} \rfloor} - e_{r+1-\lfloor \frac{r}{2} \rfloor} \right\}$ to get optimal bounds. However, the optimal bounds do not give in (8) an exponent of q better than -1 for $(\lambda_1, ..., \lambda_r) = (-1, 0, ..., 0, 1)$.

Let Γ be a cocompact lattice in G_0 satisfying the condition

(C) for any $x \in G/K$ and any $\gamma \in \Gamma$ different from 1, the distance from x to γx along the 1-skeleton of G/K is > 2.

The quotient $X = \Gamma\backslash G/K$ is the set of vertices of a simplicial complex whose faces are the quotient by Γ of the faces of the building G/K. Since $\Gamma \subseteq G_0$, we have an obvious type function type : $X \to \mathbb{Z}/r\mathbb{Z}$, and thanks to the condition (C), X has no multiple edges. For any non-empty subset $I \subseteq \mathbb{Z}/r\mathbb{Z}$ we write $X_I = \Gamma\backslash Y_I$ and we identify X_I with the set of faces of type I in X. We note that for $i \in \mathbb{Z}/r\mathbb{Z}$, $X_{\{i\}} \stackrel{\text{def}}{=} \Gamma\backslash G_i/K$ is the subset of vertices of type i in X and is of cardinality $\frac{1}{r}|X|$.

For any $(\lambda_1, ..., \lambda_r) \in \Lambda$ we have a normalized Hecke operator $H_{\lambda_1,...,\lambda_r}$ acting on $\ell^2(\Gamma\backslash G/K)$ in the following way: we view elements of $\ell^2(\Gamma\backslash G/K)$ as Γ-invariant functions on G/K and $H_{\lambda_1,...,\lambda_r}(f)(x)$ is the average of the values of f on the vertices $y \in G/K$ which satisfy $\sigma(x, y) = (\lambda_1, ..., \lambda_r)$. For $\xi, \eta \in \ell^2(\Gamma\backslash G/K)$ considered as K-invariants vectors in $L^2(\Gamma\backslash G)$, endowed with the representation π of G by right translations, we have

$$\langle \eta, H_{\lambda_1,...,\lambda_r}\xi\rangle = \langle \eta, \pi(D(\lambda_1, ..., \lambda_r))\xi\rangle. \tag{9}$$

The subspace of G^+-invariant vectors in $L^2(\Gamma\backslash G)$ is $L^2(\Gamma\backslash G)^{G^+} = L^2(F^\times/(F^\times)^r \det(\Gamma))$ because $G/G^+ = F^\times/(F^\times)^r$ is abelian. Let H be the orthogonal complement of $L^2(\Gamma\backslash G)^{G^+}$ in $L^2(\Gamma\backslash G)$. The representation (H, π) does not have any nonzero G^+-invariant vector and we will apply inequality (8) to it. For any set Z endowed with a finite measure we write $L^2_0(Z)$ for the hyperplane of $L^2(Z)$ which is orthogonal to the constant function 1. For $(\lambda_1, ..., \lambda_r) \in \Lambda_0$, $H_{\lambda_1,...,\lambda_r}$ acts diagonally on the direct sum decomposition $\ell^2(\Gamma\backslash G/K) = \bigoplus_{i\in\mathbb{Z}/r\mathbb{Z}} \ell^2(\Gamma\backslash G_i/K)$ and the estimate (8), together with (9), implies the operator norm bound

$$\|H_{\lambda_1,...,\lambda_r}\|_{\mathcal{L}\left(\ell^2_0(\Gamma\backslash G_i/K)\right)} \leqslant \frac{3}{q}, \tag{10}$$

for any $i \in \mathbb{Z}/r\mathbb{Z}$ and for any nonzero $(\lambda_1, ..., \lambda_r) \in \Lambda_0$. To justify (10) it remains to check that for any $i \in \mathbb{Z}/r\mathbb{Z}$ and any $f \in \ell^2_0(\Gamma\backslash G_i/K)$, the extension of f by 0 to $\Gamma\backslash G/K$, considered as an element of $L^2(\Gamma\backslash G)$, belongs to H, i.e., is orthogonal to $L^2(\Gamma\backslash G)^{G^+}$. Indeed $G^+K = G_0$, since we have $\det(K) = \mathcal{O}_F^\times/(\mathcal{O}_F^\times)^r$ and an exact sequence

$$1 \to \mathcal{O}_F^\times/(\mathcal{O}_F^\times)^r \to F^\times/(F^\times)^r \to \mathbb{Z}/r\mathbb{Z} \to 0.$$

Proposition 5.1. *Let $I \subseteq \mathbb{Z}/r\mathbb{Z}$ be a non-empty set. For $i \in I$, fix $\alpha_i > 0$ and a subset $Z_i \subseteq X_{\{i\}}$ of cardinality $\geqslant \alpha_i \left|X_{\{i\}}\right|$. Then the proportion of elements of $(x_i)_{i\in I} \in X_I$ satisfying $x_i \in Z_i$ for every $i \in I$ is at least*

$$\prod_{i\in I} \alpha_i - \frac{2(|I| - 1)}{\sqrt{q}}.$$

Proof. Proposition 5.1 is obvious for $|I| = 1$, and it follows by induction on $|I|$ due to the following lemma. □

Lemma 5.2. *Fix $I \subseteq \mathbb{Z}/r\mathbb{Z}$ of cardinality $\geqslant 2$, and $i \in I$. Write $I' = I \setminus \{i\}$. For $\alpha, \alpha' > 0$, $Z \subseteq X_{\{i\}}$ of cardinality $\geqslant \alpha \left|X_{\{i\}}\right|$ and $Z' \subseteq X_{I'}$ of cardinality $\geqslant \alpha' |X_{I'}|$, the proportion of $(x_j)_{j\in I} \in X_I$ satisfying $x_i \in Z$ and $(x_j)_{j\in I'} \in Z'$ is at least*

$$\alpha\alpha' - \frac{2}{\sqrt{q}}.$$

Proof. Let $T : \ell^2(X_{\{i\}}) \to \ell^2(X_{I'})$ be the following averaging operator: for any $f \in \ell^2\left(X_{\{i\}}\right)$ and $(x_j)_{j\in I'} \in X_{I'}$, $T(f)((x_j)_{j\in I'})$ is the average of $f(x_i)$ over $x_i \in X_{\{i\}}$ such that $(x_j)_{j\in I}$ belongs to X_I. We normalize the norms of $\ell^2(X_{\{i\}})$ and $\ell^2(X_{I'})$ such that the constant function 1 has norm 1. We denote by $T_0 : \ell^2_0(X_{\{i\}}) \to \ell^2_0(X_{I'})$ the restriction of T to the hyperplane orthogonal to the constant function 1.

Note that $T^*T : \ell^2(X_{\{i\}}) \to \ell^2(X_{\{i\}})$ is also an averaging operator, and in fact it is an average of the Hecke operators $H_{\lambda_1,\dots,\lambda_r}$ for $(\lambda_1,\dots,\lambda_r) \in \Lambda_0$ such that $\lambda_r \leqslant \lambda_1 + 2$. In this average, the coefficient of $H_{0,\dots,0} = \mathrm{Id}$ is $\leqslant q^{-1}$. Indeed, let $\widetilde{i} \in \mathbb{Z}$ be a lifting of i and let $\widetilde{i}_-, \widetilde{i}_+$ be the biggest integer $< \widetilde{i}$ (resp. the smallest integer $> \widetilde{i}$) whose images i_-, i_+ belong to I'. Then for any $(x_j)_{j\in I'} \in X_{I'}$, the number of $x_i \in X_{\{i\}}$ such that $(x_j)_{j\in I}$ belongs to X_I is exactly $\left|\mathrm{Gr}\left(\widetilde{i} - \widetilde{i}_-, \widetilde{i}_+ - \widetilde{i}_-\right)(\mathbb{F}_q)\right|$, the number of sub-$\mathbb{F}_q$-vector spaces of dimension $\widetilde{i} - \widetilde{i}_-$ in $\mathbb{F}_q^{\widetilde{i}_+ - \widetilde{i}_-}$. The coefficient of $H_{0,\dots,0}$ is $\left|\mathrm{Gr}\left(\widetilde{i} - \widetilde{i}_-, \widetilde{i}_+ - \widetilde{i}_-\right)(\mathbb{F}_q)\right|^{-1}$, and it is clear that $\left|\mathrm{Gr}\left(\widetilde{i} - \widetilde{i}_-, \widetilde{i}_+ - \widetilde{i}_-\right)(\mathbb{F}_q)\right| \geqslant q$. By (10) we have

$$\|T_0^* T_0\|_{\mathcal{L}\left(\ell^2_0(X_{\{i\}})\right)} \leqslant \frac{1}{q} + \frac{3}{q} = \frac{4}{q},$$

implying that $\|T_0\|_{\mathcal{L}\left(\ell^2_0(X_{\{i\}}), \ell^2_0(X_{I'})\right)} \leqslant 2q^{-\frac{1}{2}}$. Therefore

$$\left|\langle T(\mathbf{1}_Z), \mathbf{1}_{Z'}\rangle - \frac{|Z| \cdot |Z'|}{\left|X_{\{i\}}\right| \cdot |X_{I'}|}\right| \leqslant 2q^{-\frac{1}{2}} \sqrt{\frac{|Z| \cdot |Z'|}{\left|X_{\{i\}}\right| \cdot |X_{I'}|}} \leqslant 2q^{-\frac{1}{2}}, \tag{11}$$

completing the proof of Lemma 5.2. □

Remark 5.1. If $X = \Gamma\backslash G/K$ is assumed to be a Ramanujan complex (see [22]) we can improve the estimate in (10), but at the end, in (11), we get the same exponent $-\frac{1}{2}$.

Proposition 5.1 has the following corollary.

Corollary 5.3. *Let $I \subseteq \mathbb{Z}/r\mathbb{Z}$ be of cardinality $\geqslant 2$. Let $\mathcal{X}_I$ be the $|I|$-uniform hypergraph with vertices $X = \Gamma\backslash G/K$ and with hyperedges the set X_I of faces of X of type I. Then for q large enough (as a function of $|I|$ alone) the overlap parameter of $\mathcal{X}_I$ is bounded below by a positive constant depending only on $|I|$.*

Proof. Let $f : X \to \mathbb{R}^{|I|-1}$ be an injection. By the main result of [27] (which is a strengthening of Theorem 1.2), applied to the sets $f\left(X_{\{i\}}\right) \subseteq \mathbb{R}^{|I|-1}$ for $i \in I$, we obtain sets $P_i \subseteq f\left(X_{\{i\}}\right)$ with $|P_i| \geqslant c''_{|I|-1}\left|X_{\{i\}}\right|$, for $i \in I$, such that all the simplices with vertices in each of the P_i have a point in common. Here $c''_{|I|-1}$ is a constant depending only on $|I|$. By an application of Proposition 5.1 to the sets $Z_i = f^{-1}(P_i)$ with $\alpha_i = c''_{|I|-1}$, we see that if

$$q > \frac{16(|I|-1)^2}{\left(c''_{|I|-1}\right)^{2|I|}}$$

then the overlap parameter of $\Gamma\backslash G/K$ is at least $\frac{1}{2}\left(c''_{|I|-1}\right)^{|I|}$. □

Lemma 5.4. *For any integer r and for any non-archimedean local field F, there exists a cocompact lattice Γ in G_0 which satisfies condition (C), and a sequence of finite index subgroups $\Gamma_n \subseteq \Gamma$ with $\lim_{n\to\infty} |\Gamma/\Gamma_n| = \infty$.*

Proof. We start by choosing any cocompact arithmetic subgroup $\widetilde{\Gamma}$ in $G = PGL_r(F)$. The existence of a cocompact arithmetic subgroup is well-known: use division algebras over function fields when the characteristic of F is finite and unitary groups over number fields when the characteristic of F is 0. In the case of characteristic 0, it is a particular case of a theorem of Borel and Harder [7]; see also Example 5.1.4, Corollary 5.12 and the remark after it in [6], for a short proof. Since $G/G_0 = \mathbb{Z}/r\mathbb{Z}$, $\overline{\Gamma} = \widetilde{\Gamma} \cap G_0$ is of finite index in $\widetilde{\Gamma}$. Since $\overline{\Gamma}$ is a cocompact lattice in G_0, the elements $\gamma \in \overline{\Gamma} \setminus \{1\}$, such that there exists $x \in G/K$ with $d(x, \gamma x) \leqslant 2$, form a finite number of conjugacy classes. By its arithmetic nature, $\overline{\Gamma}$ embeds in the product of its finite quotients. Therefore there is a finite index subgroup $\Gamma \subseteq \overline{\Gamma}$ which satisfies condition (C) and a sequence of finite index subgroups Γ_n of Γ with $\lim_{n\to\infty} |\Gamma/\Gamma_n| = \infty$. □

Corollary 5.3, applied (for $r \geqslant 3$) to the lattices Γ_n of Lemma 5.4, yields highly overlapping families of $|I|$-uniform hypergraphs. In the particular case where $I = \mathbb{Z}/r\mathbb{Z}$, these hypergraphs are finite quotients of the building of $PGL_r(F)$ and their hyperedges are the images of the chambers of the building.

Proof of Theorem 1.5. Theorem 1.5 is a consequence of Corollary 5.3, applied to the Ramanujan complexes constructed in [21] (which are based on a lattice construction from [12]), together with their description in [21] as the clique complexes corresponding to the Cayley graphs associated to certain (explicitly defined) generators of $PGL_r(F)$, where $F = \mathbb{F}_p((t))$. We just need to ensure that the corresponding lattices are sublattices of G_0 (so that they preserve the type), as well as that condition (C) is satisfied. This is true for arbitrarily large m due to Corollary 6.8 of [21], or equivalently the case $r = d$ of Theorem 7.1 in [21]. Alternatively, one can consider the construction of the Ramanujan complexes in [31], specifically the second extreme distinguished case of Corollary 36 in [31]. □

6. A Szemerédi-type theorem for infinite hypergraphs

In a measurable space, an *atom* is a measurable set which has positive measure and contains no set of smaller but positive measure. A measure which has no atoms is called *non-atomic*. A basic result of Sierpinski [32] states that if μ is a non-atomic measure, then for any measurable set A and any b with $0 \leqslant b \leqslant \mu(A)$, there is a measurable subset B of A with $\mu(B) = b$. Given a measurable space on a set V and a measure μ, the h-fold product measurable space is generated by all sets $B_1 \times \cdots \times B_h$, with $B_1, \ldots, B_h$ measurable subsets of V, and the product measure μ^h is the unique measure on this space given by $\mu^h(B_1 \times \cdots \times B_h) = \mu(B_1) \cdots \mu(B_h)$.

The aim of this section is to establish a Szemerédi-type theorem for infinite hypergraphs with a non-atomic measure on their vertex sets. The statement will be used in the next section, for the proof of Theorem 1.8.

Given a finite or infinite h-uniform hypergraph $G = (V, E)$, we say that an h-tuple $(V_1, \ldots, V_h)$ of disjoint subsets of V is *homogeneous* with respect to G if either all elements of $V_1 \times \ldots \times V_h$ are hyperedges of G or none of them are. If, in addition, we have a finite non-atomic measure μ on V, we say that G is (c, μ)-*structured* provided that for all disjoint measurable subsets $S_1, \ldots, S_h \subseteq V$, there exist measurable $Y_i \subseteq S_i$ for $1 \leqslant i \leqslant h$ with $\mu(Y_i) \geqslant c\mu(S_i)$ such that the h-tuple $(Y_1, \ldots, Y_h)$ is homogeneous. The following theorem generalizes a result in [28], where the case $h = 2$ was settled. The proof is based on an idea of Komlós.

Theorem 6.1. *For any $c, \varepsilon > 0$ and for any positive integer h, there is $K = K(h, c, \varepsilon)$ such that the following statement is true. If $G = (V, E)$ is a (c, μ)-structured h-uniform hypergraph, where μ is a non-atomic measure μ, then for each $k \geqslant K$ there is a partition $V = V_1 \cup \ldots \cup V_k$ of V into k parts of equal measure such that all but at most an ε-fraction of the h-tuples $(V_{i_1}, \ldots, V_{i_h})$ are homogeneous.*

In the sequel, let $G = (V, E)$ be a fixed h-uniform hypergraph with a finite non-atomic measure μ on V. The *measure* of G, denoted by $\mu(G)$, is the product measure μ^h of the set of h-tuples $(v_1, \ldots, v_h) \in V^h$ with $\{v_1, \ldots, v_h\}$ an edge of G (we assume throughout that this set is μ^h-measurable). Define the *edge-density* $d(G)$ of G to be $\frac{\mu(G)}{\mu(V)^h}$. For measurable vertex subsets $V_1, \ldots, V_h$, define

$$\nu(V_1, \ldots, V_h) \stackrel{\text{def}}{=} \mu^h\Big((v_1, \ldots, v_h) \in V_1 \times \cdots \times V_h :\ \{v_1, \ldots, v_h\} \in E\Big), \tag{12}$$

and the *edge density*

$$d(V_1, \ldots, V_h) \stackrel{\text{def}}{=} \frac{\nu(V_1, \ldots, V_h)}{\mu(V_1) \cdots \mu(V_h)}.$$

An h-tuple of disjoint subsets $(X_1, \ldots, X_h)$ of vertices in an h-uniform hypergraph is said to be (γ, δ)-*superregular* if for any subsets $Y_i \subseteq X_i$ with $\mu(Y_1) \cdots \mu(Y_h) \geqslant \gamma\mu(X_1) \cdots \mu(X_h)$, we have $d(Y_1, \ldots, Y_h) \geqslant \delta$.

The following lemma shows that the vertex set of any dense h-uniform hypergraph can be partitioned into h parts of equal measure such that the sub-hypergraph formed by all of its edges that contain one point from each part is still relatively dense. The analogue of this statement for finite hypergraphs is well known and very easy to prove, as a uniformly random partition of the vertex set into equal parts will almost surely work. Here we do not have the leisure of taking such a uniform random partition.

Lemma 6.2. *Let $G = (V, E)$ be an h-uniform hypergraph, μ a finite non-atomic measure on V. Then there is a partition $V = V_1 \cup \ldots \cup V_h$ into measurable subsets, each of measure $\mu(V)/h$, such that $d(V_1, \ldots, V_h) \geqslant \frac{d(G)}{2}$.*

Proof. Write $t = h\lceil h/d(G)\rceil$. Arbitrarily partition V into t subsets $V = U_1 \cup \ldots \cup U_t$, each of measure $\mu(V)/t$. This can be done since μ is non-atomic. The product measure of the set of h-tuples of vertices from distinct U_i is at least

$$\prod_{i=0}^{h-1}\left(1 - \frac{i}{t}\right)\mu(V) \geqslant \left(1 - \frac{\binom{h}{2}}{t}\right)\mu(V)^h \geqslant \left(1 - \frac{h^2}{2t}\right)\mu(V)^h \geqslant \left(1 - \frac{d(G)}{2}\right)\mu(V)^h.$$

Hence, the edge density of the sub-hypergraph of G with vertices in different U_i is at least $d(G) - d(G)/2 = d(G)/2$. We randomly partition $V = V_1 \cup \ldots \cup V_h$, where each V_i is a union of t/h of the U_j, with each such partition being equally likely. For $U_{j_1}, \ldots, U_{j_h}$ with $j_1 < \ldots < j_h$, the probability that $U_{j_i} \subseteq V_i$ for all $1 \leqslant i \leqslant h$ is at least h^{-h}. Hence, by linearity of expectation and the fact $\mu(V_1) \cdots \mu(V_h) = h^{-h}\mu(V)^h$, the expected value of $d(V_1, \ldots, V_h)$ is at least $d(G)/2$. It follows that there is a partition $V = V_1 \cup \ldots \cup V_h$ into parts of equal measure with $d(V_1, \ldots, V_h) \geqslant d(G)/2$. □

The next lemma shows that if an h-tuple of sets is not (γ, δ)-superregular, then we can find subsets of large measure such that the edge density between them is significantly larger than the edge density between the original h-tuple.

Lemma 6.3. *Let $G = (V, E)$ be an h-uniform hypergraph and μ a finite non-atomic measure on V. If a collection of h measurable vertex subsets $(W_1, \ldots, W_h)$ with $d(W_1, \ldots, W_h) = c$ is not (δ, γ)-superregular, then there are subsets $Z_i \subseteq W_i$ for $1 \leqslant i \leqslant h$ such that*

$$\mu(Z_1) \ldots \mu(Z_h) \geqslant \frac{\delta\gamma}{2^h}\mu(W_1) \cdots \mu(W_h)$$

and

$$d(Z_1, \ldots, Z_h) \geqslant c + (c - 2\delta)\frac{\gamma}{1-\gamma}.$$

Proof. Since $(W_1, \ldots, W_h)$ is not (δ, γ)-superregular, there exist subsets $Y_i \subseteq W_i$ for every $1 \leqslant i \leqslant h$ with

$$\mu(Y_1) \cdots \mu(Y_h) \geqslant \gamma\mu(W_1) \cdots \mu(W_h) \quad \text{and} \quad d(Y_1, \ldots, Y_h) < \delta. \tag{13}$$

The sum of all the 2^h terms $\nu(T_1, \ldots, T_h)$ (where ν is defined in (12)) with $T_i = Y_i$ or $T_i = W_i \setminus Y_i$ equals $\nu(W_1, \ldots, W_h) = c\mu(W_1) \cdots \mu(W_h)$. The sum of $\nu(T_1, \ldots, T_h)$ over all such terms with not all $T_i = Y_i$ and with

$$\mu(T_1) \cdots \mu(T_h) \geqslant \frac{\delta\gamma}{2^h}\mu(W_1) \cdots \mu(W_h)$$

is therefore greater than

$$\left(c - 2^h\frac{\delta\gamma}{2^h}\right)\mu(W_1) \cdots \mu(W_h) - \delta\mu(Y_1) \cdots \mu(Y_h) \geqslant c\mu(W_1) \cdots \mu(W_h) - 2\delta\mu(Y_1) \cdots \mu(Y_h).$$

Here we used both inequalities in (13). Also, the sum of $\mu(T_1) \cdots \mu(T_h)$ over these terms is at most $\mu(W_1) \cdots \mu(W_h) - \mu(Y_1) \cdots \mu(Y_h)$. By averaging, if $a_1 + \cdots + a_k \geqslant A$ and $b_1 + \cdots + b_k \leqslant B$ with all a_i, b_i positive, then there is $1 \leqslant i \leqslant k$ such that $\frac{a_i}{b_i} \geqslant \frac{A}{B}$. Hence, there are $T_1, \ldots, T_h$ with $T_i = Y_i$ or $T_i = W_i \setminus Y_i$ and not all $T_i = Y_i$, with $\mu(T_1) \cdots \mu(T_h) \geqslant 2^{-h}\delta\gamma\mu(W_1) \cdots \mu(W_h)$ and

$$\begin{aligned} d(T_1, \ldots, T_h) = \frac{\nu(T_1, \ldots, T_h)}{\mu(T_1) \cdots \mu(T_h)} &\geqslant \frac{c\mu(W_1) \cdots \mu(W_h) - 2\delta\mu(Y_1) \cdots \mu(Y_h)}{\mu(W_1) \cdots \mu(W_h) - \mu(Y_1) \cdots \mu(Y_h)} \\ &= c + (c - 2\delta)\frac{\mu(Y_1) \cdots \mu(Y_h)}{\mu(W_1) \cdots \mu(W_h) - \mu(Y_1) \cdots \mu(Y_h)} \geqslant c + (c - 2\delta)\frac{\gamma}{1-\gamma}, \end{aligned}$$

as required. □

By repeated application of Lemma 6.3, we obtain the following result, which shows that a dense hypergraph contains a superregular h-tuple of sets of large measure.

Lemma 6.4. *For $\gamma, \delta > 0$ and a positive integer h, there is $\alpha = \alpha(\gamma, \delta, h)$ such that the following holds. If $G = (V, E)$ is an h-uniform hypergraph, μ is a finite non-atomic measure on V with $d(G) \geqslant 8\delta$, then there is an h-tuple $(X_1, \ldots, X_h)$ of disjoint measurable vertex subsets, which is (γ, δ)-superregular and satisfies $\mu(X_1) \cdots \mu(X_h) \geqslant \alpha \mu(V)^h$.*

Proof. Define
$$\alpha = \frac{1}{h^h} \left(\frac{\delta\gamma}{2^h} \right)^{\frac{2}{\gamma} \log_2(1/\delta)}.$$
By Lemma 6.2, there is a partition $V = V_1 \cup \ldots \cup V_h$ into parts of equal measure such that $d \stackrel{\text{def}}{=} d(V_1, \ldots, V_h) \geqslant d(G)/2$. We will repeatedly apply Lemma 6.3, starting with the sets $V_1, \ldots, V_h$, until we get a (γ, δ)-superregular h-tuple.

If $(V_1, \ldots, V_h)$ is not (γ, δ)-superregular, then we can find $V_1^1 \subseteq V_1, \ldots, V_h^1 \subseteq V_h$ with
$$\mu(V_1^1) \cdots \mu(V_h^1) \geqslant \frac{\delta\gamma}{2^h} \mu(V_1) \cdots \mu(V_h)$$
and
$$d\left(V_1^1, \ldots, V_h^1\right) \geqslant d + (d - 2\delta)\frac{\gamma}{1-\gamma} > d\left(1 + \frac{\gamma}{2}\right).$$
After k iterations, we either have found a (γ, δ)-superregular h-tuple, or we find $V_1^k, \ldots, V_h^k$ with
$$\mu(V_1^k) \cdots \mu(V_h^k) \geqslant \left(\frac{\delta\gamma}{2^h} \right)^k \mu(V_1) \cdots \mu(V_h)$$
and
$$d\left(V_1^k, \ldots, V_h^k\right) > d\left(1 + \frac{\gamma}{2}\right)^k.$$
This cannot continue for more than $k_0 = \frac{2}{\gamma} \log_2(1/d)$ iterations, as otherwise we would produce an h-tuple of density more than 1, a contradiction. Thus, at some step $k \leqslant k_0$, we find an h-tuple of sets $X_i \stackrel{\text{def}}{=} V_i^k$ with
$$\mu(X_1) \cdots \mu(X_h) \geqslant \left(\frac{\delta\gamma}{2^h} \right)^k \mu(V_1) \cdots \mu(V_h) = \frac{1}{h^h} \left(\frac{\delta\gamma}{2^h} \right)^k \mu(V)^h \geqslant \alpha \mu(V)^h,$$
which is (γ, δ)-superregular. □

An h-uniform hypergraph $H = (V, E)$ is *h-partite* if there is a partition $V = V_1 \cup \ldots \cup V_h$ such that every edge has exactly one vertex in each V_i. For a vertex set V, and a collection $\mathcal{C}$ of h-tuples $A_1, \ldots, A_h$ of vertex subsets of V, define the hypergraph $H(\mathcal{C})$ on V, which is the union of the complete h-partite h-uniform hypergraphs with parts $A_1, \ldots, A_h$.

Lemma 6.5. *For $c, \varepsilon > 0$ and a positive integer h, there is $L = L(h, c, \varepsilon)$ such that the following statement is true. If $G = (V, E)$ is a (c, μ)-structured h-uniform hypergraph with μ a finite non-atomic measure, then there is a collection $\mathcal{C}$ of at most L homogeneous h-tuples of vertex subsets such that the density of $H(\mathcal{C})$ is at least $1 - \varepsilon$.*

Proof. Let H_0 denote the complete h-uniform hypergraph on V, and let $\gamma = c^h$, $\delta = \varepsilon/8$, $\alpha = \alpha(\gamma, \delta, h)$ as in Lemma 6.4, $\beta = h!\delta c^h \alpha$, and $L = \beta^{-1}$. Suppose that for some $i \geqslant 1$, we have already defined H_{i-1}. If G has an h-tuple $A_1, \ldots, A_h$ of disjoint subsets which is homogeneous and such that the hypergraph which consists of those edges of H_{i-1} that have one vertex in each A_ℓ has edge density at least β, then let H_i denote the sub-hypergraph of H_{i-1} obtained by deleting all edges with one vertex in each A_ℓ. Otherwise, we stop. This process will clearly terminate in $j \leqslant \beta^{-1}$ steps with an h-uniform hypergraph H_j as $d(H_i) \leqslant 1 - i\beta$ for each i.

We next show that $d(H_j) < \varepsilon$. Indeed, otherwise by Lemma 6.4, H_j has a collection of h disjoint vertex subsets $(X_1, \ldots, X_h)$ which is (γ, δ)-superregular with

$$\mu(X_1) \cdots \mu(X_h) \geqslant \alpha \mu(V)^h.$$

Since G is (c, μ)-structured, there are subsets $A_i \subseteq X_i$ for $1 \leqslant i \leqslant h$ with $\mu(A_i) \geqslant c\mu(X_i)$ such that $(A_1, \ldots, A_h)$ is homogeneous with respect to G. As $\gamma = c^h$,

$$\mu(A_1) \cdots \mu(A_h) \geqslant \gamma \mu(X_1) \cdots \mu(X_h),$$

and $(X_1, \ldots, X_h)$ is (γ, δ)-superregular, we have $d(A_1, \ldots, A_h) \geqslant \delta$, and hence the hypergraph which consists of those edges of H_j that have one vertex in each A_ℓ has edge density at least

$$h!\delta \cdot \frac{\mu(A_1) \cdots \mu(A_h)}{\mu(V)^h} \geqslant h!\delta c^h \cdot \frac{\mu(X_1) \cdots \mu(X_h)}{\mu(V)^h} \geqslant h!\delta c^h \alpha = \beta.$$

However, this contradicts the fact that since the construction terminated at the jth step, there are no such subsets $A_1, \ldots, A_h$. $\square$

For an h-uniform hypergraph $G = (V, E)$ and a vertex partition $\mathcal{P} = \{V = V_1 \cup \ldots \cup V_k\}$, the *homogeneous hypergraph* $G_{\mathcal{P}}$ is the h-uniform hypergraph on V where $(v_1, \ldots, v_h) \in V_{i_1} \times \cdots \times V_{i_h}$ is an edge if and only if $(V_{i_1}, \ldots, V_{i_h})$ is homogeneous with respect to G. Note that if $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then $d(G_{\mathcal{P}'}) \geqslant d(G_{\mathcal{P}})$. Given a collection $\mathcal{C}$ of h-tuples $(A_1, \ldots, A_h)$ which are homogeneous in G, define the partition $\mathcal{P}$ of V into $(h+1)^{|\mathcal{C}|}$ parts, where each part consists of those vertices in the same A_i (or in none of the A_i) for each h-tuple in $\mathcal{C}$. The hypergraph $G_{\mathcal{P}}$ contains the hypergraph $H(\mathcal{C})$. Hence, we have the following corollary of Lemma 6.5 with $M = (h+1)^L$.

Corollary 6.6. *For any $c, \varepsilon > 0$ and for any positive integer h, there is $M = M(h, c, \varepsilon)$ such that the following statement is true. If $G = (V, E)$ is a (c, μ)-structured h-uniform hypergraph with a finite non-atomic measure μ on its vertex set, then there is a partition $\mathcal{P}$ of V into at most M parts such that $d(G_{\mathcal{P}}) \geqslant 1 - \varepsilon$.*

Proof of Theorem 6.1. Let $M = M(h, c, \frac{\varepsilon}{2})$ as in Corollary 6.6 and K be the smallest integer which is at least $2Mh\varepsilon^{-1}$. Fix $k \geqslant K$. By Corollary 6.6, there is a partition $\mathcal{P}'$ of V into at most M parts such that $d(G_{\mathcal{P}'}) \geqslant 1 - \frac{\varepsilon}{2}$. If V_i is a part of $\mathcal{P}'$, arbitrarily partition V_i into parts of measure $\mu(V)/k$ and one remaining piece of measure at most $\mu(V)/k$. Let W be the union of the remaining parts, so $\mu(W) \leqslant M\mu(V)/k$, and arbitrarily partition W into parts of measure $\mu(V)/k$. We have thus produced a partition $\mathcal{P}$ into k parts of equal measure, and we next show that this partition satisfies the assertion of Theorem 6.1. The edge density of the hypergraph of h-tuples in V^h that contain a vertex in W is at most

$$1 - \left(1 - \frac{\mu(W)}{\mu(V)}\right)^h \leqslant h\frac{\mu(W)}{\mu(V)} \leqslant \frac{hM}{k} \leqslant \frac{\varepsilon}{2}.$$

Thus $d(G_{\mathcal{P}}) \geqslant d(G_{\mathcal{P}'}) - \frac{\varepsilon}{2} \geqslant 1 - \varepsilon$, and hence the partition $\mathcal{P}$ satisfies the conclusion of Theorem 6.1. $\square$

7. A partition result—Proof of Theorem 1.8

Before proving Theorem 1.8 in its full generality, we give a simple argument for the special case $d = 2$, which provides a good upper bound on the constant $K(\varepsilon, 2)$.

Proposition 7.1. *Let $\varepsilon > 0$ and $k \geqslant \frac{12}{\varepsilon} + 1$. For any finite Borel measure μ on $\mathbb{R}^2$ with respect to which every line has measure 0, and for any point $q \in \mathbb{R}^d$, there is a partition $\mathbb{R}^2 = A_1 \cup \ldots \cup A_k$ into k measurable parts of equal measure, such that all but at most an ε-fraction of the triples A_h, A_i, A_j are homogenous with respect to q.*

Proof. Partition $\mathbb{R}^2$ radially around q into k cones $A_1, \ldots, A_k$ of equal measure. Notice that a triple A_h, A_i, A_j is not homogeneous with respect to q if and only if one of them intersects the reflection of another about q. Since the number of such triples of cones is at most $2k(k-2)$, the fraction of nonhomogeneous triples cannot exceed $2k(k-2)/\binom{k}{3} = 12/(k-1) \leqslant \varepsilon$, which completes the proof. $\square$

Next we turn to the proof of Theorem 1.8 in the general case. We break the proof into four lemmas. For the first one, we recall Radon's theorem, which states that any set of $d+2$ points in $\mathbb{R}^d$ can be partitioned into two sets whose convex hulls have nonempty intersection (see [16]).

Lemma 7.2. *Let $v_1, \ldots, v_{d+1} \in \mathbb{R}^d$. A point $q \in \mathbb{R}^d$ belongs to the simplex with vertex set $V = \{v_1, \ldots, v_{d+1}\} \subseteq \mathbb{R}^d$ if and only if for each nonempty proper subset $X \subseteq V$, there is a hyperplane passing through q which separates X from $V \setminus X$.*

Proof. In one direction the statement is clear, as there is a hyperplane through any internal point separating any proper subset of the vertex set from its complement. In the other direction, suppose q is not in the simplex, and consider the set $V \cup \{q\}$ of $d + 2$ points. By Radon's theorem, there is a partition $V \cup \{q\} = A \cup B$ such that the convex hull of A and the convex hull of B have a point p in common. Suppose without loss of generality that $q \in A$. Clearly, $|A| > 1$, because q is outside of the simplex. But then $A \setminus \{q\}$ and B cannot be separated by a hyperplane passing through q. $\square$

We need the following version of the ham sandwich theorem (see, e.g., [24]).

Lemma 7.3. *Let $S_1, \ldots, S_{d-1}$ be measurable subsets of $\mathbb{R}^d$ and $q \in \mathbb{R}^d$. There is a hyperplane through q that partitions each S_i into two parts of equal measure.*

By repeated application of Lemma 7.3, and then using Lemma 7.2, we obtain:

Lemma 7.4. *For any positive integer d, there is $c_d > 0$ satisfying the following condition. Let μ be a finite Borel measure on $\mathbb{R}^d$ with respect to which every hyperplane has measure 0, let $S_1, \ldots, S_{d+1}$ be measurable subsets of $\mathbb{R}^d$, and let $q \in \mathbb{R}^d$. Then there exist $Y_i \subseteq S_i$ for all $1 \leqslant i \leqslant d+1$ such that $\mu(Y_i) \geqslant c_d\mu(S_i)$ and $Y_1, \ldots, Y_{d+1}$ are homogeneous with respect to q.*

Proof. Consider an arbitrary labeling $X_1, \ldots, X_{2^{d+1}-2}$ of the nonempty proper subsets of $[d + 1] = \{1, \ldots, d + 1\}$. We describe an iterative process for constructing the desired sets $Y_1, \ldots, Y_{d+1}$.

Let $S_j^0 = S_j$, for every $j \in [d + 1]$. After completing step i, we will have subsets $S_j^i \subseteq S_j$ with $\mu(S_j^i) \geqslant 2^{-i}\mu(S_j)$ such that at least one of the following two conditions is satisfied:

(1) there is a hyperplane through q such that $S_1^i \cup S_2^i \cup \ldots \cup S_{d+1}^i$ lies entirely on one of its sides, or
(2) for every $1 \leqslant k \leqslant i$, there is a hyperplane through q that separates the sets $\{S_j^i\}_{j \in X_k}$ from the sets $\{S_j^i\}_{j \in [d+1] \setminus X_k}$.

The proof shows that we can take $c_d = 2^{2-2^{d+1}}$. Notice that the inductive hypothesis holds vacuously at the end of step 0. Suppose we have already completed step i and we wish to proceed to the next step.

If condition (1) is satisfied, then we simply let $S_j^{i+1} = S_j^i$. Thus, we may suppose that condition (2) is satisfied. Let $a \in X_{i+1}$ and $b \in [d+1] \setminus X_{i+1}$. We apply Lemma 7.3 to the $d-1$ sets S_j^i with $j \in [d+1] \setminus \{a, b\}$. There is a hyperplane H containing q that separates each such S_j^i into parts of equal measure. The hyperplane H partitions S_a^i into two subsets. Let S_a^{i+1} be the subset of larger measure. Similarly, H partitions S_b^i into two subsets. Let S_b^{i+1} be the subset of larger measure. If S_a^{i+1} and S_b^{i+1} are on the same side of H, then let S_j^{i+1} for $j \in [d+1] \setminus \{a, b\}$ be the subset of S_j^i consisting of those points on the same side of H as S_a^{i+1} and S_b^{i+1}. In this case, we have $\mu(S_j^{i+1}) \geqslant \frac{1}{2}\mu(S_j^i) \geqslant 2^{-i-1}\mu(S_j)$, the first of the two desired properties holds, and we have completed step $i+1$. Otherwise, for $j \in X_{i+1} \setminus \{a\}$, let S_j^{i+1} be the subset of S_j^i consisting of those points on the same side of H as S_a^{i+1}, and for $j \in [d+1] \setminus (X_{i+1} \cup \{b\})$, let S_j^{i+1} be the subset of S_j^i consisting of those points on the same side of H as S_b^{i+1}. By construction, we have $\mu(S_j^{i+1}) \geqslant \frac{1}{2}\mu(S_j^i) \geqslant 2^{-i-1}\mu(S_j)$, the second of the two desired properties holds, and we have completed step $i+1$.

We may therefore assume that we finish the iterative process, and in the end we have sets $Y_j = S_j^{2^{d+1}-2}$ with $\mu(Y_j) \geqslant 2^{2-2^{d+1}}\mu(S_j)$ for $1 \leqslant j \leqslant d+1$, and for each $1 \leqslant k \leqslant 2^{d+1}-2$, there is a hyperplane through q that separates the sets $\{Y_j\}_{j \in X_k}$ from the sets $\{Y_j\}_{j \in [d+1] \setminus X_k}$. By Lemma 7.2, this implies that every simplex with one vertex in each Y_j contains q. □

We are now ready to complete the proof of Theorem 1.8. Given a point $q \in \mathbb{R}^d$, define the hypergraph H_q with vertex set $\mathbb{R}^d$ as the set of all $(d+1)$-tuples of points whose convex hulls contain q. Lemma 7.4 states that for each finite measure μ on $\mathbb{R}^d$ such that every hyperplane has measure 0, the hypergraph H_q is (c_d, μ)-structured. Theorem 1.8 then follows from Theorem 6.1.

8. Optimal sparse constructions in space—Proof of Theorem 1.7

In this section, we deduce Theorem 1.7 from Corollary 1.9. Let $H = (V, E)$ be a $(d+1)$-uniform hypergraph. The *edge density*

$$\rho(H) \stackrel{\text{def}}{=} \frac{|E|}{\binom{|V|}{d+1}}$$

is the fraction of $(d+1)$-tuples that are hyperedges of H. For vertex subsets $V_1, \ldots, V_{d+1} \subseteq V$, the edge density $\rho(V_1, \ldots, V_{d+1})$ is defined as the fraction of $(d+1)$-tuples in $V_1 \times \ldots \times V_{d+1}$ that are hyperedges of H.

Proof of Theorem 1.7. Let $t = 8d^2/\varepsilon$, so that for any equipartition of the vertex set of a complete $(d+1)$-uniform hypergraph into $k \geqslant t$ parts, all but at most an $\frac{\varepsilon}{8}$-fraction of its hyperedges have their vertices in different parts. Indeed, the fraction of hyperedges with one vertex in each part is

$$\left(\frac{n}{k}\right)^{d+1} \cdot \frac{\binom{k}{d+1}}{\binom{n}{d+1}} \geqslant \prod_{i=1}^{d}\left(1-\frac{i}{k}\right) \geqslant 1-\frac{d^2}{k} \geqslant 1-\frac{\varepsilon}{8}.$$

Let $K = K(\varepsilon/8, d)$ be the constant from Corollary 1.9, and $k = \max\{K, t\}$.

Let $r(d,\varepsilon)$ be sufficiently large so that for any $r \geqslant r(d,\varepsilon)$ and n a sufficiently large multiple of $d+1$, there is an r-regular $(d+1)$-uniform hypergraph H on n vertices, whose hyperedges are uniformly distributed in the sense that for any disjoint vertex subsets $V_1, \ldots, V_{d+1} \subseteq V(H)$ with $|V_i| \geqslant \frac{n}{k}$ for $1 \leqslant i \leqslant d+1$,

$$\left|\frac{\rho(V_1,\ldots,V_{d+1})}{\rho(H)} - 1\right| \leqslant \frac{\varepsilon}{4}. \tag{14}$$

The existence of an integer $r(d,\varepsilon)$ and a hypergraph H with the above properties follows from the standard fact that an n-vertex r-regular $(d+1)$-uniform hypergraph H chosen uniformly at random from all n-vertex r-regular $(d+1)$-uniform hypergraphs, meets the requirements for large enough r with probability tending to 1 as $n \to \infty$.

Let $P \subseteq \mathbb{R}^d$ satisfy $|P| = n$. Since n is sufficiently large, there is a point q that is in at least $c(d) - \frac{\varepsilon}{4}$ fraction of the simplices with vertices in P. Since $k \geqslant K$, by Corollary 1.9, there is an equipartition $P = P_1 \cup \ldots \cup P_k$ such that all but at most an $\frac{\varepsilon}{8}$-fraction of the $(d+1)$-tuples $P_{i_1}, \ldots, P_{i_{d+1}}$ are homogenous with respect to q. Since $k \geqslant t$, all but at most an $\frac{\varepsilon}{8} + \frac{\varepsilon}{8} = \frac{\varepsilon}{4}$-fraction of the $(d+1)$-tuples of points of P have their vertices in $d+1$ different parts of the partition, and these parts are homogeneous. Since q is in at least a fraction $c(d) - \frac{\varepsilon}{4}$ of the simplices with vertices in P, at least a fraction $c(d) - \frac{\varepsilon}{2}$ of the $(d+1)$-tuples of points of P span a simplex containing q and having its vertices in $d+1$ different parts of the partition such that these parts are homogeneous.

Let $f : V(H) \to P$ be an arbitrary bijection between the vertices of H and P. Write $V_i = f^{-1}(P_i)$ for $1 \leqslant i \leqslant k$. Note that if $P_{i_1}, P_{i_2}, \ldots, P_{i_{d+1}}$ are homogeneous with respect to q such that there is a simplex containing q with one vertex in each of these parts, then necessarily all the simplices with one vertex in each of these parts contains q. By (14), for all $d+1$ parts $V_{i_1}, \ldots, V_{i_{d+1}}$, the hyperedge density in H between these parts is at least $\left(1 - \frac{\varepsilon}{4}\right)\rho(H)$. Putting this together with the previous paragraph, we get that at least a fraction $\left(c(d) - \frac{\varepsilon}{2}\right)\left(1 - \frac{\varepsilon}{4}\right) \geqslant c(d) - \varepsilon$ of the hyperedges of H induce simplices containing q. We conclude that $c(H) \geqslant c(d) - \varepsilon$. □

9. Sparse hypergraphs with nearly maximal overlap number

Here we prove the following result, which implies Theorem 1.6.

Theorem 9.1. *Let d and Δ be positive integers and $\varepsilon > 0$. If $n \geqslant 2^9\varepsilon^{-3}d^9\Delta^3$ and H is a $(d+1)$-uniform hypergraph with n vertices, maximum degree Δ, and without isolated vertices, and P is a set of n points in $\mathbb{R}^d$ such that no point in $\mathbb{R}^d$ is in a fraction more than c of the simplices with vertices in P, then there is a bijection $f : V(H) \to P$ such that no point of $\mathbb{R}^d$ is in a fraction more than $c + \varepsilon$ of the simplices whose vertices are the image by f of a hyperedge of H.*

We actually show that almost surely we may take f to be a uniform random bijection.

We shall use below Azuma's inequality (see, e.g., [1]), which asserts that if $c = X_0, \dots, X_n$ is a martingale with $|X_{a+1} - X_a| \leqslant t$ for all $0 \leqslant a \leqslant n-1$, then

$$\Pr\left[|X_n - c| > \lambda\sqrt{n}\right] < 2e^{-\frac{\lambda^2}{2t^2}}. \tag{15}$$

Let H and F be hypergraphs each with the same number of vertices. For a bijection $f : V(H) \to V(F)$, let m_f denote the number of hyperedges of H whose image is a hyperedge of F.

Lemma 9.2. *Let H and F be k-uniform hypergraphs each with n vertices such that H has maximum degree Δ. Then the probability that for a random bijection $f : V(H) \to V(F)$, the number m_f deviates from $|E(H)| \cdot |E(F)|/\binom{n}{k}$ by more than $\lambda\sqrt{n}$ is at most $2e^{-\frac{\lambda^2}{2(2k+1)^2\Delta^2}}$.*

Proof. Let $e_1, \dots, e_{|E(H)|}$ denote the hyperedges of H. For a hyperedge e_j, let $X(e_j)$ be the indicator random variable of the event that the image of e_j by f is a hyperedge of F. That is, $X(e_j) = 1$ if $f(e_j)$ is a hyperedge of F, and $X(e_j) = 0$ otherwise. Let Y denote the random variable counting the number of hyperedges of H whose image is a hyperedge of F, so $Y = \sum_{j=1}^{|E(H)|} X(e_j)$. Each hyperedge of H has a probability $|E(F)|/\binom{n}{k}$ of being mapped by f to a hyperedge of F. By linearity of expectation, the expected value of Y is $\mathbb{E}[Y] = |E(H)| \cdot |E(F)|/\binom{n}{k}$.

Let $V(H) = [n] = \{1, \dots, n\}$. For $a = 0, \dots, n$, let $X_a(e_j)$ be the probability of the event that the image of e_j by f is a hyperedge of F after picking $f(1), \dots, f(a)$, and let $Y_a \stackrel{\text{def}}{=} \sum_{j=1}^{|E(H)|} X_a(e_j)$ denote the expected value of Y after picking $f(1), \dots, f(a)$. So Y_0 denotes the expected value of m_f, which is $|E(H)| \cdot |E(F)|/\binom{n}{k}$.

By construction, $\{Y_a\}_{a=0}^n$ is a martingale. We next give an upper bound on

$$|Y_{a+1} - Y_a| \leqslant \sum_{j=1}^{|E(H)|} |X_{a+1}(e_j) - X_a(e_j)|,$$

so as to apply Azuma's inequality (15). For the $\leqslant \Delta$ hyperedges e_j that contain $a+1$, we bound $|X_{a+1}(e_j) - X_a(e_j)| \leqslant 1$. For those hyperedges with all vertices in $\{1, \dots, a\}$, we have $X_{a+1}(e_j) = X_a(e_j)$. Let e_j be a hyperedge which does not contain $a+1$, and contains a vertex which is more than $a+1$. Let $i_1, \dots, i_h$ be the vertices of e_j that are at most a, so $h < k$ as e_j contains a vertex which is more than $a+1$. Let $w(e_j) = k - h$ denote the number of vertices of e_j which are greater than $a+1$. All of these vertices are in $\{a+2, \dots, n\}$, and therefore $n - a \geqslant k - h + 1$. It follows that

$$\begin{aligned} n - a - w(e_j) = n - a - k + h &= \frac{n-a}{k} + \left(1 - \frac{1}{k}\right)(n-a) - k + h \\ &\geqslant \frac{n-a}{k} + \left(1 - \frac{1}{k}\right)(k - h + 1) - k + h = \frac{n-a-1}{k} + \frac{h}{k} \geqslant \frac{n-a-1}{k}. \end{aligned} \tag{16}$$

Let Z denote the number of hyperedges of F that contain $f(i_1), \dots, f(i_h)$ and whose remaining vertices are in $V(F) \setminus \{f(1), \dots, f(a)\}$, and Z' denote the number of hyperedges of F that contain $f(i_1), \dots, f(i_h)$ and whose remaining vertices are in $V(F) \setminus \{f(1), \dots, f(a), f(a+1)\}$.

We have $Z \geqslant Z' \geqslant Z - \binom{n-a-1}{k-h-1}$, as $f(i_1), \ldots, f(i_h), f(a+1)$ are in at most $\binom{n-a-1}{k-h-1}$ hyperedges of F whose remaining vertices are in $V(F) \setminus \{f(1), \ldots, f(a), f(a+1)\}$. We also have $Z \leqslant \binom{n-a}{k-h}$. Note that $X_a(e_j) = \frac{Z}{\binom{n-a}{k-h}}$ and $X_{a+1}(e_j) = \frac{Z'}{\binom{n-a-1}{k-h}}$. Hence,

$$\begin{aligned}
|X_a(e_j) - X_{a+1}(e_j)| &= \left| \frac{Z}{\binom{n-a}{k-h}} - \frac{Z'}{\binom{n-a-1}{k-h}} \right| \\
&\leqslant Z\left(\binom{n-a-1}{k-h}^{-1} - \binom{n-a}{k-h}^{-1}\right) + (Z - Z')\binom{n-a-1}{k-h}^{-1} \\
&\leqslant \left(\frac{n-a}{n-a-k+h} - 1\right) + \frac{k-h}{n-a-k+h} \\
&= \frac{2(k-h)}{n-a-k+h} \\
&= 2\frac{w(e_j)}{n-a-w(e_j)},
\end{aligned}$$

where the first inequality is the triangle inequality, the second inequality follows from substituting in $Z \leqslant \binom{n-a}{k-h}$ and $Z - Z' \leqslant \binom{n-a-1}{k-h-1}$.

We have $\sum w(e_j) \leqslant \Delta(n-a-1)$, where the sum is over all hyperedges that contain a vertex greater than $a+1$, as each vertex has degree at most Δ. Hence, the sum of $|X_a(e_j) - X_{a+1}(e_j)|$ over all vertices that contain a vertex greater than $a+1$ is at most

$$2\sum \frac{w(e_j)}{n-a-w(e_j)} \leqslant 2\Delta k,$$

where we used (16). Putting this altogether, we have $|Y_{a+1} - Y_a| \leqslant \Delta(2k+1)$. By Azuma's inequality (15), the probability that $Y_n = Y = m_f$ differs from $X_0 = |E(H)| \cdot |E(F)|/\binom{n}{k}$ by more than $\lambda\sqrt{n}$ is at most $2e^{-\frac{\lambda^2}{2(2k+1)^2\Delta^2}}$. □

We use the following well known fact (see [30], pp. 43–52).

Lemma 9.3. *Any set P of n points in $\mathbb{R}^d$ determine at most n^{d^2} regions, such that any two points in the same region are in the same collection of simplices with vertices in P.*

The two previous lemmas are all the tools we need to complete the proof of Theorem 9.1.

Proof of Theorem 9.1: Since H does not have isolated vertices, the number $|E(H)|$ of hyperedges of H is at least $n/(d+1)$. Let P be a set of n points in $\mathbb{R}^d$ such that no point $q \in \mathbb{R}^d$ is in more than a fraction c of the simplices whose vertices are in P. By Lemma 9.3, P determines at most n^{d^2} regions, such that any two points in the same region are in the same collection of simplices with vertices in P. Let q be a representative point for a region, $c_q \leqslant c$ denote the fraction of simplices with vertices in P that contain q, and let F_q denote the hypergraph on P consisting of all simplices with vertices in P that contain the point q.

Let $f : V(H) \to P$ be a bijection taken uniformly at random. By Lemma 9.2, the probability that the fraction of hyperedges of H which map to simplices containing q is at least $c_q + \varepsilon$ is at most $2e^{-\frac{\lambda^2}{2(2d+3)^2\Delta^2}}$, where $\lambda\sqrt{n} = \varepsilon|E(H)| \geqslant \varepsilon n/(d+1)$. However, there are at most n^{d^2} such hypergraphs F_q, and since $n \geqslant 2^9\varepsilon^{-3}d^9\Delta^3$, the probability is at most

$$n^{d^2}e^{-\frac{\lambda^2}{2(2d+3)^2\Delta^2}} \leqslant \exp\left(d^2\log n - \frac{(\varepsilon\sqrt{n}/(d+1))^2}{2(2d+3)^2\Delta^2}\right) \leqslant \exp\left(d^2\log n - \frac{\varepsilon^2 n}{2(2d+3)^4\Delta^2}\right) = o(1)$$

that there is a point in $\mathbb{R}^d$ contained in more than a fraction $c+\varepsilon$ of the hyperedges of H. Thus, with high probability, a uniformly random bijection has the desired property. $\square$

10. Concluding remarks

An alternative proof of Theorem 1.3 was given by Gromov [17], based on an application of Garland's vanishing theorem to finite quotients of certain Bruhat-Tits building. Gromov's argument fails for $d > 2$, and does not yield the sharp constant $\frac{2}{9}$ as in Theorem 1.3, yet his construction has some remarkable stronger properties which we now describe.

Gromov investigated [17] the role of the fact that the edges of the triangles in the Boros-Füredi theorem are assumed to be straight line segments. He showed that it suffices to replace "straight lines" by sufficiently regular Jordan arcs. Gromov's construction based on Garland's theorem enjoys this stronger property as well (see Section 2.10 in [17]): one just needs the associated mapping from Gromov's simplicial complex to $\mathbb{R}^2$ to be continuous, the image of each edge to be nowhere dense in $\mathbb{R}^2$, and that its restriction to each face is at most r-to-1 (in which case the resulting bounds depend on r). At present it is unknown whether for $d \geqslant 3$ there exist arbitrarily large bounded degree d-dimensional simplicial complexes which are highly overlapping with respect to non-affine embeddings into $\mathbb{R}^d$.

An inspection of the construction of Section 4.2 reveals that, if the graph G has girth greater than $2d$, then the resulting bounded degree d-dimensional highly overlapping simplicial complexes admit a continuous and piecewise affine retraction onto their 1-skeleton. It follows that for these complexes, if we replace "simplices" by "generalized simplices" whose edges are allowed to be continuous and piecewise affine arcs rather than straight lines, then the conclusion that there must be a point in a constant fraction of these generalized simplices, corresponding to an embedding of their vertex set into $\mathbb{R}^d$, would fail. Thus, the situation in the sparse setting is subtle, and passing from the case of affine mapping to more general continuous mappings is not automatic.

It follows for instance from [4] and [27] that for any system of at least constant times n^3 triangles induced by a set of n points in the plane in general position, there is a point covered by at least a positive fraction of all triangles. In the present paper, we studied sparse systems of triangles with similar properties. Another closely related question is the following. For any positive integers n and $1 < t < \binom{n}{3}$, determine the largest number m such that for any system of at least t triangles induced by a set of n points in the plane, there is a point contained in at least m triangles. See [2], [25]. The best known general lower bound is roughly t^3/n^6, but for most values of the parameters this is probably a very weak bound.

Acknowledgements. We thank Gábor Tardos for many useful conversations and ideas, and Yves Benoist for his help on Section 5.

References

[1] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[2] B. Aronov, B. Chazelle, H. Edelsbrunner, L. J. Guibas, M. Sharir, and R. Wenger. Points and triangles in the plane and halving planes in space. *Discrete Comput. Geom.*, 6(5):435–442, 1991.

[3] C. M. Ballantine. Ramanujan type buildings. *Canad. J. Math.*, 52(6):1121–1148, 2000.

[4] I. Bárány. A generalization of Carathéodory's theorem. *Discrete Math.*, 40(2-3):141–152, 1982.

[5] B. Bekka, P. de la Harpe, and A. Valette. *Kazhdan's property (T)*, volume 11 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2008.

[6] Y. Benoist. Five lectures on lattices in semisimple Lie groups. 2004. To appear in the collection "Séminaires et congrès" of the SMF. Preprint available at `http://www.math.u-psud.fr/~benoist/prepubli/04lattice.pdf`.

[7] A. Borel and G. Harder. Existence of discrete cocompact subgroups of reductive groups over local fields. *J. Reine Angew. Math.*, 298:53–64, 1978.

[8] E. Boros and Z. Füredi. The number of triangles covering the center of an n-set. *Geom. Dedicata*, 17(1):69–77, 1984.

[9] B. Bukh. A point in many triangles. *Electron. J. Combin.*, 13(1):Note 10, 3 pp. (electronic), 2006.

[10] B. Bukh, J. Matoušek, and G. Nivasch. Stabbing simplices by points and flats. 2006. Preprint availabel at arXiv:0804.4464, to appear in Discrete Comput. Geom.

[11] D. I. Cartwright, P. Solé, and A. Żuk. Ramanujan geometries of type $\tilde{A}_n$. *Discrete Math.*, 269(1-3):35–43, 2003.

[12] D. I. Cartwright and T. Steger. A family of $\tilde{A}_n$-groups. *Israel J. Math.*, 103:125–140, 1998.

[13] J. G. Ceder. Generalized sixpartite problems. *Bol. Soc. Mat. Mexicana (2)*, 9:28–32, 1964.

[14] G. Davidoff, P. Sarnak, and A. Valette. *Elementary number theory, group theory, and Ramanujan graphs*, volume 55 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2003.

[15] D. Dubhashi and D. Ranjan. Balls and bins: a study in negative dependence. *Random Structures Algorithms*, 13(2):99–124, 1998.

[16] J. Eckhoff. Helly, Radon, and Carathéodory type theorems. In *Handbook of convex geometry, Vol. A, B*, pages 389–448. North-Holland, Amsterdam, 1993.

[17] M. Gromov. Singularities, expanders and topology of maps, part 2: From combinatorics to topology via algebraic isoperimetry. 2008. To appear in Geom. Funct. Anal.

[18] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006.

[19] W.-C. W. Li. Ramanujan hypergraphs. *Geom. Funct. Anal.*, 14(2):380–399, 2004.

[20] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[21] A. Lubotzky, B. Samuels, and U. Vishne. Explicit constructions of Ramanujan complexes of type $\tilde{A}_d$. *European J. Combin.*, 26(6):965–993, 2005.

[22] A. Lubotzky, B. Samuels, and U. Vishne. Ramanujan complexes of type $\tilde{A}_d$. *Israel J. Math.*, 149:267–299, 2005. Probability in mathematics.

[23] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[24] J. Matoušek. *Using the Borsuk-Ulam theorem*. Universitext. Springer-Verlag, Berlin, 2003. Lectures on topological methods in combinatorics and geometry, Written in cooperation with Anders Björner and Günter M. Ziegler.

[25] G. Nivasch and M. Sharir. Eppstein's bound on intersecting triangles revisited. *J. Combin. Theory Ser. A*, 116(2):494–497, 2009.

[26] H. Oh. Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants. *Duke Math. J.*, 113(1):133–192, 2002.

[27] J. Pach. A Tverberg-type result on multicolored simplices. *Comput. Geom.*, 10(2):71–76, 1998.

[28] J. Pach and J. Solymosi. Structure theorems for systems of segments. In *Discrete and computational geometry (Tokyo, 2000)*, volume 2098 of *Lecture Notes in Comput. Sci.*, pages 308–317. Springer, Berlin, 2001.
[29] M. S. Pinsker. On the complexity of a concentrator. In *7th International Telegraffic Conference*, pages 318/1–318/4. 1973.
[30] G. Polya. *Induction and analogy in mathematics. Mathematics and plausible reasoning, vol. I.* Princeton University Press, Princeton, N. J., 1954.
[31] A. Sarveniazi. Explicit construction of a Ramanujan $(n_1, n_2, \ldots, n_{d-1})$-regular hypergraph. *Duke Math. J.*, 139(1):141–171, 2007.
[32] W. Sierpiński. Sur les fonctions d'ensemble additives et continues. *Fund. Math.*, 3:240–246, 1922.
[33] T. Steger. Local fields and buildings. In *Harmonic functions on trees and buildings (New York, 1995)*, volume 206 of *Contemp. Math.*, pages 79–107. Amer. Math. Soc., Providence, RI, 1997.

Jacob Fox: Department of Mathematics, Princeton University. Email: jacobfox@math.princeton.edu.
Research supported by an NSF Graduate Research Fellowship and a Princeton Centennial Fellowship.

Mikhail Gromov: IHES and Courant Institute, NYU. Email: gromov@cims.nyu.edu.

Vincent Lafforgue: CNRS, Université Paris 7 Denis Diderot. Email: vlafforg@math.jussieu.fr.

Assaf Naor: Courant Institute, NYU. Email: naor@cims.nyu.edu. Research supported by NSF grants CCF-0635078 and CCF-0832795, BSF grant 2006009, and the Packard Foundation.

János Pach: EPFL, CUNY, and Courant Institute, NYU. Email: pach@cims.nyu.edu. Supported by NSF Grant CCF-08-30272, and by grants from NSA, PSC-CUNY, the Hungarian Research Foundation OTKA, and BSF.