Large Simple d-Cycles in Simplicial Complexes

Roy Meshulam [*] Ilan Newman [†] Yuri Rabinovich[‡]

August 15, 2023

Abstract

We show that the size of the largest simple d-cycle in a simplicial d-complex K is at least a square root of K's density. This is a higher dimensional analogue of a classical result of Erdős and Gallai [7] for graphs.

1 Introduction

Let $G = (V, E)$ be a (finite) simple graph. A classical result of Erdős and Gallai [7] asserts that if $|E| > \frac{k(|V|-1)}{2}$, then G contains a simple cycle of length $> k$. In this paper we study the analogous question for higher dimensional simplicial complexes.

A (homological) d-cycle in a d-dimensional simplicial complex is a set of d-faces with coefficients in a ring R, whose *boundary* is 0. A d-cycle on a set of faces C is *simple* if every cycle that is supported on C is either trivial, or has full support on C. In graphs the choice of the ring R is not important as far as the structure of the simple cycles is concerned. In higher dimensions it is of importance. In this paper we assume that R is an arbitrary field $\mathbb{F}$. One advantage of working over a field is that it introduces a matroidal structure on the set of d-faces of a complex, where a subset of d-faces is *independent* if it supports no nontrivial d-cycles. Our results exploits the combinatorial structure of d-cycles, and does not dependent on the choice of $\mathbb{F}$.

Let $c(G)$ denote the size of the maximum simple cycle in G. The graph-theoretic lower bound of Erdős and Gallai [7] can be interpreted in two somewhat different ways. The first interpretation is that $c(G)$ is linear in $D = 2|E|/|V|$, the average degree of G. The second interpretation is that $c(G)$ is linear in $|E|/\mathrm{rank}(G)$, and even in $\max_{G' \subseteq G} |E(G')|/\mathrm{rank}(G')$, where the rank of a graph is the size of a maximum acyclic subset of edges in it. The latter interpretation is more suitable for a generalization, and we shall pursue it for the most part of the paper. It will also imply a generalization of the former interpretation.

Let $\gamma(G) = \max_{G' \subseteq G} |E(G')|/\mathrm{rank}(G')$, where the maximum is taken over all subgraphs of G. (To avoid discussing degeneracies, we assume here $0/0 = 0$.) $\gamma(G)$ is a standard graph theoretic parameter which measures the maximum local density of G. It appears e.g., in Nash-Williams Theorem (see [4]), where it determines exactly the minimum number of subforests of G required to cover $E(G)$. Nash-Williams

[*]Department of Mathematics, Technion, Haifa 32000, Israel. E-mail: meshulam@technion.ac.il. Supported in part by Israel Science Foundation, grant 326/16.

[†]Department of Computer Science, University of Haifa, Haifa, Israel. E-mail: ilan@cs.haifa.ac.il. Supported in part by the Israel Science Foundation, grant 497/17.

[‡]Department of Computer Science, University of Haifa, Haifa, Israel. E-mail: yuri@cs.haifa.ac.il.

theorem states that this number is precisely $\lceil\gamma(G)\rceil$. This was generalizes to matroids, and hence to simplicial complexes. A classical result of Rado (see e.g., [13]) asserts that the minimum number of independent sets in a (loopless) matroid M required to cover M is $\lceil\gamma(M)\rceil$, where $\gamma(M) = \max_{A\subset E(M)} |A|/\text{rank}(A)$. Moreover, $\gamma(M)$ is efficiently computable.

In this paper we obtain lower bounds on $c(M)$, the size of the largest simple cycle of M, in terms of $\gamma(M)$. Our main result is the following.

Theorem *Let K be a simplicial complex containing nontrivial d-cycles. Let $f_\ell(K)$ be the number of ℓ-simplices in K. Then, K contains a simple d-cycle of size at least* $\sqrt{2/(d+1)} \cdot \sqrt{f_d(K)/f_{d-1}(K)} - 1$.

The paper contributes to the rapidly evolving study of the combinatorics of simplicial complexes in the context of their homological and homotopical properties. Let us mention, e.g., [1, 10, 11, 6]. The latter elegant paper [6] deals with a higher dimensional analogue of another extremal problem from graph theory, the Moore bound (see, e.g., [3]), and shows that dense simplicial complexes contain *small* cycles, providing almost tight quantitative upper and lower bounds.

2 Preliminaries

We use here standard notations from the area of combinatorial simplicial complexes. We also need some basic facts from Matroid theory. For completeness we presents these notations and fact here. The reader feeling at home in these areas is advised to skip this section.

Simplicial complexes: A d-dimensional simplex, abbreviated as d-simplex, is a set $\sigma \subseteq [n]$. A simplicial complex K is a collection of simplices over $[n]$ that is closed under containment, i.e., if $\sigma \in K$, then so are all the subsets of σ, aka the faces of σ. The dimension of K is the largest dimension over all its simplices. Some of the complexes discussed in this paper are *pure* d-dimensional complexes, i.e., all the maximal faces of K are all of the same dimension d. Maximal faces are called *facets*.

Let K be a finite d-dimensional simplicial complex on the vertex set V and let $\mathbb{F}$ be a field. Let $\prec$ be a fixed linear order on V. We orient each simplex in K according to this order, i.e., $\sigma = [v_0, \ldots, v_i]$ if $v_0 \prec \cdots \prec v_i$. Let $K^{(i)}$ denote the set of oriented i-dimensional faces of K, and let $f_i(K) = |K^{(i)}|$. Let $C_i(K;\mathbb{F})$ be the space of i-chains of K, where a chain is a formal sum of $\mathbb{F}$-weighted oriented i-simplices in $K^{(i)}$. The *boundary* of d-simplex $\sigma = [v_0, \ldots, v_d]$ is $\partial_d(\sigma) = \sum_{j=0}^{d}(-1)^j[v_0, \ldots, v_{j-1}, v_{j+1}, \ldots v_d]$. This linearly extends to the *boundary map* $\partial_d : C_d(K;\mathbb{F}) \to C_{d-1}(K;\mathbb{F})$.

Having defined the linear map ∂_d, define $Z_d(K;\mathbb{F}) = \ker \partial_d$ to be the linear space of d-*cycles* in K, and $B_{d-1}(K;\mathbb{F}) = \text{Im}\, \partial_d$ to be the linear space of $(d-1)$-*boundaries* in K.

Let $C = \sum_{i=1}^{m} \alpha_i\sigma_i \in Z_d(X;\mathbb{F})$ be a non-empty d-cycle in K, where for $i = 1, \ldots, m$, $\alpha_i \neq 0$, and the σ_i's are distinct d-simplices. The *support* of C is $\text{supp}(C) = \{\sigma_1, \ldots, \sigma_m\}$. A non-empty d-cycle $C = \sum_{i=1}^{m} \alpha_i\sigma_i$ is *simple* if the set $\{\partial_d\sigma_1, \ldots \partial_d\sigma_m\} \setminus \{\partial_d\sigma_i\}$ is linearly independent over $\mathbb{F}$ for every i, $1 \leq i \leq m$. Equivalently, C is simple if its *support* does not strictly contain the support of any other non-trivial d-cycle in K.

Matroids: We only list here some of the notions of the matroid theory that are essential for the forthcoming discussion. For a systematic and detailed exposition of the matroid theory see, e.g., [13].

Let $M = (E, \mathcal{I})$ be a matroid on a finite ground set E, with $\mathcal{I}$ its collection of independent sets. The independent sets are closed under inclusion and satisfy the exchange property, i.e., if $I, J \in \mathcal{I}$ and $|J| > |I|$, then there exists $j \in J \setminus I$ such that $I \cup \{j\} \in \mathcal{I}$.

Some common examples of matroids are the *graphic matroids* and the *linear matroids*. Given an undirected graph G, the corresponding graphic matroid is $(E[G], \mathcal{F})$, where $E[G]$ are the edges of G, and $\mathcal{F}$ are the acyclic subsets of $E[G]$. A linear matroid $(V, \mathcal{B})$ has as a ground set a finite subset V of some linear space over a field $\mathbb{F}$, and $\mathcal{B}$ is formed by all the linearly independent subsets of V.

The *rank* of $A \subseteq E$, denoted $\mathrm{rank}(A)$, is the size of the maximum independent set in it. The rank of M is defined to be $\mathrm{rank}(E)$. A *circuit* is a minimally dependent set.

The *span* (also called the *closure*) of A in M, denoted by $\mathrm{span}(A)$, is the maximal subset of E containing A, and satisfying $\mathrm{rank}(\mathrm{span}(A)) = \mathrm{rank}(A)$. In other words, in addition to A, $\mathrm{span}(A)$ contains all $e \in E$ such that there exists a circuit in $A \cup \{e\}$ containing e.

A matroid M is called *loopless* if for every $e \in E(M)$, $\{e\}$ is not a circuit. It has no *double edges* if for every $e, f \in E(M)$, $\{e, f\}$ is not a circuit. A loopless matroid without double edges is called *simple*.

Let $\sim$ be the binary relation on $E(M)$, where $e \sim f$ if there exists a circuit containing both. This relation is an equivalence relation. The equivalence classes $\{E_i\}$ of this relation are called the *connected components* of M. It holds that $\sum_i \mathrm{rank}(E_i) = \mathrm{rank}(E)$. Furthermore, for all i, $\mathrm{span}(E_i) = E_i$. The above properties can be summarized by saying that M is a *direct sum* of M_i's, where M_i is the submatroid of M on the connected component E_i. The M_i's are called the *components* of M. The matroid M is called *connected* if it has a single component.

A *minor* of a matroid M, just like a graph minor, is a matroid obtained from M by series of element *deletions* and *contractions*. Deletion of $e \in E = E(M)$ results in the matroid $M \setminus e = (E \setminus \{e\}, \mathcal{I}')$ where $\mathcal{I}' = \{I \subseteq E \setminus \{e\} | \ I \in \mathcal{I}\}$. Contraction of e results in the matroid $M/e = (E \setminus \{e\}, \mathcal{I}')$ where $\mathcal{I}' = \{I \subseteq E \setminus \{e\} \mid I \cup \{e\} \in \mathcal{I}\}$.

Let A be a subset of E. Contracting A results in a matroid M/A on $E \setminus A$, where for each circuit C of M, the set $C \setminus A \subseteq E(M/A)$ splits into a disjoint union of circuits, possibly loops, in M/A. In the other direction, if C' is a circuit of M/A, then $C' \cup A$ contains a circuit in M. For any set $A \subset E(M)$, it holds that $\mathrm{rank}(M) = \mathrm{rank}(A) + \mathrm{rank}(M/A)$.

Let K be a simplicial complex. The simplicial matroid $(K^{(d)}, \mathcal{A})$ with respect to a field $\mathbb{F}$ has the d-faces of K as its ground set, and $\mathcal{A}$ is formed by all subsets $A \subseteq K^{(d)}$ that support no nontrivial d-cycles over $\mathbb{F}$. A simplicial matroid $(K^{(d)}, \mathcal{A})$ over $\mathbb{F}$ is isomorphic to a linear matroid over $\mathbb{F}$, and thus is $\mathbb{F}$-*representable*. The representation is given by $\phi : K^{(d)} \mapsto \mathbb{F}^{f_{d-1}(K)}$, where the vectors are indexed by the $(d-1)$-faces $K^{(d-1)}$, and, for a d-face σ and a $(d-1)$-face τ in K, and the value of τ's coordinate of $\phi(\sigma)$ is the coefficient of τ in $\partial_d \sigma$.

3 Approaches and Results

We present in Section 3.2 a basic approach that bounds $C(M)$ in terms of $\gamma(M)$ for general matroids. Then, in Section 4, we discuss simplicial complexes and prove our main result.

Let M be a matroid. Recall that $c(M)$ denotes the size of a maximum circuit in M, and $\gamma(M) = \max_{A \subset E(M)} |A|/\mathrm{rank}(A)$. As before, we treat $0/0$ as 0.

3.1 Using Forbidden Minors: $\mathbb{F}_q$-representable Matroids

Observe that the class of graphs G with $c(G) < k$ is precisely the class of graphs lacking C_k, the size-k cycle, as a minor. Indeed, the deletion does not create new cycles, while the contraction only shrinks or leaves alone the existing ones. This observation could be employed to obtain a weaker version of [7] by using classical results (see, e.g., [14]) about the density of graphs lacking a size-k minor.

As with graphs, and for the same reason, the class of matroids M with $c(M) < k$ is precisely the class of matroids lacking as minor(s) the matroids associated with size-k circuits. Observe that the latter are all isomorphic to U_k^{k-1}, a matroid on a ground set of size k, where all the subsets of size $< k$ are independent. It is also the graphic matroid associated with the graph C_k. This allows us to use the following hard result of Geelen and Whittle [8] (see also [9]):

Theorem 1. **[8]** *Let M be a simple $\mathbb{F}_q$-representable matroid (equivalently, a linear matroid over a field of size q) lacking a graphic minor of a graph on k vertices. Then,*

$$\gamma(M) \; < \; q^{q^{3k}} \; .$$

Corollary 2. *For M as above,* $c(M) \; > \; \frac{1}{3} \log_q \log_q \gamma(M)$.

While this does establish a weak lower bound on $c(M)$ in terms of $\gamma(M)$ for, say, binary matroids, the bound can be considerably strengthened; see below. For infinite fields like $\mathbb{Q}$ it yields nothing.

3.2 Using Seymour's Lemma: General Matroids

The following theorem by Seymour will be used.

Theorem 3. **(Th. 3.4 in [5])** *Let M be a connected loopless matroid, $|E(M)| > 1$, and let C be a maximum size circuit in M. Then, size of the maximum circuit in the matroid M/C obtained from M by contracting C, is strictly less than $|C|$.*

For matroids with loops, $\gamma(M)$ is not interesting, as formally $\gamma(M) = \infty$. For loopless matroids, as far as proving lower bounds on $c(M)$ in terms of $\gamma(M)$ goes, we may restrict our attention to connected matroids. Indeed, assume that M has connected components $\{M_i\}_1^\ell$. Then, it holds that $\gamma(M) = \max_i \gamma(M_i)$. The direction "$\geq$" is obviously true. For the direction "$\leq$", let $K \subseteq M$ be the subset of elements on which $\gamma(M)$ is achieved, and let $K_i = K \cap E(M_i)$. One has

$$\gamma(M) \; = \; \frac{|K|}{\text{rank}(K)} \; = \; \frac{\sum_i |K_i|}{\sum_i \text{rank}(K_i)} \; \leq \; \max_i \frac{|K_i|}{\text{rank}(K_i)} \; \leq \; \max_i \gamma(M_i) \; .$$

Since a circuit of M_i is also a circuit of M, it trivially holds that $c(M) \geq c(M_i)$. Thus, a lower bound on $c(M_i)$ in terms of $\gamma(M_i)$ is also a lower bound on $c(M)$. In fact, $c(M) = \max_i c(M_i)$, since any circuit of M lies entirely in some M_i.

Definition 1. *Given a connected loopless matroid M with $|E(M)| > 1$, we define the following decomposition process of M, described by a tree $\mathcal{T}_M$:*

** Each vertex x of $\mathcal{T}_M$ will have an associated pair (M_x, C_x), where M_x is a connected loopless minor of M, and C_x is a maximum size circuit of M_x (an arbitrary choice is taken if there are several max-size circuits).*

** The matroid associated with the root vertex is the original M.*

** The children of the vertex x in $\mathcal{T}_M$ correspond to the components of M_x/C_x after the removal of the loops. If there are no non-empty components, that is,* $\text{rank}(M_x/C_x) = 0$, *then x is a leaf of $\mathcal{T}_M$.*

The following claim establishes some basic properties of $\mathcal{T}_M$.

Claim 4.

(i) $\mathcal{T}_M$ *is well defined (given the choice of* C_x *at every vertex* x*).*

(ii) *For any* x, y*, where* y *is the parent of* x *in* $\mathcal{T}_M$*, it holds that* $|C_y| > |C_x|$*.* *Consequently, the depth of* $\mathcal{T}_M$ *is smaller than* $c(M) - 1$*.*

(iii) $\sum_{x \in \mathcal{T}_M} (|C_x| - 1) \;=\; \mathrm{rank}(M)$.

Proof.

(i) One needs to verify that for any vertex x created in the process of $\mathcal{T}_M$'s generation, M_x has a circuit. Observe that by the definition of $\mathcal{T}_M$, every M_x is nonempty, connected and loopless. Thus, the only concern is that the ground set of M_x may consist of a single element. Call a vertex $y \in \mathcal{T}_M$ good if $|E(M_y)| > 1$. Let us show by induction that all $x \in \mathcal{T}_M$ are good.

By the initial assumption, $|E(M)| > 1$, and so the root vertex is good. Let y be a good vertex, with a child x in $\mathcal{T}_M$. The corresponding M_x is nonempty, and so it contains some element e. Then $e \in E(M_y)$ as well, and since y is good, and M_y is loopless and connected, there is a circuit C in M_y containing e and some other elements. In M_y/C_y, the set C splits into disjoint circuits, possibly loops. Thus e is contained in some circuit in M_y/C_y. Keeping in mind that M_x is a loopless component of M_y/C_y, the conclusion follows.

(ii) An immediate consequence of Theorem 3, and the fact that $|C_x| > 1$ for all $x \in \mathcal{T}_M$.

(iii) We claim that for *every* vertex z of $\mathcal{T}_M$ it holds that $\sum_{x \in \mathcal{T}_{M_z}} (|C_x| - 1) \;=\; \mathrm{rank}(M_z)$, where the sum is taken over all the vertices of the subtree $\mathcal{T}_{M_z}$ of $\mathcal{T}_M$ rooted at z. This extends the original statement, which claims this only for the root of $\mathcal{T}_M$.

The proof is by a bottom-up induction on the structure of $\mathcal{T}_M$. When z is a leaf, the rank of M_z/C_z is 0, or, equivalently, $\mathrm{span}(C_z) = E(M_z)$. Since C_z is minimally dependent, this implies $\mathrm{rank}(M_z) = |C_z| - 1$. Consider now a vertex $z \in \mathcal{T}_M$ with children $\{y_i\}_{i=1}^{\ell}$. By inductive assumption, the statement holds for all y_i's. Keeping in mind that M_z/C_z is a direct sum of its components, i.e., M_{y_i}'s and the removed loops. Using the fact that for any set A, $\mathrm{rank}(M) = \mathrm{rank}(A) + \mathrm{rank}(M/A)$, we conclude that:

$$\begin{aligned}\mathrm{rank}(M_z) \;&=\; \mathrm{rank}(C_z) \;+\; \mathrm{rank}(M_z/C_z) \;=\; (|C_y| - 1) \;+\; \sum_i \mathrm{rank}(M_{y_i}) \\ &=\; (|C_z| - 1) \;+\; \textstyle\sum_{x \in \mathcal{T}_{M_{y_i}}} (|C_y| - 1) = \; \textstyle\sum_{x \in \mathcal{T}_{M_z}} (|C_x| - 1)\,. \qquad \square\end{aligned}$$

Recall that our goal is to relate $c(M)$ to $\gamma(M)$, the maximum local density in M. The local density profile $\{s_M(i)\}$ of M, that is defined below, will play a key role in the forthcoming discussion.

Definition 2. *For a matroid* M*, and an integer* $i \geq 0$*, let*

$$s_M(i) \;=\; \max\big\{\, |A| : A \subset E(M),\ \mathrm{rank}(A) \leq i \big\}.$$

The following theorem is the central result of this section:

Theorem 5. *Let* M *be a loopless matroid with* $c(M) = k > 1$*. Then,*

$$\gamma(M) \;\leq\; s_M\Big((k-1)k/\,2\Big).$$

Proof. Let N be the submatroid of M on which $\gamma(M)$ is achieved. That is, $|E(N)| = \gamma(M) \cdot \mathrm{rank}(N)$. As we have seen, N is w.l.o.g., connected, loopless, with $|E(N)| > 1$. Let $r = c(N)$. We shall prove that

$$|E(N)| \;\le\; s_N\Big((r-1)r/\,2\Big) \cdot \mathrm{rank}(N),$$

This would imply the original statement, as $s_M(*)$ dominates $s_N(*)$, $r \le k$, and $s_M(*)$ is monotone nondecreasing.

Let $\mathcal{T}_N$ be the decomposition tree of N, and consider an element $e \in E(N)$. Tracing its "life" in $\mathcal{T}_N$, we conclude that the set of vertices $x \in \mathcal{T}_N$ such that $e \in E(N_x)$, constitutes a path P_z from the root to a vertex $z \in \mathcal{T}_N$, such that $e \in \mathrm{span}(\cup_{x \in P_z} C_x)$. Indeed, since e made it to N_z but not to its children, it was eliminated either as a part of C_z, or as a loop in N_z/C_z. This happens if and only if $e \in \mathrm{span}(C_z)$ in N_z. Now, in general, $(N/A)/B \cong N/(A \cup B)$, and thus N_z is a component of $(N/ \cup_{x \in P_z \setminus \{z\}} C_x)$ containing e. Hence, $e \in \mathrm{span}(C_z)$ in N_z implies that $e \in \mathrm{span}(\cup_{x \in P_z} C_x)$ in N.

Keeping in mind that all the elements of N get eventually eliminated during the decomposition process described by $\mathcal{T}_N$, the previous discussion leads to the conclusion

$$\bigcup_{z \in \text{ leaves of } \mathcal{T}_N} \mathrm{span}\left(\bigcup_{x \in P_z} C_x\right) \;=\; E(N)\,. \tag{1}$$

Now, for any $z \in \mathcal{T}_N$,

$$\mathrm{rank}\left(\mathrm{span}\left(\bigcup_{x \in P_z} C_x\right)\right) \;=\; \mathrm{rank}\left(\bigcup_{x \in P_z} C_x\right) \;=\; \sum_{x \in P_z} |C_x| - 1\,. \tag{2}$$

By Theorem 3, the size of C_x drops down by at least 1 every time one moves down the tree. Therefore, for any $z \in \mathcal{T}_N$, it holds that $\sum_{x \in P_z} (|C_x| - 1) \;\le\; (r-1) + (r-2) + \ldots + 2 \;<\; (r-1)r/2$. Keeping in mind the definition of $s_M(*)$, this implies that

$$\left|\mathrm{span}\left(\bigcup_{x \in P_z} C_x\right)\right| \;\le\; s_N\Big((r-1)r/2\Big). \tag{3}$$

In view of Claim 4 (iii), the number of vertices of $\mathcal{T}_N$, and in particular the number of leaves there, is at most $\mathrm{rank}(N)$. Combining this with Equations (1) and (3) we conclude that

$$|E(N)| \;\le\; \sum_{z:\, \text{leaves of } \mathcal{T}_N} \left|\mathrm{span}\left(\bigcup_{x \in P_z} C_x\right)\right| \;\le\; \mathrm{rank}(N) \cdot s_N\Big((r-1)r/2\Big),$$

as desired. □

For an application of Theorem 5, consider the case when M is a $\mathbb{F}_q$-representable simple matroid, i.e., it is isomorphic to a linear matroid over $\mathbb{F}_q$. In this case $s_M(r) \le q^r$. By Theorem 5,

Corollary 6. *For a $\mathbb{F}_q$-representable simple matroid M, $\gamma(M) \;\le\; q^{(c(M)-1)\cdot c(M)/2}$. Consequently, $c(M) \;>\; \sqrt{2 \log_q \gamma(M)}$.*

This is a considerable improvement over Corollary 2. The lower bound on $c(M)$ is tight up to a square root and negligible additive terms. To see this, consider a k-dimensional linear space over $\mathbb{F}_q$ with 0 removed. For the corresponding matroid M, $\gamma(M) = (q^k - 1)/k$, while $c(M) = k + 1$.

4 Back to Simplicial Complexes

The general results obtained in the previous section apply to simplicial complexes. Let us first relate the matroidal notation used in the previous section to that of simplicial complexes.

Let K be a d-dimensional simplicial complex, and let $\mathbb{F}$ be a field. The boundary mapping ∂_d : $C_d(K;\mathbb{F}) \to C_{d-1}(K;\mathbb{F})$ maps d-chains of K to $(d-1)$-chains. Its kernel is $Z_d(K,\mathbb{F})$, the d-cycles, and its image is $B_{d-1}(K,\mathbb{F})$, the $(d-1)$-boundaries. Both are linear spaces over $\mathbb{F}$.

The rank function rank_d that introduces a matroidal structure on $K^{(d)}$ is defined as follows. For $A \subseteq K^{(d)}$, its rank is:

$$\mathrm{rank}_d(A) \;=\; \mathrm{rank}_{\mathbb{F}}\{\partial_d\sigma \mid \sigma \in A\} \;=\; \dim_{\mathbb{F}}\mathrm{span}\{\partial_d\sigma \mid \sigma \in A\},$$

where $\mathrm{rank}_{\mathbb{F}}$ is the usual linear-algebraic rank function of sets of vectors, and span above is the usual linear-algebraic span. Note that this is consistent of writing for $A \subseteq K^{(d)}$, that its matroidal span is $\mathrm{span}_M(A) = \{\sigma \in K^{(d)} \mid \sigma \in \mathrm{span}_{\mathbb{F}}\{\partial_d\sigma \mid \sigma \in A\}\ \}$. We will avoid using the subscript M in span_M in what follows, as we do not use the linear algebraic span anymore. This results in an $\mathbb{F}$-representable matroid $M(K^{(d)})$. In particular, $\mathcal{I}$, the independent sets in $M(K^{(d)})$, are the the sets of d-simplices in K that support no nontrivial d-cycles, and equivalently $A \subseteq K^{(d)}$ is independent if $\mathrm{rank}_d(A) = |A|$.

We shall use $c_d(K)$ and $\gamma_d(K)$ to denote the size of the largest circuit, and the value of the parameter γ in the above matroid $M(K^{(d)})$, respectively. Note that $c_d(K)$ coincides with the size of the largest simple d-cycle in K, as defined in Section 2. Slightly abusing the notation, we shall use $\mathrm{rank}_d(K)$ to denote $\mathrm{rank}_d(M(K^{(d)}))$. It is the dimension of $B_{d-1}(K,\mathbb{F})$. Recall that the f-vector of K is $(f_d, f_{d-1}, \ldots, f_0)$ where $f_i = |K^{(i)}|$.

In order to employ Theorem 5, we shall estimate $s_d(t)$, the maximum possible size of a family $A \subseteq K^{(d)}$ of d-simplices, with $\mathrm{rank}_d(A) \le t$. Let us first cite the Kruskal-Katona theorem in the relaxed form of Lóvász:

Theorem 7. (Lóvász [12]) *Among all simplicial complexes K with a fixed $f_d(K)$ on a finite ground set V equipped with a total order, the minimum value of f_{d-1} is achieved on K_0 whose d-faces* compressed *with respect to the colexicographic order the subsets of V of size $(d+1)$. Furthermore:*

$$f_d(K) \;=\; \binom{x}{d+1} \qquad\Longrightarrow\qquad f_{d-1}(K) \;\ge\; f_{d-1}(K_0) \;\ge\; \binom{x}{d}.$$

Since removing the faces of K that are not subfaces of $K^{(d)}$ has no effect neither on f_d, nor on $\mathrm{rank}_d(K)$, and since after such removal it trivially holds that $\mathrm{rank}_d(K) \ge \frac{1}{d+1} \cdot f_{d-1}(K)$, Theorem 7 yields an upper bound on f_d in terms of $\mathrm{rank}_d(K)$ that is already sufficient for our needs. A stronger result is provided by:

Theorem 8. (implicit in [2] [1]) *Among all simplicial complexes K with a fixed $f_d(K)$ on a finite ground set V equipped with a total order, the minimum value of* $\mathrm{rank}_d(K)$ *over any field $\mathbb{F}$ is achieved on the same K_0 as in Theorem 7. Moreover,* $\mathrm{rank}_d(K_0)$ *is the number of d-simplices in K_0 that contain the smallest vertex v_1 of V.*

Further investigation of the combinatorial structure of the latter set based on Theorem 7, yields

$$f_d(K) \;=\; \binom{x}{d+1} \qquad\Longrightarrow\qquad \mathrm{rank}_d(K) \;\ge\; \binom{x-1}{d}.$$

[1] This was also independently discovered by N. Linial and Y. Peled, private communication, by a direct argument involving the combinatorial shifting technique.

Corollary 9. $s_d(t) \;\leq\; t^{1+\frac{1}{d}}$.

Proof. The conclusion of Theorem 8 can be restated, given the monotonicity of the functions involved, in the form

$$\mathrm{rank}_d(K) \;=\; \binom{y}{d} \qquad\Longrightarrow\qquad f_d(K) \;\leq\; \binom{y+1}{d+1}.$$

Let $t = \binom{y}{d}$. Then, $f_d(K) \;\leq\; \binom{y+1}{d+1} \;=\; t \cdot \frac{y+1}{d+1}$. By a standard estimation of binomial coefficients, it holds that $\left(\frac{y}{d}\right)^d \leq t \leq \left(\frac{ey}{d}\right)^d$. Thus, $f_d(K) \;\leq\; t \cdot t^{1/d} \cdot \frac{d}{d+1} \cdot \frac{y+1}{y} \;\leq\; t^{1+1/d}$. $\square$

Combining the above corollary with Theorem 5 yields:

Theorem 10. *For any simplicial complex K containing nontrivial d-cycles,* $c_d(K) \;\geq\; (2\gamma_d(K))^{\frac{1}{2}-\frac{1}{2(d+1)}}$.

The rest of this section is dedicated to strengthening this theorem. The goal is to establish the following:

Theorem 11. *Let K be a simplicial complex containing nontrivial d-cycles. Then,*

$$f_d(K) \;>\; \frac{d+1}{2} \cdot k(k+1) \cdot \mathrm{rank}_d(K) \quad\Longrightarrow\quad K \textit{ contains a simple } d\textit{-cycle of size} > k.$$

Consequently, $c_d(K) \geq \sqrt{2/(d+1) \cdot \gamma_d(K)} - 1$.

Proof. We shall use the tree $\mathcal{T}_K = \mathcal{T}_{M(K^{(d)})}$ as in Definition 1, with the notation introduced in the proof of Theorem 5. We warn the reader, however, that contraction of a cycle from a simplicial matroid does not necessarily results in a simplicial matroid.

For $A \subseteq K^{(d)}$, let $A^{(i)} \subseteq K^{(i)}$ denote the union of all the i-faces of the σ's in A. The *starred closure* $\mathrm{cl}^*(A)$ of A is defined as follows:

$$\mathrm{cl}^*(A) \;=\; \{\, \sigma \in K^{(d)} \mid \sigma^{(d-1)} \subseteq A^{(d-1)} \,\}\,.$$

Observe that $\mathrm{span}(A) \subseteq \mathrm{cl}^*(A)$. The containment can be strict. For example, consider the set $A = \{(1,x,y),(2,y,z),(3,z,x)\}$. Then, (x,y,z) belongs to $\mathrm{cl}^*(A)$, but not to $\mathrm{span}(A)$.

Let P_x be, as before, the path from the root of $\mathcal{T}_K$ to the vertex x in it. Define

$$S_x \;=\; \bigcup_{v \in P_x \setminus \{x\}} C_v \;\subseteq\; K^{(d)}\,.$$

In particular, for the root vertex r, $S_r = \emptyset$.

Claim 12. *The following holds:*

$$\bigcup_{x \in \mathcal{T}_M} \mathrm{span}(S_x \cup C_x) \setminus \mathrm{span}(S_x) \;\;=\;\; K^{(d)}\,. \tag{4}$$

$$\sum_{x \in \mathcal{T}_M} |\mathrm{cl}^*(S_x \cup C_x)| - |\mathrm{cl}^*(S_x)| \;\;\geq\;\; |K^{(d)}| \;\;=\;\; f_d(K)\,. \tag{5}$$

Proof. We start with (4). As we have seen earlier in the proof of Theorem 5, for every $\sigma \in K^{(d)}$, the set of all vertices v such that $\sigma \in E(M_v)$ forms a path P_x from the root to some vertex $x \in \mathcal{T}_K$, where x, the lowest vertex of P_x, is the place where σ is terminated. Since every σ gets eventually terminated at a uniquely defined vertex, and $\mathrm{span}(S_x \cup C_x) \setminus \mathrm{span}(S_x)$ contains precisely the set of all σ's terminated at x, the statement follows. Moreover, the outer union is disjoint.

We proceed with establishing (5). The left-hand-side of it can be rewritten in the following form:

$$\sum_{x \in \mathcal{T}_M} |\mathrm{cl}^*(S_x \cup C_x)| - |\mathrm{cl}^*(S_x)| \;=\; \sum_{\sigma \in K^{(d)}} \sum_{x \in \mathcal{T}_M} \mathbb{1}_\sigma(\mathrm{cl}^*(S_x \cup C_x)) - \mathbb{1}_\sigma(\mathrm{cl}^*(S_x)) \;=$$

$$= \sum_{\sigma \in K^{(d)}} \sum_{x \in \mathcal{T}_M} \mathbb{1}_\sigma\Big(\mathrm{cl}^*(S_x \cup C_x) \setminus \mathrm{cl}^*(S_x)\Big),$$

where $\mathbb{1}_\sigma$ is the indicator function of σ. Consider a fixed $\sigma \in K^{(d)}$. From (4) we know that there exist vertices $v \in \mathcal{T}_K$ such that $\sigma \in \mathrm{span}(S_v \cup C_v) \subseteq \mathrm{cl}^*(S_v \cup C_v)$. Let x be a topmost vertex (there could be many) for which $\sigma \in \mathrm{cl}^*(S_x \cup C_x)$. By definition of x, $\sigma \notin \mathrm{cl}^*(S_x)$, and thus x contributes 1 to the inner sum of the above expression corresponding to the fixed σ. Hence, (5) follows. $\square$

Next, we bound the inner term $|\mathrm{cl}^*(S_x \cup C_x)| - |\mathrm{cl}^*(C_x)|$ from (5) in the following manner. Notice that if $\sigma \in \mathrm{cl}^*(S_x \cup C_x) \setminus \mathrm{cl}^*(S_x)$, then there must exists a $(d-1)$-face τ of σ, such that $\tau \notin S_x^{(d-1)}$. Consequently, $\tau \in C_x^{(d-1)}$. Since all the vertices of σ belong to $(S_x \cup C_x)^{(0)}$, it follows that

$$|\mathrm{cl}^*(S_x \cup C_x)| - |\mathrm{cl}^*(C_x)| \;\leq\; |C_x^{(d-1)}| \cdot |(S_x \cup C_x)^{(0)}| \,. \tag{6}$$

Since C_x is a d-cycle, every $(d-1)$-face in it is adjacent to two or more d-faces of C_x, while every d-face is adjacent to $(d+1)$ d-faces. Thus, $|C_x^{(d-1)}| \leq \frac{d+1}{2} \cdot |C_x|$.

Consider now $|(S_x \cup C_x)^{(0)}|$. We claim that $|(S_x \cup C_x)^{(0)}| \leq |S_x \cup C_x|$. To see this, consider the bipartite graph $G = (L, R; E)$ where $L = (S_x \cup C_x)^{(0)}$, $R = (S_x \cup C_x)$ and $(v, \sigma) \in E$ if v is a vertex of σ, i.e., $v \in \sigma^{(0)}$. The degree in G of every $\sigma \in R$ is $(d+1)$ by definition of σ. We will show that the degree of every $v \in L$ is at least $(d+1)$. This immediately implies that $|L| \leq |R|$, as needed.

Clearly, a vertex of a d-cycle in $K^{(d)}$ must belong to at least $(d+1)$ d-simplices in it. Thus, it suffices to show that every d-simplex of $(S_x \cup C_x)$ is contained in some d-cycle $C \subseteq S_x \cup C_x$ in $K^{(d)}$. Arguing by induction (on the depth of x in $\mathcal{T}$), it suffices to show this for simplices in C_x.

Indeed, by definition, C_x is a circuit in $K^{(d)}/S_x$. Let I_x be a maximal independent set in S_x. Since $\mathrm{span}(I_x) = \mathrm{span}(S_x)$, it follows that C_x is a circuit in $K^{(d)}/I_x$ as well. Thus, $\mathrm{rank}(I_x \cup C_x) = \mathrm{rank}(I_x) + \mathrm{rank}((I_x \cup C_x)/I_x) = |I_x| + (|C_x| - 1) = |(I_x \cup C_x)| - 1$, implying that $(I_x \cup C_x)$ is not acyclic, and so it contains a d-cycle in $K^{(d)}$. The same argument implies that for any $\sigma \in C_x$, $((I_x \cup C_x) \setminus \{\sigma\})$ is acyclic.

Let $I_x^* \subseteq I_x$ be the minimal subset of I_x such that $(I_x^* \cup C_x)$ is not acyclic. Observe that removing any σ from $(I_x^* \cup C_x)$ makes it acyclic. Therefore $(I_x^* \cup C_x)$ is a simple d-cycle in $K^{(d)}$. This shows that every d-simplex in C_x is indeed contained in some d-cycle in $(C_x \cup S_x)$, concluding the argument establishing $|(S_x \cup C_x)^{(0)}| \leq |S_x \cup C_x|$.

Finally, since $S_x \cup C_x = \bigcup_{v \in P_x} C_v$, Claim 4 implies that $|S_x \cup C_x| \;\leq\; k + (k-1) + \ldots + 1 \;=\; k(k+1)/2$, where $k = c_d(K)$.

Combining (6) with the subsequent observations implies that:

$$|\mathrm{cl}^*(S_x \cup C_x)| - |\mathrm{cl}^*(S_x)| \;\leq\; \frac{d+1}{2} \cdot |C_x| \cdot |(S_x \cup C_x)| \;\leq\; \frac{d+1}{4} \cdot (k+1)k \cdot |C_x| \,. \tag{7}$$

Combining (5), (7), and using Claim 4, one arrives at

$$
\begin{aligned}
f_d(K) \;&\leq\; \sum_{x\in\mathcal{T}_M} |\mathrm{cl}^*(S_x \cup C_x)| - |\mathrm{cl}^*(S_x)| \;\leq\; \sum_{x\in\mathcal{T}_M} \frac{d+1}{4}\cdot k(k+1)\cdot |C_x|\,. \\
&\leq\; \frac{d+1}{2}\cdot k(k+1)\cdot \sum_{x\in\mathcal{T}_M} (|C_x|-1) \;\leq\; \frac{d+1}{2}\cdot k(k+1)\cdot \mathrm{rank}_d(K)\,.
\end{aligned}
$$

The contrapositive of this inequality is the desired statement. This completes the proof of the theorem. □

Now let K be a simplicial complex with $c_d(K) > 0$. Since $f_{d-1} \geq \mathrm{rank}_d(K)$, Theorem 11 implies the following lower bound on $c_d(K)$ in terms of its density:

Theorem 13. *For K as above,*

$$
c_d(K) \;\geq\; \sqrt{\frac{2}{d+1}\cdot\frac{f_d(K)}{f_{d-1}(K)}} \;-\; 1\,.
$$

Open Problems The most intriguing open problem raised by this paper is the tightness of the above lower bounds. While cliques are extremal graphs for the Erdős-Gallai problem, in higher dimensions the complete d-dimensional simplicial complex K_n^d on n vertices has $c_d(K_n^d) = (1-o(1))\binom{n-1}{d}$ and $\gamma_d(K_n^d) = \frac{n}{d+1}$, and so $c_d(K_n^d) = \Theta(\gamma_d(K_n^d))^d$.

The situation is unclear for small $\gamma_d(K)$'s as well. Our lower bounds are trivial when $1 < \gamma_d(K) \leq d^3$, since if $\gamma_d(K) > 1$, i.e., K is not acyclic, then $c_d(K) \geq d+1$. What happens in this range?

Acknowledgements We are very grateful to Jan van den Heuvel from the London School of Economics for an insightful discussion which led to the formulation of the main question addressed in this paper.

References

[1] L. Aronshtam, N. Linial, T. Luczak and R. Meshulam, Collapsibility and vanishing of top homology in random simplicial complexes, *Discrete and Computational Geometry*, **49**(2013), 317-334.

[2] A. Björner, G. Kalai, An extended Euler-Poincare theorem, *Acta Mathematica*, **161(1)**(1988), 279-303.

[3] B. Bollobás, Modern Graph Theory, Springer Verlag, New York, 1998.

[4] R. Diestel, Graph Theory, Springer, 2017.

[5] G. Ding, B. Oporowski and J. Oxley, On infinite antichains of matroids, *J. Combin. Theory Ser. B*, **63**(1995) 21-40.

[6] D. Dotterrer, L. Guth, M. Kahle, 2-Complexes with Large 2-Girth, *Discrete & Computational Geometry*, **59(2)**(2018) 383-412.

[7] P. Erdős and T. Gallai, On maximal paths and circuits of graphs, *Acta Math. Acad. Sci. Hungar*, **10**(1959) 337-356.

[8] J. Geelen and G. Whittle, Cliques in dense $GF(q)$-representable matroids, *J. Combin. Theory Ser. B.*, **87**(2003) 264-269.

[9] J. Geelen, Small cocircuits in matroids, *European J. Combin.*, **32**(2011) 795-801.

[10] N. Linial and Y. Peled, On the phase transition in random simplicial complexes, *Annals of Mathematics*, **184**(2016), 745-773.

[11] T. Kaufman, D. Kazhdan, A. Lubotzky, Isoperimetric Inequalities for Ramanujan Complexes and Topological Expanders, *Geometric and Functional Analysis*, **26**(2016), 250-287.

[12] L. Lovász, Combinatorial problems and exercises. Corrected reprint of the 1993 second edition. AMS Chelsea Publishing, Providence, RI, 2007.

[13] J. Oxley, Matroid theory. Second edition. Oxford Graduate Texts in Mathematics, 21, Oxford University Press, Oxford, 2011.

[14] A. Thomason, The extremal function for contractions of graphs, *Math. Proc. Cambridge Philos. Soc.*, **95** (1984), pp. 261-265.