

# Small Even Covers, Locally Decodable Codes and Restricted Subgraphs of Edge-Colored Kikuchi Graphs

Jun-Ting Hsieh\*   Pravesh K. Kothari †   Sidhanth Mohanty‡   David Munhá Correia§  
Benny Sudakov§

## Abstract

Given a  $k$ -uniform hypergraph  $H$  on  $n$  vertices, an even cover in  $H$  is a collection of hyperedges that touch each vertex an even number of times. Even covers are a generalization of cycles in graphs and are equivalent to linearly dependent subsets of a system of linear equations modulo 2. As a result, they arise naturally in the context of well-studied questions in coding theory and refuting unsatisfiable  $k$ -SAT formulas. Analogous to the irregular Moore bound of Alon, Hoory and Linial [3], Feige conjectured [8] an extremal trade-off between the number of hyperedges and the length of the smallest even cover in a  $k$ -uniform hypergraph. This conjecture was recently settled up to a multiplicative logarithmic factor in the number of hyperedges [12, 13]. These works introduce the new technique that relates hypergraph even covers to cycles in the associated *Kikuchi* graphs. Their analysis of these Kikuchi graphs, especially for odd  $k$ , is rather involved and relies on matrix concentration inequalities.

In this work, we give a simple and purely combinatorial argument which recovers the best known bound for Feige’s conjecture for even  $k$ . We also introduce a novel variant of a Kikuchi graph which together with this argument improves the logarithmic factor in the best known bounds for odd  $k$ . As an application of our ideas, we also give a purely combinatorial proof of the improved lower bounds [4] on 3-query binary linear locally decodable codes.

## 1 Introduction

A set  $S$  of edges in a hypergraph  $\mathcal{H}$  is an *even cover* if

$$\bigoplus_{E \in S} E := \{v \in v(\mathcal{H}) : v \text{ belongs to an odd number of edges in } S\} = \emptyset.$$

Equivalently,  $S$  is an even cover if  $\sum_{E \in S} v_E = \mathbf{0}$  over  $\mathbb{F}_2$ , where  $v_E$  denotes the characteristic vector of  $E$ . In this work, we are interested in understanding the extremal trade-offs between the *size* of a  $k$ -uniform hypergraph  $H$  and *girth*, i.e., the number of edges in the shortest even cover in it.

**Even Covers and Linear Dependencies** Even covers in  $k$ -uniform hypergraphs correspond to linearly dependent subsets of a system of  $k$ -sparse (i.e., each equation has exactly  $k$  non-zero coefficients) linear equations over  $\mathbb{F}_2$ . To see why, let us associate a variable  $x_v$  for each  $v \in H$  and the  $|E|$ -sparse equation  $\sum_{v \in E} x_v = b_E$  for  $b_E \in \mathbb{F}_2$  with each edge  $E$  in  $H$ . Then, observe that for any even covers  $S$ , the left hand sides of the equations corresponding to  $E \in S$  add up to 0 and are thus linearly dependent. Thus, size vs girth trade-offs for  $k$ -uniform hypergraphs correspond to the largest possible size  $\ell$  of the minimum linear dependency in a system of linear equations with  $m$  equations in  $n$  variables.

---

\*juntingh@cs.cmu.edu, Carnegie Mellon University

†kothari@cs.princeton.edu, Princeton University. Supported by NSF CAREER Award #2047933, NSF #2211971, an Alfred P. Sloan Fellowship, and a Google Research Scholar Award.

‡sidhanth@csail.mit.edu, MIT

§Department of Mathematics, ETH, Zürich, Switzerland. Research supported in part by SNSF grant 200021\_196965.

Emails: {david.munhacanascorreia, benjamin.sudakov}@math.ethz.ch.

Such  $k$ -sparse linear equations arise naturally as the parity check equations for *low-density parity check* error correcting codes. The size vs girth trade-offs for  $k$ -uniform hypergraphs thus correspond to rate vs distance trade-offs for such codes.

By the equivalence between even covers and linear dependencies, it is clear that every hypergraph with  $m \geq n + 1$  hyperedges must have an even cover of length at most  $n + 1$  and this is clearly tight. The natural question is then: How does the trade-off between  $m$  and the maximum possible girth look as  $m$  increases beyond  $n + 1$ ?

**Size vs Girth Trade-offs** When  $H$  is a 2-uniform hypergraph, an even cover is simply an even subgraph, i.e., a subgraph with all vertices of even degree. Such a subgraph is a union of edge-disjoint cycles in  $H$ . An even cover with smallest number of edges must in fact be a simple cycle and thus, its length must be the girth of the graph  $H$ . The extremal trade-off between the size of  $H$  (i.e., the number of edges in the graph) and its girth was conjectured by Bollobás and confirmed by Alon, Hoory and Linial [3] who proved the *irregular Moore bound* — every graph on  $n$  vertices with average degree  $d$  has girth at most  $2\lceil \log_{d-1}(n) \rceil$ . It is an outstanding open problem whether the constant 2 in this bound can be further improved (for the best known constant, see [17]).

For  $k$ -uniform hypergraphs with  $k > 2$ , the size vs girth trade-offs were first studied by Naor and Verstraete [20] through applications to rate vs distance trade-offs for LDPC codes discussed above. They showed that every  $H$  with  $m \geq n^{k/2} \log^{O(1)}(n)$  hyperedges on  $n$  vertices must contain an even cover of length  $O(\log n)$ . The  $\log^{O(1)}(n)$  factor was further improved to a  $O(\log \log n)$ -factor in a subsequent work of Feige [8]. For  $k = 2$ , this recovers a coarse version of the irregular Moore bound. For  $k > 2$ , however, there is an interesting regime between the two extreme thresholds of  $m = n + 1$  (with maximum possible girth of  $n + 1$ ) and  $m \sim n^{k/2} \log^{O(1)}(n)$  (with maximum possible girth of  $O(\log n)$ ).

**Feige’s Conjecture** In 2008, Feige [8] formulated a conjecture about this in-between regime that suggests a smooth interpolation between the two extremes noted above.

**Conjecture 1.1.** *Fix any  $k \in \mathbb{N}$ . Then, there exists a sufficiently large  $C > 0$  such that for sufficiently large  $n \in \mathbb{N}$  and every  $\ell \in \mathbb{N}$ , every  $k$ -uniform hypergraph  $H$  with  $m \geq Cn(\frac{n}{\ell})^{k/2-1}$  has an even cover of length at most  $O(\ell \log_2 n)$ .*

The quantitative behavior above can be verified for random hypergraphs (up to a multiplicative factor of  $\log(n)$  in  $m$ ). Indeed, Feige’s conjecture was based on the hypothesis that random hypergraphs are approximately extremal for the purpose of avoiding short even covers. The motivation for this conjecture comes from the question of showing existence of (and/or efficiently finding) polynomial size *refutation witnesses* — easily verifiable witnesses of unsatisfiability of — randomly chosen  $k$ -SAT formulas parameterized by the number of clauses. Feige’s conjecture implies that the result of Feige, Kim and Ofek [9] that showed that random 3-SAT formulas with  $m \geq O(n^{1.4})$  clauses admit a polynomial size refutation witness with high probability will extend to the significantly more general setting of *smoothed* 3-SAT formulas also studied by Feige in addition to simplifying the construction and arguments based on the second moment method in [9].

Until recently, not much was known about Feige’s conjecture except for the work of Alon and Feige [2] that showed a suboptimal version for the case of  $k = 3$  and that of Feige and Wagner [10] that built an approach to the problem of even covers by viewing them as an instance of generalized girth problems about hypergraphs. In 2022, Guruswami, Kothari and Manohar [12] proved Feige’s conjecture up to an additional loss of  $\log^{2k}(n)$  multiplicative factor in  $m$  via a spectral argument applied to the *Kikuchi* graph, a graph with an appropriate algebraic structure, built from the given hypergraph. Their argument was simplified and tightened to reduce the loss down to a  $O(\log n)$  multiplicative factor in  $m$  by Hsieh, Kothari and Mohanty in [13]. In this work, as we will soon discuss, we give a substantially simpler, purely combinatorial argument that recovers their result and improves the logarithmic factors for hypergraphs of odd uniformity.

**Linear Locally Decodable Codes** A binary error correcting code is a map  $C : \{0, 1\}^m \rightarrow \{0, 1\}^n$  where we view the input as a  $m$ -bit “message” and the output as an  $n$ -bit codeword. By a slight abuse of notation, we use  $C$  to also denote the set of all *codewords*:  $\{y \mid \exists x \in \{0, 1\}^m, C(x) = y\}$ . We say that  $C$  is *linear* if, when viewing the input and output as  $\mathbb{F}_2^m$  and  $\mathbb{F}_2^n$ , respectively,  $C$  is a  $\mathbb{F}_2$ -linear map. A code  $C$  is called  $(q, \delta)$ -locally decodable (LDC) if, in addition, it admits a *local decoding* algorithm. Such a local decoding algorithm takes as input any target message bit  $i$  for  $1 \leq i \leq m$  and a corrupted codeword  $y$  such that  $\text{dist}(y, C(x)) \leq \delta$  for some  $x \in \{0, 1\}^m$  where  $\text{dist}$  counts the fraction of coordinates that  $y$  and  $C(x)$  differ. The goal of the algorithm is to access at most  $q$  locations in  $y$  and output  $x_i$  correctly with high probability over the choice of the  $q$  locations. In other words, the local decoder can decode any bit of the message by reading at most  $q$  locations of the received corrupted codeword.

Locally decodable codes are intensely investigated in computer science (see the survey [24] for background and applications) with applications to probabilistically checkable proofs, private information retrieval [23], and worst-case to average-case reductions in computational complexity theory. They also have deep connections with additive combinatorics and incidence geometry [5]. We are typically concerned with codes that are locally decodable with very few queries, such as  $q = 2$  or  $3$ , and the fundamental question is the smallest possible  $n = n(m)$  such that there is a  $(q, \delta)$ -binary LDC  $C : \{0, 1\}^m \rightarrow \{0, 1\}^n$ . Classical results have essentially completely resolved the case of  $q = 2$  and we know that a blocklength of  $n \leq 2^{O(m)}$  (for a constant  $\delta$ ) can be achieved by Hadamard codes with a matching lower bound [11, 15]. The case of  $q = 3$  already presents with wide gaps where until recently, the best known lower bound [11, 15] was  $n \geq \tilde{O}(m^2)$  while the best known construction [7, 22] gives a 3-query binary linear code with  $n \leq \exp(\exp(O(\sqrt{\log m \log \log m})))$ . Recently, using spectral refutations via Kikuchi matrices, Alrabiah et. al. [4] improved the quadratic bound above to obtain a lower bound of  $n \geq m^3/\text{poly log } m$  for 3-query binary, locally decodable codes.

## 1.1 Our Results

The arguments in the most recent works on Feige’s conjecture and locally decodable codes are involved and in particular, require the use of matrix concentration inequalities. Our main contribution is a short, purely self-contained combinatorial argument that recovers their result. This technique might be useful for other similar questions and we illustrate it below, by an application to another well studied problem. Our arguments for Feige’s conjecture also utilize Kikuchi graphs introduced in the above prior works. We then introduce a new variant of a Kikuchi graph that, when combined with our combinatorial argument, improves on their results for the case of odd  $k$ .

**Theorem 1.2.** *For all  $k$ , there is a sufficiently large  $C$  such that the following holds for all sufficiently large  $n$ .*

- (i) *If  $k$  is even, then every  $k$ -uniform  $n$ -vertex hypergraph with at least  $Cn(n/l)^{k/2-1} \cdot \log n$  edges contains an even cover of size  $O(l \log n)$ .*
- (ii) *If  $k$  is odd and  $l \leq n/\log^2 n$ , then every  $k$ -uniform  $n$ -vertex hypergraph with at least  $Cn(n/l)^{k/2-1} \cdot (\log n)^{\frac{1}{k+1}}$  edges contains an even cover of size  $O(l \log n)$ .*

We remark that in the above theorem for odd  $k$ , we require only a very mild restriction on  $l$  that  $l \leq n/\log^2 n$ , which is purely a technical artefact of our proof. In fact, if  $l \geq n/100 \log n$ , Feige’s conjecture holds trivially from a standard linear algebra argument (see Lemma 2.4). For  $n/\log^2 n \leq l \leq n/100 \log n$ , then one can use the some arguments as in the proof of the above theorem (namely Section 4.4) to show the same bound as in part (i) for this range.

The ideas we developed to prove the above theorem naturally extend to the setting of *linear* binary locally decodable codes. We can use an altered version of these arguments to show the following result previously obtained by Alrabiah et. al [4]. As in the above results, their proof involves spectral arguments

on signed adjacency matrices of Kikuchi graphs based on matrix concentration inequalities. The proof we give is simple and purely combinatorial. Although the work in [4] deals also with non-linear codes, we remark that all known constructions of locally decodable codes including the ones discussed in the previous section are linear.

**Theorem 1.3.** *Let  $C : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$  be a linear map that gives a 3-query locally decodable code with distance  $\delta > 0$ . Then,  $m \leq Kn^{1/3} \log n$  for  $K = 10^7/\delta^2$ .*

**Finding structures in edge-colored graphs** We finally remark that the results in this paper, which we discussed above, are proven by reducing both our results to the one of finding a subgraph, satisfying certain restrictions, in an edge-coloured graph. Recall that an edge-colored graph is a graph along with a coloring of the edges so that no vertex has two or more edges of same color incident on it. Finding structures in edge-colored subgraphs is a powerful method in Combinatorics and has been used to study several well-known questions. For example, the famous conjecture of Ringel from 1963 says that the edges of the complete graph  $K_{2n+1}$  can be decomposed into copies of any tree on  $n$  vertices. It was observed by Kőtzig that this problem can be reduced to showing that a certain edge-colouring of  $K_{2n+1}$  contains a rainbow copy of any tree on  $n$  vertices. The existence of such rainbow trees was recently established in [19].

Another famous application is the resolution of the Ryser-Brualdi-Stein conjecture which was open for more than 60 years until it was recently solved by Montgomery [18]. It states that every Latin square  $n \times n$  contains a transversal (i.e., a collection of cells which do not share the same row, column or symbol) of size  $n - 1$ . This can be reduced to finding a rainbow matching missing only one colour in any proper edge-colouring of the complete bipartite graph  $K_{n,n}$ . Finally, a well-studied problem in additive combinatorics studies the additive dimension of sets with small doubling, in any group. This problem had been solved for abelian groups by Sanders [21]. For general groups, this question was reduced by Alon et. al [1] to a well-known problem (see [14]) of finding a rainbow cycle in a properly edge-coloured graph with sufficiently many edges, for which Alon et. al provide an almost tight bound.

## 2 Preliminaries

### 2.1 Notation and definitions

**Definition 2.1.** All logarithms are taken base 2. Given a graph  $G$ , we let  $d(G)$  denote its average degree. Let  $\mathcal{H}$  be a hypergraph. We let  $V(\mathcal{H})$  denote its vertex set and  $E(\mathcal{H})$  its edge-set. Similarly, we let  $v(\mathcal{H})$  denote the size of its vertex set and  $e(\mathcal{H})$  denote the number of edges. Given a hypergraph  $\mathcal{H}$ , a *bucket* is a pair  $(\mathcal{E}, X)$  where  $\mathcal{E}$  is a set of edges in  $\mathcal{H}$  which all contain the set of vertices  $X \subseteq V(\mathcal{H})$ . Given a hypergraph  $\mathcal{H}$ , an  $(m, t)$ -*bucket decomposition* is a partition of the edges of  $\mathcal{H}$  into buckets  $(\mathcal{E}_1, X_1), (\mathcal{E}_2, X_2), \dots$  such that  $|\mathcal{E}_i| = m$  and  $|X_i| = t$ .

### 2.2 Standard tools

In this section, we collect some lemmas which will be useful throughout the paper. The first is the standard tool used for finding subhypergraphs with large minimum degree.

**Lemma 2.2.** *Let  $\mathcal{H}$  be a hypergraph on  $n$  vertices and  $nd$  edges. Then, there exists an induced sub-hypergraph  $\mathcal{H}' \subseteq \mathcal{H}$  with at least  $nd/2$  edges and minimum degree at least  $d/2$ .*

**Proof.** We perform the following process. Start with  $\mathcal{H}' := \mathcal{H}$  and while it has a vertex  $v$  with degree less than  $d/2$  in  $\mathcal{H}'$ , remove it from  $\mathcal{H}'$ . Notice that at any point,  $\mathcal{H}'$  has at least  $e(\mathcal{H}) - nd/2 \geq nd/2$  edges and thus, the process must stop. The final  $\mathcal{H}'$  is then a sub-hypergraph with at least  $nd/2$  edges and minimum degree at least  $d/2$ .  $\square$

The next lemma is simply an observation on finding trivial bucket decompositions in hypergraphs.

**Lemma 2.3.** *Let  $\mathcal{H}$  be a hypergraph on  $n$  vertices and  $nd$  edges. Then, there exists an induced sub-hypergraph  $\mathcal{H}' \subseteq \mathcal{H}$  with at least  $nd/2$  edges and a  $(d/2, 1)$ -bucket decomposition.*

**Proof.** We do the following process. First, set  $\mathcal{H}' := \emptyset$ . While  $e(\mathcal{H}') < nd/2$ , consider the hypergraph  $\mathcal{H} \setminus \mathcal{H}'$ ; it has at least  $nd/2$  edges and thus there is a vertex  $v$  with degree at least  $d/2$ ; then, take a collection  $\mathcal{E} \subseteq \mathcal{H} \setminus \mathcal{H}'$  of  $d/2$  edges containing  $v$  and add them to  $\mathcal{H}'$ . Note that at the end of the process the hypergraph  $\mathcal{H}'$  is as desired since the collections  $\mathcal{E}$  taken at each step give the  $(d/2, 1)$ -bucket decomposition.  $\square$

The next lemma is a simple application of linear algebra in order to show that hypergraphs with sufficiently many edges contain even covers (with no size restriction).

**Lemma 2.4.** *Let  $\mathcal{H}$  be an  $n$ -vertex hypergraph with at least  $n + 1$  edges. Then,  $\mathcal{H}$  contains an even cover.*

**Proof.** For each edge  $E$  in the hypergraph consider its characteristic vector  $v_E$  in  $\mathbb{Z}_2^n$ . Note that since we have at least  $n + 1$  vectors, they must be linearly dependent, implying thus an even cover as desired.  $\square$

The next lemma concerns finding coloured cycles in edge-coloured graphs. Although a variant of it already appeared in [14], we include a the proof for sake of completeness.

**Lemma 2.5.** *Let  $G$  be an  $n$ -vertex graph and  $C$  a set of colours so that each edge in  $G$  is assigned a set of  $s$  colours in  $C$ . Suppose that for every vertex  $v \in G$  and colour  $c \in C$ , the number of edges incident on  $v$  whose assigned set of colours contains  $c$  is at most  $d(G)/20s \log n$ . Then,  $G$  contains a closed walk, of size at most  $2 \log n$ , such that some colour appears exactly once.*

**Proof.** Let  $l := \log n$  and  $r := d(G)/20s \log n$ . For sake of contradiction, suppose  $G$  contains no such closed walk of size  $2l$ . We will double-count the number of rainbow paths in  $G$  of size  $l$  - by *rainbow* we mean that no colour is assigned to more than one edge of the path. In order to give a lower bound on the number of rainbow paths, note first that Lemma 2.2 implies that  $G$  contains a subgraph  $G' \subseteq G$  with minimum degree at least  $d' = d(G)/4$ . We can then take a vertex  $v \in G'$  and greedily count the number of rainbow paths in  $G'$  of the form  $vv_2v_3 \dots v_{l+1}$ . Indeed, we have at least  $d'$  options for  $v_2$ ; then, given the assumption on  $G$  that every colour is incident to a vertex in at most  $r$  edges and that  $G$  has no rainbow cycle of size at most  $2l$ , we have at least  $d' - sr$  options for  $v_3$ ; with the same reasoning we have at least  $d' - 2sr$  options for  $v_4$  and so on. Concluding the number of such rainbow paths is at least  $d' \cdot (d' - sr) \cdot \dots \cdot (d' - sr(l - 1)) > (0.8d')^l$ , since  $d' - sr(l - 1) > d' - d(G)/20 \geq 0.8d'$ .

On the other hand, since there is no closed walk as described above, we can upper bound the number of rainbow paths as follows. For each pair of vertices  $x, y$  (there are at most  $n^2$  of them), observe that if there is a rainbow path between them using a set of colours  $S \subset C$ , then any other rainbow path between  $x$  and  $y$  must use the same set of colours  $S$ . Now, the set  $S$  has size  $sl$  and thus, we can upper bound the number of rainbow paths  $xz_1z_2 \dots z_ly$  of size  $l$  between  $x$  and  $y$  as follows. Choosing  $z_1$  can be done by choosing a colour  $c_1 \in S$  and then choosing an edge incident on  $v$  which contains the colour  $c_1$ . By assumption, there are at most  $slr$  such choices and the choice of the rest of the vertices  $z_2, \dots, z_{l-1}$  can be done in the same way. Concluding, there are at most  $(slr)^l$  such paths. Therefore, there are at most  $n^2 \cdot (slr)^l = (4slr)^l = (0.8d')^l$  rainbow paths of size  $l$ , which is a contradiction given the previous paragraph.  $\square$

### 2.3 Combinatorial Characterization of Locally Decodable Codes

In this section, we recall standard results that reduce proving a lower bound on the blocklength of a linear locally decodable code to establishing the existence of a special kind of even cover in a properly edge colored hypergraph.

We will use the following standard reduction to a locally decodable code in the *normal* form. We direct the reader to the monograph [24] for a more general result that holds even for non-linear codes. We use  $\oplus$  to denote addition modulo 2 (i.e., over  $\mathbb{F}_2^n$ ) in the following.

**Proposition 2.6** (LDCs in normal form, see Theorem 8.1 in [6]). *Let  $C : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$  be a binary, linear 3-query locally decodable code with distance  $\delta > 0$ . Then, there is a collection of 3-uniform hypergraph matchings  $H_1, H_2, \dots, H_m$  of size  $\geq \delta n/6$  one for each message bit, such that for every  $x \in \mathbb{F}_2^m$ , for every  $i \in [m]$  and every hyperedge  $e \in H_i$ ,  $\bigoplus_{j \in e} C(x)_j = x_i$ .*

As an immediate corollary of the normal form of a locally decodable code, we obtain the following useful combinatorial characterization:

**Lemma 2.7.** *Let  $C : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$  be a binary, linear 3-query locally decodable code with distance  $\delta > 0$  and let  $H_1, H_2, \dots, H_m$  be the associated matchings as in Proposition 2.6 of size  $\geq \delta n$ . Then, for any even cover  $\{e_1, e_2, \dots, e_t\}$  in the multi-hypergraph  $\cup_{i \leq [m]} H_i$ , each  $H_i$  contributes an even number of hyperedges.*

**Proof.** Suppose not and say  $\{e_1, e_2, \dots, e_t\}$  is an even cover that contains an odd number of hyperedges from, say,  $H_1$ . Then, choose  $x = (1, 0, \dots, 0) \in \mathbb{F}_2^m$  and observe that on the one hand,  $\bigoplus_{i \leq t} \bigoplus_{j \in e_i} C(x)_j = 0$  because  $C_i$ s form an even cover. But, on the other hand,  $\bigoplus_{i \leq t} \bigoplus_{j \in e_i} C(x)_j = C(x)_1 = 1$  since  $H_1$  contributes an odd number of edges to  $\{e_1, e_2, \dots, e_t\}$  and for every edge  $e \in H_1$ ,  $\bigoplus_{j \in e} C(x)_j = x_1$ . This is a contradiction.  $\square$

By thinking of each edge in  $H_i$  as having the color  $i$ , the hypergraph  $\cup_{i \leq m} H_i$  is a properly edge colored hypergraph. Thus, in light of Lemma 2.7, Theorem 1.3 is implied by the following theorem that we will establish in this work.

**Theorem 2.8.** *For all sufficiently small  $\alpha > 0$ , there is  $K := 10^7/\alpha^2 > 0$  such that the following holds for all sufficiently large  $n$ . Let  $\mathcal{H}$  be a 3-uniform  $n$ -vertex hypergraph which is properly edge-coloured with  $Kn^{1/3} \log n$  colours so that each colour has at least  $\alpha n$  edges. Then,  $\mathcal{H}$  contains an even cover such that some colour appears exactly once.*

### 3 Even $k$ and some ideas of the proof

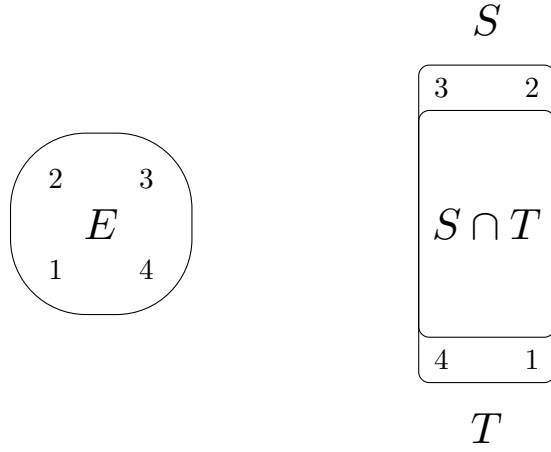
#### 3.1 Feige's problem for even $k$

To illustrate our main ideas, we start this section by giving a very short solution of Feige's problem up to a logarithmic factor for all even  $k$ . In the next section we will discuss how to use the methods inspired by this proof for the odd values of  $k$ .

**Proof of Theorem 1.2 (i).** Let  $\mathcal{H}$  be a hypergraph satisfying the assertion of the theorem. By choosing constant  $C = 10^k k^{k/2-1}$  we can assume that  $l \geq k$  and that  $\mathcal{H}$  has at least  $10^k n (n/l)^{k/2-1} \log n$  edges. We define an edge-coloured graph  $G$ , called the *Kikuchi graph*, as follows.

- The vertex set of  $G$  consists of all  $l$ -element subsets of the vertex of  $\mathcal{H}$ , i.e.,  $V(G) := \binom{[n]}{l}$ .
- We define an edge  $S \leftrightarrow T$  for two  $S, T \in V$  if there exists an edge  $E \in \mathcal{H}$  such that  $S \oplus T = E$  and  $|S \cap E| = |T \cap E| = k/2$ . Moreover, we colour the edge  $S \leftrightarrow T$  in  $G$  with colour  $E$ .

Let us make some important remarks about  $G$ . First, note that the colouring of its edges is *well-defined* and *proper*. Indeed, the first is the case since the colour of the edge  $S \leftrightarrow T$  is uniquely defined by  $S \oplus T$ . The latter also holds since given  $S \in V$  and  $E \in \mathcal{H}$ , there exists at most one  $T \in V$  such that  $S \oplus T = E$ . We now show the following claim.



**Figure 1:** An illustration of the Kikuchi graph defined above with an edge  $E = \{1, 2, 3, 4\}$  of  $\mathcal{H}$  and an edge  $S \leftrightarrow T$  of  $G$  coloured with  $E$ .

**Claim 3.1.**  $G$  contains a closed walk of length  $O(l \log n)$  such that some colour appears only once.

**Proof.** To prove this we apply Lemma 2.5. Note that  $G$  has  $N := \binom{n}{l}$  vertices and each edge  $E \in \mathcal{H}$  creates at least  $\frac{1}{2} \cdot \binom{k}{k/2} \cdot \binom{n-k}{l-k/2}$  edges of  $G$ , since we need to choose the intersections  $S \cap E, T \cap E$  (which are disjoint and of size  $k/2$ ) and then the set  $S \setminus E = T \setminus E$ . Therefore, the number of edges in  $G$  is  $e(G) \geq \frac{1}{2} \cdot e(\mathcal{H}) \binom{k}{k/2} \binom{n-k}{l-k/2}$ . Since  $N = \binom{n}{l}$  we have  $\log N \leq \log n^l = l \log n$  and

$$N = \binom{n}{l} = \binom{n-k}{l-k/2} \cdot \frac{(n-k+1) \cdots n}{(l-k/2+1) \cdots l \cdot (n-l) \cdots (n-l-k/2+1)} \leq \binom{n-k}{l-k/2} \cdot (4n/l)^{k/2},$$

where we used that  $l \leq n/\log n$  by Lemma 2.4 and that  $l \geq k$ . Therefore

$$\begin{aligned} e(G) &\geq \frac{1}{2} \cdot e(\mathcal{H}) \binom{k}{k/2} \binom{n-k}{l-k/2} \geq \frac{1}{2} \cdot 10^k n (n/l)^{k/2-1} \log n \cdot \binom{k}{k/2} \cdot \binom{n-k}{l-k/2} \\ &\geq 5^k (4n/l)^{k/2} \binom{n-k}{l-k/2} (l \log n) > 10N \log N. \end{aligned}$$

Therefore, since  $G$  is properly-coloured, Lemma 2.5 implies that it contains the desired type of closed walk of size at most  $2 \log N \leq 2l \log n$ .  $\square$

To finish, we can take the closed walk  $S_1, S_2, \dots, S_r$  in  $G$  given above and consider the collection  $\mathcal{C}$  of edges in  $\mathcal{H}$  which appear an odd number of times as colours in this walk - that is, recalling the definition of  $G$ , those edges  $E \in \mathcal{H}$  such that  $E = S_i \oplus S_{i+1}$  (the indices being taken modulo  $r$ ) for an odd number of  $1 \leq i \leq r$ . Since  $\bigoplus_{1 \leq i \leq r} (S_i \oplus S_{i+1}) = \emptyset$ , we have that  $\bigoplus_{E \in \mathcal{C}} E = \emptyset$  and so  $\mathcal{C}$  is an even cover of size at most  $r = O(l \log n)$ , as desired.  $\square$

### 3.2 Feige's problem for odd $k$ and LDCs

We now briefly discuss the methods for proving Theorem 1.2 in full generality. In the previous section, in order to solve Feige's problem for even  $k$ , we defined an edge-coloured graph  $G$  based on the hypergraph  $\mathcal{H}$ . Crucially, the graph  $G$  had the following properties.

1.  $G$  has many edges and is properly edge-coloured.
2. A closed walk in  $G$  in which some colour appears once implies an even cover in  $\mathcal{H}$ .

Then we just applied Lemma 2.5 to find a the desired even cover in  $\mathcal{H}$ .

### Even covers with odd $k$

In the case of odd  $k$  our general strategy will be similar. We will also want to define an edge-coloured graph  $G$  based on our hypergraph  $\mathcal{H}$  such that some appropriate versions of the properties above hold. For this, we will define two different variants of the Kikuchi graph defined in the previous section. The first version is given in Section 4.4 and has already appeared in previous works (see [12] and [13]) and can alone, be used to show a bound as in part (i) of Theorem 1.2 for odd values of  $k$ . To get the improved part (ii) of Theorem 1.2 for odd  $k$ , another variant is necessary. In Section 4.5, we introduce then the *flower Kikuchi graph*.

Before using these Kikuchi graphs however, we will need to *clean* our hypergraph  $\mathcal{H}$  using the tools in Section 4. Specifically, we will be able to reduce our problem to considering a hypergraph  $\mathcal{H}$  with a nice bucket decomposition and co-degree assumptions (see Lemma 4.5). Then, based on the outcome of this Lemma, we will either use the Kikuchi graph on  $\mathcal{H}$  defined in Section 4.4 or the new flower Kikuchi graph.

In the first case, the graph is edge-coloured, but each edge receives now a pair of colours  $(C, C')$ , where each colour  $C$  is an edge in  $\mathcal{H}$ . Due to the additional assumptions we now have on the hypergraph  $\mathcal{H}$  we can show that the Kikuchi graph contains a subgraph  $G' \subseteq G$  with the following properties similar to the ones before.

1.  $G'$  has many edges and is such that every vertex is incident to at most  $d(G')/80 \log |G'|$  edges containing a given colour.
2. A closed walk in  $G'$  in which some colour appears once implies an even cover in  $\mathcal{H}$ .

Then, like before, applying Lemma 2.5 implies the desired even cover in  $\mathcal{H}$ . As mentioned above, finding  $G'$  is only possible because of the additional assumptions on  $\mathcal{H}$ . In Proposition 4.9 we state the crucial hypergraph property which implies the existence of such a  $G'$ .

In the second case, the flower Kikuchi graph  $G$  is edge-coloured, with each edge receiving a colour  $C$  which is an edge of  $\mathcal{H}$ . Like in the previous case, we can also then perform a (simpler) cleaning procedure to find a subgraph subgraph  $G' \subseteq G$  with the same properties as those mentioned in the beginning of the section:  $G'$  has many edges and is properly coloured; a closed walk in  $G'$  in which some colour appears once implies an even cover in  $\mathcal{H}$ . Then, applying Lemma 2.5 implies the desired even cover in  $\mathcal{H}$ .

### Locally decodable codes

The problem of LDCs is very similar. As discussed in the introduction, we can reformulate this problem as finding a special even cover in an edge-coloured hypergraph (see Theorem 2.8). The only difference here is that the edges of the hypergraph are properly colored and we want to find an even cover where some color is unique. We again will first clean our hypergraph using Lemma 4.1. Then we will show that the given Kikuchi graph defined in Section 4.4 will satisfy the property in Proposition 4.9 and thus gives the desired even cover in  $\mathcal{H}$ .

## 4 Cleaning tools, the Kikuchi graph and even covers

In this section we present two tools for cleaning hypergraphs. As we mentioned in the previous section, the goal is to find hypergraphs with *nice* bucket decompositions and co-degree conditions. Here and later in the paper we will always assume that  $k$  is odd.

### 4.1 Cleaning I: A general tool for hypergraphs

First, we need the following simple observation for general hypergraphs.



**Lemma 4.1.** *Let  $\mathcal{H}$  be a  $k$ -uniform hypergraph and  $j \leq k$ . Then, for all  $t$  and  $e \leq e(\mathcal{H})$ , there is a  $\mathcal{H}' \subseteq \mathcal{H}$  such that one of the following holds.*

1.  $e(\mathcal{H}') \geq e(\mathcal{H}) - e$  and every set  $R$  of size  $t$  has  $\deg_{\mathcal{H}'}(R) < m$ .
2.  $\mathcal{H}'$  has a  $(m, t)$ -bucket decomposition and  $e(\mathcal{H}') \geq e$ .

**Proof.** Starting with  $\mathcal{H}_0 := \mathcal{H}$ , perform the following procedure. While  $\mathcal{H}_0$  has a set  $R$  of size  $t$  with  $\deg_{\mathcal{H}_0}(R) \geq m$ , then select  $m$  such edges in  $\mathcal{H}_0$  which contain  $R$  and remove them from  $\mathcal{H}_0$ . When this process stops, if  $e(\mathcal{H}_0) \geq e(\mathcal{H}) - e$ , then by taking  $\mathcal{H}' = \mathcal{H}_0$  we have the first case. Otherwise by taking  $\mathcal{H}' = \mathcal{H} - \mathcal{H}_0$  we have the second case.  $\square$

## 4.2 Cleaning II: Hypergraphs with high co-degrees have small even covers

Next, we will prove a particular statement concerning Feige's problem. This will later allow us to reduce the general problem to considering a hypergraph with some co-degree assumptions.

**Lemma 4.2.** *Let  $\mathcal{H}$  be an  $n$ -vertex  $k$ -uniform hypergraph which has a  $(2, \frac{k+1}{2})$ -bucket decomposition. Then, for sufficiently large constant  $C$ , if  $e(\mathcal{H}) \geq C \cdot n (n/l)^{\frac{k-3}{2}} \cdot \log n$ , it contains an even cover of size  $O(l \log n)$ .*

**Proof.** Let us denote the  $(2, \frac{k+1}{2})$ -bucket decomposition by  $(\mathcal{E}_1, X_1), (\mathcal{E}_2, X_2), \dots$  and denote  $\mathcal{E}_i$  as  $\{Y_i, Z_i\}$ . By the pigeonhole principle we can find some  $j \geq \frac{k+1}{2}$  such that considering only the buckets with  $|Y_i \cap Z_i| = j$ , will reduce the size of the hypergraph by at most a  $2/k$  factor. Now, we construct a new  $2(k-j)$ -uniform multihypergraph  $\mathcal{G}$  on the same vertex set as  $\mathcal{H}$  consisting of the edges  $Y_i \oplus Z_i$ . The key observation which is easy to check from this definition is the following.

**Observation 4.3.** *An even cover in  $\mathcal{G}$  implies an even cover of twice the same size in  $\mathcal{H}$ .*

The above implies that we need only to show that  $\mathcal{G}$  contains an even cover of size  $O(l \log n)$ . For this, note that since  $2(k-j)$  is even and  $e(\mathcal{G}) \geq e(\mathcal{H})/k \geq (C/k) \cdot n (n/l)^{\frac{k-3}{2}} \cdot \log n \geq (C/k) \cdot n (n/l)^{(k-j)-1} \cdot \log n$ . Thus we can directly imply Theorem 1.2 (i).  $\square$

The above discussion together with the proof of Lemma 4.1 (applied with  $j = \frac{k+1}{2}$ ) immediately imply the following corollary.

**Corollary 4.4.** *For sufficiently large constant  $C$ , every  $n$ -vertex  $k$ -uniform hypergraph  $\mathcal{H}$  without an even cover of size  $O(l \log n)$  has a subhypergraph  $\mathcal{H}' \subseteq \mathcal{H}$  such that the following hold.*

- $e(\mathcal{H}') \geq e(\mathcal{H}) - C \cdot n (n/l)^{\frac{k-3}{2}} \cdot \log n$ .
- Every set  $S$  of size at least  $\frac{k+1}{2}$  has  $\deg_{\mathcal{H}'}(S) \leq 1$ .

## 4.3 Cleaning III: A general tool for hypergraphs

In this section we present a tool for finding a subhypergraph with both a nice bucket decomposition and co-degree assumptions. Given the lemmas in the previous section, we will need only to apply it to  $k$ -uniform hypergraphs with odd  $k$  in which no two edges share more than  $k/2$  vertices.

**Lemma 4.5.** *Let  $\mathcal{H}$  be an  $n$ -vertex  $k$ -uniform hypergraph with  $nd$  edges and consider any function  $m : [k] \rightarrow \mathbb{R}_{>0}$  with  $m(1) = d/10k$ . Then, there exists a subhypergraph  $\mathcal{H}_0 \subseteq \mathcal{H}$  with at least  $nd/2k$  edges and a  $1 \leq t \leq k/2$  such that the following hold.*

- 1) For all  $S \subseteq V(\mathcal{H}_0)$  with  $|S| > t$ , we have that  $\deg_{\mathcal{H}_0}(S) < m(|S|)$ .

2) If  $t = 1$ , then  $\mathcal{H}_0$  has minimum degree at least  $m(1)$ . If  $t \geq 2$ , then  $\mathcal{H}_0$  has a  $(m(t), t)$ -bucket decomposition.

**Proof.** Set  $\mathcal{H}_2, \dots, \mathcal{H}_{k/2} := \emptyset$  and  $\mathcal{H}' := \mathcal{H}$  and repeat the following procedure as long as possible. Let  $r \geq 2$  be maximal such that there exists some  $S \subseteq V(\mathcal{H}')$  with  $|S| = r$  with  $\deg_{\mathcal{H}'}(S) \geq m(r)$ . Take exactly  $m(r)$  edges in  $\mathcal{H}'$  that contain  $S$ , remove them from  $\mathcal{H}'$  and add them to  $\mathcal{H}_r$ . The process stops when such an  $r$  does not exist.

Now, suppose first that this process goes until  $e(\mathcal{H}') < e(\mathcal{H})/2$ . Then, we can stop the process at the first step in which that occurs. Note then that by pigeonholing, one of the hypergraphs  $\mathcal{H}_r$  will have at least  $nd/2k$  edges. We then define  $\mathcal{H}_0 := \mathcal{H}_r$  and observe that it satisfies the desired conditions.

Otherwise, the process stops while  $e(\mathcal{H}') \geq e(\mathcal{H})/2 = nd/2$ . Note that since the process stopped, it must be that such an  $r$  no longer exists and thus,  $\deg_{\mathcal{H}'}(S) < m(|S|)$  for all sets  $S$  with  $|S| \geq 2$ . Now, by Lemma 2.2,  $\mathcal{H}'$  has a subgraph, which we take to be  $\mathcal{H}_0$ , with at least  $nd/4$  edges and minimum degree at least  $d/4$ . Clearly, the conditions now are satisfied for  $t = 1$ .  $\square$

#### 4.4 The Kikuchi graph and even covers

We will now introduce the variant of the *Kikuchi graph* for an edge-coloured hypergraph with a given  $(m, t)$ -bucket decomposition which we already mentioned in Section 3.2.

**Definition 4.6.** Let  $\mathcal{H}$  be an  $n$ -vertex edge-coloured  $k$ -uniform hypergraph with an  $(m, t)$ -bucket decomposition with buckets  $\{(\mathcal{E}_i, X_i)\}_{1 \leq i \leq p}$ . Given an  $l$ , we define the  $l$ -Kikuchi graph of  $\mathcal{H}$  and this bucket decomposition to be an edge-coloured graph  $G = (V, E)$  as follows. Let  $[n] \times [2]$  denote two disjoint copies of  $[n]$ , whose vertices are colored red and blue in order to distinguish between these sets. The vertex set  $V(G)$  consists of all subsets of  $[n] \times [2]$  of size  $l$ . Each such set  $S \in V$  is viewed as  $(S^{(1)}, S^{(2)})$ , where  $S^{(1)}, S^{(2)} \subseteq [n]$  are *red* and *blue* respectively. For each  $i \in [p]$  and each ordered pair  $(C, C')$  of edges  $C, C' \in \mathcal{E}_i$ , let  $\tilde{C}^{(1)}$  be  $\tilde{C} := C \setminus X_i$  colored red and  $\tilde{C}'^{(2)}$  be  $\tilde{C}' := C' \setminus X_i$  colored blue. We add an edge between  $S, T \in V$ , denoted  $S \xleftrightarrow{(C, C')} T$ , if  $S \oplus T = \tilde{C}^{(1)} \oplus \tilde{C}'^{(2)}$  and one of the following holds.

- $|\tilde{C}^{(1)} \cap S^{(1)}| = |\tilde{C}'^{(2)} \cap T^{(2)}| = \lceil \frac{k-t}{2} \rceil$  and  $|\tilde{C}'^{(2)} \cap S^{(2)}| = |\tilde{C}^{(1)} \cap T^{(1)}| = \lfloor \frac{k-t}{2} \rfloor$ .
- $|\tilde{C}^{(1)} \cap S^{(1)}| = |\tilde{C}'^{(2)} \cap T^{(2)}| = \lfloor \frac{k-t}{2} \rfloor$  and  $|\tilde{C}'^{(2)} \cap S^{(2)}| = |\tilde{C}^{(1)} \cap T^{(1)}| = \lceil \frac{k-t}{2} \rceil$ .

Further, the edge  $S \leftrightarrow T$  is said to be *associated* with the edges  $C, C'$  in  $\mathcal{H}$  and is colored by the pair  $(c, c')$  where  $c$  is the colour of  $C$  in  $\mathcal{H}$  and  $c'$  is the colour of  $C'$ .

Note that contrary to the Kikuchi graph introduced in Section 3.1, the construction here is more complicated. We split the vertex set  $V$  into a red copy of  $[n]$  and a blue copy of  $[n]$ . This is needed because the pairs  $C, C'$  in our definition of buckets may have intersection larger than  $t$ , meaning that  $|C \oplus C'| = |\tilde{C} \oplus \tilde{C}'| < 2(k-t)$ . Taking  $\tilde{C}, \tilde{C}'$  to be subsets of two different copies of  $[n]$  automatically makes  $\tilde{C}^{(1)}, \tilde{C}'^{(2)}$  disjoint, so that  $|S \oplus T| = |\tilde{C}^{(1)} \oplus \tilde{C}'^{(2)}| = 2(k-t)$ . Note also that we allow a vertex of the Kikuchi graph  $S \subseteq [n] \times [2]$  to contain two copies of some element in  $[n]$  with different colors.

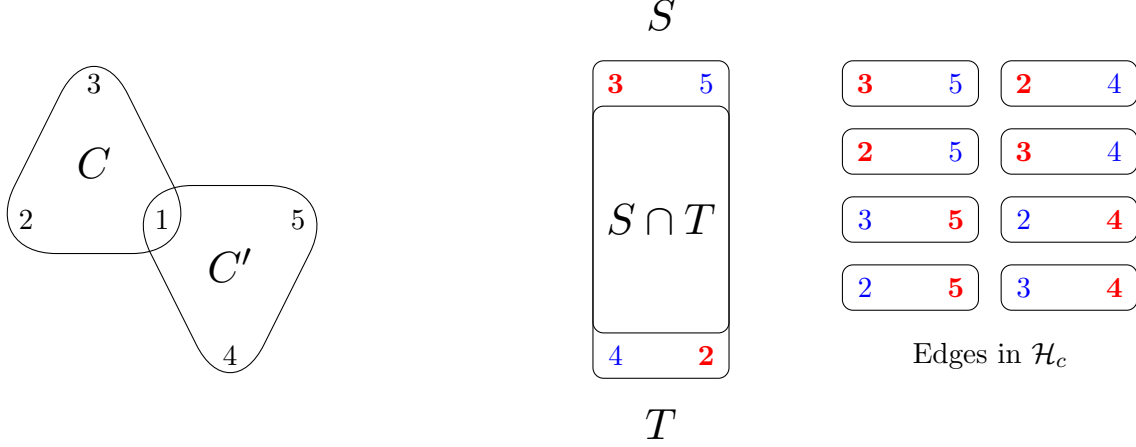
Given the above Kikuchi graph, we also define associated hypergraphs  $\mathcal{H}_c$  which will be crucial for our proof. Indeed, in Proposition 4.9 we will show that if these hypergraphs have edges that are rather uniformly distributed, that is, that there is no set of vertices with too large co-degree, then  $\mathcal{H}$  must contain an even cover of small size.

**Definition 4.7.** For each colour  $c$  of  $\mathcal{H}$ , we define the  $(k-t)$ -uniform multihypergraph  $\mathcal{H}_c$  with vertex set  $[n] \times [2]$  as follows: for each  $c$ -coloured edge  $C$  in  $\mathcal{H}$  belonging to some bucket  $\mathcal{E}_i$  and an edge  $C' \neq C$  also in  $\mathcal{E}_i$ , we put in  $\mathcal{H}_c$  all edges  $F$  such that one of the following holds.

- $\{|\tilde{C}^{(1)} \cap F^{(1)}|, |\tilde{C}'^{(2)} \cap F^{(2)}|\} = \{\lceil \frac{k-t}{2} \rceil, \lfloor \frac{k-t}{2} \rfloor\}$ .

- $\{|\tilde{C}^{(2)} \cap F^{(2)}|, |\tilde{C}'^{(1)} \cap F^{(1)}|\} = \{\lfloor \frac{k-t}{2} \rfloor, \lceil \frac{k-t}{2} \rceil\}$ .

Note that, if  $S \xleftrightarrow{(C,C')} T$  is an edge of the Kikuchi graph  $G$  and  $C, C'$  have colors  $c, c'$  respectively, then both  $S$  and  $T$  contain in addition to common part  $S \cap T$  an edge from  $\mathcal{H}_c \cap \mathcal{H}_{c'}$ .



**Figure 2:** An illustration of Definitions 4.6 and 4.7. On the left, we have a pair of edges  $C, C'$  which both belong to a bucket. Then, on its right is depicted an edge  $S \xleftrightarrow{(C,C')} T$  created by this pair. Then, on the further right are illustrated all edges of  $\mathcal{H}_c$  (where  $c$  is the colour of the edge  $C$ ) which are created in Definition 4.7 by the pair  $(C, C')$ .

Let us now note two important properties of the  $l$ -Kikuchi graph  $G$ .

**Observation 4.8.** *If  $G$  contains a closed walk whose edges are overall associated to a multi-set  $\mathcal{H}'$  of edges in  $\mathcal{H}$ , then it is such that  $\bigoplus_{C \in \mathcal{H}'} C = \emptyset$ . In particular, if there is some colour which appears an odd number of times in  $\mathcal{H}'$ , then  $\mathcal{H}$  contains an even cover of size at most  $e(\mathcal{H}')$  in which some colour appears an odd number of times. Further,  $G$  has average degree at least*

$$e(\mathcal{H})(m-1) \cdot \frac{\binom{k-t}{\lfloor \frac{k-t}{2} \rfloor}^2 \cdot \binom{2n-2k+2t}{l-k+t}}{\binom{2n}{l}} \geq e(\mathcal{H})(m-1) \cdot \left(\frac{l}{2n}\right)^{k-t}.$$

**Proof.** We will briefly explain the average degree computation. Indeed, the number of ordered pairs  $(C, C')$  of edges in the same bucket is  $e(\mathcal{H})(m-1)$ . Now, with the pair  $(C, C')$  fixed, choosing an edge  $S \xleftrightarrow{(C,C')} T$  can be viewed as a two-step choice. First, we choose the intersections  $\tilde{C}_1 \cap S^{(1)}, \tilde{C}_1 \cap T^{(1)}, \tilde{C}_2 \cap S^{(2)}, \tilde{C}_2 \cap T^{(2)}$  so that these intersections sizes are as required in Definition 4.6 - there are  $\binom{k-t}{\lfloor \frac{k-t}{2} \rfloor}^2$  such choices. Once those intersections are chosen, the next step is to choose a set  $S \cap T = S \setminus F = T \setminus F'$  of size  $l - k + t$  out of  $2n - 2k + 2t$  vertices. Combining all of these, we get that the number of edges in  $G$  is  $e(\mathcal{H})(m-1) \cdot \binom{k-t}{\lfloor \frac{k-t}{2} \rfloor}^2 \cdot \binom{2n-2k+2t}{l-k+t} \binom{2n}{l}$ . Since  $G$  has  $\binom{2n}{l}$  vertices, we have the desired average degree.  $\square$

As anticipated by the above observations, we can use Kikuchi graphs to find even covers in hypergraphs.

**Proposition 4.9.** *Let  $k$  be odd and  $\mathcal{H}$  be an edge-coloured  $n$ -vertex  $k$ -uniform hypergraph with  $nd$  edges and an  $(m, t)$ -bucket decomposition. Let  $G$  be the  $l$ -Kikuchi graph of  $\mathcal{H}$  and suppose that the following holds.*

- For all colours  $c$ , edges  $F \in \mathcal{H}_c$  and  $0 \leq j \leq k - t$ , the number of edges  $F' \in \mathcal{H}_c$  with  $|F' \cap F| = j$  is at most  $(d(G)/2000 \log v(G)) \cdot (n/l)^{k-t-j}$ .

Then,  $\mathcal{H}$  contains an even cover of size at most  $2 \log v(G)$  in which some colour appears an odd number of times.

**Proof.** Our aim is to use Lemma 2.5. However, we might not be able to apply it directly to the Kikuchi graph  $G$  (which by Observation 4.8 would imply an even cover in  $\mathcal{H}$ ). This is because  $G$  might not have a well-behaved colouring - it could well be that there are vertices and colours which do not satisfy the condition of Lemma 2.5.

Instead, it turns out that we will be able to delete some edges from  $G$  and get a subgraph  $G' \subseteq G$  which does satisfy the necessary conditions. Let  $d(G)$  denote the average degree of  $G$  and for each ordered pair  $(C, C')$  of edges in  $\mathcal{H}$  with colours  $c, c'$ , we delete all edges  $S \xleftrightarrow{C, C'} T$  such that one of  $S$  or  $T$  is adjacent to more than  $d(G)/80 \log v(G)$  edges with one of the colours  $c, c'$ . Let  $G'$  be the resulting graph. Note that if we can show that  $d(G') \geq d(G)/2$ , then Lemma 2.5 (with  $s = 2$ ) implies that  $G$  contains a closed walk of size at most  $2 \log v(G)$  in which some colour appears once. By Observation 4.8, this implies an even cover in  $\mathcal{H}$  of size at most  $2 \log v(G)$  as desired.

The following claim clearly implies that  $d(G') \geq d(G)/2$ . Let  $C_1, C_2$  be any two edges of  $\mathcal{H}$  (with colours  $c_1, c_2$  respectively) in the same bucket  $\mathcal{E}_i$ .

**Claim 4.10.** *At least half of the edges with colour-pair  $(C_1, C_2)$  are in  $G'$ .*

**Proof.** To prove this claim it is enough to show that uniformly chosen edge with color pair  $(C_1, C_2)$  is deleted with probability at most  $1/2$ . Note that with  $C_1, C_2$  fixed, choosing a uniformly random edge  $S \xleftrightarrow{(C_1, C_2)} T$  can be viewed as a two-step uniform choice. First, we choose at random the intersections  $\tilde{C}_1 \cap S^{(1)}, \tilde{C}_1 \cap T^{(1)}, \tilde{C}_2 \cap S^{(2)}, \tilde{C}_2 \cap T^{(2)}$  so that these intersections sizes are as required in Definition 4.6. This will, for  $S$  (and  $T$ ), fix an edge  $F \in \mathcal{H}_{c_1} \cap \mathcal{H}_{c_2}$  which is contained in  $S$  (and another edge  $F'$  in  $\mathcal{H}_{c_1} \cap \mathcal{H}_{c_2}$  contained in  $T$ ). Secondly, once those intersections are chosen, the next step is to choose a uniformly random set  $S \cap T = S \setminus F = T \setminus F'$  of size  $l - k + t$  out of  $2n - 2k + 2t$  vertices.

Now, we will upper bound the probability that such a random edge is not in  $G'$ . We will deal first with the case that  $S$  is incident to more than  $d(G)/80 \log v(G)$  other edges with the colour  $c_1$  (and thus, the edge  $S \leftrightarrow T$  will not be in  $G'$ ) and will show that a random edge has this property with probability at most  $1/8$ . The other four cases ( $S$  incident to too many edges of colour  $c_2$ ,  $T$  incident to too many edges of colour  $c_1$  and  $T$  incident to too many edges of colour  $c_2$ ) are analogous and so, we will in the end have a total probability of deletion of at most  $4/8 = 1/2$ . Suppose then that  $(C_3, C_4)$  are two other edges both in the same bucket and one of them has colour  $c_1$  - without loss of generality, assume it is  $C_3$ . Suppose there exists an edge  $S \xleftrightarrow{C_3, C_4} T'$  for some other  $T'$ . Then,  $S$  must contain some additional edge in  $\mathcal{H}_{c_1}$  - let us denote this by  $E$ . For each such possible  $E$ , since  $S \setminus F$  is chosen uniformly, the probability that  $S$  contains it is at most

$$\frac{\binom{2n-2k+2t}{l-|E \setminus F|-k+t}}{\binom{2n-2k+2t}{l-k+t}} \leq \left( \frac{l-k+t}{2n-2k+2t-(l-k+t)} \right)^{|E \setminus F|} \leq (0.51l/n)^{|E \setminus F|},$$

using the fact that in all our proofs we can assume that  $l \leq n/\log n$ . Furthermore, by assumption there are at most  $(d(G)/2000 \log v(G)) \cdot (n/l)^{k-t-j}$  edges of  $\mathcal{H}_{c_1}$  with  $|E \cap F| = j$ . Therefore, the expected number of such edges  $S \xleftrightarrow{C_3, C_4} T'$  is at most

$$(d(G)/2000 \log v(G)) \cdot \sum_{1 \leq j \leq k-t} (n/l)^{k-t-j} \cdot (0.51l/n)^{k-t-j} \leq d(G)/800 \log v(G).$$

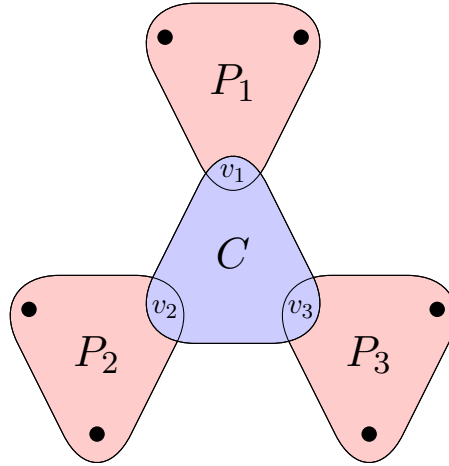
Concluding, by Markov's inequality, that the probability that a random edge is deleted because of one of its endpoints and one of its colors is at most  $0.1 < 1/8$ , as desired.  $\square$

$\square$

#### 4.5 The Flower Kikuchi graph and even covers

We will now introduce a *new* variant of the *Kikuchi graph* defined in the previous sections. Just as before, we will always consider odd uniformity  $k$ . Let  $\mathcal{H}$  be an  $n$ -vertex  $k$ -uniform hypergraph with  $nd$  edges and minimum degree  $\delta$  and such that every set of size  $k - 1$  is contained in at most one edge. Suppose that the edges of  $\mathcal{H}$  are coloured in *red* and *blue*. Suppose also that for each vertex  $v$  in  $\mathcal{H}$ , we fix a collection  $\mathcal{E}_v$  of precisely  $\delta$  edges containing  $v$ .

**Definition 4.11.** A *flower gadget*  $F = (C, P_1, \dots, P_k)$  in  $\mathcal{H}$  is a collection of  $k + 1$  hyperedges in  $\mathcal{H}$  with a *center* edge  $C = \{v_1, \dots, v_k\}$  and  $k$  *petal* edges  $P_1, \dots, P_k$  which are disjoint and such that each  $P_i$  is contained in  $\mathcal{E}_{v_i}$  with  $P_i \cap C = \{v_i\}$ .  $F$  is a *good flower gadget* if  $C$  is *blue* and all edges  $P_i$  are *red*. See Figure 3 for an illustration.



**Figure 3:** An illustration of a good flower gadget with  $k = 3$ .

**Definition 4.12.** The *flower  $l$ -Kikuchi graph*  $G$  of  $\mathcal{H}$  has vertex set  $V := \binom{[n]}{l}$  and its edges are defined and coloured as follows. For each good flower gadget  $F = (C, P_1, \dots, P_k)$ , define an edge between  $S, T \in V$ , denoted by  $S \xleftrightarrow{C} T$  and coloured with  $C$ , if  $S \oplus T$  is equal to  $P_1 \oplus \dots \oplus P_k \oplus C$ , and  $|S \cap (P_i \setminus v_i)| = |T \cap (P_i \setminus v_i)| = \frac{k-1}{2}$  for all  $i$ .

Now, let us note some inconsistencies with this colouring.

**Claim 4.13.** An edge  $S \leftrightarrow T$  can possibly be created by at most  $2^{k^3}$  good flower gadgets.

**Proof.** To prove this we only need to use the property that no set of size  $k - 1$  is contained in more than one edge of the hypergraph  $\mathcal{H}$ . Indeed, observe that for a good flower gadget  $\{C, P_1, \dots, P_k\}$  to create this edge, it must be that  $|(S \setminus T) \cap A| = |(T \setminus S) \cap A| = \frac{k-1}{2}$  for all  $A \in \{P_1 \setminus \{v_1\}, \dots, P_k \setminus \{v_k\}\}$ . Since  $|S \setminus T| = |T \setminus S| = \frac{k(k-1)}{2}$ , there are at most  $\binom{k(k-1)/2}{(k-1)/2}^{2k} \leq 2^{k^3}$  ways to define these sets  $(S \setminus T) \cap A, (T \setminus S) \cap A$ , giving us the sets  $P_1 \setminus v_1, \dots, P_k \setminus v_k$ . Since the hypergraph is such that no set of size  $k - 1$  is contained in more than one edge, through each of these sets there is a unique edge, giving us  $P_1, \dots, P_k$  (up to permutation of names which does not matter). Then the vertex set  $(P_1 \cup \dots \cup P_k) \setminus (S \cup T)$  gives  $C$ .  $\square$

Given the above claim, we fix the colouring arbitrarily by colouring  $S \xleftrightarrow{C} T$  with only one of the at most  $2^{k^3}$  possible options. We make note of some other crucial (but simple to check) properties of  $G$  just as we did previously in Section 4.4 with Observation 4.8.

**Observation 4.14.** *If  $G$  contains a closed walk and we denote by  $\{C^{(j)}, P_1^{(j)}, \dots, P_k^{(j)}\}$  the good flower gadgets which create the edges of the walk, then  $\bigoplus_j (P_1^{(j)} \oplus \dots \oplus P_k^{(j)} \oplus C^{(j)}) = \emptyset$ . In particular, if there is some (blue) edge which appears an odd number of times among the  $C^{(j)}$ , then  $\mathcal{H}$  contains an even cover whose size is of the same order as the length of the walk. Also,  $G$  has  $N := \binom{n}{l}$  vertices and letting  $m$  denote the number of good flower gadgets in  $\mathcal{H}$ , then  $G$  has average degree*

$$\frac{m \cdot \binom{n-k(k-1)}{l-k(k-1)/2}}{2^{k^3} N} \geq \frac{m}{10^{k^3}} \cdot \left(\frac{l}{n}\right)^{k(k-1)/2}.$$

**Proof.** We will briefly explain the average degree computation. Each good flower gadget creates at least  $\binom{n-k(k-1)}{l-k(k-1)/2}$  edges since after choosing the intersections  $S \cap (P_i \setminus \{v_i\}), T \cap (P_i \setminus \{v_i\})$  we are left choosing  $S \cap T$  which can be any set of size  $l - \frac{k(k-1)}{2}$  from  $V(\mathcal{H}) \setminus \bigcup_i (P_i \setminus \{v_i\})$ , which has size  $n - k(k-1)$ . Further, as discussed above, each edge is possibly created by at most  $2^{k^3}$  good gadgets, and therefore,  $G$  has at least  $\frac{m \cdot \binom{n-k(k-1)}{l-k(k-1)/2}}{2^{k^3}}$  edges. The last estimate can be verified by considering separately cases when  $l \leq k^2$  or  $l > k^2$  and recalling that we can assume that  $l \leq n/\log n$ .  $\square$

Much like in the previous section, we can show an analogous of Proposition 4.9, that if the flower  $l$ -Kikuchi graph has some 'nice' properties, then the underlying hypergraph contains a small even cover. Since we will not need the full power of such a statement, in the next section, we will only prove a simpler version of it which we use to establish Theorem 1.2 (ii).

## 5 Even covers in hypergraphs

In this section, we will apply the previous tools to prove Theorems 1.2 (ii) and 2.8.

### 5.1 Proof of Theorem 2.8

Take  $K := 10^7/\alpha^2$ , let  $\mathcal{H}$  be such a hypergraph and let  $l := n^{1/3}$ . We can assume without loss of generality that we have  $d := Kl \log n$  colours, each with precisely  $\alpha n$  edges, so that our hypergraph has  $\alpha nd$  edges. We will want to apply Proposition 4.9 in order to find such a special even cover. For this, let us first note that one of the following holds.

1. There is a  $\mathcal{H}' \subseteq \mathcal{H}$  with  $e(\mathcal{H}') \geq e(\mathcal{H})/2$  with a  $(K \log n, 2)$ -bucket decomposition.
2. There is a hypergraph  $\mathcal{H}' \subseteq \mathcal{H}$  with  $e(\mathcal{H}') \geq e(\mathcal{H})/4$  such that it has a  $(\alpha d/4, 1)$ -bucket decomposition and every set  $R$  of size two has  $\deg_{\mathcal{H}'}(R) \leq K \log n$ .

Indeed, this follows directly by applying first Lemma 4.1 with  $t = 2$  and  $m = K \log n$ . This gives that the first item holds or there is a  $\mathcal{H}' \subseteq \mathcal{H}$  with  $e(\mathcal{H}') \geq e(\mathcal{H})/2$  and such that every set  $R$  of size two has  $\deg_{\mathcal{H}'}(R) \leq K \log n$ . For this second case, we can then apply Lemma 2.3 to find the  $(\alpha d/4, 1)$ -bucket decomposition in the second item. Now we will show that in both cases above, Proposition 4.9 holds for the  $l$ -Kikuchi graph of the corresponding hypergraph  $\mathcal{H}'$  which for convenience we redenote as  $\mathcal{H}$ . Let us first suppose that (1.) holds and let us consider the  $l$ -Kikuchi graph  $G$  of  $\mathcal{H}$  and the given bucket decomposition. Using that  $K = 10^7/\alpha^2$  and  $v(G) = \binom{2n}{l} \leq (2n)^l$ , by Observation 4.8 (with  $k = 3$  and  $t = 2$ ), the  $l$ -Kikuchi graph  $G$  has average degree at least  $(\alpha nd/2) \cdot (K \log n - 1) \cdot (l/2n) \geq 8000dl \log(2n) \geq 8000d \cdot \log v(G)$ . Also note that for each colour  $c$ , the multihypergraph  $\mathcal{H}_c$  is 1-uniform

so, consists of single vertices in  $[n] \times [2]$ . Thus, by Proposition 4.9 we need only to show that given a colour  $c$ , an  $F \in \mathcal{H}_c$  and a  $0 \leq j \leq 1$  the number of  $F' \in \mathcal{H}_c$  with  $|F' \cap F| = j$  is at most  $4d \cdot (n/l)^{1-j} \leq d(G)/(2000 \log v(G)) \cdot (n/l)^{1-j}$ . For  $j = 0$ , the number of such  $F'$  is at most four times the number of edges in  $\mathcal{H}_c$  which is  $4(K \log n)n = 4d \cdot (n/l)$ . To see this, recall that the hypergraph  $\mathcal{H}$  is properly coloured and therefore at each of its vertices there is only one bucket of size  $K \log n$  containing an edge  $C$  of color  $c$  - every other edge  $C'$  in this bucket together with  $C$  gives then 4 edges of  $\mathcal{H}_c$  - indeed there are four edges of  $\mathcal{H}_c$  satisfying one of the items in Definition 4.7; these consist of the red and blue copies of the vertex in  $C \setminus C'$  and the vertex in  $C' \setminus C$ . For  $j = 1$ , let the edge  $C \in E(\mathcal{H})$  of color  $c$  be the one in Definition 4.7 forming  $F$ . The edge  $F' \in \mathcal{H}_c$  with  $F' = F$  can be obtained in two ways: first from some edge  $C'$  of  $\mathcal{H}$  incident to the same vertex as singleton the (uncolored)  $F$ ; this gives at most the maximum degree of  $\mathcal{H}$ , i.e.,  $d$  edges; we can also obtain  $F'$  from any edge  $C' \in \mathcal{H}$  from the same bucket as  $C$ . This gives at most  $K \log n$  additional edges. So in total  $d + K \log n \leq 2d$ .

Now assume that (2.) holds and recall that  $l = n^{1/3}$ ,  $d = Kl \log n$  and  $v(G) \leq (2n)^l$  where  $G$  is the  $l$ -Kikuchi graph of  $\mathcal{H}$  and the given bucket decomposition. In this case  $\mathcal{H}_c$  is a multigraph for every color  $c$ . The Observation 4.8 (with  $k = 3$  and  $t = 1$ ) implies that the  $l$ -Kikuchi graph has average degree at least

$$d(G) \geq (\alpha d/4) \cdot (\alpha d/4 - 1) \cdot (l/2n)^2 \geq (\alpha K)^2/100 \cdot (l^3/n) \cdot l \cdot (\log n)^2 \geq 10^5 K \cdot \log n \cdot \log v(G).$$

Hence  $d(G)/2000 \log v(G) \geq 50K \log n$ . Take  $F = (u, v) \in \mathcal{H}_c$  which corresponds (as in Definition 4.7) to a pair of edges  $C, C'$  of  $\mathcal{H}$  from some bucket such that  $C$  has color  $c$  and w.l.o.g contains  $u$  and  $C'$  contains  $v$ . We need to verify that for  $0 \leq j \leq 2$  the number of  $F' \in \mathcal{H}_c$  with  $|F' \cap F| = j$  is at most  $50K \log n (n/l)^{2-j}$ , i.e., satisfies the condition in Proposition 4.9 (with  $k = 3$  and  $t = 1$ ). For  $j = 0$ , this number is the number of edges in  $\mathcal{H}_c$ , which is at most  $8n$  times the size of each bucket, that is,  $8n \cdot (\alpha d/4) = 2\alpha K n l \log n \leq 2K \log n (n/l)^2$ . This is because each bucket has  $\alpha d/4$  and so if it contains an edge of colour  $c$  then it produces  $\alpha d/4$  pairs of edges with one of colour  $c$ . Each such pair then contributes 8 edges to  $\mathcal{H}_c$  (see Figure 3). Furthermore, since the coloring of  $\mathcal{H}$  is proper there are at most  $n$  buckets containing an edge of color  $c$ .

For  $j = 1$ , note that the number of  $F'$  is at most 16 times the maximum degree of  $\mathcal{H}$  plus 16 times the size of each bucket which is at most  $16d + 16\alpha d/4 \leq n/l$ . To see this note that we can get  $F'$  in two ways; one is by taking the unique edge in  $\mathcal{H}$  of color  $c$  through either  $u$  or  $v$  and combining it with any other edge from the same bucket (giving a pair of edges for Definition 4.7); the other is by considering any edge of  $\mathcal{H}$  containing either  $u$  or  $v$  and combining it with an edge of color  $c$  in the same bucket (giving a pair of edges in  $\mathcal{H}$  for Definition 4.7); like before, in each case, each pair produces 8 edges of  $\mathcal{H}$ . Finally, for  $j = 2$ , note that  $F' \in \mathcal{H}_c$  with  $F' = F$  can be obtained by taking a unique edge  $C$  of color  $c$  through  $u$  (or  $v$ ) and combining it with any edge in the same bucket which contains  $v$ . This edge must intersect the vertex of  $C - \{u\}$  which is incident to every other edge in the bucket containing  $C$  and so, since the codegree of every pair of vertices is at most  $K \log n$  this gives at most  $2K \log n$  options.  $\square$

## 5.2 Proof of Theorem 1.2 (ii)

Let  $\mathcal{H}$  be a  $k$ -uniform  $n$ -vertex hypergraph with  $nd$  where  $d \geq C (n/l)^{k/2-1} (\log n)^{\frac{1}{k+1}}$  for some sufficiently large constant  $C$  which depends on  $k$ . We take  $n$  large enough so that any fixed power of  $\log n$  which appears in the proof is bigger than any constant dependent on  $k$  appearing. Note first that using Corollary 4.4 we can assume that  $\mathcal{H}$  is such that every set of more than  $\frac{k-1}{2}$  vertices (since  $k$  is odd) has co-degree at most 1. Without loss of generality let us now assume that  $d = C (n/l)^{k/2-1} (\log n)^{\frac{1}{k+1}}$ .

We apply Lemma 4.5 with  $m$  defined as  $m(1) = d/10k$  and  $m(t) = (d/10k)^{\frac{k/2-t}{k/2-1}} \cdot (\log n)^{1-\frac{1}{k+1}}$  for  $t \geq 2$ . This gives us a sub-hypergraph  $\mathcal{H}_0$  with at least  $nd/2k$  edges and a  $t \leq \frac{k-1}{2}$  with the given properties in the lemma. Let us split the remainder of the proof into the cases that  $t > 1$  and  $t = 1$ .

**Case 1:  $t > 1$**

In this case, we know that  $\mathcal{H}_0$  has an  $(m(t), t)$ -bucket decomposition and  $\deg_{\mathcal{H}_0}(S) < m(|S|)$  for all sets  $S$  with  $|S| > t$ . We will show that this hypergraph satisfies the conditions of Proposition 4.9, which will immediately give an even cover as desired. Let then  $G$  be the  $l$ -Kikuchi graph for  $\mathcal{H}_0$  and let  $m := m(t)$ . By Observation 4.8, it has average degree  $d(G) \geq e(\mathcal{H}_0)(m-1)(l/2n)^{k-t}$ . Recalling that  $G$  has  $v(G) \leq \binom{2n}{l} \leq (2n)^l$  vertices, we have that

$$\begin{aligned} d(G) &\geq \frac{nd}{2k} \cdot m \cdot (l/2n)^{k-t} \geq \frac{nd}{2k} \cdot (d/10k)^{\frac{k/2-t}{k/2-1}} \cdot (\log n)^{1-\frac{1}{k+1}} \cdot (l/2n)^{k-t} \\ &\geq (\log n)^{1/k^2} \cdot (l \log(2n)) \geq (\log n)^{\frac{1}{(k+1)^2}} \cdot \log v(G). \end{aligned} \quad (1)$$

Now, let us redefine  $\mathcal{H}$  as  $\mathcal{H}_0$ . View  $\mathcal{H}$  as an edge-coloured hypergraph by colouring each edge with its own unique colour and let us verify the conditions of Proposition 4.9. For each edge  $C \in \mathcal{H}$ , the hypergraph  $\mathcal{H}_C$  is then formed by taking every edge  $C'$  in the same bucket as  $C$  and putting all edges  $F$  in  $\mathcal{H}_C$  which satisfy one of the items in Definition 4.7 - in particular, part of  $F$  (either  $\lceil \frac{k-t}{2} \rceil$  or  $\lfloor \frac{k-t}{2} \rfloor$ ) will be contained in either  $C^{(1)}$  or  $C^{(2)}$  and the rest will be contained in either  $C'^{(1)}$  or  $C'^{(2)}$ . Depending on parity of  $k-t$ , each such edge  $C'$  contributes to at least  $2^{\lfloor \frac{k-t}{2} \rfloor}$  and at most  $4^{\lfloor \frac{k-t}{2} \rfloor} \leq 2^{2k}$  edges for  $\mathcal{H}_C$ .

Let us fix  $C$  an edge in  $\mathcal{H}$ ,  $F \in \mathcal{H}_C$  and  $0 \leq j \leq k-t$ . Let us denote the bucket of  $\mathcal{H}$  containing  $C$  by  $(X_i, \mathcal{E}_i)$ . We bound the number of  $F' \in \mathcal{H}_C$  with  $|F' \cap F| = j$  as follows. Since  $F \in \mathcal{H}_C$ , there is some edge  $C'$  in the same bucket as  $C$  such that one of the items in Definition 4.7 applies. Let us assume without loss of generality that  $\{|\tilde{C}^{(1)} \cap F^{(1)}|, |\tilde{C}^{(2)} \cap F^{(2)}|\} = \{\lceil \frac{k-t}{2} \rceil, \lfloor \frac{k-t}{2} \rfloor\}$ . Fix a pair  $j_1, j_2$  with  $j = j_1 + j_2$  and let us consider those  $F'$  with  $j_1 = |F'^{(1)} \cap F^{(1)}| \leq \lceil \frac{k-t}{2} \rceil$  and  $j_2 = |F'^{(2)} \cap F^{(2)}| \leq \lfloor \frac{k-t}{2} \rfloor$ . We first fix for  $F'$  one of the items in Definition 4.7 which applies. Let us first count those with  $\{|\tilde{C}^{(1)} \cap F'^{(1)}|, |\tilde{C}^{(2)} \cap F'^{(2)}|\} = \{\lceil \frac{k-t}{2} \rceil, \lfloor \frac{k-t}{2} \rfloor\}$  for some edge  $C''$  in the same bucket as  $C$  and  $C'$ .

First, we have at most  $2^k$  options for the choice of  $F'^{(1)}$  since it must be a subset of  $C^{(1)}$ . For the choice of  $F'^{(2)}$ , note that after choosing an appropriate edge  $C''$ , we have again at most  $2^k$  options for the choice of  $\tilde{C}''^{(2)} \cap F'^{(2)}$ . But now the choice of  $C''$  is restricted to the fact that  $X_i \cup (F'^{(2)} \cap F^{(2)})$  (which is here considered naturally as a set of vertices in  $[n]$  and not in  $[n] \times [2]$ ) must be contained in  $C''$  and thus, since  $|X_i \cup (F'^{(2)} \cap F^{(2)})| = t + j_2$ , the conditions of Lemma 4.5 imply that there are at most

$$2^k \cdot \max_{|R|=t+j_2} \deg_{\mathcal{H}'}(R) \leq 2^k \cdot m(t+j_2) \leq 2^k \cdot (d/10k)^{\frac{k/2-t-j_2}{k/2-1}} \cdot (\log n)^{1-\frac{1}{k+1}}$$

choices for such an edge  $C''$ , where the extra  $2^k$  factor serves as a gross upper bound for the number of possible subsets  $X_i \cup (F'^{(2)} \cap F^{(2)}) \subseteq F$ . Combining all these consideration, we have at most

$$\begin{aligned} 2^{3k} \cdot (d/10k)^{\frac{k/2-t-j_2}{k/2-1}} \cdot (\log n)^{1-\frac{1}{k+1}} &\leq \log n \cdot (n/l)^{k/2-t-j_2} \\ &\leq \log n \cdot (n/l)^{k-t-j-\frac{1}{2}} \end{aligned}$$

options for such  $F'$ . Notice that we used in the previous inequality that  $k/2-t-j_2 \leq (k/2+j_1)-t-j \leq (k/2+\lceil \frac{k-t}{2} \rceil)-t-j \leq (k/2+\lceil \frac{k-2}{2} \rceil)-t-j \leq (k/2+\frac{k-1}{2})-t-j \leq k-t-j-\frac{1}{2}$ , since  $t \geq 2$ .

Now, the arguments and the upper bound are analogous in the other cases for  $F'$  (that is, for all possible values of  $j_1, j_2$  and whether  $\{|\tilde{C} \cap F'^{(1)}|, |\tilde{C}'' \cap F'^{(2)}|\} = \{\lceil \frac{k-t}{2} \rceil, \lfloor \frac{k-t}{2} \rfloor\}$  or  $\{|\tilde{C} \cap F'^{(2)}|, |\tilde{C}'' \cap F'^{(1)}|\} = \{\lceil \frac{k-t}{2} \rceil, \lfloor \frac{k-t}{2} \rfloor\}$ ). Since there are at most  $2k^2$  such cases, by using (1) and choosing  $C$  large enough as a function of  $k$ , we conclude that the number of choices for  $F'$  is in total at most

$$2k^2 \cdot \log n \cdot (n/l)^{k-t-j-\frac{1}{2}} \leq (d(G)/2000 \log v(G)) \cdot (n/l)^{k-t-j},$$



as desired, where we used that  $\log n \leq \sqrt{n/l}$ . Hence, the property of Proposition 4.9 is satisfied and therefore  $\mathcal{H}$  contains an even cover of size at most  $2 \log v(G) = O(l \log n)$ , as desired.

**Case 2:**  $t = 1$

In this case we know that  $\mathcal{H}_0$  has minimum degree at least  $d/10k$  and  $\deg_{\mathcal{H}_0}(S) < m(|S|) = (d/10k)^{\frac{k/2-|S|}{k/2-1}}$ .  $(\log n)^{1-\frac{1}{k+1}} \leq (n/l)^{k/2-|S|} \cdot (\log n)^{1-1/k^2}$  for all sets  $S$  with  $|S| > 1$  and  $m(1) = d/10k$ . As before, let us redefine  $\mathcal{H}$  as  $\mathcal{H}_0$ . For each vertex  $v$ , let  $\mathcal{E}_v$  denote a collection of precisely  $d/10k$  edges of  $\mathcal{H}$  incident on  $v$ . Now, let us independently and uniformly at random colour each edge of  $\mathcal{H}$  as either red or blue. We then have the following.

**Claim 5.1.** *With positive probability, there are at least  $nd^{k+1}/(100k)^k$  good flower gadgets.*

**Proof.** We show that the number of flower gadgets in  $\mathcal{H}$  is at least  $nd^{k+1}/(20k)^k$ . Then, since each flower gadget is also good with probability  $1/2^{k+1}$  then the expected number of good flower gadgets is at least  $nd^{k+1}/(50k)^k$  and further, with positive probability this holds.

Now, let us count the number of flower gadgets  $\{C, P_1, \dots, P_k\}$ . We have  $e(\mathcal{H}) \geq nd/2k$  options for the edge  $C = \{v_1, \dots, v_k\}$ . Then, since  $P_1 \in \mathcal{E}_{v_1}$  and it is disjoint to  $C \setminus \{v_1\}$ , there are  $|\mathcal{E}_{v_1}| - |C| \cdot m(2) > d/10k - km(2)$  options for the edge  $P_1$ . Subsequently,  $P_2$  must be an edge belonging to  $\mathcal{E}_{v_2}$  which is disjoint to  $(C \cup P_1) \setminus \{v_2\}$  and therefore, we have at least  $|\mathcal{E}_{v_2}| - |C \cup P_1| \cdot m(2) > d/10k - 2km(2)$  options for  $P_2$ . Continuing in this manner, we have that the number of flower gadgets is at least

$$(nd/2k) \prod_{1 \leq i \leq k} (d/10k - ikm(2)).$$

Since  $m(2) = O_k(1) \cdot (n/l)^{k/2-2} \cdot (\log n) < d/100k^3$ , this product is at least  $nd^{k+1}/(20k)^k$ , as desired.  $\square$

Now, consider the flower  $l$ -Kikuchi graph  $G$  of  $\mathcal{H}$ . By the previous claim and Observation 4.14, it has average degree at least  $\frac{nd^{k+1}/(100k)^k}{10k^3} \cdot \left(\frac{l}{n}\right)^{k(k-1)/2} \geq \frac{nd^{k+1}}{100k^3} \cdot \left(\frac{l}{n}\right)^{k(k-1)/2} \geq \frac{C^{k+1}}{100k^3} \cdot l \log n \geq C \log N$ . Our aim is now to apply Lemma 2.5 to the graph  $G$ . Indeed, by Observation 4.14, a closed walk in  $G$  such that some colour appears only once implies an even cover in  $\mathcal{H}$  of the same order. However, it might be that the graph  $G$  does not satisfy the conditions of Lemma 2.5 and so, we need to first do a cleaning procedure similar to what was done in the proof of Proposition 4.9. Indeed, we delete all edges  $S \overset{C}{\leftrightarrow} T$  of  $G$  such that one of  $S$  or  $T$  are incident to another edge of colour  $C$ . Let  $G'$  denote the resulting graph and note the following.

**Lemma 5.2.**  *$G'$  has average degree at least  $(C/2) \log N$ .*

**Proof.** Fix a good flower gadget  $F = (C, P_1, \dots, P_k)$  and subsets  $L_1, \dots, L_k$  and  $R_1, \dots, R_k$  such that each  $|R_i \cap (P_i \setminus \{v_i\})|$  and  $|L_i \cap (P_i \setminus \{v_i\})|$  are equal to  $\frac{k-1}{2}$ . Next, we choose a uniformly random edge  $S \overset{C}{\leftrightarrow} T$  of  $G$  such that  $S \cap (P_i \setminus \{v_i\}) = L_i$  and  $T \cap (P_i \setminus \{v_i\}) = R_i$  for all  $i \in [k]$ . As in the definition, we use  $v_i$  to denote the vertex at which  $P_i$  intersects  $C$ . We will show that the probability that  $S \overset{C}{\leftrightarrow} T$  is deleted is at most  $1/2$ , which gives the desired outcome.

In order for  $S$  to be incident to another edge of colour  $C$ , there must be  $i \in [k]$  and a hyperedge  $P \neq P_i$  in  $\mathcal{E}_{v_i}$  such that  $|(P \setminus \{v_i\}) \cap S| = \frac{k-1}{2}$ . For each such  $P$ , since the intersection  $S \cap (P_i \setminus \{v_i\}) = L_i$  is already chosen, and further, the set  $S' := S \setminus ((P_1 \setminus \{v_1\}) \cup \dots \cup (P_k \setminus \{v_k\}))$  is a uniformly random set of size  $l - \frac{k(k-1)}{2}$  in  $V(\mathcal{H}) \setminus ((P_1 \setminus \{v_1\}) \cup \dots \cup (P_k \setminus \{v_k\}))$ . Therefore, the probability that  $|(P \setminus \{v_i\}) \cap S| = \frac{k-1}{2}$  is equal to the probability that the set  $S'$  intersects  $P \setminus P_i$  in  $\frac{k-1}{2} - |(P \setminus \{v_i\}) \cap (P_i \setminus \{v_i\})| = \frac{k-1}{2} - |P \cap P_i| + 1$  vertices. Since  $S'$  is chosen uniformly, this probability is at most

$$2^k \cdot \frac{\binom{n-(k-1)(k+\frac{1}{2})+|P \cap P_i|-1}{l-\frac{(k+1)(k-1)}{2}+|P \cap P_i|-1}}{\binom{n-k(k-1)}{l-\frac{k(k-1)}{2}}} \leq 2^k \cdot (l/10n)^{(k-1)/2-|P \cap P_i|+1},$$

where the  $2^k$  is an upper bound for the number of intersections between  $S'$  and  $P \setminus P_i$ . Now we count for each  $j \geq 1$ , the number of such  $P$  with  $|P \cap P_i| = j$ . For  $j = 1$ , this is clearly at most  $|\mathcal{E}_{v_i}| = d/10k = m(1) \leq (n/l)^{k/2-1} \cdot (\log n)^{1-1/k^2}$ . Moreover, for  $j \geq 2$ , this is at most  $\max_{|X|=j} \deg_{\mathcal{H}}(X) \leq m(j) \leq (n/l)^{k/2-j} \cdot (\log n)^{1-1/k^2}$ . Further, by the assumption that no two edge in  $\mathcal{H}$  intersect in more than  $k/2$  vertices, this number is 0 for  $j \geq \frac{k+1}{2}$ . Hence, combining these observations, the probability of  $S$  having other  $C$ -colored incident edges is at most

$$\begin{aligned} 2^k \sum_{1 \leq j \leq \frac{k-1}{2}} m(j) \cdot (l/10n)^{(k-1)/2-j+1} &\leq \sum_{1 \leq j \leq \frac{k-1}{2}} (n/l)^{k/2-j} \cdot (\log n)^{1-1/k^2} \cdot (l/n)^{(k-1)/2-j+1} \\ &\leq k \sqrt{\frac{l}{n}} \cdot (\log n)^{1-1/k^2} \leq 1/4, \end{aligned}$$

where we used that  $l \leq n/(\log n)^2$  and that  $n$  is sufficiently large. Applying the same argument to  $T$  gives the desired outcome.  $\square$

To finish, note that the edge-colouring in  $G'$  is now a proper edge-colouring and further, it has average degree at least  $20 \log N$  and thus, Lemma 2.5 implies that  $G$  contains a closed walk of size  $O(\log N)$  such that some colour appears exactly once. By Observation 4.14, this implies an even cover in  $\mathcal{H}$  of size  $O(l \log n)$ , as desired.  $\square$

## 6 Concluding Remarks

In this work, we gave a simple proof without the use of matrix concentration inequalities that recovers and improves on two recent developments on finding even covers in edge-colored hypergraphs. To conclude, we would like to point out two natural directions for further progress:

- **Even Covers in Hypergraphs:** Theorem 1.2 is short of Feige's original conjecture [8] that asks to show that every  $k$ -uniform hypergraph with  $m \geq Cn(n/l)^{k/2-1}$  hyperedges for some absolute constant  $C > 0$  contains an even cover of size  $O(l \log n)$ . We believe that a strengthening of our techniques via some appropriate generalization of flower Kikuchi graphs might be enough to resolve this remaining slack. On the flip side, finding explicit constructions of hypergraphs with  $m = Cn(n/l)^{k/2-1}$  hyperedges that avoid  $c \ell \log n$  length even covers for some constant  $c > 0$  remains a natural open question.
- **Lower bounds on Locally Decodable Codes:** Theorem 2.8 shows an almost cubic lower bound on binary linear codes with constant distance and three query local decodability. This is still significantly far from the best known construction of three query LDCs due to Yekhanin and Efremenko that achieves a sub-exponential length. Finding methods that could prove a super-polynomial (or at least super-cubic) lower bounds is an outstanding open question. We note that in a recent work (also based on Kikuchi matrices) an exponential lower bound [16] (improving on the prior best cubic lower bounds above) was in fact obtained for the stronger setting of three query linear *Locally Correctable Codes* – a stricter variant of locally decodable codes where every bit of the codeword can be decoded by reading only 3-bits of the received corrupted codeword. Their methods, however, strongly exploit the additional structure offered by the local correction property.

## References

- [1] N. Alon, M. Bucić, L. Saueremann, D. Zakharov and O. Zamir. Essentially tight bounds for rainbow cycles in proper edge-colourings. arXiv:2309.04460, 2023.

- [2] N. Alon, U. Feige. On the power of two, three and four probes. *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*, 346–354, 2009.
- [3] N. Alon, S. Hoory and N. Linial. The Moore Bound for Irregular Graphs. *Graphs Comb.*, 18(1):53–57, 2002.
- [4] O. Alrabiah, V. Guruswami, P. Kothari and P. Manohar. A Near-Cubic Lower Bound for 3-Query Locally Decodable Codes from Semirandom CSP Refutation. *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, ACM, 1438–1448, 2023.
- [5] Z. Dvir. Incidence Theorems and Their Applications. arXiv:1208.5073, 2012.
- [6] Z. Dvir. Locally Decodable Codes. *Lecture Notes, Princeton University*, 2016.
- [7] K. Efremenko. 3-query locally decodable codes of subexponential length. *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, ACM, 39–44, 2009.
- [8] U. Feige. Small linear dependencies for binary vectors of low weight. *Building Bridges: Between Mathematics and Computer Science*, Springer, 283–307, 2008.
- [9] U. Feige, J. H. Kim and E. Ofek. Witnesses for non-satisfiability of dense random 3CNF formulas. *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings*, IEEE Computer Society, 497–508, 2006.
- [10] U. Feige and T. Wagner. Generalized girth problems in graphs and hypergraphs. *preprint*.
- [11] O. Goldreich, H. Karloff, L. Schulman and L. Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. *Computational Complexity*, 15(3):263–296, 2006.
- [12] V. Guruswami, P. Kothari and P. Manohar. Algorithms and certificates for Boolean CSP refutation: smoothed is no harder than random. *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, ACM, 678–689, 2022.
- [13] T. Hsieh, P. Kothari and S. Mohanty. A simple and sharper proof of the hypergraph Moore bound. *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, 2324–2344, 2023.
- [14] P. Keevash, D. Mubayi, J. Verstraëte and B. Sudakov. Rainbow Turán problems. *Combin. Probab. Comput.*, 16(1):109–126, 2007.
- [15] I. Kerenidis and R. Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *Journal of Computer and System Sciences*, 69(3):395–420, 2004.
- [16] P. Kothari and P. Manohar. An Exponential Lower Bound for Linear 3-Query Locally Correctable Codes. arXiv:2311.00558, 2023.
- [17] F. Lazebnik, V. Ustimenko and A. Woldar. A new series of dense graphs of high girth. *Bulletin of the AMS*, 32:73–79, 1995.
- [18] R. Montgomery. A proof of the Ryser-Brualdi-Stein conjecture for large even  $n$ . arXiv:2310.19779, 2023.
- [19] R. Montgomery, A. Pokrovskiy and B. Sudakov. A proof of Ringel’s conjecture. *Geom. Funct. Anal.*, 31(3):663–720, 2021.
- [20] A. Naor and J. Verstraëte. Parity check matrices and product representations of squares. *Combinatorica*, 28(2):163–185, 2008.
- [21] T. Sanders. On a theorem of Shkredov. *Online J. Anal. Comb.*, 5:4, 2010.
- [22] S. Yekhanin. Towards 3-query locally decodable codes of subexponential length. *Journal of the ACM (JACM)*, 55(1):1–16, 2008.
- [23] S. Yekhanin. Locally Decodable Codes and Private Information Retrieval Schemes. *Information Security and Cryptography*, Springer, 2010.
- [24] S. Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, Now Publishers, Inc., 6(3):139–255, 2012.