AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **82(1)** (2022), Pages 1–20

Minimally connected r-uniform hypergraphs

Mark Budden

Department of Mathematics and Computer Science
Western Carolina University
Cullowhee, NC 28723, U.S.A.
mrbudden@email.wcu.edu

Josh Hiller

Department of Mathematics and Computer Science
Adelphi University
Garden City, NY 11530-0701, U.S.A.
johiller@adelphi.edu

Andrew Penland

Department of Mathematics and Computer Science
Western Carolina University
Cullowhee, NC 28723, U.S.A.
adpenland@email.wcu.edu

Abstract

In this paper, we define a construction process for minimally connected r-uniform hypergraphs, which captures the intuitive notion of building a hypergraph piece-by-piece, and a numerical invariant called the tightness, which is independent of the construction process used. Using these tools, we prove some fundamental properties of minimally connected hypergraphs. We also give bounds on their chromatic numbers and provide some results involving hyperedge colorings. We show that every connected r-uniform hypergraph contains a minimally connected spanning subhypergraph and provide a polynomial-time algorithm for identifying such a subhypergraph.

ISSN: 2202-3518

1 Introduction

Trees, and in particular, spanning trees, offer a very useful tool for studying connectivity and have several interesting algorithmic properties in the context of other graphs. In the graph case, trees are connected acyclic graphs. Equivalently, they are minimally connected. When one transitions to the setting of r-uniform hypergraphs, these definitions are no longer equivalent. For instance, it is well-known that every connected graph has a spanning tree and that this statement is not true for r-uniform hypergraphs in general.

The distinction between graphs and other r-uniform hypergraphs with regard to spanning trees has important considerations in theoretical computer science. If one needs a spanning tree in a graph, standard algorithms such as that of Prim [16] or Kruskal [14] will do the job in low-degree polynomial time. For 3-uniform hypergraphs, an algorithm due to Lovász [15] will also determine the existence of a spanning tree in polynomial time. A subsequent, more efficient polynomial time algorithm for the same problem is due to Gabow and Stallman [13]. However, Andersen and Fleischner [1] showed that the general problem of determining whether or not a hypergraph has a spanning tree is NP-complete, even for relatively restricted classes, such as linear hypergraphs in which each vertex is contained in at most three hyperedges, or 4-uniform hypergraphs which have some vertex in common to all hyperedges. Andersen and Fleischner [1] quote this last fact as an unpublished result of Carsten Thomassen.

Another important distinction between graphs and hypergraphs arises in the equivalence of certain definitions of a spanning tree. In graph theory, every spanning minimally connected subgraph is a tree, but the analagous statement does not hold for hypergraphs. This distinction between spanning trees and minimally connected subhypergraphs appears when generalizing certain results from graphs to hypergraphs. For example, in 2014, Chartrand, Johns, McKeon, and Zhang [7] proved that a connected graph G has its rainbow connection number equal to its size if and only if G is a tree. In order to prove an analogue of this result in the setting of hypergraphs, Carpentier, Liu, Silva and Sousa [6] were forced to consider the more general class of minimally connected hypergraphs. In this paper, we examine other aspects of hypergraph theory where minimally connected hypergraphs are necessary to prove results that typically concern trees in graphs.

In the next section, we give the formal definition of an r-uniform tree and provide a very simple demonstration that for every $r > 2$, there exist connected hypergraphs which do not admit spanning r-uniform trees. In Section 3, we consider structural properties and existence theorems for minimally connected hypergraphs, introducing a numerical invariant called tightness, associated to a hypergraph. In Section 4, we examine chromatic numbers of minimally connected hypergraphs. Section 5 of this paper deals with other connectivity issues related to minimally connected hypergraphs. We conclude with some open questions and directions for future research.

2 Definitions and background

In this section, we provide the definitions, elementary examples, and concepts that are necessary to derive the results of this paper. If S is any set, then we denote by $\binom{S}{r}$ the set of all r-element subsets of S. An r-hypergraph H consists of two sets: a non-empty set $V(H)$ called the *vertex set*, and a set $E(H) \subseteq \binom{V(H)}{r}$, called the set of *hyperedges*. When $r = 2$, our definition coincides with that of a graph. When the r-uniform hypergraph being considered is clear from the context, we may write V and E in place of $V(H)$ and $E(H)$, respectively.

The *size* of an r-uniform hypergraph is $|E|$ while its *order* is $|V|$. The *degree* of a vertex $v \in V$ is the number of hyperedges that contain v. Throughout the remainder of this paper, we will focus on r-uniform hypergraphs with $r > 2$. An r-uniform hypergraph H is called *finite* if $V(H)$ is finite. We will only consider finite r-uniform hypergraphs and often, just refer to them as hypergraphs. For clarity, when we refer to graphs, we will call them 2-graphs.

A *Berge path* consists of a sequence of k distinct vertices $v_1, v_2, \ldots, v_k$ and $k-1$ distinct hyperedges $e_1, e_2, \ldots, e_{k-1}$ such that $v_i, v_{i+1} \in e_i$ for all $i \in \{1, 2, \ldots, k-1\}$. A *Berge cycle* is formed if there is a hyperedge e_k that includes both v_1 and v_k. Observe in these two definitions that for $r > 2$, the hyperedges contain vertices other than $v_1, v_2, \ldots, v_k$, which are not assumed to be distinct (unless they are in a common hyperedge). A Berge path is a *loose path* if for $i < j$,

$$|e_i \cap e_j| = \begin{cases} 0 & \text{if } j \neq i+1 \\ 1 & \text{if } j = i+1. \end{cases}$$

Observe that although vertices in a Berge path may be repeated, all vertices contained in a loose path are necessarily distinct.

An r-uniform hypergraph is called *connected* if for every distinct pair of vertices u and v, there exists some Berge path that contains both u and v. An r-uniform hypergraph that is not connected is called *disconnected*. A connected subhypergraph H' of an r-uniform hypergraph H is called a *component of* H if H' is not a proper subhypergraph of any connected subhypergraph of H. It follows that when H is connected, it contains a single component.

We say that an r-uniform hypergraph H is *minimally connected* if the removal of any hyperedge (while retaining all vertices) disconnects H. For example, consider the hypergraphs in Figure 1. Every hyperedge in the first hypergraph contains some vertex of degree one, but this is not the case in the second hypergraph.

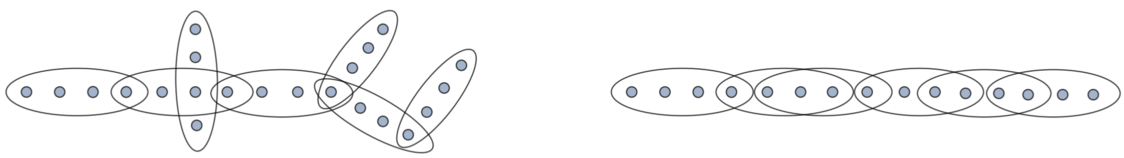

Figure 1: Two minimally connected 4-uniform hypergraphs.

Observe that if every hyperedge of a connected r-uniform hypergraph contains at least one vertex of degree 1, then it is minimally connected. The second hypergraph in Figure 1 demonstrates that the converse to this statement is false. Figure 2 shows two other minimally connected hypergraphs, both of which are Berge cycles.

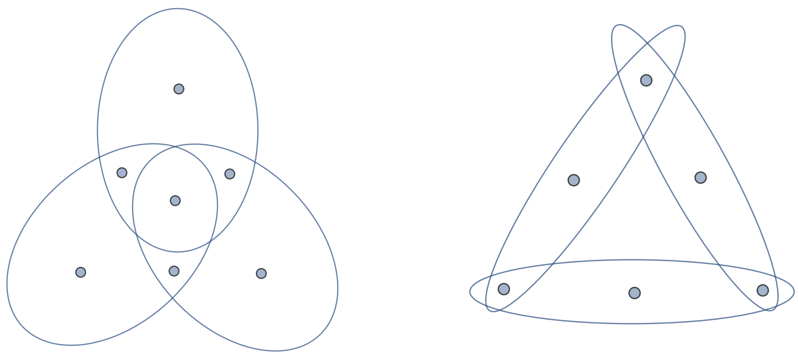

Figure 2: Two minimally connected hypergraphs that are also Berge cycles.

Of course, having defined minimally connected hypergraphs, one must inquire about the appropriate definition of a hypergraph tree. We now give several possible definitions.

Definition 2.1. The following definitions of r-uniform trees are equivalent:

(1) T is a connected r-uniform hypergraph that does not contain any Berge cycles.

(2) T is an r-uniform hypergraph that can be formed hyperedge-by-hyperedge with each new hyperedge intersecting the previous hypergraph at exactly one vertex. That is, each new hyperedge requires the creation of exactly $r-1$ new vertices.

(3) T is a connected r-uniform hypergraph in which the removal of any hyperedge (keeping all vertices) results in a hypergraph with exactly r connected components.

(4) T is an r-uniform hypergraph in which there exists a unique loose path between any pair of distinct vertices.

(5) T is a connected r-uniform hypergraph in which the size $|E|$ and order $|V|$ satisfy $|V| = (r-1)|E| + 1$.

The equivalence of (1)–(4) can be found in Theorem 2.1 of [5]. We will wait until Theorem 3.5 in Section 3 to complete the proof that (5) is also a suitable definition. An important observation is that from (3), it immediately follows that every r-uniform tree is minimally connected (in that the removal of any hyperedge disconnects the hypergraph). Of course, the hypergraphs given in Figure 2 and the second hypergraph given in Figure 1 show that not all minimally connected hypergraphs are trees.

As we begin our investigation into minimally connected hypergraphs, we will emphasize how they can differ from r-uniform trees. The examples given so far demonstrate that if a hyperedge-by-hyperedge construction of a minimally connected hypergraph exists, then it must allow for the intersection of a new hyperedge with more than one vertex (and even more than one hyperedge) in the previous hypergraph. Since some minimally connected hypergraphs are Berge cycles, there can also be multiple Berge paths between a pair of distinct vertices. All connected 2-graphs are known to contain spanning trees, but the following proposition shows that it is quite easy to construct r-uniform hypergraphs with $r > 2$ in which this is not true. Let us first define an *r-uniform hypergraph spanning tree* for an r-uniform hypergraph H to be an r-uniform tree T with $V(T) = V(H)$ and $E(T) \subseteq E(H)$.

Proposition 2.2. *For every $r > 2$ there exists a connected r-uniform hypergraph which does not admit an r-uniform spanning tree.*

Proof. Fix $r > 2$ and consider the r-uniform hypergraph on $r + 1$ vertices with two hyperedges: $e_1 = v_1v_2 \cdots v_r$ and $e_2 = v_2v_3 \cdots v_{r+1}$. Then the hypergraph obtained is connected and not a tree, but the removal of either hyperedge disconnects the hypergraph. □

3 Structural properties and existence theorems

In this section, we analyze the underlying structure of minimally connected r-uniform hypergraphs. We start by defining a construction process, which leads to the notion of the tightness sum of such a hypergraph. Our attention then turns to possible sizes and an algorithm for finding a spanning minimally connected subhypergraph for any connected r-uniform hypergraph.

3.1 Construction processes for hypergraphs

The definition we offered for a minimally connected r-uniform hypergraph can be a little fastidious to work with when proving basic properties of hypergraphs. Hence, we wish to offer an equivalent constructive definition of a minimally connected r-uniform hypergraph that will serve our purposes nicely. In order to make this notion precise, we formally define a constructive process.

Definition 3.1. Let H be an r-uniform hypergraph of size m. A *constructive process* $\mathcal{P}$ for H is a finite sequence of hypergraphs H_i, $1 \leq i \leq m$ satisfying the following properties:

- $H_m = H$,

- H_j is a subhypergraph of H_k for $j \leq k$,
- each H_i has exactly i hyperedges, and
- for each j, $E(H_{j+1}) = E(H_j) \bigcup \{e_j\}$ for some $e_j \in E(H)$.

We say that a constructive process $\mathcal{P}$ for a hypergraph H is *connected* if and only if each H_i is connected for all i. We say that a constructive process $\mathcal{P}$ for a hypergraph H is *minimally connected* if and only if each H_i is minimally connected. The following two theorems highlight the significance of constructive processes as a tool for studying connected and minimally connected r-uniform hypergraphs.

Theorem 3.2. *Let H be an r-uniform hypergraph. Then H is connected if and only if H has a connected constructive process.*

Proof. If H has a connected constructive process, then H must be connected since $H_m = H$. To prove the forward implication, we use induction on the size of H. If a hypergraph is connected and consists of a single hyperedge e, then the constructive process consists only of H_1, which is necessarily connected. Now assume that all connected r-uniform hypergraphs of size at most $m-1$ have connected constructive processes and let H be a connected r-uniform hypergraph of size m. Select a hyperedge e and remove it from H. There exist two possibilities: $H - e$ is either connected or disconnected. If it is connected, then by the inductive hypothesis, it has a connected constructive process $H_1, H_2, \ldots, H_{m-1}$. It follows that $H_1, H_2, \ldots, H_{m-1}, H$ is a connected constructive process for H. In the case where $H - e$ is disconnected, let $C_1, C_2, \ldots, C_k$ be its components that have size at least one. By the inductive hypothesis, each C_i has a connected constructive process. Denote the constructive process for C_i by

$$C_i^1, C_i^2, \ldots, C_i^{e_1},$$

where $e_1 + e_2 + \cdots + e_k = m - 1$. Let $C_1 + e$ be the connected hypergraph formed by adding hyperedge e (and all vertices in e that are not already in C_1) to the connected hypergraph C_1. It follows that

$$C_1^1, C_1^2, \ldots, C_1^{e_1}, C_1 + e, C_2^1, C_2^2, \ldots, C_2^{e_2}, \ldots, C_k^1, C_k^2, \ldots, C_k^{e_k}$$

is a connected constructive process for H, completing the proof of the theorem. □

Theorem 3.3. *Let H be an r-uniform hypergraph. Then H is minimally connected if and only if H has a minimally connected constructive process.*

Proof. If an r-uniform hypergraph H has a minimally connected constructive process, then H must be minimally connected since $H_m = H$. To prove the forward implication, suppose that H a is minimally connected r-uniform hypergraph. In particular, H is connected, so it can be constructed hyperedge-by-hyperedge with each resulting hypergraph being connected along the way. Let H_i be the resulting hypergraph after the first i hyperedges $e_1, e_2, \ldots, e_i$ have been added. Note that if H_i is not minimally connected, then there exists some hyperedge whose removal does not disconnect H_i, and hence, would not disconnect H. □

We have shown that there is a constructive process $\mathcal{P}$ that one can use to obtain a minimally connected hypergraph. Suppose that H is a minimally connected r-uniform hypergraph of size m. Denote the hyperedges in this constructive process by e_1, e_2, …, e_m and let H_i be the resulting minimally connected hypergraph after hyperedge e_i has been added. Note that with the addition of each hyperedge, at least one new vertex must be introduced. Otherwise, when adding in a hyperedge e_i that only uses existing vertices, the hypergraph H_{i-1} would have to have been disconnected, contradicting our assumption about the constructive process. It is also worth observing that the addition of any e_i cannot prevent the removal of a previous hyperedge e_j ($j < 1$) from disconnecting the hypergraph (although it may change the number of resulting components).

3.2 The notion of tightness and some applications

A constructive process $\mathcal{P}$ produces a sequence t_1, t_2, …, t_{m-1} of *tightnesses* given by

$$t_i = |V(H_i) \cap e_{i+1}|,$$

where $1 \le t_i \le r-1$ and $1 \le i \le m-1$. Although the constructive process that we have described is not unique for a given connected hypergraph H, we will show in the following theorem that the sum of the tightnesses is independent of the construction chosen.

Theorem 3.4. *Let $\mathcal{P}$ and $\mathcal{P}'$ be constructive processes for a connected r-uniform hypergraph H of size m with tightness sequences*

$$t_1, t_2, \ldots, t_{m-1} \quad \textit{and} \quad t_1', t_2', \ldots, t_{m-1}',$$

respectively. Then the tightness sums

$$t_{\mathcal{P}} := \sum_{i=1}^{m-1} t_i \quad \textit{and} \quad t_{\mathcal{P}'} := \sum_{i=1}^{m-1} t_i'$$

satisfy $t_{\mathcal{P}} = t_{\mathcal{P}'}$.

Proof. Let H be a connected r-uniform hypergraph of size m and suppose that $\mathcal{P}$ and $\mathcal{P}'$ are constructive processes for H with tightness sequences

$$t_1, t_2, \ldots, t_{m-1} \quad \text{and} \quad t_1', t_2', \ldots t_{m-1}',$$

respectively. Then for $1 \le i \le m-1$, the addition of e_{i+1} in $\mathcal{P}$ requires the addition of $r - t_i$ new vertices. An analogous statement can be made for $\mathcal{P}'$. If H has order n, then

$$n = r + \sum_{i=1}^{m-1} (r - t_i) = r + \sum_{i=1}^{m-1} (r - t_i'),$$

from which it follows that

$$\sum_{i=1}^{m-1} t_i = \sum_{i=1}^{m-1} t'_i,$$

completing the proof of the theorem. □

Thus, the tightness sum of a connected r-uniform hypergraph H is independent of the specific choice of constructive process, so we denote by t_H the tightness sum for any constructive process of H. Letting $t_1, t_2, \ldots, t_{m-1}$ be a tightness sequence for any constructive process for H, the order of H is given by

$$|V(H)| = r + \sum_{i=1}^{m-1}(r - t_i) = rm - t_H.$$

The following theorem will prove the equivalence of (2) and (5) in Definition 2.1.

Theorem 3.5. *Let H be a connected r-uniform hypergraph of size m. Then H is an r-uniform tree if and only if $t_H = m - 1$. Equivalently, H is an r-uniform tree if and only if it has order $|V(H)| = (r-1)|E(H)| + 1$.*

Proof. Every r-uniform tree of size m is minimally connected and has order $r + (m-1)(r-1)$. Thus, we need only show that if H is a connected r-uniform hypergraph of size m and order $r + (m-1)(r-1)$, then H is a tree. We prove the contrapositive to this statement. Suppose that H is a connected r-uniform hypergraph that is not a tree. Then H can be constructed hyperedge-by-hyperedge, with the resulting hypergraph being connected at each stage of the construction. Since H is connected, $t_i \geq 1$ for all $1 \leq i \leq m-1$. If H is not a tree, some tightness $t_i \geq 2$. So, the sum of the tightnesses of H satisfies $t_H \geq m$, giving a maximal order of $r + (m-1)(r-1) - 1$. □

Denote by $S_n^{(r)}$ the *r-uniform star of order n* consisting of $r-1$ vertices in the intersection of all hyperedges (called the *center*), along with each hyperedge containing a single vertex of degree 1. This definition agrees with the more general definition of a star given in [4]. Observe that such a star is minimally connected since the removal of any hyperedge leaves the vertex of degree 1 isolated. See Figure 3 for an example of a 6-uniform star.

Theorem 3.6. *For $r \geq 3$, a minimally connected r-uniform hypergraph H of order n and size m is an r-uniform star if and only if*

$$t_H = (m-1)(r-1).$$

Proof. If a minimally connected hypergraph H is a star, then every constructive process for H requires each tightness $t_i = r - 1$, giving $t_H = (m-1)(r-1)$. To prove the converse, we will use induction on m. When $m = 1$, $t_H = 0$ and H consists of a single hyperedge, which is trivially a star. Now suppose that H is a minimally connected hypergraph of size $m > 1$ in which $t_H = (m-1)(r-1)$ and

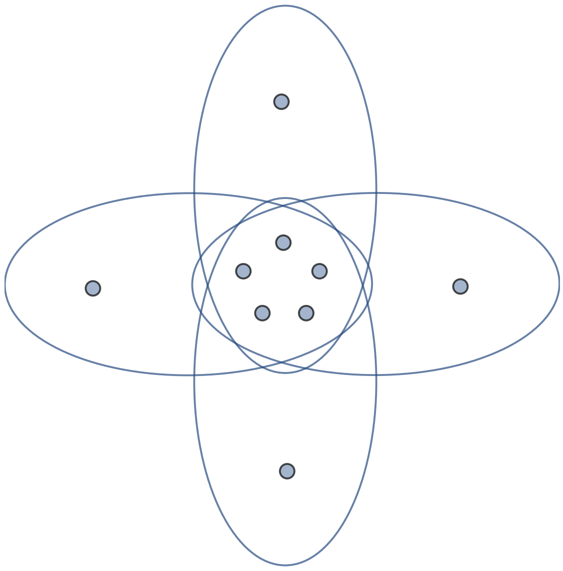

Figure 3: A 6-uniform star of order 9.

that all minimally connected hypergraphs of size $m-1$ with tightness sum equal to $(m-2)(r-1)$ are necessarily stars. As a minimally connected hypergraph, H has a minimally connected constructive process $\mathcal{P}$. Let e_m be the last hyperedge added in such a process and consider the hypergraph H' formed by removing e_m (and all resulting isolated vertices) from H. Since $t_H = (m-1)(r-1)$, all tightnesses t_i in $\mathcal{P}$ are equal to $r-1$. Hence, the removal of e_m removes only one isolated vertex and we find that $|V(H)| = |V(H')| + 1$. So,

$$\begin{aligned} t_{H'} &= r(m-1) - |V(H')| \\ &= r(m-1) - (|V(H)| - 1) \\ &= r(m-1) - (rm - t_H - 1) \\ &= r(m-1) - (rm - (m-1)(r-1) - 1) \\ &= (m-2)(r-1). \end{aligned}$$

By the inductive hypothesis, H' is a star. When adding back in hyperedge e_m to form H, including any vertex x of degree one from H' will prevent the hyperedge e_i that includes x from disconnecting H if it is removed, contradicting the assumption that H is minimally connected. Thus, the only vertices that can be included in the intersection of e_m with H' are those of degree greater than one in H'. It follows that H is a star. □

3.3 The possible sizes of minimally connected subhypergraphs

We now focus our attention on finding bounds for the size of a minimally connected r-uniform hypergraph. Let H be a connected hypergraph of order $n \geq r+1$. From Theorem 3.3, each minimally connected spanning hypergraph M can be constructed hyperedge-by-hyperedge, with each resulting hypergraph being minimally connected, and this process results in a (finite) sequence of tightnesses $t_1, t_2, \ldots, t_{m-1}$, where $1 \leq t_i \leq r-1$ and m is the size of M. The maximum number of hyperedges M can contain occurs when M is $(r-1)$-tight (i.e., $t_1 = t_2 = \cdots = t_{m-1} = r-1$). In this case, we find that M contains $m = n - r + 1$ hyperedges. For example, consider the star in Figure 3.

The minimum number of hyperedges that M can contain occurs when M is a tree, or is close to being a tree (with say, only one tightness not equal to 1). If M is an r-uniform tree of order n and size m, then

$$m = \frac{n-1}{r-1}.$$

If M is not a tree, but is close to being a tree, then suppose that

$$n - 1 = (m-1)(r-1) + k, \quad \text{where} \quad 1 \leq k \leq r-2.$$

It follows that

$$m = \frac{n-1}{r-1} + \frac{r-1-k}{r-1} = \left\lceil \frac{n-1}{r-1} \right\rceil.$$

Putting together these upper and lower bounds, we have shown the following.

Theorem 3.7. *Let H be a connected r-uniform hypergraph of order n with minimally connected spanning subhypergraph M of size m. Then*

$$\frac{n-1}{r-1} \leq m \leq n - r + 1.$$

Theorem 3.7 hints at a question: for a given order n is there a minimally connected r-uniform hypergraph of size m for every permissible value of m? The following result shows that this is the case.

Theorem 3.8. *Fix an order $n \geq r$ and let $m \in \mathbb{N}$ satisfy*

$$\frac{n-1}{r-1} \leq m \leq n - r + 1.$$

Then there exists a minimally connected r-uniform hypergraph of size m and order n.

Proof. Let $z = n - r + 1 - m$. We will define a constructive process which culminates in an r-uniform hypergraph of size m and order n. Let us choose r vertices to form e_1 and label these $v_1, v_2, \ldots, v_r$. We then proceed with two cases based on the value of z:

Case 1: If $z \geq r-2$, then define $a, b \in \mathbb{Z}$ such that $z = a(r-2)+b$. For $1 \leq i \leq a$ let $t_i = 1$, $t_{a+1} = r-b-1$, and $t_j = r-1$ for $a+1 < j < m-1$. For e_2 let $e_1 \cap e_2 = \{v_1\}$. For e_k $(2 < k < m)$ pick any t_{i-1} vertices from $v_2, v_3, \ldots, v_r$. Then every edge, except possibly e_1 contains a vertex of degree 1. If e_1 is removed then e_2 is disconnected from the rest of the hypergraph. Thus the resulting hypergraph is of size m, order n, and is minimally connected.

Case 2: If $z < r-2$, then let $t_1 = r-z-1$. Let $e_1 \cap e_2 = \{v_1, v_2, \ldots, v_{t_1}\}$ create $m-2$ edges $e_3, e_4, \ldots, e_m$ and for every $2 < i \leq m$, let $e_1 \cap e_i = \{v_1, v_2, \ldots, v_{r-1}\}$. Thus every edge has at least one vertex of degree 1, and so the resulting hypergraph is of size m, order n, and is minimally connected. □

Theorem 3.8 leads to an interesting observation. The complete hypergraph $K_n^{(r)}$ contains minimally connected subhypergraphs of size m for every permissible value of m. This fact stands in stark contrast to the context of 2-graphs, where it is well known that the size of a spanning tree is completely determined by the order of the parent graph.

3.4 Existence of an algorithm for minimally connected spanning subhypergraphs

One motivation for this paper was to generalize a ubiquitous result in the study of trees: every connected graph contains a spanning tree. We now provide an analogous result for r-uniform hypergraphs and describe a polynomial-time algorithm for finding such an subhypergaph.

Proposition 3.9. *Every connected r-uniform hypergraph contains a spanning minimally connected subhypergraph.*

Proof. Let H be a connected r-uniform hypergraph. If H is not minimally connected, then there exists some hyperedge e_1 whose removal does not disconnect H. Let H_1 be the hypergraph formed by removing e_1 from H. If H_1 is not minimally connected, then repeat this process. As H has a finite number of hyperedges, the process must eventually terminate with a minimally connected hypergraph H_i that spans the vertices in H. □

In this subsection we discuss an algorithm for finding a minimally connected spanning subhypergraph of an r-uniform hypergraph. In order to make the discussion clear, we start with some necessary definitions.

An *algorithm* is an unambiguous, step-by-step process that takes input and terminates with some output. If f and g are functions from $\mathbb{N}$ to itself, we say that $g(n)$ is $O(f(n))$ if there exist positive constants c and n_0 such that $0 \leq g(n) \leq cf(n)$ for all $n \geq n_0$. For an algorithm A, we let $T_A(n)$ denote the maximum number of steps that it takes A to terminate on an algorithm of size n; note that T_A is a function from $\mathbb{N}$ to $\mathbb{N}$. Given an algorithm A and a function $f(n)$, we say that an algorithm A is $O(f(n))$ if the function T_A is $O(f(n))$. If there exists a polynomial $p(n)$ such that

A is $O(p(n))$, we say that A is *polynomial time*. For a more thorough definition of these concepts, we would direct the curious reader towards the standard [8, Chapter 2]. Throughout the rest of this section, assume that $n = |V(H)|$, where H is an r-uniform hypergraph.

Gallo, Longo, Pallottino, and Nguyen [10] provide an algorithm $\mathbf{Visit}(H, v)$ which takes as input a hypergraph H and a vertex $v \in V(H)$, and returns all vertices in H that are in the same connected component as v. The **Visit** algorithm is $O(s)$, where

$$s = \sum_{e \in E(H)} |e|.$$

If we assume that H is an r-uniform hypergraph, then there are at most $\binom{n}{r}$ hyperedges, and each $e \in E(H)$ has $|e| = r$. Recall that if r is a fixed constant, $\binom{n}{r}$ is $O(n^r)$ when viewed as a function of n. Thus, in the case of an r-uniform hypergraph, we have $s = O(n^r)$, and **Visit** becomes a polynomial time algorithm. Notice that since being in the same connected component is an equivalence relation, it is not hard to see that an r-uniform hypergraph H is connected if and only if $\mathbf{Visit}(H, v)$ returns the entire set $V(H)$ for any $v \in V(H)$.

Theorem 3.10. *For any value of r, there exists an algorithm that takes as input a connected r-uniform hypergraph H and returns M, a minimally connected spanning subhypergraph of H. If r is viewed as a fixed constant, this algorithm runs in polynomial time on the number of vertices in H.*

Proof. Consider the following algorithm: Let H^* be a copy of H. Begin by choosing an arbitrary vertex $v^* \in H$. For each edge e in $E(H^*)$, run $\mathbf{Visit}(H^* - \{e\}, v^*)$ to determine if e can be removed without disconnecting H^*. If yes, remove e from $E(H^*)$. Note that $H^* - \{e\}$ can only be connected if $H^* - \{e\}$ will contain all of $V(H^*)$. So at each stage, H^* remains a spanning subhypergraph of H. This algorithm will terminate with H^* as a minimally connected subhypergraph of H, since if an edge of $E(H^*)$ could be removed without disconnecting H^*, that edge would have been removed at the stage it was considered. Since H is an r-uniform hypergraph, there are at most $\binom{n}{r}$ hyperedges, which is $O(n^r)$, and for each hyperedge we call $\mathbf{Visit}(H, v^*)$ which we know is an $O(n^r)$ algorithm. Hence the algorithm is $O(n^{2r})$, which is a polynomial in r. □

It is worth emphasizing that these results provide a sharp contrast between spanning minimally connected subhypergraphs and spanning trees. For instance, for a 4-uniform hypergraph on n vertices, we have just shown that there must be a spanning minimally connected subhypergraph, and in fact there is an $O(n^8)$ algorithm to find one. As discussed in the Introduction, 4-uniform hypergraphs do not always have spanning trees, and no known polynomial time algorithm is known to find such a spanning tree if it does exist. We would hazard the following intuitive explanation for the potential gap in complexity between the problems. Our algorithm is able to consider each edge one at a time and check whether removal will result in a disconnected subhypergraph. The problem of finding a spanning tree adds the additional difficulty of avoiding a cycle, which is a more global property based on collections of edges rather than single edges.

4 Chromatic numbers of minimally connected hypergraphs

Chromatic numbers give insight into the connectivity of a 2-graph or hypergraph. Let χ_w and χ_s denote the weak and strong chromatic numbers, respectively. That is, $\chi_w(H)$ is the minimum number of colors needed to color the vertices of H so that no hyperedge is monochromatic and $\chi_s(H)$ is the minimum number of colors needed to color the vertices of H so that every pair of adjacent vertices receive different colors. When these concepts are restricted to the case of 2-graphs, they both agree with that of the chromatic number.

Theorem 4.1. *If H is a minimally connected r-uniform hypergraph, then $\chi_w(H) = 2$.*

Proof. Let H be a minimally connected r-uniform hypergraph with m edges. From Theorem 3.3, we know that a minimally connected hypergraph can be constructed hyperedge-by-hyperedge with a connected hypergraph each step of the way and such that each new hyperedge requires the addition of a new vertex. Suppose that H_i is the connected hypergraph formed after adding hyperedge e_i ($1 \leq i \leq m$). We proceed by induction on m to prove that $\chi_w(H) = 2$. H_1 consists of a single hyperedge, so its vertices can be trivially 2-colored. Now suppose that H_i can be weakly 2-colored. When adding e_{i+1} to construct H_{i+1}, there are two possibilities: the vertices in $E(H_i) \cap e_{i+1}$ are all the same color or receive both colors 1 and 2. In the former case, give a new vertex added with e_{i+1} the other color. In the latter case, a new vertex can receive either color. In both cases, we find that H_{i+1} can be weakly 2-colored, and hence, $\chi_w(H) = 2$. □

While r-uniform trees with size $m \geq 1$ have strong chromatic number $\chi_s(T) = r$ (this is a simple inductive exercise to confirm), we are unable to provide such precise limitations on the strong chromatic number of hypergraphs that are only assumed to be minimally connected. The following theorem demonstrates a method for finding minimally connected hypergraphs with arbitrarily large strong chromatic numbers. Note that the following statement is not true for minimally connected 2-graphs (i.e., 2-uniform trees), so one must assume $r \geq 3$.

Theorem 4.2. *For all natural numbers $n \geq r \geq 3$, there exists a minimally connected r-uniform hypergraph with strong chromatic number equal to n.*

Proof. We begin with a complete graph K_n of order n, which has size $m = \frac{n(n-1)}{2}$ and chromatic number $\chi(K_n) = n$. From this graph, we form an r-uniform hypergraph $H_n^{(r)}$ that is minimally connected by replacing each edge ab in K_n with an r-uniform hyperedge $e_i = abx_1^i x_2^i \cdots x_{r-2}^i$, where x_j^i are new vertices (with $1 \leq i \leq m$ and $1 \leq j \leq r-2$) that all have degree one. The hypergraph $H_n^{(r)}$ is minimally connected since the removal of e_i leaves each x_j^i disconnected from the rest of the hypergraph.

Also, $H_n^{(r)}$ requires at least n colors in any strong proper coloring since adjacent vertices in K_n are still adjacent in $H_n^{(r)}$. Hence, $\chi_s(H_n^{(r)}) \geq n$. On the other hand, since $n \geq r$, there are enough colors from the original proper coloring of K_n to color the vertices $x_1^i, x_2^i, \ldots, x_{r-2}^i$ distinct from one another. Thus, $\chi_s(H_n^{(r)}) \leq n$, completing the proof. □

The construction in the above proof provides, for any natural number n, a means of producing a minimally connected r-uniform hypergraph with strong chromatic number equal to n. As an example, consider Figure 4. In this figure, $H_4^{(3)}$ is constructed from K_4. The hypergraph $H_4^{(3)}$ contains six hyperedges, the deletion of any hyperedge results in a disconnected hypergraph containing an isolated vertex, and $\chi_s(H_4^{(3)}) = 4$.

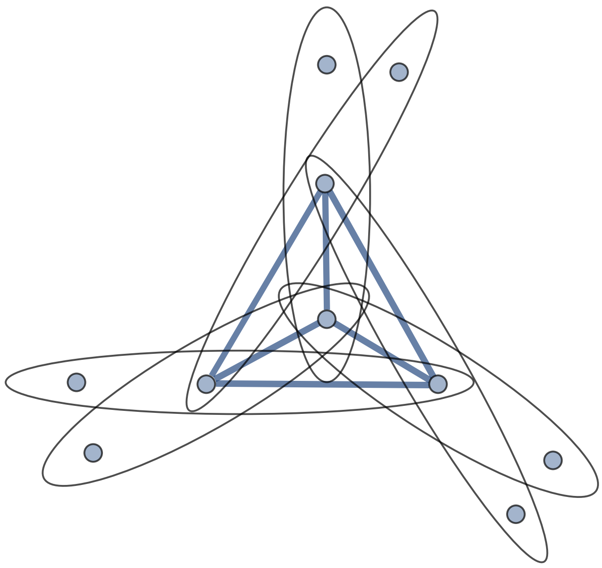

Figure 4: Using a K_4 to construct a minimally connected 3-uniform hypergraph $H_4^{(3)}$ satisfying $\chi_s(H_4^{(3)}) = 4$.

Theorem 4.3. *For $r \geq 3$, let H be a minimally connected r-uniform hypergraph of size $m \geq 2$ with minimally connected constructive process $\mathcal{P}$ having tightness sequence $t_1, t_2, \ldots, t_{m-1}$. Then*

$$\chi_s(H) \leq r + t_H - t_1 - m + 2.$$

Proof. We proceed by induction on $m \geq 2$. When $m = 2$, $t_H = t_1$, and it is easily seen that $\chi_s(H) = r$. Now, suppose that the result is true for all minimally connected r-uniform hypergraphs of size $m - 1$ and let H be a minimally connected r-uniform hypergraph of size m. Let H' be the minimally connected hypergraph formed by removing e_m (and all isolated vertices) from H. Then by the inductive hypothesis,

$$\chi_s(H') \leq r + \sum_{i=1}^{m-2} t_i - t_1 - (m-1) + 2.$$

Fix a strongly proper vertex coloring of H'. When adding in e_m, the only vertices that may need new colors are those that are the same color in $V(H') \cap e_m$. So, at most, $t_{m-1} - 1$ new colors are needed, from which it follows that

$$\chi_s(H) \leq r + \sum_{i=1}^{m-1} t_i - t_1 - m + 2 = r + t_H - t_1 - m + 2,$$

completing the proof of the theorem. □

Observe that we can optimize the bound in Theorem 4.3 by picking a constructive process in which t_1 is maximal. We encourage the reader to check that when $r \geq 3$, the inequality proved in Theorem 4.3 is tight for all r-uniform stars of order 2.

5 Some hyperedge coloring results

In this section we will examine colorings of the hyperedges of complete r-uniform hypergraphs. A *t-coloring* of an r-uniform hypergraph H is a function

$$c : E(H) \longrightarrow \{1, 2, \dots, t\}$$

that assigns colors to the hyperedges of H. We do not assume that such a coloring is proper, nor do we assume that c is surjective. A subgraph H' of H is called *rainbow* if all of the hyperedges in H' receive different colors. We call *H rainbow connected with respect to c* if for every pair of distinct vertices $u, v \in V(H)$, there exists a rainbow Berge path connecting u to v. Observe that every connected hypergraph H is rainbow connected with respect to some coloring as one could always choose the coloring in which every hyperedge of H receives a different color. The *rainbow connected number* $rc(H)$ of a connected r-uniform hypergraph H is defined to be the minimal number of colors t such that H is rainbow connected with respect to some t-coloring.

Since every connected r-uniform hypergraph H is spanned by a minimally connected subhypergraph M, the edges of M can each receive a different color, providing a rainbow Berge path between every distinct pair of vertices. Hence, from Theorem 3.7, it follows that

$$rc(H) \leq n - r + 1.$$

Let $K_n^{(r)}$ denote the complete r-uniform hypergraph of order n. A *Gallai t-coloring* of $K_n^{(r)}$ is a t-coloring of the hyperedges in $K_n^{(r)}$ such that no rainbow $K_{r+1}^{(r)}$-subhypergraph exists. Thus, when $t \leq r + 1$ all t-colorings are Gallai t-colorings.

The following theorem is a nice generalization of a result concerning Gallai colorings of graphs from Gyárfás and Simonyi [13].

Theorem 5.1. *Let $r \geq 3$. Then every Gallai $(r+1)$-coloring of $K_n^{(r)}$ contains a color that spans a connected r-uniform hypergraph using all n vertices.*

Proof. We proceed by induction on n. When $n = r+1$, at least two hyperedges in $K_{r+1}^{(r)}$ are the same color by the definition of a Gallai coloring, and hence, span the complete hypergraph. Now assume the theorem is true for $n \geq r+1$ and consider a Gallai coloring of $K_{n+1}^{(r)}$. Let $\{x_1, x_2, \dots, x_{n+1}\}$ be the vertices of this $K_{n+1}^{(r)}$ and define k_i to be the subhypergraph formed by removing x_i from the $K_{n+1}^{(r)}$, for each $i \in \{1, 2, \dots, n+1\}$. Thus, we have a total of $n+1$ complete Gallai $(r+1)$-colored hypergraphs where $n+1 \geq r+2$. By the inductive hypothesis, each k_i contains a spanning color, and by the pigeonhole principle, at least two of the k_is are spanned by the same color. This color necessarily spans all of $K_{n+1}^{(r)}$. $\square$

Combining Theorem 3.9 with Theorem 5.1, it follows that every Gallai $(r+1)$-coloring of $K_n^{(r)}$ contains a monochromatic minimally connected spanning subhypergraph.

We now turn our attention to the colorings of complete hypergraphs involving only two colors, but first, we must recall the definitions of diameter and complement. If u and v are any two distinct vertices in a connected r-uniform hypergraph H, then the *distance from u to v*, denoted $d(u,v)$, is the minimum number of hyperedges contained in any Berge path connecting u to v. The *diameter* $\mathrm{diam}(H)$ is then defined by

$$\mathrm{diam}(H) := \max\{d(u,v) \mid u, v \in V(H)\}.$$

For every r-uniform hypergraph H, there is a corresponding r-uniform hypergraph $\overline{H}$, called the *complement of H*, such that $V(\overline{H}) = V(H)$ and $E(\overline{H}) = \binom{V(H)}{r} - E(H)$.

Results involving 2-colorings in the setting of 2-graphs are plentiful. In particular, it has been noted that Erdős and Rado claimed that for any graph G, either G or $\overline{G}$ is connected [3]. Rephrasing this result in terms of colorings, every 2-coloring of K_n contains a color that spans all n vertices. More precisely, it was shown by Bialostocki, Dierker, and Voxman [2] that if a graph G is not connected, then $\overline{G}$ is connected and $\mathrm{diam}(\overline{G}) \leq 2$. When $r \geq 3$, we offer a generalization of this fact in Theorem 5.4. But first we derive some helpful results.

Theorem 5.2. *If $r \geq 3$ and H is a disconnected r-uniform hypergraph of order $n \geq r$, then $\overline{H}$ is connected and every subset $\{x_1, x_2, \dots, x_{r-1}\}$ of distinct vertices in $V(H)$ is contained in some hyperedge in $\overline{H}$.*

Proof. Suppose that H is a disconnected r-uniform hypergraph of order $n \geq r$. Let $S = \{x_1, x_2, \dots, x_{r-1}\}$ be a subset distinct vertices in $V(H)$. We consider two cases.

Case 1: Assume that two distinct vertices x_i and x_j are in different connected components in H. Then if y is any vertex not in S, then $x_1 x_2 \cdots x_{r-1} y \in E(\overline{H})$.

Case 2: Assume that C_1 and C_2 are distinct connected components in H with $x_1, x_2, \dots, x_{r-1} \in C_1$. If y is any vertex in C_2, then $x_1 x_2 \cdots x_{r-1} y \in E(\overline{H})$.

In both cases, we find that $\overline{H}$ is connected and every subset $\{x_1, x_2, \ldots, x_{r-1}\}$ of vertices is contained in some hyperedge in $\overline{H}$. □

From Theorem 5.2, observe that whenever H is disconnected, it follows that $\overline{H}$ is connected and $\mathrm{diam}(\overline{H}) = 1$. Furthermore, using this theorem with Theorem 3.9, we obtain the following corollary.

Corollary 5.3. *Let $r \geq 3$. In every 2-coloring of the hyperedges in $K_n^{(r)}$, there exists a monochromatic spanning minimally connected subhypergraph.*

Theorem 5.4. *If $r \geq 3$ and H is a connected r-uniform hypergraph of order $n \geq r$ with $\mathrm{diam}(H) \geq 2$, then $\mathrm{diam}(\overline{H}) = 1$.*

Proof. Assume that H has diameter at least 2. Then by Corollary 5.3 we can assume that H is connected. We may choose $x, y \in V(H)$ to be nonadjacent vertices. Since H is connected and has at least $r > 2$ vertices then there are $r - 2$ vertices $z_1, z_2, \ldots, z_{r-2}$ such that for some $e \in E(H)$ $\{x, z_1, z_2, \ldots, z_{r-2}\} \subset e$. Then the hyperedge $e_1 = \{x, z_1, z_2, \ldots, z_{r-2}, y\} \notin E(H)$. Thus $e_1 \in E(\overline{H})$. Therefore x and y are adjacent in $\overline{H}$, proving that $\mathrm{diam}(\overline{H}) = 1$. □

6 Conclusions and future directions

This paper has been an investigation into hypergraph generalizations of trees. The specific generalization which we explored (minimally connected hypergraphs) stresses the role that trees play in results concerning connectivity. While minimally connected hypergraphs work well to expand upon many theorems for 2-graphs, the fact that the size of a minimally connected spanning tree is not necessarily determined by the parent graph shows that many open problems still exist. We conclude by listing several open problems that we deem worthy of future investigation.

1. Is there an efficient algorithm that would provide, for arbitrary weights, the optimum cost minimally connected spanning subhypergraph? The answer to this question could have serious ramifications, not just for bioinformatics (where punning algorithms have already been applied), but also in realms as distinct as physics and banking. Algorithms for exact and approximate minimum spanning trees in hypergraphs are already known (e.g., see [9], [11], and [12]). We could ask for conditions under which a greedy approach yields the optimal result, and the running time of an optimal algorithm for finding it. Figure 5 shows that a greedy approach of taking the hyperedges in order of least weight can be suboptimal. Such an approach algorithm would yield a minimally connected subhypergraph with hyperedges *cde*, *bde*, and *ade*, for a total weight of 5, whereas the hyperedges *cde* and *abc* would form a minimally connected subhypergraph with a lower total weight of 4.

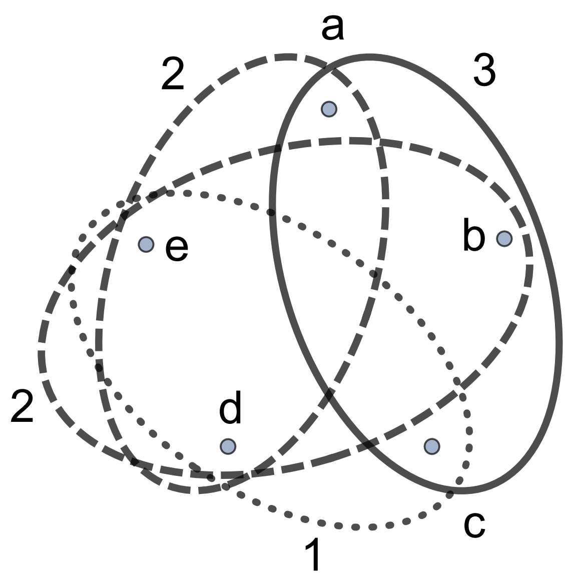

Figure 5: A weighted 3-uniform hypergraph for which the greedy approach does not produce a minimally connected subhypergraph of minimum total edge weight.

2. Recall that the Ramsey number $R(H_1, H_2; r)$ of two r-uniform hypergraphs H_1 and H_2 is defined to be the least natural number p such that every 2-coloring of hyperedges of $K_p^{(r)}$, using say, red and blue, results in a red subhypergraph isomorphic to H_1 or a blue subhypergraph isomorphic to H_2. A connected r-uniform hypergraph H of order m is called *n-good* if

$$R(H, K_n^{(r)}; r) = (m-1)\left(\left\lceil \frac{n}{r-1} \right\rceil - 1\right) + t(K_n^{(r)}),$$

where $\lceil \cdot \rceil$ is the ceiling function and $t(K_n^{(r)})$ is the minimum number of vertices in any color class of a weak vertex coloring of H. The fact that this number is a lower bound for the given Ramsey number was proved in Theorem 3.1 of [5]. So, showing that an r-uniform hypergraph is n-good follows from proving that this number is also an upper bound. In [5], it was conjectured that all r-uniform trees are n-good and infinitely-many examples of n-good 3-uniform trees were given. It was also shown that the minimally connected 3-uniform cycle $C_4^{(3)}$ of length 2 and order 4 is 4-good, but is not 5-good. What are the conditions under which a minimally connected r-uniform hypergraph is n-good? It is worth noting that if a connected r-uniform hypergraph is n-good, then so is every minimally connected spanning subhypergraph.

3. In Section 4, we considered the weak and strong chromatic numbers of minimally connected r-uniform hypergraphs, but other chromatic numbers can be considered when $r \geq 4$. More generally, define the *k-chromatic number* $\chi_k(H)$

of an r-uniform hypergraph H to be the minimum number of colors needed to color the vertices of $V(H)$ so that every hyperedge contains vertices using at least k distinct colors. It follows that

$$\chi_w(H) = \chi_2(H) \quad \text{and} \quad \chi_s(H) = \chi_r(H).$$

Can one determine $\chi_k(H)$, when $2 < k < r$ (assuming $r \geq 4$)? We saw in Section 4 that $\chi_w(H) = 2$ when H is minimally connected, but for every $n \geq r$, there exists a minimally connected r-uniform hypergraph H such that $\chi_s(H) = n$. For $r \geq 4$, what is the smallest value of k for which χ_k is bounded by a constant for all minimally connected r-uniform hypergraphs?

Acknowledgements

We wish to thank the anonymous referees for their careful reading of our manuscript and for insightful comments that greatly improved our paper.

References

[1] L. Andersen and H. Fleischner, The NP-completeness of finding A-trails in Eulerian graphs and finding spanning trees in hypergraphs, *Discrete Appl. Math.* **59**(3) (1995), 203–214.

[2] A. Bialostocki, P. Dierker and W. Voxman, Either a graph or its complement is connected: a continuing saga, (2001) (unpublished manuscript).

[3] J. Bierbrauer and A. Gyárfás, On (n, k)-colorings of complete graphs, Eighteenth Southeastern Int. Conf. on Combinatorics, Graph Theory, and Computing (Boca Raton, FL 1987), *Congr. Numer.* **58** (1987), 123–139.

[4] M. Budden, J. Hiller and A. Rapp, Hypergraph Ramsey numbers involving paths, *Acta Univ. Apulensis Math. Inform.* **48** (2016), 75–87.

[5] M. Budden and A. Penland, Trees and n-good hypergraphs, *Australas. J. Combin.* **72** (2018), 329–349; (Corrigendum, *ibid.* **75**, 171–173).

[6] R. Carpentier, H. Lui, M. Silva and T. Sousa, Rainbow connection for some families of hypergraphs, *Discrete Math.* **327** (2014), 40–50.

[7] G. Chartrand, G. Johns, K. McKeon and P. Zhang, Rainbow connection in graphs, *Math. Bohem.* **133** (2008) 85–98.

[8] T. Cormen, C. Leiserson and D. Shamir, "Introduction to Algorithms," MIT Press: Cambridge, 1997.

[9] H. Gabow and M. Stallmann, An augmenting path algorithm for linear matroid parity, *Combinatorica* **6**(2) (1986), 123–150.

[10] G. Gallo, G. Longo, S. Pallottino and S. Nguyen, Directed hypergraphs and applications, *Discrete Appl. Math.* **42**(2-3) (1993), 177–201.

[11] J. Gao, Q. Zhao, W. Ren, A. Swami, R. Ramanathan and A. Bar-Noy, Dynamic shortest path algorithms for hypergraphs, *IEEE/ACM Transactions on Networking* **23**(6) (2015), 1805–1817.

[12] A. Goodall and A. de Mier, Spanning trees of 3-uniform hypergraphs, *Adv. Appl. Math.* **47**(4) (2011), 840–868.

[13] A. Gyárfás and G. Simonyi, Edge colorings of complete graphs without tricolored triangles, *J. Graph Theory* **46**(3) (2004), 211–216.

[14] J. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.* **7** (1956), 48–50.

[15] L. Lovász, Matroid matching and some applications, *J. Combin. Theory Ser. B* **28**(2) (1980), 208–236.

[16] R. Prim, Shortest connection networks and some generalizations, *Bell Systems Technical Journal* **36**(6) (1957), 1389–1401.

(Received 19 May 2020; revised 8 Nov 2021)