

Equicovering Subgraphs of Graphs and Hypergraphs

Ilkyoo Choi	Jaehoon Kim*	Amelia Tebbe
Mathematics Department University of Illinois Illinois, U.S.A.	Mathematics Department University of Illinois Illinois, U.S.A.	Mathematics Department University of Illinois Illinois, U.S.A.
<code>icho4@illinois.edu</code>	<code>kim805@illinois.edu</code>	<code>tebbe2@illinois.edu</code>

Douglas B. West[†]
Mathematics Department
Zhejiang Normal University and University of Illinois
Illinois, U.S.A.
`west@math.uiuc.edu`

Submitted: Dec 31, 2013; Accepted: Mar 9, 2014; Published: Mar 17, 2014
Mathematics Subject Classifications: 05C35, 05C65

Abstract

As a variation on the t -Equal Union Property (t -EUP) introduced by Lindström, we introduce the t -Equal Valence Property (t -EVP) for hypergraphs: a hypergraph satisfies the t -EVP if there are t pairwise edge-disjoint subhypergraphs such that for each vertex v , the degree of v in all t subhypergraphs is the same. In the t -EUP, the subhypergraphs just have the same sets of vertices with positive degree. For both the 2-EUP and the 2-EVP, we characterize the graphs satisfying the property and determine the maximum number of edges in a graph not satisfying it. We also study the maximum number of edges in both k -uniform and general hypergraphs not satisfying the t -EVP.

1 Introduction

We consider conditions for a hypergraph to have t edge-disjoint subhypergraphs whose local behavior at vertices is somehow “the same”. Such structures may be useful for

*Research partially supported by the Arnold O. Beckman Research Award of the University of Illinois at Urbana-Champaign.

[†]Research partially supported by NSA Award H98230-10-1-0363.

fault-tolerance of communication protocols. Lindström [12] introduced a precise notion of such a property. The *degree* (or *valence*) of a vertex v in a hypergraph H , written $d_H(v)$, is the number of edges containing v . The *span* of H , written $S(H)$, is the set of vertices with positive degree.

Definition 1.1. A hypergraph H satisfies the *t-Equal Union Property (t-EUP)* if H has t edge-disjoint distinct subhypergraphs H_1, \dots, H_t with the same span. If the spans are equal but H_1, \dots, H_t are not edge-disjoint, then H satisfies the *weak t-EUP*.

When the t -EUP holds, the union of the edges is the same for each subhypergraph. Note that the hypergraph whose edges are the lines in the Fano plane does not satisfy the 2-EUP. Lindström [12] studied the number of edges needed in an n -vertex hypergraph to guarantee the t -EUP. Tverberg [14] gave an algebraic proof of Lindström's result.

Theorem 1.2. [12] *n-vertex hypergraphs with more than $n(t-1)$ edges satisfy the t-EUP.*

A hypergraph is *k-uniform* if every edge has size k . For $t = 2$ and ordinary graphs (2-uniform hypergraphs) the bound is sharp, since when n is odd an n -cycle has n edges but does not satisfy the 2-EUP. On the other hand, the existence of an even cycle is clearly sufficient, and we show in Theorem 2.1 that a graph satisfies the 2-EUP if and only if it has a even cycle or has a component containing two odd cycles. The Fano plane is a 3-uniform hypergraph with seven vertices and seven edges that does not satisfy the 2-EUP, thereby providing another sharpness example.

It is natural to require more than equal span for the common behavior in each subgraph. We introduce the *t-Equal Valence Property (t-EVP)*, more restrictive than the t -EUP.

Definition 1.3. A hypergraph H satisfies the *t-Equal Valence Property (t-EVP)* if H has t edge-disjoint distinct subhypergraphs H_1, \dots, H_t such that $d_{H_1}(v) = \dots = d_{H_t}(v)$ for each $v \in V(H)$. If the valence condition holds but H_1, \dots, H_t are not edge-disjoint, then H satisfies the *weak t-EVP*.

Every hypergraph having t perfect matchings satisfies the t -EVP; this includes t -regular t -edge-colorable graphs and the hypergraphs of resolvable designs (such as affine planes).

Example 1.4. *Projective planes satisfy the 2-EUP but not the 2-EVP.* A sufficient (but not necessary) condition for the t -EUP is existence of a t -coloring of the edges so that each vertex has at least one incident edge of each color. Although the Fano plane has no such coloring, larger projective planes do.

A *projective plane* of order q is a $(q+1)$ -uniform hypergraph with $q^2 + q + 1$ points and $q^2 + q + 1$ edges (called *lines*), in which every two lines have exactly one common

point, every two points appear in exactly one common line, and there exist four points among which no three appear in one line. Such hypergraphs exist whenever q is a power of a prime.

Given a projective plane, let S be a fixed set of three points not on a common line. For each line L , give color a to L , where $|L \cap S| \equiv a \pmod 2$. It is easy to check that each point appears in lines of both colors.

On the other hand, if a projective plane of order q satisfies the 2-EVP, then we have vertices v_1, \dots, v_n (where $n = q^2 + q + 1$) and edge-disjoint subhypergraphs H and H' such that $d_H(v_i) = d_{H'}(v_i) = d_i$ for some numbers d_1, \dots, d_n . Since the hypergraph is uniform and the degree sum is the same, $|E(H)| = |E(H')|$; let H and H' both have s edges.

We claim first that $s^2 = \sum_{i=1}^n d_i^2$. Both sides equal $\sum |e \cap e'|$, summed over $e \in E(H)$ and $e' \in E(H')$. Each term equals 1, and there are s^2 terms. On the other hand, since v_i belongs to d_i edges in H and d_i edges in H' , each vertex v_i contributes exactly d_i^2 times.

Next, we claim that $s(s-1) = \sum_{i=1}^n d_i(d_i-1)$. Both sides equal $\sum |e_i \cap e_j|$, summed over ordered pairs of distinct edges in H . Again all terms equal 1, and there are $s(s-1)$ terms. Since vertex v_i appears in d_i edges, it contributes $d_i(d_i-1)$ times.

Subtracting now yields $s = \sum_{i=1}^n d_i$, but $\sum_{i=1}^n d_i = s(q+1)$, since each edge has size $q+1$. We conclude $q=0$, but there is no projective plane of order 0.

We study the extremal problem posed by Lindström and the analogous extremal problem for equal valence, considering also the restrictions to k -uniform hypergraphs. Let $U(n, t)$ denote the maximum number of edges in an n -vertex hypergraph not satisfying the t -EUP. Let $U_k(n, t)$ denote the maximum number of edges in a k -uniform n -vertex hypergraph not satisfying the t -EUP. Let $U'(n, t)$ and $U'_k(n, t)$ denote the analogous maximum numbers of edges for avoiding the weak t -EUP (we avoid trivialities by not allowing two subhypergraphs to have exactly the same edge set). For the t -EVP, use V instead of U to define the corresponding extremal problems. By definition, $A(n, t) \geq A_k(n, t)$ for $A \in \{U, U', V, V'\}$.

n -vertex graphs	t -EUP	weak t -EUP	t -EVP	weak t -EVP
k -uniform hypergraphs	$U_2(n, t)$	$U'_2(n, t)$	$V_2(n, t)$	$V'_2(n, t)$
general hypergraphs	$U_k(n, t)$	$U'_k(n, t)$	$V_k(n, t)$	$V'_k(n, t)$
	$U(n, t)$	$U'(n, t)$	$V(n, t)$	$V'(n, t)$

Like the 2-EUP, the 2-EVP has a simple characterization for graphs; we show in Theorem 2.3 that a graph satisfies the 2-EVP if and only if it contains an even circuit (a closed trail of even length). The characterizations in Section 2 yield $U_2(n, 2) = U'_2(n, 2) = n$ and $V_2(n, 2) = V'_2(n, 2) = \lfloor \frac{4}{3}n \rfloor - 1$. In Section 3, we prove $V'_2(n, 3) = \lfloor \frac{4}{3}n \rfloor$ and $\left(t - 1 + \frac{1}{2(t-1)}\right)(n-1) - (t-2)^2 \leq V_2(n, t) \leq (2 + o(1))(t-1)n$. In Section 4, we consider k -uniform hypergraphs. We prove $V_k(n, 2) \in O(n)$ and $V(n, 2) \in O(n \log n)$. We also prove $V(n, t), V_k(n, t) \in O\left(n^2 \left(\frac{\log n}{\log \log \log n}\right)^2\right)$ and obtain a lower bound for $V_k(n, 2)$.

We close this introduction with several simple observations used in later proofs.

Observation 1. *The weak 2-EVP and the 2-EVP are equivalent. Thus $V(n, 2) = V'(n, 2)$, and $V_k(n, 2) = V'_k(n, 2)$ for $k \geq 2$.*

Proof. If subhypergraphs H_1 and H_2 witness the weak 2-EVP, then $H_1 - E(H_1) \cap E(H_2)$ and $H_2 - E(H_1) \cap E(H_2)$ witness the 2-EVP, since each vertex loses the same number of incident edges from H_1 and H_2 . \square

Observation 2. $U'(n, t) < n + \log_2 t$ and $U'_k(n, t) < \log_2 \left(\sum_{i=k}^n \binom{n}{i} \right) + \log_2 t$.

Proof. A hypergraph with m edges has 2^m distinct subhypergraphs. An n -vertex hypergraph H with $\lceil n + \log_2 t \rceil$ edges has at least $t2^n$ distinct subhypergraphs. Since there are only 2^n choices for the span of a subhypergraph of H , by the pigeonhole principle there are t distinct subhypergraphs with the same span. Thus H satisfies the weak t -EUP.

The second statement uses the fact that the span of nonempty k -uniform subhypergraphs has size at least k . \square

Observation 3. $V_k(n + 1, t) \geq V_{k-1}(n, t)$ for $k \geq 3$.

Proof. Let H be a $(k - 1)$ -uniform hypergraph not satisfying the t -EVP. Form H' from H by adding one vertex x belonging to every edge. If H'_1, \dots, H'_t witness the t -EVP in H' , then deleting x from each edge in each H'_i yields H_1, \dots, H_t witnessing the t -EVP in H . \square

Observation 4. *If H_1 and H_2 have equal degree at every vertex of G , then their symmetric difference contains no cut-edge of G .*

Proof. Let e be a cut-edge of G contained in H_1 but not H_2 . Let G' be a component of $G - e$. The sum of the degrees of the subgraphs of H_1 and H_2 in G' differ by 1. Hence one of them is a graph with odd degree sum, which does not exist. \square

Observation 5. *If H_1, \dots, H_t witness the weak t -EVP in G , with degree 0 at v or degree $d_G(v)$ at v , then deleting v yields graphs H'_1, \dots, H'_t witnessing the weak t -EVP in $G - v$.*

Proof. Since the usage of edges incident to v is the same in each H_i , ignoring those edges also contributes the same degree from all t subgraphs at each of the remaining vertices. \square

2 The 2-EUP and the 2-EVP for Graphs

Our graphs have no repeated edges or loops. A *walk* in a graph is a list of vertices in which every two consecutive vertices are adjacent (edges may repeat). A *trail* is a walk that traverses each edge at most once (vertices may repeat). A walk or trail is *closed* if its first and last vertices are the same. A *circuit* is a closed trail. A *cycle* in a graph is a connected subgraph in which every vertex has degree 2. The *length* of a walk, trail, circuit, or cycle is the number of edges it traverses, and the object is *even* or *odd* as its length is even or odd.

We characterize the graphs satisfying the 2-EUP (Theorem 2.1) and the graphs satisfying the 2-EVP (Theorems 2.3–2.4). From the characterizations, we determine $U_2(n, 2)$, $U'_2(n, 2)$, $V_2(n, 2)$, and $V'_2(n, 2)$.

Theorem 2.1. *A graph G satisfies the 2-EUP if and only if G has an even cycle or has a component containing two odd cycles.*

Proof. (Sufficiency) Alternating edges along an even cycle gives two subgraphs witnessing the 2-EUP. Now, let G have a component containing odd cycles C_1 and C_2 . If C_1 and C_2 share an edge, then G has an even cycle. Otherwise, let P be a shortest path joining $V(C_1)$ and $V(C_2)$, with endpoints $u \in V(C_1)$ and $v \in V(C_2)$. Let T be a trail that starts at u , follows C_1 , moves from u to v through P , and follows C_2 . Alternating edges along T gives two subgraphs with vertex set $V(C_1) \cup V(C_2) \cup V(P)$.

(Necessity) Let G_1 and G_2 be two subgraphs of G witnessing the 2-EUP, and take $u_0u_1 \in E(G_1) - E(G_2)$. Since u_1 must be covered by both $E(G_1)$ and $E(G_2)$, there exists $u_1u_2 \in E(G_2)$. Iteratively, we find $u_iu_{i+1} \in E(G_1)$ and $u_{i+1}u_{i+2} \in E(G_2)$ for even i . Since G is finite, $u_k = u_{k+r}$ for some r and k . Taking the first such repetition, u_k, \dots, u_{k+r-1} form a cycle. If r is even, then G has an even cycle.

If r is odd, then by symmetry we may assume that u_ku_{k+1} and u_ku_{k+r-1} are both in $E(G_1)$. Since G_1 and G_2 have the same span, we find v_1 such that $u_kv_1 \in E(G_2)$. We now generate v_1, v_2, \dots as above, with $v_jv_{j+1} \in E(G_1)$ and $v_{j+1}v_{j+2} \in E(G_2)$ for odd j , until we obtain $v_l = v_{l+s}$ or $v_l = u_s$ for some l and s with $s > k$; consider the first such occurrence. (In particular, we ignore u_0, \dots, u_{k-1} , as if they had never been defined.)

If $v_l = v_{l+s}$, then v_l, \dots, v_{l+s-1} form an even cycle when s is even. If s is odd, then $[v_l, \dots, v_{l+s-1}]$ and $[u_k, \dots, u_{k+r-1}]$ form two odd cycles in the same component of G .

If $v_l = u_s$, then the edges we have discovered form three internally disjoint paths joining u_k and u_s . The union of some two of them is an even cycle. \square

Corollary 2.2. *If a graph G does not satisfy the 2-EUP, then $|E(G)| \leq n$. Equality holds if and only if every component of G has at most one cycle and all cycles in G have odd length. Thus $U_2(n, 2) = n$.*

A *nontrivial graph* is a graph having an edge.

Theorem 2.3. *A graph G satisfies the 2-EVP if and only if it has an even circuit.*

Proof. (Sufficiency) If G has an even circuit, then alternating edges along it provides two subgraphs witnessing the 2-EVP.

(Necessity) Let H_1 and H_2 be two subgraphs witnessing the 2-EVP in G . Since H_1 and H_2 have the same number of edges and have equal degree at each vertex, their union H is a nontrivial graph whose components have an even number of edges and have even degree at each vertex. Hence H decomposes into cycles. If any such cycle has even length, then G contains an even circuit.

Otherwise, all cycles in the decomposition of a nontrivial component of H have odd length, and there are an even number of them. Since a connected union of cycles has no cut-edge, two such odd cycles C and C' are joined by two edge-disjoint paths P and P' in H . The endpoints of P and P' on C are connected by two paths of opposite parity along C , and the same is true along C' . Hence the union of P , P' , and appropriate paths connecting them in C and C' is a subgraph traversed by an even circuit. \square

A *cactus* is a connected graph in which every edge is in at most one cycle. A *strict odd cactus* is a cactus in which every cycle has odd length and no two cycles share a vertex.

Theorem 2.4. *A graph G fails to satisfy the 2-EVP if and only if every component is a strict odd cactus.*

Proof. (Sufficiency) Theorem 2.3 applies, since a strict odd cactus contains no even circuit.

(Necessity) We use induction on n , where $n = |V(G)|$. Let G be an n -vertex graph not satisfying the 2-EVP. For $n \leq 3$, the statement is true. For larger n , since every tree is a strict odd cactus, we may assume that G has a cycle C ; avoiding the 2-EVP requires C to be odd. If $G = C$, then G is a strict odd cactus. Otherwise, both $G - V(C)$ and C have fewer than n vertices and do not satisfy the 2-EVP. By the induction hypothesis, each component of $G - V(C)$ is a strict odd cactus. If at most one edge joins $V(C)$ to any such component H , then again every component of G is a strict odd cactus.

Finally, consider edges xy and $x'y'$ with $x, x' \in V(C)$ and $y, y' \in V(H)$. Combining these two edges with a y, y' -path in H yields an x, x' -path that is disjoint from C except for its endpoints. Now adding one of the paths from x to x' along C completes an even circuit, witnessing the 2-EVP for G . Hence this possibility does not occur, and every component of G is a strict odd cactus. \square

Theorem 2.5. *If an n -vertex graph G does not satisfy the 2-EVP, then $|E(G)| \leq \lfloor \frac{4}{3}n \rfloor - 1$. Also, the bound is sharp, so $V_2(n, 2) = \lfloor \frac{4}{3}n \rfloor - 1$.*

Proof. By Theorem 2.4, every component of G is a strict odd cactus. A largest such graph on n vertices is connected. If G has k cycles, then $k \leq \lfloor n/3 \rfloor$, since each cycle has length at least 3 and the cycles share no vertices. Also, deleting one edge from each cycle yields a tree. Hence $|E(G)| = n - 1 + k \leq \lfloor \frac{4}{3}n \rfloor - 1$. Equality holds for any strict odd cactus having $\lfloor n/3 \rfloor$ triangles and j vertices not in triangles, where $n \equiv j \pmod{3}$ with $j \in \{0, 1, 2\}$. \square

3 The t -EVP for Graphs

In this section we study $V_2(n, t)$. We give an upper bound in Theorem 3.5 using a theorem of Alon, Friedland, and Kalai [3]. In Theorem 3.6, we construct a sequence of graphs to provide a lower bound. We close the section by determining $V_2'(n, 3)$ exactly in Theorem 3.12. We begin with two well-known facts.

Fact 3.1. Every graph G has a bipartite subgraph H with $|E(H)| \geq \frac{1}{2}|E(G)|$.

Fact 3.2. [5] For $k \geq 2$, there exists a prime number p such that $k \leq p < 2k$.

For $q \in \mathbb{N}$, a q -divisible graph is a graph where every vertex degree is a multiple of q .

Theorem 3.3. [3] *If q is a prime power, then every n -vertex graph G with $|E(G)| \geq (q - 1)n + 1$ contains a q -divisible subgraph.*

Lemma 3.4. *Every t -divisible bipartite graph G satisfies the t -EVP.*

Proof. For $v \in V(G)$, let $c(v) = d_G(v)/t$. Form a graph G' by expanding each vertex $v \in V(G)$ into an independent set of size $c(v)$ in which each vertex inherits t of the edges incident to v . The graph G' is t -regular and bipartite, so G' decomposes into t edge-disjoint perfect matchings. These perfect matchings correspond to t edge-disjoint subgraphs H_1, \dots, H_t of G , each having degree $c(v)$ at v . Thus H_1, \dots, H_t witness the t -EVP for G . \square

Theorem 3.5. *If t is a prime power, then $V_2(n, t) \leq 2(t - 1)n$. In general, $V_2(n, t) \leq 2.4(t - 1)n$ and $V_2(n, t) \leq (2 + o(1))(t - 1)n$.*

Proof. Let G be an n -vertex graph. If $|E(G)| > m$ implies that G satisfies the t -EVP, then $V_2(n, t) \leq m$.

Suppose first that t is a prime power and $|E(G)| > 2(t - 1)n$. By Fact 3.1, G has a spanning bipartite subgraph H such that $|E(H)| > (t - 1)n$. By Theorem 3.3, H contains a t -divisible subgraph Q . By Lemma 3.4, Q satisfies the t -EVP, and hence G does also.

Consider general t and $|E(G)| > 2.4(t - 1)n$. By the preceding paragraph, it suffices to find a prime power q between t and $1.2t$. By Fact 3.2 (Bertrand's Postulate), there is a

prime number between t and $2t - 1$. For $t \geq 25$, Nagura [13] proved that there is a prime between t and $1.2t$. For $2 \leq t \leq 24$, there fails to be a prime in the desired range for $t \in \{4, 8, 14, 24\}$. However, for these values there is a prime power in the desired range.

Meanwhile, Iwaniec and Pintz [8] proved that there exists $t_0 \in \mathbb{N}$ such that for $x > t_0$, there is a prime between $x - x^{23/42}$ and x . This yields $V_2(n, t) \leq (2 + o(1))(t - 1)n$ in the same way as above. The value $23/42$ was further reduced to $.525$ in [4]. \square

Alon, Friedland, and Kalai [3] conjectured that the conclusion of Theorem 3.3 holds even when q is not a power of a prime. This would improve the upper bound on $V_2(n, t)$ from $2.4(t - 1)n$ to $2(t - 1)n$.

A graph is k -degenerate if every subgraph has a vertex of degree at most k . In a $(t - 1)$ -degenerate graph, every nontrivial subgraph has a vertex that cannot occur in the span of t edge-disjoint subgraphs, so $(t - 1)$ -degenerate graphs cannot satisfy the t -EVP. There is an n -vertex $(t - 1)$ -degenerate graph with as many as $\binom{t}{2} + (t - 1)(n - t)$ edges, so $V_2(n, t) \geq (t - 1)(n - t/2)$. We improve the leading coefficient.

Theorem 3.6. $V_2(n, t) \geq (t - 1 + \frac{1}{2t-2})(n - 1) - (t - 2)^2$.

Proof. The k -fold wheel $W_{k,r}$ is obtained from the complete bipartite graph $K_{k,r}$ by adding the edges of a cycle through the r vertices of degree k , giving them degree $k + 2$. Call the vertices of degree r centers, the vertices of degree $k + 2$ rim vertices, and the edges joining central and rim vertices spokes. In any copy of $W_{k,r}$, specify two distinct rim vertices as the head and tail. Note that $W_{k,r}$ has $k + r$ vertices and $r(k + 1)$ edges.

For $t \geq 2$, let B_1, \dots, B_ℓ be copies of $W_{t-2,t+1}$. Let z_i denote the tail of B_i (and z_0 denote the head of B_1). For $1 \leq i < \ell$, merge z_i with the head of B_{i+1} . Add edges joining all centers of B_i to all centers of B_{i+1} , for $1 \leq i < \ell$. Let G_t^ℓ be the resulting graph.

Note that G_t^ℓ has $\ell(2t - 2) + 1$ vertices and $\ell(t + 1)(t - 1) + (\ell - 1)(t - 2)^2$ edges. Writing $n = \ell(2t - 2) + 1$, the number of edges becomes $(t - 1 + \frac{1}{2t-2})(n - 1) - (t - 2)^2$. Hence it suffices to show that G_t^ℓ does not satisfy the t -EVP.

Suppose that H_1, \dots, H_t are subgraphs of G_t^ℓ witnessing the t -EVP. Let $H = H_1 \cup \dots \cup H_t$; each vertex degree in H is a multiple of t . Say that the edges of H are “used”. The edges of G_t^ℓ that are not spokes form a $(t - 2)$ -degenerate graph, so some spoke must be used.

Each rim vertex other than z_0, \dots, z_ℓ has degree t . Let i be the least index such that some spoke in B_i is used. If only spokes incident to z_i are used, then some center of B_i has degree between 1 and $t - 1$. Hence some spoke not incident to z_i is used. Since rim vertices have degree t in B_i and form a connected subgraph, it follows that all edges of B_i are used except possibly some spokes incident to z_i .

If any spoke incident to z_i is used, then its other endpoint has degree between $t + 1$ and $2t - 1$ in H ($t + 1$ rim vertices and up to $t - 2$ centers in B_{i+1}). Hence no spoke of B_i at z_i is used, but all other edges of B_i are used.

Now z_i has two neighbors in H that are rim vertices of B_i , plus up to t neighbors in B_{i+1} . Hence $d_H(z_i) = t$. This requires the rim edges of B_i at z_i to belong to H_a and H_b with $a \neq b$. The remaining vertices of B_i have neighbors only in $V(B_i)$ in H , and hence H_a must form a matching there. However, $|V(B_i)| = 2t - 1$, so no such matching is possible.

When $n \not\equiv 1 \pmod{2(t-1)}$, we may add one small complete graph and possibly one copy of K_t to G_t^ℓ (where $\ell = \lfloor n/(2t-2) \rfloor$) to reach n vertices. \square

The remainder of this section determines $V'_2(n, 3)$. We start with the lower bound.

Lemma 3.7. *A graph containing at most one even circuit does not satisfy the weak 3-EVP.*

Proof. Suppose that G satisfies the weak 3-EVP, witnessed by three distinct nontrivial subgraphs H_1, H_2, H_3 having equal degrees at each vertex. Since H_1 and H_2 are distinct and have the same vertex degrees, the symmetric difference of H_1 and H_2 is nontrivial, has even degree at each vertex, and has an even number of edges in each component (the same number from both H_1 and H_2). Hence each nontrivial component can be expressed as an even circuit. The same is true for H_1 and H_3 , which cannot have the same components since H_2 and H_3 are distinct. Hence G contains at least two even circuits. \square

Lemma 3.8. $V'_2(n, 3) \geq \lfloor \frac{4}{3}n \rfloor$ for $n > 3$.

Proof. Let j be the congruence class of n modulo 3, with $j \in \{0, 1, 2\}$. Let G be a maximal strict odd cactus with n vertices. As computed in Theorem 2.4, G has $\lfloor \frac{4}{3}n \rfloor - 1$ edges and no even circuit; also G has $\lfloor n/3 \rfloor$ triangles and j vertices not in triangles. Further restrict G when $n \equiv 2$ so that the vertices not in triangles are leaves with a common neighbor.

In each case, we add one edge to G . When $j = 1$, add one edge joining the vertex v that lies in no triangle to a vertex of a nearest triangle; this creates one even circuit (of length 4). When $j = 2$, add one edge joining the two leaves not in triangles, creating one even circuit (of length 6). By Lemma 3.7, the resulting graph does not satisfy the weak 3-EVP.

We construct an explicit example G' for $n = 6$, consisting of disjoint triangles on abc and xyz plus edges ax and by . Every proper induced subgraph of G' has at most one even circuit and does not satisfy the weak 3-EVP, by Lemma 3.7. Hence by Lemma 5 subgraphs witnessing it have degree 1 at c and z and have degree 1 or 2 at the other

vertices. If all have degree 1, then the subgraphs are perfect matchings, but G' has only two perfect matchings. If all have degree 2, then they omit a perfect matching, and again there are only two.

Hence we may assume that exactly two vertices have degree 2. By symmetry, consider cases for the degrees of x and y , with $d(x) \geq d(y)$. If $d(x) > d(y)$, then xz must appear in each subgraph. By reasoning as in Lemma 5, $G' - z$ then satisfies the weak 3-EVP, which it does not. If $d(x) = d(y)$, then in each case a subgraph with the specified degrees is determined by choosing xz or yz , so again there are only two.

For larger multiples of 3, we use this 6-vertex subgraph G' plus a maximal strict odd cactus \hat{G} on $n - 6$ vertices joined to G' by a single cut-edge e . If there exist H_1, H_2, H_3 witnessing the weak 3-EVP in G , then any two of them have equal degrees at all vertices. By Observation 4, the symmetric difference of any two cannot contain the cut-edge e . Hence all three agree on e . Hence $G - e$ satisfies the weak 3-EVP. This requires the weak 3-EVP to hold in \hat{G} or G' , which it does not. \square

The upper bound needs several lemmas.

Lemma 3.9. *Let G be a graph.*

(a) *If G has $\lceil \log_2 t \rceil$ edge-disjoint even circuits, then G satisfies the weak t -EVP.*

(b) *If G has three edge-disjoint trails with the same endpoints whose lengths have the same parity, then G satisfies the weak 3-EVP.*

Proof. (a) Let $s = \lceil \log_2 t \rceil$, so $2^s \geq t$, and let C_1, \dots, C_s be the given edge-disjoint even circuits. There are two ways to take alternating edges on each circuit. Hence at least t distinct subgraphs of $C_1 \cup \dots \cup C_s$ have the same degrees at all vertices.

(b) Consider three edge-disjoint u, v -trails. If all have even length, then take alternating edges in each trail, starting at u in two of the trails. Each way of doing this yields the same degree at all vertices. The three ways to do it yield the weak 3-EVP.

If all have odd length, then the union of two form an even circuit. Take alternating edges from the other so that u and v are not covered, plus alternating edges along the circuit. The three ways to choose a pair of trails and two ways to choose edges along the circuit yield six subgraphs having the same degree at all vertices. Thus G satisfies the weak 6-EVP. \square

Lemma 3.10. *If G has a path through w, x, y, z such that $d_G(x) = d_G(y) = 2$, then G satisfies the weak t -EVP if and only if G' satisfies the weak t -EVP, where $G' = G - \{x, y\} \cup wz$.*

Proof. If H'_1, \dots, H'_t witness the t -EVP in G' , then modify each H'_i as follows. If $wz \in E(H'_i)$, then put $wx, yz \in E(H_i)$ and $xy \notin E(H_i)$, leaving the usage of all other edges

the same. If $wz \notin E(H'_i)$, then put $xy \in E(H_i)$ and $wx, yz \notin E(H_i)$, again leaving other edges as in H'_i . In each H_i , the degree is 1 at x and at y , and at all other vertices it is the same as in H'_i , so H_1, \dots, H_t witness the t -EVP in G' .

For the converse, let H_1, \dots, H_t witness the t -EVP in G . If each subgraph has degree 1 at both x and y , then the transformation above can be reversed. If the subgraphs all have degree 0 or degree 2 at x or y , then they all have the same usage of each edge in $\{wx, xy, yz\}$. Now letting $H'_i = H_i - \{x, y\}$ for all i yields subgraphs witnessing the t -EVP in G' (they all omit the edge wz). \square

We apply Lemma 3.10 to subdivisions of K_4 that arise in the proof of the upper bound.

Lemma 3.11. *Let G be a subdivision of K_4 , and call the paths joining branch vertices “threads”. If G has an odd cycle C through three branch vertices, and the cycles through exactly two threads of C all have even length, then G satisfies the weak 3-EVP.*

Proof. By Lemma 3.10, we may assume that every thread has length 1 or 2.

When all threads in C have length 1, the three threads incident to the remaining vertex z all have the same parity. If they are single edges, then use the decomposition of K_4 into three matchings M_1, M_2, M_3 . Otherwise, the threads at z each have length 2. Each subgraph H_i uses one edge incident to z and a path of length 4 joining the other two neighbors of z .

If only one thread in C has odd length, then the lengths of the two threads from its endpoints to z have the same parity. If they have length 1, then we can treat that cycle as C above. If they have length 2, then the other thread at z has length 1; now exactly two threads (with no shared endpoints) have odd length, and our graph G is obtained from K_4 by subdividing the edges of a 4-cycle.

The 8-cycle has two crossing chords; call them uv and $u'v'$. Pairing uv with each of the two perfect matchings in the 8-cycle generates two subgraphs with the same vertex degrees. A third such subgraph uses $u'v'$ and the edges at u and v other than uv . Hence G satisfies the weak 3-EVP. \square

Theorem 3.12. $V'_2(n, 3) = \lfloor \frac{4}{3}n \rfloor$. *That is, every n -vertex graph with more than $\lfloor \frac{4}{3}n \rfloor$ edges satisfies the weak 3-EVP, and this is sharp.*

Proof. We need only prove the upper bound. Let G be an n -vertex graph with $\lfloor \frac{4}{3}n \rfloor + 1$ edges. By using the component with largest average degree, we may take G to be connected.

Let G' be a subgraph of G with the most edges that contains no even circuit. Adding a cut-edge joining components does not create a circuit. Hence G' is a strict odd cactus

spanning all of $V(G)$, by Theorem 2.3. By Theorem 2.4 and Theorem 2.5, $|E(G')| \leq \lfloor \frac{4}{3}n \rfloor - 1$. Thus G' omits at least two edges of G ; choose $e_1, e_2 \in E(G) - E(G')$. By the definition of G' , adding e_i to G' creates an even circuit C_i . If C_1 and C_2 share no edge, then G satisfies the weak 4-EVP, by Lemma 3.9(a). Thus, we may assume that they share an edge.

Shrinking the odd cycles of G' into vertices yields a tree T . Call the vertices of T *nodes* to distinguish them from vertices of G' ; they correspond to vertices or odd cycles in G' . Note that $C_1 - e_1$ and $C_2 - e_2$ shrink to paths P_1 and P_2 in T . Since C_1 and C_2 are not edge-disjoint, P_1 and P_2 intersect in T . If the intersection is one node of T , then C_1 and C_2 intersect in a single nontrivial path in G . Since both circuits have even length, the endpoints of this path are joined by three edge-disjoint trails of the same parity. Hence G satisfies the weak 3-EVP, by Lemma 3.9(b).

Hence we may assume that P_1 and P_2 intersect in a nontrivial path P' in T ; let u and w be its endnodes. Let U and W be the subgraphs of G' corresponding to u and w , they may be single vertices or odd cycles in G' . Let Q_0 be the subgraph of G' corresponding to P' . Let x_0 and y_0 be the vertices of U and W with largest degree in Q_0 , and let Q'_0 be an x_0, y_0 -path through Q_0 .

For $i \in \{1, 2\}$, let Q_i be the trail obtained from C_i by deleting $E(Q_0)$; note that $e_i \in E(Q_i)$. Let x_i and y_i be the endpoints of Q_i in U and W , respectively.

If Q_2 visits an odd cycle \hat{C} in G' , then let x and y be the vertices in U and W that are endpoints of $C_2 - E(C_1)$. By traversing \hat{C} appropriately, we obtain a trail whose length has the same parity as the two x, y -trails along C_1 , and Lemma 3.9(b) applies. Hence we may assume that Q_2 (and similarly Q_1) visits no such cycle.

Now combining Q_1 and Q_2 with appropriate traversals through U and W may yield an even circuit containing Q_1 and Q_2 . If this happens, and Q'_0 visits an odd cycle in G' , then again Lemma 3.9(b) applies.

Hence two cases remain. In one case, $x_0 = x_1 = x_2$ and $y_0 = y_1 = y_2$, these are the only vertices of U and W , and Q_1 and Q_2 are paths with opposite parity. Now $Q_1 \cup Q_2$ is an odd cycle that shares vertices with no odd cycles in G' . Letting e' be the edge of Q'_0 incident to U , we now have $G' - e' \cup \{e_1, e_2\}$ as a subgraph of G having no even circuit, contradicting the maximality of G' .

In the remaining case, the interior nodes of P' are single vertices in G' , not odd cycles. The x_1, x_0 -path through U in C_1 and the x_2, x_0 -path through U in C_2 may or may not overlap. If they overlap, then taking the complement within U for each of these paths makes them disjoint but changes the parity of the circuits. However, if the same thing happens within W , then performing the complementations at both ends of P' leaves us with two even cycles C'_1 and C'_2 in G whose intersection is the single path P' , and Lemma 3.9(b) applies. The same holds if complementation was not needed on either end.

We are left with the case where C_1 and C_2 both contain the x_1, x_2 -path through U that

avoids x_0 , while in W they contain the edge-disjoint y_0, y_1 -path and y_0, y_2 -path. Deleting the edges and interior vertices of the y_1, y_2 -path (if any exist), we obtain a subdivision of K_4 . Three of the threads form the odd cycle U . Each of C_1 and C_2 is a union of four threads in this subdivision with total length even, omitting one of the threads in U . It follows that the circuit in the subgraph obtained by omitting the remaining thread in U also has even length. Hence Lemma 3.11 applies, and G satisfies the weak 3-EVP. \square

4 The Weak t -EVP and t -EVP for Hypergraphs

In this section, we obtain further upper bounds. For the case $t = 2$, recall that the 2-EVP and weak 2-EVP are equivalent. Write \lg for \log_2 .

Theorem 4.1. *If $k \geq 3$, then $V'_k(n, t) < (\lg k + 3 \lg \lg k)n + \lg(t - 1)$. In fact, more than this many edges forces the weak t -EVP for any hypergraph whose average edge-size is at most k .*

Proof. It suffices to prove that an n -vertex hypergraph H with m edges of average size at most k satisfies the weak t -EVP, where $m = \lceil (\lg k + 3 \lg \lg k)n + \lg(t - 1) \rceil$. Let $V(H) = \{v_1, \dots, v_n\}$, and let $d_i = d_H(v_i)$. For every subhypergraph H' , we have $d_{H'}(v_i) \leq d_i$. Hence there are at most $\prod_{i=1}^n (d_i + 1)$ degree lists of subhypergraphs. Since $\sum_{i=1}^n d_i = km$, it follows that $\prod_{i=1}^n (d_i + 1) \leq \left(\frac{km}{n} + 1\right)^n$.

With m edges in H , there are 2^m subhypergraphs of H . If $2^m > (t - 1) \left(\frac{km}{n} + 1\right)^n$, then some t subhypergraphs have the same degree list. With $u = m/n$, we need $2^u > (t - 1)^{1/n}ku + 1$. It suffices to have $u \geq \lg k + 3 \lg \lg k + \lg(t - 1)$. \square

The constant 3 can be decreased to a constant $1 + c_k$ with $c_k > 0$ by carefully considering lower-order terms. As $k \rightarrow \infty$, actually $c_k \rightarrow 0$, but quite slowly. For $t = 2$, still $c_k > 0.01$ when $k = 10^{42}$. Below we show c_k and m for small values of k when $t = 2$.

k	3	4	5	6	7	8
c_k	1.9389	1.1314	0.8569	0.7162	0.6296	0.57041
m	$< 3.5377n$	$< 4.1314n$	$< 4.5787n$	$< 4.9364n$	$< 5.23417n$	$< 5.48904n$

When we say nothing about the average size of edges, Theorem 4.1 still gives an upper bound by setting $k = n$.

Corollary 4.2. $V'(n, t) < n(\lg n + 3 \lg \lg n) + \lg(t - 1)$.

The upper bound in Theorem 4.1 is not valid for $V_k(n, t)$ when $t \geq 3$ because the t subhypergraphs found with the same degree lists need not have the same pairwise

intersections. This means there is no set of edges to discard from them all that preserves equality of degrees. To avoid this problem, we use the fact that large families of sets contain large subfamilies with common pairwise intersections.

Definition 4.3. A family of r sets is an r - Δ -system if the intersections of any two sets in the family are the same. Let $f(k, r)$ be the least integer q such that every k -uniform family of q sets contains an r - Δ -system. Let $g(m, r)$ be the least integer q such that every family of q subsets of an m -element set contains an r - Δ -system.

Theorem 4.4 (Kostochka [11]). *For $r \geq 3$ and $\alpha > 1$, there exists $D(r, \alpha)$ such that $f(k, r) \leq D(r, \alpha)k! \left(\frac{(\log \log \log k)^2}{\alpha \log \log k}\right)^k$.*

Theorem 4.5 (Erdős and Szemerédi [7]). $g(m, 3) \leq 2^{(m-1)\sqrt{m}}$.

In proving this upper bound, Erdős and Szemerédi used $f(k, 3) \leq (1 + o(1))k!$, which was the best known upper bound at the time. Kostochka's later improvement in Theorem 4.4 yields a better upper bound for $g(m, r)$ when r is fixed and m is large. The proof is the same as that in [7], just invoking the improved bound on $f(k, r)$ at the appropriate point.

Lemma 4.6. *For $r \geq 3$ and $\epsilon > 0$, there exists $N(r, \epsilon)$ such that*

$$g(m, r) < 2^{m-(1-\epsilon)\sqrt{m}} \log \log \log m$$

for $m \geq N(r, \epsilon)$.

Theorem 4.7. *If $t \geq 3$ and $\epsilon > 0$, then $V(n, t) < (4 + \epsilon)n^2 \left(\frac{\log n}{\log \log \log n}\right)^2$ for $n \geq N(t, \epsilon)$, where $N(t, \epsilon)$ is defined as in Lemma 4.6.*

Proof. Let H be an n -vertex hypergraph with $|E(H)| \geq (4 + \epsilon)n^2 \left(\frac{\log n}{\log \log \log n}\right)^2$ and $n \geq N(t, \epsilon)$. We show that H satisfies the t -EVP.

As in Theorem 4.1, let $V(H) = \{v_1, \dots, v_n\}$ and $d_i = d_H(v_i)$. For every subhypergraph H' , we have $d_{H'}(v_i) \leq d_i$. Again there are at most $\prod_{i=1}^n (d_i + 1)$ degree lists of subhypergraphs. Since $0 \leq d_i \leq m$ for all i , it follows that $\prod_{i=1}^n (d_i + 1) \leq (m + 1)^n$.

With m edges in H , there are 2^m subhypergraphs of H . Let $q = \left\lceil \frac{2^m}{(m+1)^n} \right\rceil$; there are distinct subhypergraphs H_1, \dots, H_q having the same degree list.

Since $m = (4 + \epsilon)n^2 \left(\frac{\log n}{\log \log \log n}\right)^2$, Theorem 4.6 yields

$$\begin{aligned} \log q &> m - n \log(m + 1) > m - n \log \left(n^2 \left(\frac{\log n}{\log \log \log n} \right)^2 \right) \\ &= m - 2n \log n - n \log \left(\left(\frac{\log n}{\log \log \log n} \right)^2 \right) \\ &> m - (1 - \epsilon)\sqrt{m} \log \log \log m - o(\sqrt{m}) \\ &> m - (1 - \epsilon)\sqrt{m} \log \log \log m > \log g(m, t). \end{aligned}$$

Thus, $q \geq g(m, t)$, which implies that among H_1, \dots, H_q , each viewed as a subset of the set of m edges in H , there exists a Δ -system of size t . That is, we obtain t subhypergraphs that are pairwise disjoint except for the same set of edges shared by all of them. Deleting that common subset yields t subhypergraphs witnessing the t -EVP. \square

The same argument yields a bound when the average size of edges is at most k . As in Theorem 4.1, use $\prod_{i=1}^n (d_i + 1) \leq \binom{km}{n} + 1$.

Theorem 4.8. *If $t \geq 3$, $k \geq 2$, and $\epsilon > 0$, then $V_k(n, t) < (1 + \epsilon)n^2 \left(\frac{\log n}{\log \log \log n}\right)^2$ for $n \geq \max\{k, N(t, \epsilon)\}$.*

If the upper bound in Lemma 4.6 can be improved to $(2 - \epsilon_r)^n$, then we obtain an upper bound of $C_t n \log k$ for Theorem 4.8 and an upper bound of $C_t n \log n$ for Theorem 4.7, where C_t is a constant depending on t .

Erdős and Rado conjectured that $f(n, t) \leq (c_1(t))^n$. If this conjecture is true, then in Lemma 4.6 it yields $g(n, t) \leq 2^{n - c\sqrt{n} \log n}$, and in Theorem 4.7 it yields $V(n, t) \leq c_2(t)n^2$, where $c_1(t)$ and $c_2(t)$ are functions of t .

We close with a simple lower bound.

Lemma 4.9. *Let H' and H'' be disjoint k -uniform hypergraphs. Fix s with $s \leq \lg k$, and form a hypergraph H by adding s edges to $H' \cup H''$ so that the i th added edge has exactly 2^{i-1} vertices in $V(H')$. If both H' and H'' fail the 2-EVP, then also H fails the 2-EVP.*

Proof. Let F be the set of s added edges. Since H' and H'' fail the 2-EVP, a pair $\{H_1, H_2\}$ witnessing the 2-EVP in H must use some edge in F . Since H_1 and H_2 share no edges, and the contributions by edges of F to the total degree in $V(H')$ are distinct powers of 2, the contributions by $E(H_1) \cap F$ and $E(H_2) \cap F$ to the total degree in $V(H')$ are distinct. Since $\sum_{i=1}^{\lfloor \lg k \rfloor} 2^{i-1} < k$, the difference between the two amounts is less than k . However, the total degree contributed by edges in H' is a multiple of k (for both H_1 and H_2). Hence the total degrees in $V(H')$ for H_1 and H_2 are distinct, so they cannot witness the 2-EVP. \square

Corollary 4.10. *If $k \geq 3$, then $V_k(n + k - 3, 2) \geq \frac{11n-31}{7}$.*

Proof. By Observation 3, it suffices to prove $V_3(n, 2) \geq \frac{11n-31}{7}$.

By exhaustive computer search, we know that the 7-vertex 3-uniform hypergraph with 10 edges whose incidence matrix appears below does not satisfy the 2-EVP. Starting with r copies of this hypergraph, $r - 1$ applications of Lemma 4.9 adds $r - 1$ edges (since $\lfloor \lg 3 \rfloor = 1$) and produces a 3-uniform hypergraph with $7r$ vertices and $11r - 1$ edges. Similarly, we can add one edge for each vertex beyond a multiple of 7. Letting $n = 7r + j$

with $0 \leq j \leq 6$, we obtain $V_3(n, 2) \geq 11 \frac{n-j}{7} - 1 + j \geq \frac{11n-31}{7}$. \square

	1	2	3	4	5	6	7
$\{1, 2, 3\}$	1	1	1	0	0	0	0
$\{1, 2, 4\}$	1	1	0	1	0	0	0
$\{1, 2, 5\}$	1	1	0	0	1	0	0
$\{1, 3, 4\}$	1	0	1	1	0	0	0
$\{1, 5, 6\}$	1	0	0	0	1	1	0
$\{1, 6, 7\}$	1	0	0	0	0	1	1
$\{2, 3, 6\}$	0	1	1	0	0	1	0
$\{2, 4, 7\}$	0	1	0	1	0	0	1
$\{3, 5, 7\}$	0	0	1	0	1	0	1
$\{4, 5, 6\}$	0	0	0	1	1	1	0

Acknowledgment

We thank Robert Jamison for suggesting the study of this problem.

References

- [1] H.L. Abbot and D. Hanson, On Finite Δ -Systems, *Discrete Mathematics* 8 (1974), 1–12.
- [2] H.L. Abbot and D. Hanson, On Finite Δ -Systems II, *Discrete Mathematics* 17 (1977), 121–126.
- [3] N. Alon, S. Friedland, and G. Kalai, Regular Subgraphs of Almost Regular Graphs, *Journal of Discrete Algorithms* 37 (1984), 79–91.
- [4] R.C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes, II. *Proc. London Math. Soc. (3)* 83 (2001), 532–562.
- [5] P. Erdős, Beweis eines Satzes von Tschebyschef, *Acta Sci. Math. (Szeged)* 5 (1930–1932), 194–198.
- [6] P. Erdős and R. Rado, Intersection theorems for systems of sets, *J. London Math. Soc.* 35 (1960), 85–90.
- [7] P. Erdős and E. Szemerédi, Combinatorial Properties of Systems of Sets, *J. Combin. Theory (A)* 24 (1978), 308–313.
- [8] H. Iwaniec and J. Pintz, Primes in short intervals, *Monatsh. Math.* 98 (1984), 115–143.

- [9] D.P. Jacobs and R.E. Jamison, A note on equal unions in families of sets, *Discrete Mathematics* 241 (2001), 387–393.
- [10] D.P. Jacobs and R.E. Jamison, Polynomial recognition of equal unions in hypger-graphs with few vertices of large degree, *Journal of Discrete Algorithms* 4 (2006), 201–208.
- [11] A.V. Kostochka, An intersection Theorem for Systems of Sets, *Random Struct. Algor.* 9 (1996), 213–221.
- [12] B. Lindström, A theorem on families of sets, *J. Combin. Theory (A)* 13 (1972), 274–277.
- [13] J. Nagura, On the interval containing at least one prime number, *Proc. Japan Academy (A)* 28 (1952), 177–181.
- [14] H. Tverberg, On equal unions of sets, in: *Studies in Pure Mathematics*, L. Mirsky, ed., (Academic Press, 1971), 249–250.