A linear bound on the Manickam-Miklós-Singhi Conjecture

Alexey Pokrovskiy

Methods for Discrete Structures, Department of Mathematics, Freie Universität, Berlin, Germany

Abstract

Suppose that we have a set S of n real numbers which have nonnegative sum. How few subsets of S of order k can have nonnegative sum? Manickam, Miklós, and Singhi conjectured that for $n \geq 4k$ the answer is $\binom{n-1}{k-1}$. This conjecture is known to hold when n is large compared to k. The best known bounds are due to Alon, Huang, and Sudakov who proved the conjecture when $n \geq 33k^2$. In this paper we improve this bound by showing that there is a constant C such that the conjecture holds when $n \geq Ck$. This establishes the conjecture in a range which is a constant factor away from the conjectured bound.

Keywords: Extremal combinatorics; hypergraphs; additive combinatorics; Katona's cycle method.

1. Introduction

Suppose that we have a set of real numbers $\{x_1, \dots, x_n\}$ satisfying $x_1 + \dots + x_n \geq 0$. How few subsets $A \subset \{x_1, \dots, x_n\}$ satisfying $\sum_{a \in A} a \geq 0$ can $\{x_1, \dots, x_n\}$ have?

By choosing $x_1 = n-1$ and $x_2 = \dots = x_n = -1$ we see that the answer to this question can be at most 2^{n-1}. In fact, this example has the minimal number of nonnegative sets. Indeed, for any set $A \subset \{x_1, \dots, x_n\}$ either A or $\{x_1, \dots, x_n\} \setminus A$ must have nonnegative sum, so there are always at least 2^{n-1} nonnegative subsets in any set of numbers $\{x_1, \dots, x_n\}$ with nonnegative sum.

A more difficult problem arises if we count only subsets of fixed order. By again considering the example when $x_1 = n-1$ and $x_2 = \dots = x_n = -1$ we see that there are sets of n numbers with nonnegative sums which have only $\binom{n-1}{k-1}$ nonnegative k-sums (sums of k distinct numbers). Manickam, Miklós, and Singhi conjectured that for $n \geq 4k$ this assignment gives the least possible number of nonnegative k-sums.

Conjecture 1.1 (Manickam, Miklós, Singhi, [17, 18])**.** *Suppose that $n \geq 4k$, and we have n real numbers $x_1, \dots, x_n$ such that $x_1 + \dots + x_n \geq 0$. Then, at least $\binom{n-1}{k-1}$ subsets $A \subset \{x_1, \dots, x_n\}$ of order k satisfy $\sum_{a \in A} a \geq 0$*

Conjecture 1.1 appeared in [18] where it was phrased in terms of calculating invariants of an association scheme known as the *Johnson Scheme*. In [17], Conjecture 1.1 was phrased

Email address: alja123@gmail.com (Alexey Pokrovskiy)

in the combinatorial form in which it is stated above. In this paper we will speak only about the combinatorial version—we refer the reader to [18, 3] for more details about the association scheme version.

A motivation for the bound "$n \geq 4k$" is that for $k \geq 3$ and $n = 3k+1$ there exists an assignment of values to $x_1, \ldots, x_{3k+1}$ which results in less than $\binom{n-1}{k-1}$ nonnegative k-sums. Indeed, letting $x_1 = x_2 = x_3 = 2-3k$ and $x_4 = \cdots = x_{3k+1} = 3$ gives an assigment satisfying $x_1 + \cdots + x_{3k+1} = 0$ but having $\binom{3k-2}{k}$ nonnegative k-sums, which is less than $\binom{3k}{k-1}$ for $k \geq 3$ for this particular pair of n and k. Notice that these examples exist only when $n = 3k+1$. Thus it is possible that the bound "$n \geq 4k$" could be slightly strengthened in Conjecture 1.1. For example for $k = 3$, Chowdhury proved that Conjecture 1.1 holds with the improved bound of $n \geq 11$, and that this bound is best possible [6].

Despite the apparent simplicity of the statement of Conjecture 1.1, it has been open for over two decades. Many partial results have been proven. The conjecture has been proven for $k \leq 3$ by Manickam [16] and independently by Chiaselotti and Marino [9]. It has been proven whenever $n \equiv 0 \pmod{k}$ by Manickam and Singhi [18].

In addition, several results have been proved establishing the conjecture when n is large compared to k. Manickam and Miklós [17] showed that the conjecture holds when $n \geq (k-1)(k^k + k^2) + k$ holds. Tyomkyn [21] improved this bound to $n \geq k(4e \log k)^k \approx e^{ck \log\log k}$. Recently Alon, Huang, and Sudakov [2] showed that the conjecture holds when $n \geq 33k^2$. Subsequently Frankl [8] gave an alternative proof of the conjecture in a range of the form $n \geq 3k^3/2$, and Chowdhury, Sarkis, and Shahriari [7] gave a proof of the conjecture in the range of the form $n \geq 8k^2$.

The aim of this paper is to improve these bounds by showing that the conjecture holds in a range when n is linear with respect to k.

Theorem 1.2. *Suppose that $n \geq 10^{46}k$, and we have n real numbers $x_1, \ldots, x_n$ such that $x_1 + \cdots + x_n \geq 0$. At least $\binom{n-1}{k-1}$ subsets $A \subset \{x_1, \ldots, x_n\}$ of order k satisfy $\sum_{a \in A} a \geq 0$*

It should be noted, that after we announced this paper, a full proof of Conjecture 1.1 was announced by Blinovsky in [4].

The method we use to prove Theorem 1.2 in this paper is inspired by an averaging argument which Katona used in his proof of the Erdős-Ko-Rado Theorem in [12]. This method is often referred to as "Katona's cycle method" [13]. Before we can say how the method is useful for approaching Conjecture 1.1, we will need a few definitions.

Suppose that we have a hypergraph $\mathcal{H}$ together with an assignment of real numbers to the vertices of $\mathcal{H}$ given by $f : V(\mathcal{H}) \to \mathbb{R}$. We can extend f to the power set of $V(\mathcal{H})$ by letting $f(A) = \sum_{v \in A} f(v)$ for every $A \subseteq V(\mathcal{H})$. We say that an edge $e \in E(\mathcal{H})$ is *negative* if $f(e) < 0$, and e is *nonnegative* otherwise. We let $e_f^+(\mathcal{H})$ be the number of nonnegative edges of $\mathcal{H}$. Recall that the degree $d(v)$ of a vertex v in a hypergraph $\mathcal{H}$ is the number of edges containing v. A hypergraph $\mathcal{H}$ is d-regular if every vertex has degree d. The minimum degree of a hypergraph $\mathcal{H}$ is $\delta(\mathcal{H}) = \min_{v \in V(\mathcal{H})} d(v)$. The k-uniform complete hypergraph on n vertices is denoted by $\mathcal{K}_n^{(k)}$.

The following observation is key to our proof of Theorem 1.2.

Lemma 1.3. *Let $\mathcal{H}$ be a d-regular k-uniform hypergraph on n vertices. Suppose that for every $f : V(\mathcal{H}) \to \mathbb{R}$ satisfying $\sum_{x\in V(\mathcal{H})} f(x) \geq 0$ we have $e_f^+(\mathcal{H}) \geq d$. Then for every $f : V(\mathcal{K}_n^{(k)}) \to \mathbb{R}$ satisfying $\sum_{x\in V(\mathcal{K}_n^{(k)})} f(x) \geq 0$ we have $e_f^+(\mathcal{K}_n^{(k)}) \geq \binom{n-1}{k-1}$ (and so Conjecture 1.1 holds for this particular pair of n and k).*

Lemma 1.3 is proved by an averaging technique very similar to Katona's proof of the Erdős-Ko-Rado Theorem (see Section 2).

Lemma 1.3 shows that instead of proving the conjecture about the complete graph $\mathcal{K}_n^{(k)}$, it may be possible to find regular hypergraphs $\mathcal{H}$ which satisfy the condition in Lemma 1.3 and hence deduce the conjecture. In most classic applications of Katona's cycle method, the hypergraph $\mathcal{H}$ is taken to be a tight cycle (the set of all intervals of length k in $\mathbb{Z}_n$ for some n). The novelty of our application of the Lemma 1.3 in this paper is that our hypergraphs $\mathcal{H}$ will *not* be a tight cycles. This turns out to be crucial for producing the improved bound on the conjecture that we have in Theorem 1.2.

Lemma 1.3 motivates us to make the following definition.

Definition 1.4. *A k-uniform hypergraph $\mathcal{H}$ has the **MMS-property** if for every $f : V(\mathcal{H}) \to \mathbb{R}$ satisfying $\sum_{x\in V(\mathcal{H})} f(x) \geq 0$ we have $e^+(\mathcal{H}) \geq \delta(\mathcal{H})$.*

Conjecture 1.1 is equivalent to the statement that for $n \geq 4k$ the complete hypergraph on n vertices has the MMS-property. Lemma 1.3 shows that in order to prove Conjecture 1.1 for particular n and k, it is sufficient to find one regular n-vertex k-uniform hypergraph $\mathcal{H}$ with the MMS-property. This hypergraph $\mathcal{H}$ may be much sparser than the complete hypergraph—allowing for very different proof techniques.

Perhaps the first two candidates one chooses for hypergraphs that may have the MMS-property are matchings and tight cycles. The matching $\mathcal{M}_{t,k}$ is defined as the k-uniform hypergraph consisting of tk vertices and t vertex disjoint edges. Notice that $\mathcal{M}_{t,k}$ is 1-regular. The matching $\mathcal{M}_{t,k}$ always has the MMS-property—indeed we have that $\sum_{e\in E(\mathcal{M}_{t,k})} f(e) = \sum_{x\in\mathcal{M}_{t,k}} f(x) \geq 0$, and so one of the edges of $\mathcal{M}_{t,k}$ is nonnegative. This observation was used in [18] to prove Conjecture 1.1 whenever k divides n.

The tight cycle $\mathcal{C}_{n,k}$ is defined as the hypergraph with vertex set $\mathbb{Z}_n$ and edges formed by the intervals $\{i \pmod n, i+1 \pmod n, \ldots, i+k \pmod n\}$ for $i \in \mathbb{Z}_n$. It turns out that the tight cycles do not have the MMS-property when $n \not\equiv 0 \pmod k$. To see this for example when $k = 3$ and $n \equiv 1 \pmod k$, let $f(x) = 50, 50, 50, -101, 50, 50, -101, 50, 50, -101\ldots$ for $x = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \ldots$—i.e. $f(x) = -101$ if $x \equiv 1 \pmod 3$ and $x \neq 1$, and $x = 50$ otherwise. This hypergraph is 3-regular but the only nonnegative edge is $\{1, 2, 3\}$, so $C_{n,3}$ doesn't have the MMS-property when $n \equiv 1 \pmod 3$. Similar constructions work for all n and k with $n \not\equiv 0 \pmod k$. Despite $\mathcal{C}_{n,k}$ not having the MMS-property, tight cycles were still useful in [17] to prove Conjecture 1.1 when $n \geq (k-1)(k^k + k^2) + k$.

An interesting question, is "which hypergraphs have the MMS-property?" This question has been studied by Huang and Sudakov, who found some large classes of hypergraphs with the MMS-property [11].

The main result of this paper is showing that there exist $k(k-1)^2$-regular k-uniform hypergraphs on n vertices which have the MMS-property, for all $n \geq 10^{46}k$.

Theorem 1.5. *For $n \geq 10^{46}k$, there are $k(k-1)^2$-regular k-uniform hypergraphs on n vertices, $\mathcal{H}_{n,k}$, with the property that for every $f : V(\mathcal{H}_{n,k}) \to \mathbb{R}$ satisfying $\sum_{x \in V(\mathcal{H}_{n,k})} f(x) \geq 0$ we have $e^+(\mathcal{H}_{n,k}) \geq k(k-1)^2$.*

Combining Theorem 1.5 and Lemma 1.3 immediately implies Theorem 1.2.

Throughout this paper, we will use notation from Additive Combinatorics for sumsets $A + B = \{a + b : a \in A, b \in B\}$ and translates $A + x = \{a + x : a \in A\}$. For all standard notation we refer to [5].

The structure of this paper is as follows. In Section 2 we prove Lemma 1.3. In Section 3, we define the graphs $\mathcal{H}_{n,k}$ used in Theorem 1.5 and prove some of their basic properties. In Section 4, we prove Theorem 1.5 with the weaker bound of $n \geq 14k^4$ in order to illustrate the main ideas in the proof of Theorem 1.5. In Section 5 we prove Theorem 1.5. In Section 6, we conclude by discussing the techniques used in this paper and whether they could be used to prove Conjecture 1.1 in general.

2. Proof of the averaging lemma

Here we prove Lemma 1.3.

Proof. Suppose that we have a function $f : \{1, \ldots, n\} \to \mathbb{R}$ satisfying $\sum_{x \in \{1,\ldots,n\}} f(x) \geq 0$. Consider a random permutation σ of $\{1, \ldots, n\}$, chosen uniformly out of all permutations of $\{1, \ldots, n\}$. We define a function $f_\sigma : \{1, \ldots, n\} \to \mathbb{R}$ given by $f_\sigma : x \to f(\sigma(x))$. Clearly $\sum_{x \in \{1,\ldots,n\}} f_\sigma(x) \geq 0$. We will count $\mathbb{E}(e^+_{f_\sigma}(\mathcal{H}))$ in two different ways. For an edge $e \in \mathcal{K}_n^{(k)}$, we have

$$\mathbb{P}(\sigma(e) \in \mathcal{H}) = \frac{e(\mathcal{H})}{\binom{n}{k}} = \frac{d}{\binom{n-1}{k-1}}$$

Therefore we have

$$\mathbb{E}(e^+_{f_\sigma}(\mathcal{H})) = \sum_{\substack{e \in \mathcal{K}_n^{(k)}, \\ f(e) \geq 0}} \mathbb{P}(\sigma(e) \in \mathcal{H}) = e^+(\mathcal{K}_n^{(k)}) \frac{d}{\binom{n-1}{k-1}}$$

However, by the assumption of the lemma, $\mathbb{E}(e^+_{f_\sigma}(\mathcal{H}))$ is at least d. This gives us

$$e^+(\mathcal{K}_n^{(k)}) \geq \binom{n-1}{k-1}.$$

□

3. Construction of the hypergraphs $\mathcal{H}_{n,k}$

In this section we construct graphs $\mathcal{H}_{n,k}$ which satisfy Theorem 1.5. We also prove some basic properties which the graphs $\mathcal{H}_{n,k}$ have.

Define the clockwise interval between a and $b \in \mathbb{Z}_n$ to be $[a, b] = \{a, a+1, \ldots, b\}$. The graph $\mathcal{H}_{n,k}$ has vertex set $\mathbb{Z}_n$. We define k-edges $e(v, i, j)$ as follows:

$$e(v,i,j) = [v, v+i-1] \cup [v+i+j, v+j+k-1]$$

The edges of $\mathcal{H}_{n,k}$ are given by $e(v,i,j)$ for $v \in \mathbb{Z}_n$ and $i,j \in \{1,\ldots,k-1\}$. In other words $\mathcal{H}_{n,k}$ consists of all the double intervals of order k, where the distance between the two intervals is at most $k-1$.

Notice that the graph $\mathcal{H}_{n,k}$ is indeed $k(k-1)^2$ regular.

In order to deal with the graphs $\mathcal{H}_{n,k}$ it will be convenient to assign a particular set $E(v)$ of $O(k^2)$ edges to each vertex v. First, for each vertex v in $\mathcal{H}_{n,k}$ and $i,j \in [1,k-1]$, we will define a set of edges, $E(v,i,j)$. Then $E(v)$ will be a union of the sets $E(v,i,j)$.

The definition of the sets $E(v,i,j)$ is quite tedious. However the sets $E(v,i,j)$ are constructed to satisfy only a few properties. One property that we will need is that for fixed, v, i, j certain intervals can be formed as disjoint unions of edges in $E(v,i,j)$. See Figures 1 – 4 for illustrations of the precise configurations that we will use. Another property that we will need is that any edge $e \in \mathcal{H}_{n,k}$ is contained in at most C_1 of the sets $E(v,i,j)$ for some absolute constant C_1. See Lemmas 3.1 and 3.2 for precise statements of these two properties.

Over the next four pages we define the sets $E(v,i,j)$.

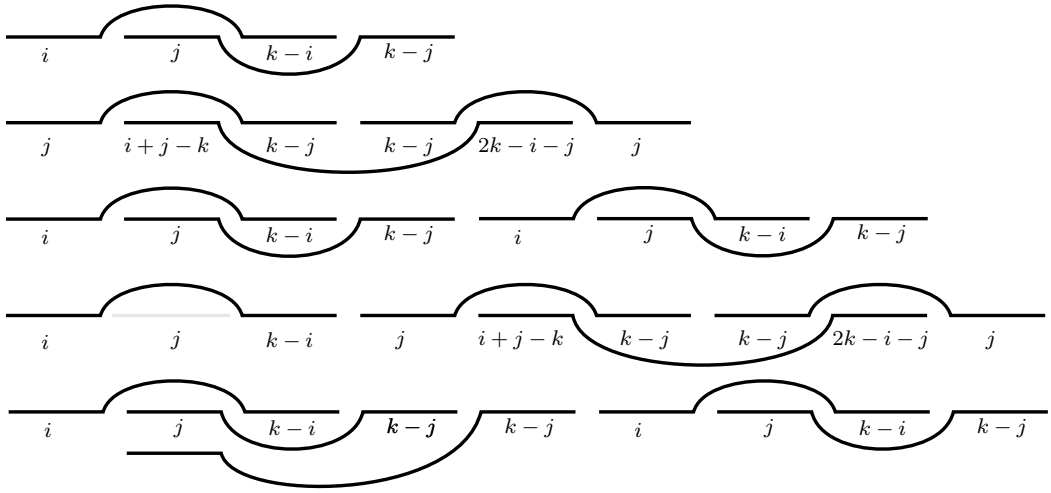

Figure 1: The edges in $E(v, i, j)$ when we have $i + j \geq k$ and $i \geq j$.

If $i + j \geq k$ and $i \geq j$, then we let

$$\begin{aligned} E(v,i,j) = \{&e(v,i,j), e(v+k+j, i, i+j-k),\\ &e(v+k+i+j, i+j-k, 2k-2i), e(v+i, j, k-i),\\ &e(v+k+i+2j, k-i, 2k-i-j), e(v+i, j, 2k-i-j),\\ &e(v+3k-j, i, j), e(v+3k-j+i, j, k-i),\\ &e(v+i, i+j-k, 2k-2i), e(v+i+j, k-i, 2k-i-j),\\ &e(v+2k, i, j), e(v+2k+i, j, k-i)\}. \end{aligned}$$

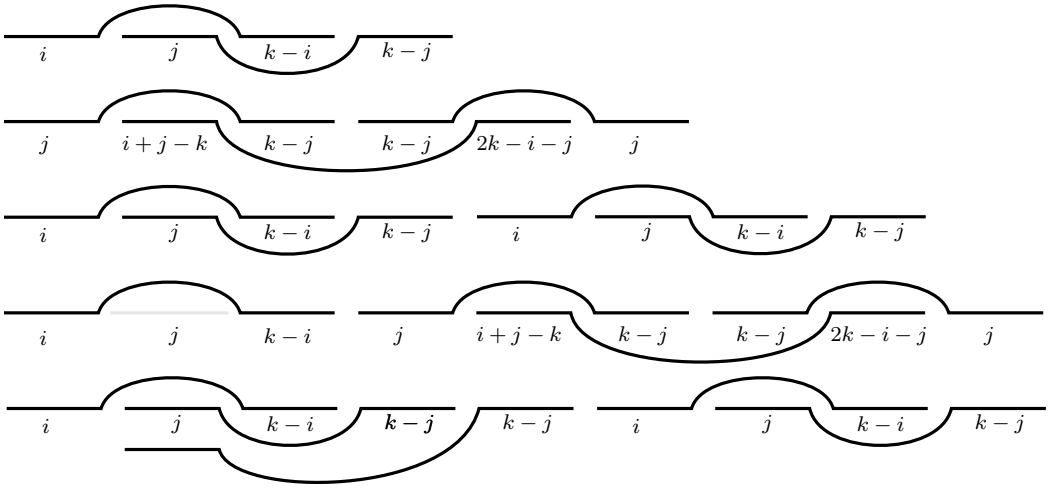

Figure 2: The edges in $E(v, i, j)$ when we have $i + j \geq k$ and $j > i$.

If $i + j \geq k$ and $j > i$, then we let

$$\begin{aligned} E(v,i,j) = \{&e(v,i,j), e(v+k+j, j, i+j-k),\\ &e(v+k+2j, i+j-k, 2k-2j), e(v+i, j, k-i),\\ &e(v+k+i+2j, k-j, 2k-i-j),\\ &e(v+i, j, 2k-i-j), e(v+3k-j, i, j),\\ &e(v+3k-j+i, j, k-i), e(v, j, i+j-k),\\ &e(v+j, i+j-k, 2k-2j), e(v+i+j, k-j, 2k-i-j),\\ &e(v+2k, i, j), e(v+2k+i, j, k-i)\}. \end{aligned}$$

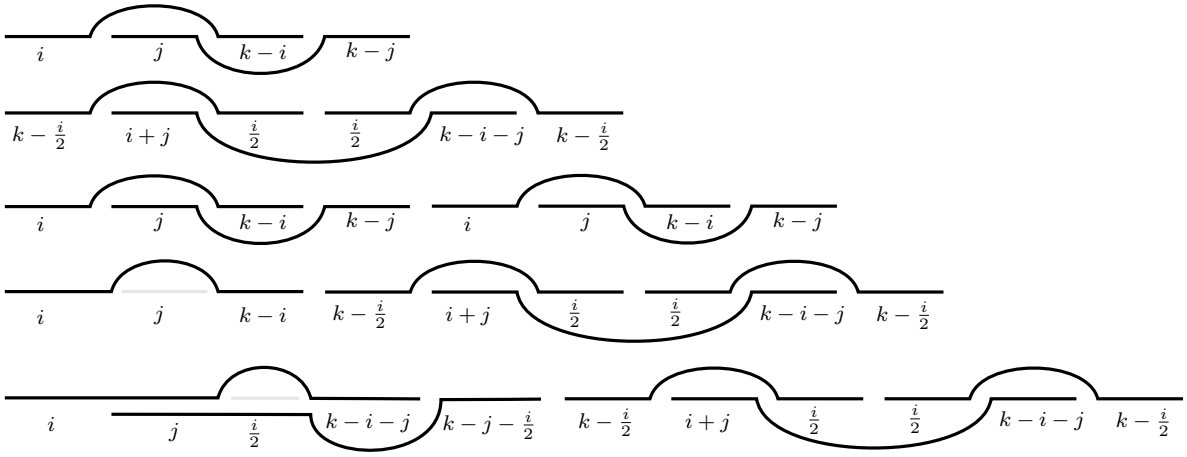

Figure 3: The edges in $E(v,i,j)$ when we have $i+j<k$ and i is even.

If $i+j<k$ and i is even, then we let

$$\begin{aligned}E(v,i,j) = \{&e(v,i,j), e(v+k+j, k-i/2, i+j),\\ &e(v+2k+j-i/2, i+j, i), e(v, i+j, i/2),\\ &e(v+2k+i+2j, i/2, k-i/2), e(v+i, j+i/2, k-i-j),\\ &e(v+2k-j, k-i/2, i+j), e(v+3k-j-i/2, i+j, i),\\ &e(v+3k+i, i/2, k-i/2), e(v, k-i/2, i+j),\\ &e(v+k-i/2, i+j, i), e(v+k+i+j, i/2, k-i-j),\\ &e(v+i, j, k-i), e(v+2k, i, j), e(v+2k+i, j, k-i)\}.\end{aligned}$$

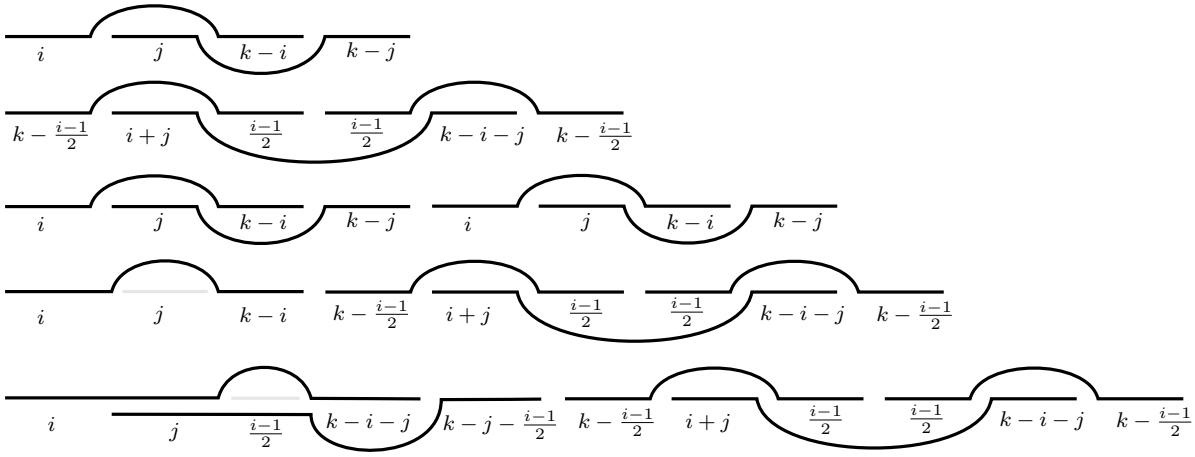

Figure 4: The edges in $E(v,i,j)$ when we have $i+j<k$ and i is odd.

If $i+j<k$ and i is odd, then we let

$$\begin{aligned}E(v,i,j)=\Big\{&e\Big(v,i,j\Big),e\Big(v+k+j,k-\frac{i-1}{2},i+j\Big),\\&e\Big(v+2k+j-\frac{i-1}{2},i+j,i\Big),e\Big(v,i+j,\frac{i-1}{2}\Big),\\&e\Big(v+2k+i+2j,\frac{i-1}{2},k-\frac{i-1}{2}\Big),e\Big(v+i,j+\frac{i-1}{2},k-i-j\Big),\\&e\Big(v+2k-j,k-\frac{i-1}{2},i+j\Big),e\Big(v+3k-j-\frac{i-1}{2},i+j,i\Big),\\&e\Big(v+3k+i,\frac{i-1}{2},k-\frac{i-1}{2}\Big),e\Big(v,k-\frac{i-1}{2},i+j\Big),\\&e\Big(v+k-\frac{i-1}{2},i+j,i\Big),e\Big(v+k+i+j,\frac{i-1}{2},k-i-j\Big),\\&e\Big(v+i,j,k-i\Big),e\Big(v+2k,i,j\Big),e\Big(v+2k+i,j,k-i\Big)\Big\}.\end{aligned}$$

We define $E^-(v,i,j)$ to be the set of edges corresponding to edges in $E(v,i,j)$, but going anticlockwise (i.e. $E^-(v,i,j) = \{\{x_1,\dots,x_k\} : \{v-(x_1-v),\dots,v-(x_k-v)\} \in E(v,i,j)\}$. For each vertex v, we let

$$E(v) = \bigcup_{i,j\in[1,k-1]} E(v,i,j) \cup E^-(v,i,j).$$

Notice that from the definition of $E(v,i,j)$, we certainly have $E(v,i,j) \leq 15$ for every $i,j \in [1,k-1]$, which implies that $|E(v)| \leq 15(k-1)^2$. Also, since $e(v,i,j) \in E(v)$ for every $i,j \in [1,k-1]$, we have that $E(v) \geq (k-1)^2$. Therefore, we have $|E(v)| = \Theta(k^2)$.

There are only two features of the sets $E(v,i,j)$ that will be needed in the proof of Theorem 1.5. One is that $E(v,i,j)$ contains special subsets of edges like the ones pictured in Figures 1 – 4. This allows us to prove the following lemma.

Lemma 3.1. *Suppose that $i,j \in [1,k-1]$ and all the edges in $E(v,i,j)$ are negative. The following hold.*

(i) $f([v,v+2k-1]) < 0$.

(ii) $f([v,v+3k-1]) < 0$.

(iii) $f([v,v+4k-1]) < 0$.

(iv) $f([v+i,v+i+j-1]) < 0 \implies f([v,v+4k+j-1]) < 0$.

(v) $f([v+i,v+i+j-1]) \geq 0 \implies f([v,v+5k-j-1]) < 0$.

Proof. Figures 1 – 4 illustrates the constructions that are used in the proof of this lemma.

(i) This follows from the fact that $e(v,i,j), e(v+i,j,k-i) \in E(v,i,j)$ and $e(v,i,j) \cup e(v+i,j,k-i) = [v,v+2k-1]$.

(ii) For $i+j \geq k$ and $i \geq j$, this follows from the fact that $e(v,i,i+j-k), e(v+i,i+j-k,2k-2i), e(v+i+j,k-i,2k-i-j) \in E(v,i,j)$ and $e(v,i,i+j-k) \cup e(v+i,i+j-k,2k-2i) \cup e(v+i+j,k-i,2k-i-j) = [v,v+3k-1]$. The other cases are similar.

(iii) This follows from the fact that $e(v,i,j), e(v+i,j,k-i), e(v+2k,i,j), e(v+2k+i,j,k-i) \in E(v,i,j)$ and $e(v,i,j) \cup e(v+i,j,k-i) \cup e(v+2k,i,j) \cup e(v+2k+i,j,k-i) = [v,v+4k-1]$.

(iv) For $i+j \geq k$ and $i \geq j$, this follows from the fact that $e(v,i,j), e(v+k+j,i,i+j-k), e(v+k+i+j,i+j-k,2k-2i), e(v+k+i+2j,k-i,2k-i-j) \in E(v,i,j)$ and $e(v,i,j) \cup e(v+k+j,i,i+j-k) \cup e(v+k+i+j,i+j-k,2k-2i) \cup e(v+k+i+2j,k-i,2k-i-j) \cup [v+i,v+i+j-1] = [v,v+4k+j-1]$. The other cases are similar.

(v) For $i+j \geq k$ and $i \geq j$, this follows from the fact that $e(v,i,j), e(v+i,j,k-i), e(v+i,j,2k-i-j), e(v+3k-j,i,j), e(v+3k-j+i,j,k-i) \in E(v,i,j)$ and also $e(v,i,j) \cup e(v+i,j,k-i) \cup e(v+i,j,2k-i-j) \cup e(v+3k-j,i,j) \cup e(v+3k-j+i,j,k-i) = [v, v+5k-j-1]$ and $e(v+i,j,k-i) \cap e(v+i,j,2k-i-j) = [v+i, v+i+j-1]$. The other cases are similar.

□

The other feature of the sets $E(v,i,j)$ that we need is that no edge is contained in too many of the sets $E(v,i,j)$. This is quantified in the following lemma. For the duration of this paper, we fix the constant $C_1 = 110$.

Lemma 3.2. *Let e be an edge in $\mathcal{H}_{n,k}$. The edge e is contained in at most C_1 of the sets $E(v,i,j) \cup E^-(v,i,j)$ for $v \in V(\mathcal{H}_{n,k})$, and $i,j \in [1,k-1]$.*

Proof. Notice that there are 55 edges mentioned in the definition of $E(v,i,j)$. For $t = 1,\ldots,55$, let $F^t(v,i,j)$ be the singleton containing the tth edge in the definition of $E(v,i,j)$, i.e. $F^1(v,i,j) = \{e(v,i,j)\}$, $F^2(v,i,j) = \{e(v+k+j,i,i+j-k)\}, \ldots, F^{55}(v,i,j) = \{e(v+2k+i,j,k-i)\}$. This definition is purely formal—for certain i and j, it is possible that an edge in $F^t(v,i,j)$ is not an edge of $\mathcal{H}_{n,k}$ (for example $F^3(v,i,j)$ contains the edge $e(v+k+i+j, i+j-k, 2k-2i)$ which is not an edge of $\mathcal{H}_{n,k}$ if $2k-2i \geq k$). Similarly it is possible for $F^t(v,i,j)$ to be empty for certain i and j—for example $F^{52}(v,i,j)$ should contain $e(v+k+i+j, \frac{i-1}{2}, k-i-j)$ which is not defined when i is even.

Clearly $E(v,i,j) \subseteq \bigcup_{t=1}^{55} F^t(v,i,j)$ holds. Also, it is straightforward to check that for fixed t, the sets $F^t(v,i,j)$ are all disjoint for $v \in V(\mathcal{H}_{n,k})$, and $i,j \in [1,k-1]$. Indeed for fixed t, if we have $e(u,a,b) \in F^t(v,i,j)$, then it is always possible to work out v,i, and j uniquely in terms of u, a, and b. These two facts, together with the Pigeonhole Principle imply that the edge e can be contained in at most 55 of the sets $E(v,i,j)$ for $v \in V(\mathcal{H}_{n,k})$, and $i,j \in [1,k]$. The lemma follows, since $C_1 \geq 2 \cdot 55 = 110$. □

A useful corollary of Lemma 3.2 is that an edge e can be contained in at most 110 of the sets $E(v)$ for $v \in V(\mathcal{H}_{n,k})$.

4. Hypergraphs of order $O(k^4)$ with the MMS-property

In this section we prove Theorem 1.5, with a weaker bound of $n \geq 14k^4$. This proof has many of the same ideas as the proof of Theorem 1.5, but is much shorter. We therefore present it in order to illustrate the techniques that we will use in proving Theorem 1.5, and hopefully aid the reader to understand that theorem.

Theorem 4.1. *For $n \geq 17k^4$, and every function $f : V(\mathcal{H}_{n,k}) \to \mathbb{R}$ which satisfies $\sum_{x \in V(\mathcal{H}_{n,k})} f(x) \geq 0$ we have $e_f^+(\mathcal{H}_{n,k}) \geq k(k-1)^2$.*

Proof. Suppose for the sake of contradiction that we have a function $f : V(\mathcal{H}_{n,k}) \to \mathbb{R}$ satisfying $\sum_{x \in V(\mathcal{H}_{n,k})} f(x) \geq 0$ such that we have $e_f^+(\mathcal{H}_{n,k}) < k(k-1)^2$.

The proof of the theorem rests on two claims. The first of these says that any sufficiently small interval I in $\mathbb{Z}_n$ is contained in a negative interval of almost the same order as I.

Claim 4.2. *Let I be an interval in $\mathbb{Z}_n$ such that $|I| \leq n - 4k$. Then there is an interval $J = [j_1, j_t]$ which satisfies the following:*

(i) $|J| \leq |I| + 4k$.

(ii) $I \subseteq J$.

(iii) $f(J) < 0$.

Proof. Without loss of generality, we may assume that I is the interval $[2k, 2km + l]$ for some $l \in [0, 2k-1]$ and $m \leq \frac{n}{2k} - 1$. First we will exhibit $2k(k-1)^2$ sets of vertex-disjoint edges covering I.

For $v \in \{0 \dots 2k-1\}$, $i, j \in \{1, \dots, k-1\}$ we let

$$\mathcal{D}(v, i, j) = \bigcup_{t=0}^{m+1} \Big(e(v + 2tk, i, j) \cup e(v + 2tk + i, j, k - i) \Big)$$

Notice that an edge $e(u, a, b)$ is contained only in the sets $\mathcal{D}(u \pmod{2k}, a, b)$ and $\mathcal{D}(u - k + b \pmod{2k}, k - b, a)$. Therefore, since there are at less than $k(k-1)^2$ nonnegative edges in $\mathcal{H}_{n,k}$, there are some v_0,i_0 and j_0 for which the set $\mathcal{D}(v_0, i_0, j_0)$ contains only negative edges. Letting $J = \bigcup \mathcal{D}(v_0, i_0, j_0) = [v_0, v_0 + 2k(m+1)]$ implies the claim. □

The second claim that we need shows that any sufficiently large interval which does not contain nonnegative edges in $\mathcal{H}_{n,d}$ must be negative.

Claim 4.3. *Let $I = [i_1, i_m]$ be an interval in $\mathbb{Z}_n$ which satisfies the following:*

(i) $|I| \geq 13k$.

(ii) *There are no nonnegative edges of $\mathcal{H}_{n,k}$ contained in I.*

We have that $f(I) < 0$.

Proof. Let $R_0 = \{v \in I : f([0, v-1]) < 0\}$ and $R_m = \{v \in I : f([v, m]) < 0\}$. Let $Q^- = \{i \in [1, k-1] : f([1, i]) < 0\}$ and $Q^+ = \{k - i \in [1, k-1] : f([1, i]) \geq 0\}$.

Since I contains only negative edges, parts (iv) and (v) of Lemma 3.1 imply that we have that $(Q^- \cup Q^+) + 4k \subseteq R_0$. Part (iii) of Lemma 3.1 implies that $4k \in R_0$. Then, parts (i) and (ii) of Lemma 3.1 imply that $(Q^- \cup Q^+ \cup \{0\}) + tk \subseteq R_0$ for any $t \in \{6, 7, \dots, \lfloor \frac{m}{k} \rfloor - 1\}$. This implies that we have $R_0 \cap [u, u + k - 1] \geq |Q^- \cup Q^+ \cup \{0\}|$ for any $u \in [6k, m - k - 1]$.

Notice that $Q^- \cup Q^+$ contains at least one element from each of the sets $\{1, k-1\}, \ldots, \left\{\left\lfloor \frac{k}{2} \right\rfloor, \left\lceil \frac{k}{2} \right\rceil\right\}$. This implies that for every $u \in \{6k, \ldots, m-k-1\}$ we have

$$|R_0 \cap [u, u+k-1]| \geq |Q^- \cup Q^+ \cup \{0\}| \geq \left\lfloor \frac{k}{2} \right\rfloor + 1 > \frac{k}{2}.$$

Similarly we obtain $|R_m \cap [u, u+k-1]| > \frac{k}{2}$ for every $u \in \{k, \ldots, m-6k\}$. By choosing $u = 6k$, we have that $|R_0 \cap [6k, 7k-1]|, |R_m \cap [6k, 7k-1]| > \frac{k}{2}$, and hence there exists some $i \in [6k, 7k-1]$ such that $i \in R_0, R_m$ hold. This gives us $f([0, m]) = f([0, i]) + f([i+1, m]) < 0$, proving the claim. □

We now prove the theorem. Suppose that every interval of order $17k$ in $\mathcal{H}_{n,k}$ contains a nonnegative edge. Since there are at least $\frac{n}{17k} \geq k^3$ such disjoint intervals in $\mathcal{H}_{n,k}$, we have at least k^3 nonnegative edges in $\mathcal{H}_{n,k}$, contradicting our initial assumption that $e_f^+(\mathcal{H}_{n,k}) < k(k-1)^2$.

Suppose that there is an interval I of order $17k$ in $\mathcal{H}_{n,k}$ which contains only negative edges. Applying Claim 4.2 to $V(\mathcal{H}_{n,k}) \setminus I)$ we obtain an interval $J \subseteq I$ such that $f(V(\mathcal{H}_{n,k}) \setminus J) < 0$ and $|J| \geq 13k$. Applying Claim 4.3 to J we obtain that $f(J) < 0$. Therefore, we have $f(V(\mathcal{H}_{n,k})) = f(J) + f(V(\mathcal{H}_{n,k}) \setminus J) < 0$ contradicting the assumption that $f(V(\mathcal{H}_{n,k})) \geq 0$ in the theorem □

It is not hard to see that Claim 4.3 would still be true if we allowed I to contain a small number of nonnegative edges. The proof of Theorem 1.5 is similar to the proof of Theorem 4.1 since it also consists of two main claims which are analogues of Claims 4.2 and 4.3. However the analogue of Claim 4.3 is much stronger since it allows for $O(k^3)$ nonnegative edges to be contained in I. This is the main improvement in the proof of Theorem 4.1 which is needed to obtain the linear bound which we have in Theorem 1.5.

5. Proof of Theorem 1.5

In this section we use ideas from Sections 3 and 4 in order to prove Theorem 1.5.

Proof of Theorem 1.5. For convenience, we fix the following constants for the duration of the proof.

$$\begin{array}{ll} C_0 = 10^{46} & \epsilon_0 = 10^{-9} \\ C_1 = 110 & \epsilon_1 = 10^{-18} \\ C_2 = 10^{16} & \epsilon_2 = 10^{-6} \\ C_3 = 28 & \epsilon_3 = 10^{-2} \\ & \epsilon_4 = 0.1 \\ & \epsilon_5 = 0.25 \end{array}$$

Let $n \geq C_0 k$, and let $\mathcal{H}_{n,k}$ be the hypergraph defined in Section 3. Recall that for any vertex $v \in V(\mathcal{H}_{n,k})$, we have $|E(v)| = \Theta(k^2)$.

Definition 5.1. *We say that a vertex v in $\mathcal{H}_{n,k}$ is **bad** if at least $\epsilon_0 k^2$ of the edges in $E(v)$ are nonnegative and **good** otherwise.*

Let $G_{\mathcal{H}}$ be the set of good vertices in $\mathcal{H}_{n,k}$.

Suppose that we have a function $f : V(\mathcal{H}_{n,k}) \to \mathbb{R}$ such that we have $e_f^+(\mathcal{H}_{n,k}) < k(k-1)^2$. We will show that $f(V(\mathcal{H}_{n,k})) < 0$ holds. The proof of the theorem consists of the following two claims.

Claim 5.2. *Let I be an interval in $\mathbb{Z}_n$ such that $|I| \leq n - 4C_2k$. There is an interval $J = [j_1, j_t]$ which satisfies the following:*

(i) $|J| \leq |I| + 4C_2k$.

(ii) $I \subseteq J$.

(iii) *Both $j_1 - 1$ and $j_t + 1$ are good.*

(iv) $f(J) < 0$.

Claim 5.3. *Let $I = [i_1, i_m]$ be an interval in $\mathbb{Z}_n$ which satisfies the following:*

(i) $C_3k \leq |I| \leq (C_3 + 4C_2)k$.

(ii) *Both i_1 and i_m are good.*

(iii) *Every subinterval of I of order k, contains at most $\epsilon_1 k$ bad vertices.*

We have that $f(I) < 0$.

Once we have these two claims, the theorem follows easily:

First suppose that no intervals in $\mathbb{Z}_n$ of order $(C_3 + 4C_2)k$ satisfies condition (iii) of Claim 5.3. This implies that there are at least $\epsilon_1 C_0 k/(C_3 + 4C_2)$ bad vertices in $\mathcal{H}_{n,k}$. Then Claim 3.2 together with the definition of "bad" implies that there are at least $\epsilon_0\epsilon_1 C_0 k^3/C_1(C_3+4C_2)$ nonnegative edges in $\mathcal{H}_{n,k}$. However, since $\epsilon_0\epsilon_1 C_0/C_1(C_3+4C_2) \geq 1$, this contradicts our assumption that $e_f^+(\mathcal{H}_{n,k}) < k(k-1)^2$.

Now, suppose that there is an interval I of order $(C_3 + 4C_2)k$ which satisfies condition (iii) of Claim 5.3. Notice that all subintervals of I will also satisfy condition (iii) of Claim 5.3. Applying Claim 5.2 to $V(\mathcal{H}_{n,k}) \setminus I$ gives an interval $J \subseteq I$ which satisfies all the conditions of Claim 5.3 and also $f(V(\mathcal{H}_{n,k}) \setminus J) < 0$. Applying Claim 5.3 to J implies that we also have $f(J) < 0$. We have $\sum_{v \in \mathcal{H}_{n,k}} f(v) = f(V(\mathcal{H}_{n,k}) \setminus J) + f(J) < 0$, contradicting our initial assumption and proving the theorem.

It remains to prove Claims 5.2 and 5.3.

Proof of Claim 5.2. Without loss of generality, we may assume that I is the interval $[0, 2km + l]$ for some $l \in [0, 2k-1]$ and $m < \frac{n}{2k} - 2C_2$. We partition $[1, 2k]$ into two sets as follows.

Definition 5.4. *For $r \in [1, 2k]$ we say that r is **unblocked** if for every $t \in [-C_2, m + C_2]$, there are some $i, j \in [1, k-1]$ such that both of the edges $e(2tk + r, i, j)$ and $e(2tk + r + i, j, k - i)$ are negative. We say that r is **blocked** otherwise.*

Notice that if r is unblocked, then for every $t_1 \in [-C_2, 0]$ and $t_2 \in [m, m+C_2]$ we have that $f([2t_1k + r, 2t_2k + r - 1]) < 0$. Therefore the claim holds unless either $2t_1k + r - 1$ or $2t_2k + r$ is bad. Therefore, for each r which is unblocked, we can assume that all the vertices in either $\{r - 1 - 2kC_2, r - 1 - 2k(C_2 - 1), \ldots, r - 1\}$ or $\{r + 2km, r + 2k(m + 1), \ldots, r + 2k(m + C_2)\}$ are bad.

To each $r \in [1, 2k]$, we assign a set of nonnegative edges, $P(r)$, as follows:

- If r is blocked, then there is some $t_r \in [-C_2, m+C_2]$, such that for every $i, j \in [1, k-1]$ one of the edges $e(2t_rk+r, i, j)$ or $e(2t_rk+r+i, j, k-i)$ is nonnegative. We let $P(r)$ be the set of these edges. Notice that this ensures that $|P(r)| \geq (k-1)^2$. Also, note that for fixed a,b,c the set $P(r)$ can contain at most one edge of the form $e(a + 2tk, b, c)$ for any $t \in [-C_2, m + C_2]$.

- If r is unblocked we know that all the vertices in either $\{r-1-2kC_2, r-1-2k(C_2-1), \ldots, r-1\}$ or $\{r + 2km, r + 2k(m + 1), \ldots, r + 2k(m + C_2)\}$ are bad. Let $P(r)$ be the set of nonnegative edges in $E(r-1-2kC_2) \cup E(r-1-2k(C_2-1)) \cup \cdots \cup E(r-1) \cup E(r + 2km) \cup E(r + 2k(m + 1)) \cup \cdots \cup E(r + 2k(m + C_2))$. Since at least C_2 of these vertices are bad, Lemma 3.2 together with the Pigeonhole Principle implies that $|P(r)| \geq \frac{C_2\epsilon_0}{C_1}k^2$.

Notice that an edge e can be in at most 2 of the sets $P(r)$ for r blocked. This is because it can be in at most one such set as an edge of the form "$e(tk + r, i, j)$" and in at most one such set and as an edge of the form "$e(tk + r + i, j, k - i)$". Therefore we have:

$$\left| \bigcup_{r \text{ blocked}} P(r) \right| \geq \sum_{r \text{ blocked}} \frac{1}{2}(k-1)^2 \tag{5.1}$$

Lemma 3.2 implies that an edge e can be in at most C_1 of the sets $P(r)$ for r unblocked. Therefore we have:

$$\left| \bigcup_{r \text{ unblocked}} P(r) \right| \geq \sum_{r \text{ unblocked}} \frac{C_2\epsilon_0}{(C_1)^2}k^2 \tag{5.2}$$

We claim that for any $s \in [1, 2k]$, we have

$$\left| \left(\bigcup_{t \in [-C_2, m+C_2]} E(s + 2tk) \right) \cap \left(\bigcup_{r \text{ blocked}} P(r) \right) \right| \leq 2|E(s)|. \tag{5.3}$$

Indeed, otherwise the Pigeonhole Principle implies that for some $r \in [1, 2k]$, t_1, t_2, $t_3 \in [-C_2, m + C_2]$, and $i, j \in [1, k - 1]$ we have three distinct edges $e(r + 2t_1k, i, j)$, $e(r + 2t_2k, i, j)$, and $e(r + 2t_3k, i, j)$ which are are all contained in $\left(\bigcup_{t \in [-C_2, m+C_2]} E(s + 2tk)\right) \cap \left(\bigcup_{r \text{ blocked}} P(r)\right)$. This means that there are some r_1, r_2, and $r_3 \in [1, 2k]$ which are blocked, such that $e(r + 2t_lk, i, j) \in P(r_l)$ holds for $l = 1$, 2 and 3. Since each r_l is blocked, all the edges in $P(r_l)$ are of the form $e(2t'k + r_l, i', j')$ or $e(2t'k + r_l + i', j', k - i')$ for some

$t' \in [-C_2, m+C_2]$ and $i', j' \in [1, k-1]$. This, together with $e(r+2t_lk, i, j) \in P(r_l)$, implies that we have $r_1, r_2, r_3 \in \{r, r-k+j\}$. This means that for some distinct $l, l' \in \{1,2,3\}$, we have $r_l = r_{l'}$, which means that both $e(r+2t_lk, i, j)$ and $e(r+2t_{l'}k, i, j)$ are contained in $P(r_l)$. However, this contradicts our definition of $P(r_l)$ for r_l blocked which allowed only one edge of the form $e(r+2tk, i, j)$ to be in $P(r_l)$ for fixed r, i and j. This shows that (5.3) holds for all $s \in [1, 2k]$.

Recall that for all vertices s we have $|E(s)| \leq C_1k^2$. This, together with (5.3) implies that we have

$$\begin{aligned}\left|\left(\bigcup_{s \text{ unblocked}} P(s)\right) \cap \left(\bigcup_{r \text{ blocked}} P(r)\right)\right| &\leq \left|\left(\bigcup_{\substack{s \text{ unblocked,} \\ t\in[-C_2,m+C_2]}} E(s+2tk)\right) \cap \left(\bigcup_{r \text{ blocked}} P(r)\right)\right| \\ &\leq \sum_{s \text{ unblocked}} 2|E(s)| \\ &\leq \sum_{s \text{ unblocked}} 2C_1k^2. \end{aligned} \tag{5.4}$$

Putting (5.1), (5.2), and (5.4) together, we obtain:

$$\begin{aligned} e_f^+(\mathcal{H}_{n,k}) &\geq \left|\bigcup_{r \text{ blocked}} P(r)\right| + \left|\bigcup_{r \text{ unblocked}} P(r)\right| - \left|\left(\bigcup_{s \text{ unblocked}} P(s)\right) \cap \left(\bigcup_{r \text{ blocked}} P(r)\right)\right| \\ &\geq \sum_{r \text{ blocked}} \frac{1}{2}(k-1)^2 + \sum_{r \text{ unblocked}} \frac{C_2\epsilon_0}{(C_1)^2}k^2 - \sum_{s \text{ unblocked}} 2C_1k^2 \\ &\geq \sum_{r \text{ blocked}} \frac{1}{2}(k-1)^2 + \sum_{r \text{ unblocked}} \frac{1}{2}k^2 \\ &\geq k(k-1)^2. \end{aligned} \tag{5.5}$$

The second last inequality follows from $\frac{C_2\epsilon_0}{(C_1)^2} - 2C_1 \geq \frac{1}{2}$. The last inequality follows from the fact that "the number of blocked vertices" + "the number of unblocked vertices" $= 2k$. However (5.5) contradicts the assumption that there are less than $k(k-1)^2$ nonnegative edges in $\mathcal{H}_{n,k}$, proving the claim. □

It remains to prove Claim 5.3.

Proof of Claim 5.3. Without loss of generality, we can assume that $I = [0, m]$ for some $m \leq (C_3 + 4C_2)k$.

Recall that we are using notation from additive combinatorics for sumsets and translates. Except where otherwise stated, sumsets will lie in $\mathbb{Z}$. For a set $A \subseteq \mathbb{Z}$, define

$$A \,(\text{mod } k) = \{b \in [0, k-1] : b \equiv a \,(\text{mod } k) \text{ for some } a \in A\}.$$

For each vertex v, we define a set of vertices $R(v)$ contained in I.

$$R(v) = \{u \in [v+1, m] : f([v, u-1]) < 0 \text{ and } u \text{ is good.}\}$$

$R(v)$ has the following basic properties.

Claim 5.5. *The following hold.*

(i) *If $u > v$ and $u \in R(v)$, we have $R(u) \subseteq R(v)$.*

(ii) *Suppose that $t \geq 2$ and we have a set $X \subseteq R(v) \cap [w, w+2k-1]$, for some vertex w. There is a subset $X' \subseteq X$, such that we have $|X'| \geq |X| - 2\epsilon_1 kt$ and $X' + t'k \subseteq R(v)$ for every $t' \in \{2, \ldots, t\}$.*

(iii) *Suppose that we have $X \subseteq [0, 2k-1]$ such that $X + t_0 k \subseteq R(0)$ for some t_0. There is a subset $X' \subseteq X \pmod k$, such that $X' + (t_0+3)k \subseteq R(0)$ and $|X'| \geq |X| - 6\epsilon_1 k$.*

(iv) *Suppose that we have $X \subseteq [w, w+k-1] \cap R(0)$ for some w. Then for any $v \geq w+2k$, we have we have $|R(0) \cap [v, v+k-1]| \geq |X| - 2\epsilon_1(v-w+k)$.*

Proof. (i) This part is immediate from the definition of $R(v)$.

(ii) First, we deal with the case when $t = 2$ or 3. The general case will follow by induction.

Suppose that we have $x \in X$. Since x is good, Lemma 3.2 implies that there are at most $\epsilon_0 C_1 k^2$ pairs i, j for which $E(x, i, j)$ contains a nonnegative edge. Therefore, since $\epsilon_0 C_1 < 1$, there must be at least one pair i_0, j_0 for which all the edges in $E(x, i_0, j_0)$ are negative. Combining this with parts (i) and (ii) of Lemma 3.1 implies that we have

$$f([v, x+2k-1]), f([v, x+3k-1]) < 0. \tag{5.6}$$

If $t = 2$ we let $X' = X \cap (G_{\mathcal{H}} - 2k)$. The identity 5.6 implies that $X' + 2k \subseteq R(v)$. By condition (iii) of Claim 5.3, we know that there are at most $2\epsilon_1 k$ bad vertices in $[w+2k, w+4k-1]$, which implies that $|X'| \geq |X| - 2\epsilon_1 k$.

Similarly, if $t = 3$ we let $X' = X \cap (G_{\mathcal{H}} - 2k) \cap (G_{\mathcal{H}} - 3k)$. The identity 5.6 implies that $X' + 2k, X' + 3k \subseteq R(v)$. By condition (iii) of Claim 5.3, we know that there are at most $3\epsilon_1 k$ bad vertices in $[w+2k, w+5k-1]$, which implies that $|X'| \geq |X| - 3\epsilon_1 k$.

Suppose that the claim holds for $t = t_0$ for some $t_0 \geq 3$. We will show that it holds for $t = t_0 + 1$. We know that there is a set $X' \subseteq X + t_0 k$, such that we have $|X'| \geq |X| - \epsilon_1 k t_0$ and $X' + t'k \subseteq R(v)$ for $t' = 2, \ldots, t_0$. Applying the $t = 2$ part of this claim to $X' + t_0 k$ we obtain a set $X'' \subseteq X'$ such that $|X''| \geq |X'| - \epsilon_1 k \geq |X| - \epsilon_1 k(t_0+1)$ and also $X'' + (t_0+1)k \subseteq R(v)$. This proves the claim by induction.

(iii) Apply part (ii) to $X + t_0 k$ with $t = 3$ to obtain a set X' with $|X'| \geq |X| - 3\epsilon_1 k$ and $X' + t_0 k + \{2k, 3k\} \subseteq R(0)$. Let $X'' = X' \pmod k$ to obtain a set satisfying $X'' \subseteq X \pmod k$ and $|X''| \geq |X \pmod k| - 3\epsilon_1 k$. We have that $X'' + t_0 + 3k = (X' \cap [0, k-1] + t_0 + 3k) \cup (X' \cap [k, 2k-1] + t_0 + 2k) \subseteq X' + t_0 + \{2k, 3k\} \subseteq R(0)$.

(iv) Apply part (ii) to X with $t = \left\lfloor \frac{v-w}{k} \right\rfloor + 1$ to obtain a set X' with $|X'| \geq |X| - \epsilon_1 \left(\left\lfloor \frac{v-w}{k} \right\rfloor + 1 \right) k$ and $X' + t'k \subseteq R(0)$ for any $t' = 2, \ldots, \left(\left\lfloor \frac{v-w}{k} \right\rfloor + 1 \right) k$. For any $x \in X'$, either $x + \left\lfloor \frac{v-w}{k} \right\rfloor k$ or $x + \left(\left\lfloor \frac{v-w}{k} \right\rfloor + 1 \right) k$ is in $[v, v+k-1] \cap R_0$, which implies that $|R(0) \cap [v, v+k-1]| \geq |X'| \geq |X| - \epsilon_1 \left(\left\lfloor \frac{v-w}{k} \right\rfloor + 1 \right) k \geq |X| - \epsilon_1(v-w+k)$.

□

To every vertex $v \in I$ and $\epsilon > 0$, we assign sets $Q^+_\epsilon(v)$, $Q^-_\epsilon(v)$, $Q_\epsilon(v) \subseteq [1, k-1]$ as follows.

$$\begin{aligned} Q^-_\epsilon(v) = \{j \in [1,k-1] : f([v+i, v+i+j-1]) < 0 \\ \text{for at least } \epsilon k \text{ numbers } i \in [1,k-1]\} \\ Q^+_\epsilon(v) = \{k-j \in [1,k-1] : f([v+i, v+i+j-1]) \geq 0 \\ \text{for at least } \epsilon k \text{ numbers } i \in [1,k-1]\} \\ Q_\epsilon(v) = Q^-_\epsilon(v) \cup Q^+_\epsilon(v) \cup \{0\}. \end{aligned}$$

$Q_\epsilon(v)$ has the following basic properties.

Claim 5.6. *The following hold.*

(i) *For any $r \in [0,k]$, we have $Q_{2\epsilon}(v) \subseteq Q_\epsilon(v-r) \cup Q_\epsilon(v-r+k)$.*

(ii) *For $\epsilon \leq \frac{1}{2}$, $x \in [1,k-1]$, and $v \in I$ either x or $k-x$ is in $Q_\epsilon(v)$.*

(iii) *For $\epsilon \leq \frac{1}{2}$ and $v \in I$, we have $|Q_\epsilon(v)| \geq \frac{1}{2}k$.*

Proof. If $j \in Q^-_{2\epsilon}(v)$, then there are at least $2\epsilon k$ numbers $i \in [1,k-1]$ for which $f([v+i, v+i+j-1]) < 0$. For every $r \in [0,k]$ the Pigeonhole Principle implies that there must either be at least ϵk numbers $i \in [1,k-1]$ for which $f([v-r+i, v-r+i+j-1]) < 0$ or at least ϵk numbers $i \in [1,k-1]$ for which $f([v-r+k+i, v-r+k+i+j-1]) < 0$. Therefore we have $Q^-_{2\epsilon}(v) \subseteq Q^-_\epsilon(v-r) \cup Q^-_\epsilon(v-r+k)$. Similarly we obtain $Q^+_{2\epsilon}(v) \subseteq Q^+_\epsilon(v-r) \cup Q^+_\epsilon(v-r+k)$ which implies part (i).

Part (ii) is immediate from the definition of $Q_\epsilon(v)$. Part (iii) follows from (ii). □

The following claim shows that for a good vertex v, there is a certain translate of $Q_{\epsilon_5}(v)$ which will nearly be contained in $R(v)$.

Claim 5.7. *For any good vertex v satisfying $0 \leq v \leq m - 5k$, there is a $Q' \subseteq Q_{\epsilon_5}(v)$ such that $|Q'| \geq |Q_{\epsilon_5}(v)| - \epsilon_2 k$ and we have*

$$Q' + 4k + v \subseteq R(v).$$

Proof. Let $T \subseteq [1,k-1]$ be the set of $j \in [1,k-1]$ for which there are at least $\epsilon_5 k$ numbers $i \in [1,k-1]$ such that $E(v,i,j)$ contains a nonnegative edge. We have at least $|T|\epsilon_5 k$ pairs $i,j \in [1,k-1]$ for which $E(v,i,j)$ contains a nonnegative edge. Since v is good, Lemma 3.2 implies that at most $\epsilon_0 C_1 k^2$ of the sets $E(v,i,j)$ contain nonnegative edges for $i,j \in [1,k-1]$. Therefore, we have $|T|\epsilon_5 k \leq \epsilon_0 C_1 k^2$. We define the set Q' as

$$Q' = \big((Q^-_{\epsilon_5}(v) \setminus T) \cup (Q^+_{\epsilon_5}(v) \setminus T) \cup \{0\}\big) \cap (G_{\mathcal{H}} - 4k).$$

First we prove $Q' + 4k + v \subseteq R(v)$. Suppose that we have $j \in Q^-_{\epsilon_5}(v) \setminus T$. From the definition of T, there are at at more than $k - 1 - \epsilon_5 k$ numbers $i \in [1, k-1]$ such that all the edges in $E(v, i, j)$ are negative. From the definition of $Q^-_{\epsilon_5}(v)$, there are at least $\epsilon_5 k$ numbers $i \in [1, k-1]$ such that $[v + i, v + i + j - 1]$ is negative. Therefore, there is some $i \in [1, k-1]$ such that all the edges in $E(v, i, j)$ are negative and also $[v + i, v + i + j - 1]$ is negative. Part (iv) of Lemma 3.1 implies that we have $f(v, v + 4k + j - 1) < 0$ and so $(Q^-_{\epsilon_5}(v) \setminus T + 4k + v) \cap G_{\mathcal{H}} \subseteq R(v)$. Similarly, using part (v) of Lemma 3.1, it is possible to show that $(Q^+_{\epsilon_5}(v) \setminus T + 4k + v) \cap G_{\mathcal{H}} \subseteq R(v)$. Finally, part (iii) of Lemma 3.1 implies that we have $(\{0\} + 4k + v) \cap G_{\mathcal{H}} \subseteq R(v)$, and hence $Q' + 4k + v \subseteq R(v)$.

Now we prove $|Q'| \geq |Q_{\epsilon_5}(v)| - \epsilon_2 k$. Since $|T| \leq \epsilon_0 C_1 k/\epsilon_5$, we must have

$$|Q_{\epsilon_5}(v) \setminus T| \geq |Q_{\epsilon_5}(v)| - \frac{\epsilon_0 C_1}{\epsilon_5} k. \tag{5.7}$$

Condition (iii) of Claim 5.3 implies that

$$|Q'| \geq |Q_{\epsilon_5}(v) \setminus T| - \epsilon_1 k. \tag{5.8}$$

Now, (5.7), (5.8) and $\epsilon_2 \geq \epsilon_0 C_1/\epsilon_5 + \epsilon_1$ imply $|Q'| \geq |Q_{\epsilon_5}(v)| - \epsilon_2 k$, proving the claim. □

Definition 5.8. *For $S \subseteq A \times B$ we define*

$$A +_S B = \{a + b : (a, b) \in S\}.$$

The following claim shows that for a certain large set S, a translate of $Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k)$ is contained in $R(0)$.

Claim 5.9. *There is a set $S \subseteq Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k)$ such that $|S| \geq |Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k)| - \epsilon_3^2 k^2$ and we have*

$$\big(Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k)\big) + 13k \subseteq R(0).$$

Proof. For every good vertex $v \in I$, Claim 5.7 combined with part (ii) of Claim 5.5 implies that there is a set $Q_v \subseteq Q_{\epsilon_5}(v)$ such that we have $Q_v + v + \{6k, 7k\} \subseteq R(v)$ and also

$$|Q_v| \geq |Q_{\epsilon_5}(v)| - (7\epsilon_1 + \epsilon_2)k. \tag{5.9}$$

Now, part (i) of Claim 5.5 implies that we have

$$\bigcup_{v \in R(0) \cap [6k, 8k-1]} R(v) \subseteq R(0). \tag{5.10}$$

Combining $Q_v + v + \{6k, 7k\} \subseteq R(v)$ with (5.10) implies that we have

$$\bigcup_{v \in (Q_0 + \{6k, 7k\})} (Q_v + v + \{6k, 7k\}) \subseteq R(0). \tag{5.11}$$

We let

$$S = \{(a,b) \in Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k) : a \in Q_0 \text{ and } b \in Q_{a+6k} \cup Q_{a+7k}\}.$$

The identity (5.11) implies that we have

$$\begin{aligned}
Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k) + 13k &= \{a+b : a \in Q_0 \text{ and } \\
&\qquad\qquad b \in (Q_{a+6k} \cup Q_{a+7k}) \cap Q_{2\epsilon_5}(7k)\} + 13k \\
&\subseteq \{a+b : a \in Q_0 \text{ and } b \in Q_{a+6k} \cup Q_{a+7k}\} + 13k \\
&= \left(\bigcup_{a \in Q_0+6k} Q_a + a + 7k\right) \cup \left(\bigcup_{a \in Q_0+7k} Q_a + a + 6k\right) \\
&\subseteq \bigcup_{a \in (Q_0 + \{6k,7k\})} (Q_a + a + \{6k, 7k\}) \\
&\subseteq R(0).
\end{aligned}$$

Now we prove $|S| \geq |Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k)| - \epsilon_3^2 k^2$. Notice that for each $a \in [0, k-1]$, part (i) of Claim 5.6 implies

$$Q_{2\epsilon_5}(7k) \subseteq Q_{\epsilon_5}(a+6k) \cup Q_{\epsilon_5}(a+7k) \text{ for all } a \in Q_{\epsilon_5}(0). \tag{5.12}$$

The identity (5.12) combined with (5.9) and $Q_v \subseteq Q_{\epsilon_5}(v)$ implies that for all $a \in [1, k-1]$ we have

$$\begin{aligned}
|(Q_{a+6k} \cup Q_{a+7k}) \cap Q_{2\epsilon_5}(7k)| &\geq |(Q_{\epsilon_5}(a+6k) \cup Q_{\epsilon_5}(a+7k)) \cap Q_{2\epsilon_5}(7k)| \\
&\qquad - (14\epsilon_1 + 2\epsilon_2)k \\
&= |Q_{2\epsilon_5}(7k)| - (14\epsilon_1 + 2\epsilon_2)k.
\end{aligned}$$

This gives us

$$\begin{aligned}
|S| &= \sum_{a \in Q_0} |(Q_{a+6k} \cup Q_{a+7k}) \cap Q_{2\epsilon_5}(v)| \\
&\geq \sum_{a \in Q_0} \Big(|Q_{2\epsilon_5}(7k)| - (14\epsilon_1 + 2\epsilon_2)k\Big) \\
&\geq \Big(|Q_{\epsilon_5}(0)| - (7\epsilon_1 + \epsilon_2)k\Big)\Big(|Q_{2\epsilon_5}(7k)| - (14\epsilon_1 + 2\epsilon_2)k\Big) \\
&\geq |Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k)| - (21\epsilon_1 + 3\epsilon_2)k^2 \\
&\geq |Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k)| - \epsilon_3^2 k^2.
\end{aligned}$$

The second last inequality follows from $|Q_{\epsilon_5}(0)|, |Q_{2\epsilon_5}(7k)| \leq k$. The last inequality follows from $\epsilon_3^2 \geq 21\epsilon_1 + 3\epsilon_2$. □

Claim 5.9 is combined with the following.

Claim 5.10. *Suppose that A and $B \subseteq \mathbb{Z}_k$, and satisfy that for any $x \in \mathbb{Z}_k$, either x or $-x \in A$ and either x or $-x \in B$. Let $S \subseteq A \times B$ be a set satisfying $|S| \geq |A \times B| - \epsilon_3^2 k^2$. We have*

$$|A +_S B| \geq \left(\frac{1}{2} + \epsilon_4\right) k.$$

When k is prime, Claim 5.10 follows from a theorem due to Lev [15][1], which itself is closely related to a theorem due to Pollard [19]. In order to prove Claim 5.10, we will need some results from additive combinatorics. We define

$$(A+B)_i = \{x \in \mathbb{Z}_k : x = a + b \text{ for at least } i \text{ distinct pairs } (a,b) \in A \times B\}.$$

Notice that we have $(A+B)_{i+1} \subseteq (A+B)_i$.

The proof of Claim 5.10 will use the following theorem due to Grynkiewicz.

Theorem 5.11 (Grynkiewicz, [10])**.** *Let A and $B \subseteq \mathbb{Z}_k$ and $t \leq k$. We have one of the following.*

(i) *The following holds.*

$$\sum_{i=1}^{t} |(A+B)_i| \geq t|A| + t|B| - 2t^2 + 1. \tag{5.13}$$

(ii) *There are sets $A' \subseteq A$ and $B' \subseteq B$ such that $|A \setminus A'| + |B \setminus B'| \leq t-1$ and we have $A' + B' = (A+B)_t$.*

We define the *stabiliser* of a set $X \in \mathbb{Z}_k$ to be $\mathrm{Stab}(X) = \{y \in \mathbb{Z}_k : y + X = X\}$. We use the following theorem due to Kneser.

Theorem 5.12 (Kneser, [14])**.** *Let A and $B \subseteq \mathbb{Z}_k$ and H the stabiliser of $A+B$ in $\mathbb{Z}_k$. We have*

$$|A+B| \geq |A+H| + |B+H| - |H|. \tag{5.14}$$

Sumsets in Claim 5.10, Theorem 5.11 and Theorem 5.12 are all in $\mathbb{Z}_k$.

Proof of Claim 5.10. Notice that since x or $-x \in A, B$, we must have $|A|, |B| \geq \frac{1}{2}k$. Our initial goal will be to show that we have

$$|(A+B)_{\epsilon_3 k}| \geq \left(\frac{1}{2} + \epsilon_4 + \epsilon_3\right) k. \tag{5.15}$$

Apply Theorem 5.11 to A and B with $t = 2\epsilon_3 k$. We split into two cases, depending on which part of Theorem 5.11 holds.

[1] A special case of Theorem 1 in [15] is that if A, B are subsets of $\mathbb{Z}_p$, and $R \subseteq A \times B$ with $|A| + |B| \leq p + \sqrt{|R|}$ and $\sqrt{|R|} \leq |A|, |B|$ then we have $|A +_{A \times B \setminus R} B| \geq |A| + |B| - 2\sqrt{|R|} - 1$. This can be used to prove Claim 5.10 for prime k (by setting $S = A \times B \setminus R$).

(i) Suppose that (5.13) holds. Since we are working over $\mathbb{Z}_k$ in this claim, we have $|(A+B)_i| \leq k$. Combining this with (5.13) implies

$$\sum_{i=\epsilon_3 k}^{2\epsilon_3 k} |(A+B)_i| \geq 2\epsilon_3 k\Big(|A|+|B|-4\epsilon_3 k\Big) + 1 - \sum_{i=1}^{\epsilon_3 k-1} |(A+B)_i|$$
$$\geq \epsilon_3 k\Big(2|A|+2|B|-(1+8\epsilon_3)k\Big).$$

This, together with $(A+B)_{i+1} \subseteq (A+B)_i$ implies that we have

$$|(A+B)_{\epsilon_3 k}| \geq 2|A|+2|B|-(1+8\epsilon_3)k.$$

The identity (5.15) follows since we have $|A|, |B| \geq \frac{1}{2}k$ and $1-8\epsilon_3 \geq 1/2+\epsilon_4+\epsilon_3$.

(ii) Suppose that we have two sets A' and B' as in part (ii) of Theorem 5.11. Apply Theorem 5.12 to the sets A' and B'.

Note that $|A \setminus A'| + |B \setminus B'| \leq t-1$ together with (5.14) and $|A|, |B| \geq \frac{1}{2}k$ implies that we have

$$\begin{aligned}
|(A+B)_{\epsilon_3 k}| &\geq |(A+B)_{2\epsilon_3 k}| \\
&= |A'+B'| \\
&\geq |A'+\operatorname{Stab}(A'+B')| + |B'+\operatorname{Stab}(A'+B')| - |\operatorname{Stab}(A'+B')| \qquad (5.16)\\
&\geq |A|+|B|-|\operatorname{Stab}(A'+B')| - 2\epsilon_3 k \\
&\geq (1-2\epsilon_3)k - |\operatorname{Stab}(A'+B')|. \qquad (5.17)
\end{aligned}$$

If $|\operatorname{Stab}(A'+B')| \leq \frac{1}{3}k$, then (5.15) follows (5.17) combined with $1-2\epsilon_3-1/3 \geq 1/2+\epsilon_4+\epsilon_3$.

Otherwise, Lagrange's Theorem implies that $\operatorname{Stab}(A'+B')$ is either all of $\mathbb{Z}_k$ or that k is even and $\operatorname{Stab}(A'+B')$ is the set of even elements of $\mathbb{Z}_k$. If $\operatorname{Stab}(A'+B') = \mathbb{Z}_k$ holds, then we have $A'+\operatorname{Stab}(A'+B') = B'+\operatorname{Stab}(A'+B') = \mathbb{Z}_k$. Substituting this into (5.16) implies that we have $|(A+B)_{\epsilon_3 k}| = k$ and so (5.15) holds.

Suppose that $\operatorname{Stab}(A'+B')$ consists of all the even elements of $\mathbb{Z}_k$. Since for every x, either x or $-x \in A$, there are at least $\frac{1}{4}k$ even elements in A, and at least $\frac{1}{4}k$ odd elements in A. Therefore, since $|A'| \geq |A| - 2\epsilon_3 k$, A' must contain an even element and an odd element. This implies that $A'+\operatorname{Stab}(A'+B') = \mathbb{Z}_k$. Similarly $B'+\operatorname{Stab}(A'+B') = \mathbb{Z}_k$. Thus (5.16) implies that we have $|(A+B)_{\epsilon_3 k}| = k$ and so (5.15) holds.

Now, we use (5.15) to deduce the claim. Let $T = (A+B)_{\epsilon_3 k} \setminus (A +_S B)$. We have $|A +_S B| + |T| \geq |(A+B)_{\epsilon_3 k}|$. Notice that from the definition of $(A+B)_{\epsilon_3 k}$ we have

$\epsilon_3 k|T| + |S| \le |A \times B|$. This, combined with (5.15) and $|S| \ge |A \times B| - \epsilon_3^2 k^2$ implies that we have

$$\begin{aligned} |A +_S B| &\ge |(A+B)_{\epsilon_3 k}| - |T| \\ &\ge |(A+B)_{\epsilon_3 k}| - \frac{1}{\epsilon_3 k}(|A \times B| - |S|) \\ &\ge |(A+B)_{\epsilon_3 k}| - \epsilon_3 k \\ &\ge \left(\frac{1}{2} + \epsilon_4\right) k. \end{aligned}$$

□

Claims 5.9 and 5.10 cannot be directly combined since sumsets in Claim 5.9 are in $\mathbb{Z}$ whereas sumsets in Claim 5.10 are in $\mathbb{Z}_k$. However, Claim 5.9 gives us a set S such that $|S| \ge |Q_{\epsilon_5}(0) \times Q_{2\epsilon_5}(7k)| - \epsilon_3^2 k^2$ and we have $\left(Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k)\right) + 13k \subseteq R(0)$. Part (iii) of Claim 5.5 implies that there is a subset $Q' \subseteq \left(Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k)\right) \pmod k$ such that $Q' + 16k \subseteq R(0)$ and we have

$$|Q'| \ge |\left(Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k)\right) \pmod k| - 3\epsilon_1 k. \tag{5.18}$$

By Claim 5.10 and part (ii) of Claim 5.6, we have

$$|\left(Q_{\epsilon_5}(0) +_S Q_{2\epsilon_5}(7k)\right) \pmod k| \ge \left(\frac{1}{2} + \epsilon_4\right) k. \tag{5.19}$$

Combining (5.18) and (5.19) implies that $|R(0) \cap [16k, 17k-1]| \ge (1/2 + \epsilon_4 - 3)\,\epsilon_1 k$. Applying part (iv) of Claim 5.5 with $X = R(0) \cap [16k, 17k-1]$ implies that for any $w \in I$ with $w \ge 18k$, we have

$$|R(0) \cap [w, w+k-1]| \ge \left(\frac{1}{2} + \epsilon_4 - \epsilon_1\left(\frac{w}{k} + 4\right)\right) k.$$

Combining this with $m \le (4C_2 + C_3)k$ gives

$$|R(0) \cap [m - 17k, m - 16k - 1]| \ge \left(\frac{1}{2} + \epsilon_4 - \epsilon_1(4C_2 + C_3 + 4)\right). \tag{5.20}$$

We can define $R^-(v) = \{u \in I \cap G_{\mathcal{H}} : f([u+1, v]) < 0\}$. By symmetry, we obtain

$$|R^-(m) \cap [m - 17k, m - 16k - 1]| \ge \left(\frac{1}{2} + \epsilon_4 - 3\epsilon_1\right) k. \tag{5.21}$$

Now, (5.20), (5.21), and $\epsilon_4 > \epsilon_1(4C_2 + C_3 + 4)$ imply that we have

$$\begin{aligned} |R(m) \cap [m - 17k, m - 16k - 1]| &> \frac{1}{2}k, \\ |R^-(m) \cap [m - 17k, m - 16k - 1]| &> \frac{1}{2}k. \end{aligned}$$

Therefore, there is some $v \in [m - 17k, m - 16k - 1]$ such that $v \in R(0)$ and $v - 1 \in R^-(m)$. By definition of $R(0)$ and $R(m)$ we obtain $f(I) < 0$. □

As mentioned before, Claims 5.2 and 5.3 imply the theorem. □

6. Remarks

In this section we discuss some further directions one might take with our approach to Conjecture 1.1.

- The constant 10^{46} in Theorem 1.5 can certainly be improved by being more careful in the proof.

  One natural modification which would substantially improve the bound is to consider repeated sumsets in the proof of Lemma 5.3. By this we mean applying versions of Lemmas 5.9 and 5.10 repeatedly in order to obtain better estimates on the number of elemets in the set $R(0)$. Although such a modification would give a substantial improvement to our bounds, the resulting bound in Theorem 1.2 would still be far from the conjectured bound of 4. Therefore we felt that it was better to present the simpler proof of Lemma 5.3 although it gives a worse bound.

  The main question is whether a better choice of hypergraphs $\mathcal{H}_{n,k}$ can lead to a proof to Conjecture 1.1. It is not clear what kind of hypergraphs one should look for. Although in the above theorem, the hypergraphs $\mathcal{H}_{n,k}$ are quite sparse, this does not seem to be crucial in the proof.

- The constant "10^{46}" cannot be reduced to "4" in Theorem 1.5 without changing the graphs $\mathcal{H}_{n,k}$. Indeed for large k, the graphs $\mathcal{H}_{5(k-1),k}$ do not have the MMS-property. To see this, consider the following function $f : V(G) \to \mathbb{R}$.

$$\begin{aligned} f(i) &= k-2 \text{ if } i \equiv 0 \pmod{k-1}, \\ f(i) &= -1 \text{ if } i \not\equiv 0 \pmod{k-1}. \end{aligned}$$

  It is easy to see that we have $\sum_{x\in V(G)} f(x) = 0$. For two vertices i and j let

$$p(i,j) = \begin{cases} \text{The number of edges of } \mathcal{H}_{5(k-1),k} \text{ containing } i \text{ and } j & \text{if } i \neq j \\ 0 & \text{if } i = j. \end{cases}$$

  The graph $\mathcal{H}_{5(k-1),k}$ has five nonnegative vertices $0, k-1, 2(k-1), 3(k-1), 4(k-1)$. An edge $e \in \mathcal{H}_{5(k-1),k}$ is nonnegative if and only if e contains at least two of these vertices. Therefore the number of nonnegative edges in $\mathcal{H}_{5(k-1),k}$ is at most

$$\frac{1}{2} \sum_{\substack{i,j\in\{0,k-1,2(k-1),\\ 3(k-1),4(k-1)\}}} p(i,j) = 5p(0,k-1) + 5p(0,2(k-1)). \tag{6.1}$$

  Notice that an edge $e(-v,i,j)$ contains both 0 and $k-1$ if and only if we have

$$i \geq v+1, \tag{6.2}$$
$$j \geq v, \tag{6.3}$$
$$i+j \geq v+k-1. \tag{6.4}$$

It's easy to check that the number of triples (v, i, j) which satisfy (6.2) – (6.4) is less than $\frac{1}{6}k^3 + o(k^3)$, which implies that $p(0, k-1) = \frac{1}{6}k^3 + o(k^3)$.

The only edges $\mathcal{H}_{5(k-1),k}$ which contain 0 and $2(k-1)$ are of the form $e(0, i, k-1)$ for some i, so we have that $p(0, 2(k-1)) = k-1$. Therefore, there are less than $\frac{5}{6}k^3 + o(k^3)$ nonnegative edges in $\mathcal{H}_{5(k-1),k}$ which is smaller than $k(k-1)^2$ for large enough k.

The above argument shows that the constant "10^{46}" in Theorem 1.5 cannot be reduced to less than 5. This shows that Conjecture 1.1 cannot be proved by the argument we used in this paper without changing the graphs $\mathcal{H}_{n,k}$ to some other construction.

- It would be interesting to apply the methods introduced in this paper to problems other than Conjecture 1.1. One of the key ideas we used in our proof was averaging over graphs other than tight cycles. A similar idea was used by Frankl and Katona [13] to give an alternative proof of the t-intersecting version of the Erdős-Ko-Rado Theorem for certain values of k and t. In this application, Frankl and Katona averaged over Steiner systems rather than tight cycles.

  There are some problems related to Conjecture 1.1 to which our methods may be applicable. One is a conjecture due to Samuels about the probability that $X_1 + \cdots + X_\ell < 1$ where $X_1, \ldots, X_\ell$ are random variables with specified means (see [1, 20] for more details). In the past similar methods have been used for attacking Samuel's Conjecture and Conjecture 1.1, so perhaps the method we introduced in this paper could be used to prove some new cases of Samuel's Conjecture.

- We conclude with the following general problem.

  Problem 6.1. *Which hypergraphs have the MMS-property?*

  This problem is probably quite hard, since a solution to it would mean a generalization of Conjecture 1.1. This problem has been studied by Huang and Sudakov in [11]. They showed that all sufficiently large hypergraphs with equal codegrees have the MMS-property (a hypergraph has equal codegrees if any pair of distinct vertices is contained in the same number of edges).

Acknowledgment

The author would like to thank Peter Allen, Ameera Chowdhury, Jan van den Heuvel, Jozef Skokan, and Benny Sudakov for their advice and discussions.

This research was supported by the LSE postgraduate research studentship scheme, and the Methods for Discrete Structures, Berlin graduate school.

References

[1] N. Alon, P. Frankl, H. Huang, V. Rödl, A. Rucinski, and B. Sudakov. Large matchings in uniform hypergraphs and the conjectures of Erdős and Samuels. *J. Combin. Theory Ser. A*, 119:1200–1215, 2012.

[2] N. Alon, H. Huang, and B. Sudakov. Nonnegative k-sums, fractional covers, and probability of small deviations. *J. Combinatorial Theory Ser. B*, 102:784–796, 2012.

[3] T. Bier. A distribution invariant for association schemes and strongly regular graphs. *Linear Algebra Appl.*, 57:105–113, 1984.

[4] V. Blinovsky. Minimal number of edges in hypergraph guaranteeing perfect fractional matching and MMS conjecture. *arXiv:1310.0989*, 2013.

[5] B. Bollobás. *Combinatorics*. Cambridge University Press, 1986.

[6] A. Chowdhury. A note on the Manickam-Miklós-Singhi conjecture. *European J. Combin.*, 35:131–140, 2014.

[7] A. Chowdhury, G. Sarkis, and S. Shahriari. The Manickam-Miklos-Singhi conjectures for sets and vector spaces. *J. Combin Theory Ser. A*, 128:84–103, 2014.

[8] P. Frankl. On the number of nonnegative sums. *J. Combinatorial Theory Ser. B*, 103:647–649, 2013.

[9] G. C. G. Marino. A method to count the positive 3-subsets in a set of real numbers with non-negative sum. *European J. Combin.*, 23:619–629, 2002.

[10] D. J. Grynkiewicz. On extending Pollard's theorem for t-representable sums. *Israel J. Math*, 177(1):413–439, 2010.

[11] H. Huang and B. Sudakov. The minimum number of nonnegative edges in hypergraphs. *Electron. J. Combin.*, 21, 2014.

[12] G. O. H. Katona. A simple proof of the Erdős-Chao Ko-Rado Theorem. *J. Combin. Theory Ser. B*, 13:183–184, 1972.

[13] G. O. H. Katona. The cycle method and its limits. In *Numbers, information and complexity*, pages 129–141. Kluwer Acad. Publ., 2000.

[14] M. Kneser. Abschätzung der asymptotischen Dichte von Summenmengen. *Math. Z*, 58:459–484, 1953.

[15] V. F. Lev. Restricted set addition in groups, ii. a generalization of the Erdos-Heilbronn conjecture. *Electron. J. Combin.*, 7, 2000.

[16] N. Manickam. *On the distribution invariants of association schemes*. PhD thesis, Ohio State University, 1986.

[17] N. Manickam and D. Miklós. On the number of non-negative partial sums of a non-negative sum. *Colloq. Math. Soc. János Bolyai*, 52:385–392, 1987.

[18] N. Manickam and N. Singhi. First distribution invariants and EKR theorems. *J. Combin. Theory Ser. A*, 48:91–103, 1988.

[19] J. M. Pollard. A generalisation of the Theorem of Cauchy and Davenport. *J. London Math. Soc.*, 8(2):460–462, 1974.

[20] S. Samuels. On a Chebyshev-type inequality for sums of independent random variables. *Ann. Math. Statist.*, 37:248–259, 1966.

[21] M. Tyomkyn. An improved bound for the Manickam-Miklós-Singhi conjecture. *European J. Combin.*, 33(1):27–32, 2012.