Complete Subgraphs of the Coprime Hypergraph of Integers III: Construction

Jan-Hendrik de Wiljes

Abstract. The coprime hypergraph of integers on \( n \) vertices \( CHI_k(n) \) is defined via vertex set \( \{1, 2, \ldots, n\} \) and hyperedge set \( \{\{v_1, v_2, \ldots, v_k+1\} \subseteq \{1, 2, \ldots, n\} : \gcd(v_1, v_2, \ldots, v_k+1) = 1\} \). In this article we present ideas on how to construct maximal subgraphs in \( CHI_k(n) \). This continues the author’s earlier work, which dealt with bounds on the size and structural properties of these subgraphs. We succeed in the cases \( k \in \{1, 2, 3\} \) and give promising ideas for \( k \geq 4 \).

Keywords. Hypergraphs on integers, clique, matching

2010 Mathematics Subject Classification. 11B75, 05C65, 05C69, 05C70

1 Introduction

In this third part of a series of three articles we continue the work in [3, 4]. References and notation can be found there. The object of interest is the uniform coprime hypergraph of integers \( CHI_k \), which has vertex set \( \mathbb{Z} \) and a \((k+1)\)-hyperedge exactly between every \( k+1 \) elements of \( \mathbb{Z} \) which have (not necessarily pairwise) greatest common divisor equal to 1. In particular we are interested in the subgraph \( CHI_k(n) \) of \( CHI_k \) which is induced by the vertex set \( [n] := \mathbb{N}\cap[1, n] \).

Our main focus lies on the vertex subsets of \( CHI_k(n) \) which induce complete subgraphs:

\[
CS_k(n) := \{ A \subseteq [n] : o_p(A) \leq k \text{ for all } p \in \mathbb{P}\},
\]

where \( o_p(A) := |A \cap \mathbb{P}| \). We define \( CS_k^{\max}(n) \) to be the set of elements from \( CS_k(n) \) which have maximal cardinality \( c_{nk}(n) \) among the elements of \( CS_k(n) \). Elements from \( CS_k^{\max}(n) \) are called maximum \((n, k)\)-shelves.

In [3] questions concerning the cardinality of maximum shelves were considered. The maximal number of prime divisors of elements in (maximum) shelves was investigated in [4]. It was shown that there exists a maximum shelf containing only elements with at most two prime divisors. Following this idea, we will construct such maximum \((n, 3)\)-shelves in this paper. Well known results from matching theory will be applied in the process.

We should mention that the “direct” approach of expanding a maximum \((n, k)\)-shelf does in general not lead to a maximum \((m, k)\)-shelf for \( m > n \). For example the maximum \( (6, 2)\)-shelf \( \{1, 2, 3, 5, 6\} \) is no subset of any maximum \((m, 2)\)-shelf for \( m \geq 9 \), since the number 6 is not contained in any of those shelves. Using results from [4] the following can be shown (we denote by \( \pi \) the prime counting function, by \( \omega(a) \) the number of prime divisors and by \( ssp(a) \) the second smallest prime divisor of \( a \):)

**Proposition 1.** Let \( k > 2, n \) with \( \pi(\sqrt{n}) - \pi(\sqrt[3]{n}) \geq 2 \) and

\[
m \geq \max_{A \in CS_k^{\max}(n)} \min_{a \in A} \frac{\omega(a)}{ssp(a)}^k
\]

be positive integers. Then no maximum \((n, k)\)-shelf is a subset of any maximum \((m, k)\)-shelf.
2 Shifting and a tool from matching theory

Similarly as in [4] we will consider a transformation which preserves the condition \( \omega_p(A) \leq k \), thus, mapping shelves onto shelves. We will use a specific order \( \prec \) on \( \mathbb{N}_{\geq 1} \) to make sure that this operation is idempotent (which facilitates the formulation). For \( a \neq b \in \mathbb{N}_{\geq 1} \) we define (using \( e_p(a) \) as the exponent of \( p \) in the prime factorization of \( a \))

\[
a \prec b :\iff \omega(a) < \omega(b) \quad \text{or} \quad (\omega(a) = \omega(b) \text{ and } e_q(a) < e_q(b)) \quad \text{with } q = \min\{p \in \mathbb{P} : e_p(a) \neq e_p(b)\}.
\]

For a \((n, k)\)-shelf \( A \) we carry out the following steps (\( p_i \) denotes the \( i \)th prime number):

- Set \( A_0 := A \),
- for every \( i \in [\pi(n)] \) (in ascending order) we pick \( e_i := \min\{a_{p_i}(A_{i-1}), \lfloor \log_{p_i} n \rfloor\} \) (with respect to \( \prec \)) smallest multiples \( a_{i_1}, a_{i_2}, \ldots, a_{i_{e_i}} \) of \( p_i \) from \( A_{i-1} \) and define

\[
A_i := (A_{i-1} \setminus \{a_{i_1}, a_{i_2}, \ldots, a_{i_{e_i}}\}) \cup \{p_i, p_i^2, \ldots, p_i^{e_i}\}.
\]

This procedure, which essentially substitutes elements of \( A \) with powers of one of its prime divisors (not necessarily the smallest), defines the shifting operation

\[
S : CS_k(n) \rightarrow CS_k(n), \quad A \mapsto A_{\pi(n)}.
\]

**Lemma 1.** For every \( A \in CS_k^{\max}(n) \) and every \( i \in [\pi(n)] \) we have \( e_i = \lfloor \log_{p_i} n \rfloor \).

**Proof.** Otherwise \( A \cup \{p_i^j\} \) would be a larger \((n, k)\)-shelf than \( A \) for some \( j \leq \lfloor \log_{p_i} n \rfloor \).

**Lemma 2.** We have \(|S(A)| = |A|\) for \( A \in CS_k(n) \) and especially \( S(CS_k^{\max}(n)) \subseteq CS_k^{\max}(n) \).

**Proof.** We have \(|\{a_{i_1}, a_{i_2}, \ldots, a_{i_{e_i}}\}| = |\{p_i, p_i^2, \ldots, p_i^{e_i}\}|\) and from the definition of \( \prec \) we get \( A_{i-1} \cap \{p_i, p_i^2, \ldots, p_i^{e_i}\} \subseteq \{a_{i_1}, a_{i_2}, \ldots, a_{i_{e_i}}\} \) for every \( i \in [\pi(n)] \). This implies \(|A_i| = |A_{i-1}|\).

We use the notation \( \mathbb{P}(x, y) := \mathbb{P} \cap \{x, y\} \) for \( x, y \in \mathbb{R} \). Since \( S(\mathbb{P}(1, n)) = \mathbb{P}(1, n) \in CS_1(n) \) and \( S(\mathbb{P}(1, n) \cup \{p^2 : p \in \mathbb{P}(1, \sqrt{n})\}) = \mathbb{P}(1, n) \cup \{p^2 : p \in \mathbb{P}(1, \sqrt{n})\} \in CS_2(n) \) and both sets cannot be expanded, as proper supersets do not belong to \( CS_1(n) \) resp. \( CS_2(n) \), we obtain

**Proposition 2.** For \( n \in \mathbb{N} \) we have \( \mathbb{P}(1, n) \in CS_1^{\max}(n) \) and \( \mathbb{P}(1, n) \cup \{p^2 : p \in \mathbb{P}(1, \sqrt{n})\} \in CS_2^{\max}(n) \).

Thus, the first interesting case is \( k = 3 \). To approach this problem we are going to use

**Lemma 3.** For every \( A \in CS_3^{\max}(n) \) the set \( S(A) \) contains only the number 1, prime powers and elements of the form \( pq \) with \( p \in \mathbb{P}(\sqrt{n}, \sqrt{n}) \) and \( q \in \mathbb{P}(\sqrt{n}, n) \).

In view of Lemmas 2 and 3 we solely need to maximize the number of elements \( pq \) with \( p \in \mathbb{P}(\sqrt{n}, \sqrt{n}) \) and \( q \in \mathbb{P}(\sqrt{n}, n) \), while using every \( p \) at most once and every \( q \) at most twice, to construct a maximum \((n, 3)\)-shelf. To find such a maximal number of products we will make use of the following graph theoretical result from Claude Berge [1]:

**Theorem B** (Berge 1957). A matching \( M \) of a graph \( G \) is a maximum matching if and only if there exists no \( M \)-augmenting path in \( G \).
3 Construction procedure for \( k = 3 \)

We build the bipartite graph \( G(n) \) (see Figure 1) in the following way:

- It has vertex multiset \( \mathbb{P}(\sqrt{n}, \sqrt{n}) \cup \mathbb{P}(\sqrt{n}, n) \cup \mathbb{P}(\sqrt{n}, n) \) and
- an edge exactly between every \( p \in \mathbb{P}(\sqrt{n}, \sqrt{n}) \) and \( q \in \mathbb{P}(\sqrt{n}, n) \) with \( pq \leq n \).

To construct a maximum matching in \( G(n) \), we use the following procedure:

1. Start with the empty matching \( M \) and draw \( G(n) \) as shown in Figure 1.

2. Take the highest (meaning the smallest) not matched element \( q \) from \( \mathbb{P}(\sqrt{n}, n) \cup \mathbb{P}(\sqrt{n}, n) \) and find the largest (also the highest) not matched element \( p \) from \( \mathbb{P}(\sqrt{n}, \sqrt{n}) \) being adjacent to \( q \). Expand \( M \) by \( \{p, q\} \).

3. Repeat the second step until it is not possible anymore.

4. Denote the resulting matching by \( M(n) \) and the set of not matched elements from \( \mathbb{P}(\sqrt{n}, \sqrt{n}) \) by \( F(n) \).

---

1 We could also view it as a set of size \( 2\pi(n) - \pi(\sqrt{n}) - \pi(\sqrt{n}) \) and label the vertices accordingly. Our way avoids having to distinguish between vertex and label, which could become confusing.
Lemma 4. For every \( n \in \mathbb{N} \) the matching \( M(n) \) is maximum in \( G(n) \).

Proof. Suppose, \( M(n) \) is not a maximum matching. From Theorem B we know that we find a shortest \( M(n) \)-augmenting path

\[ a_1, b_1, a_2, b_2, \ldots, a_r, b_r, \]

with \( a_i \in \mathbb{P}(\sqrt[3]{n}, \sqrt{n}) \), \( b_i \in \mathbb{P}(\sqrt{n}, n) \) for \( i \in [r] \). From the construction procedure follows that at least one of two adjacent vertices in \( G(n) \) are matched. Therefore, we have \( r \geq 2 \).

Since the construction method always picks the largest available element from \( \mathbb{P}(\sqrt[3]{n}, \sqrt{n}) \) and \( \{a_2, b_1\} \in M(n) \) holds, we have \( a_1 < a_2 \). From this we get

\[ a_1b_2 < a_2b_2 \leq n, \]

meaning there is an edge between \( a_1 \) and \( b_2 \) in \( G(n) \). But then

\[ a_1, b_2, a_3, b_3, \ldots, a_r, b_r \]

would be a shorter \( M(n) \)-augmenting path which yields a contradiction. 

Combining this result with Lemmas 2 and 3 yields:

**Theorem 1.** For every \( n \in \mathbb{N} \) the set

\[ \{1\} \cup \{p^i \leq n : p \in \mathbb{P}, i \in \{1, 2, 3\}\} \cup M(n) \cup \left\{a_1a_2, a_3a_4, \ldots, a_{2\lfloor s/2\rfloor-1}a_{2\lfloor s/2\rfloor}\right\}, \]

where \( F(n) = \{a_1, a_2, \ldots, a_s\} \), is a maximum \( (n, 3) \)-shelf.

By carefully examining the changes of \( M(n) \) when it is transitioned into \( M(n + 1) \) one gets

**Proposition 3.** For \( n \in \mathbb{N} \) we have

\[
\begin{align*}
\text{cn}_k(n + 1) & \in \begin{cases} 
\{\text{cn}_k(n) + 1\} & \text{if } n + 1 \in \mathbb{P}, \\
\{\text{cn}_k(n), \text{cn}_k(n) + 1\} & \text{if } \sqrt{n + 1} \in \mathbb{P} \text{ or } n + 1 = pq \text{ for } p, q \in \mathbb{P} \\
\{\text{cn}_k(n), \text{cn}_k(n) + 1, \text{cn}_k(n) + 2\} & \text{if } \sqrt{n + 1} \in \mathbb{P}, \\
\{\text{cn}_k(n)\} & \text{else.}
\end{cases}
\end{align*}
\]

The true value in the second and the third case depends on the number of neighbours of a given prime \( p \) in \( G(n) \) and on how the number of primes lying in \( F(n) \). Since both quantities are difficult to determine exactly and also vary largely, it does not seem possible to get good approximations for \( \text{cn}_k(n) \) from the presented method.

4 Construction ideas for \( k \geq 4 \)

Adjusting the previous method will in general not yield a maximum \( (n, k) \)-shelf for \( k \geq 4 \).

On the one hand this is due to the fact that the corresponding bipartite graph also contains elements from \( \mathbb{P}(\sqrt[3]{n}, \sqrt{n}) \) more than once. For example in case of \( n = 1202 \) and \( k = 4 \) this would cause the constructed matching not to be maximum (31, 37, 23, 47, 11, 59 would be an augmenting path).
On the other hand elements from $\mathbb{P}(\sqrt[n]{n}, \sqrt{n})$ are still usable, since their number of multiples would be smaller than $k$. This second problem can be dealt with using an idea from [4]. Repeatedly applying the exchange operation $E_{B,C}(A) := (A \setminus B) \cup C$, where $A, B, C \subseteq [n]$, and

**Lemma DWa** (de Wiljes [4]). Let $k, n \in \mathbb{N}$ and $A, B, C \subseteq [n]$ with $o_p(C) \leq k - o_p(A \setminus B)$ for all $p \in \mathbb{P}$. Then $E_{B,C}(A) \in CS_k(n)$.

as well as

**Lemma DWb** (de Wiljes [4]). Let $k \in \mathbb{N}_{\geq 4}$, $l \in \mathbb{N}$ and $n \in \mathbb{N}$ with $(k - 1)\pi(\sqrt{n^2}) - (2k - 1)\pi(\sqrt{n}) > kl - k$. Then for every $A \in CS_k(n)$ there are $l$ primes $q_1, q_2, \ldots, q_l \in \mathbb{P}(\sqrt[n]{n}, \sqrt{n^2})$ with $o_{q_i}(A) \leq k - 1$ for every $i \in [l]$.

from that paper yields:

**Theorem 2.** Let $k \in \mathbb{N}_{\geq 4}$ and $n \in \mathbb{N}$ with

$$(k - 1)\pi(\sqrt{n^2}) - (2k - 1)\pi(\sqrt{n}) > k^2 - k.$$

Then there exists some $A \in CS_k^{\text{max}}(n)$ of the form

$$A = \{1\} \cup \mathbb{P}(\sqrt[n]{n}, n) \cup \{p^2 : p \in \mathbb{P}(\sqrt[n]{n}, \sqrt{n})\} \cup X \cup Y \cup Z,$$

where

- $o_p(X) = k$ holds for all $p \in \mathbb{P}(1, \sqrt[n]{n})$,
- every element from $X$ is of the form $pq$ with $p \in \mathbb{P}(1, \sqrt[n]{n})$ and $q \in \mathbb{P}(\sqrt[n]{n}, n)$,
- every element from $Y$ is of the form $pq$ with $p \in \mathbb{P}(\sqrt[n]{n}, \sqrt{n})$ and $q \in \mathbb{P}(\sqrt[n]{n}, n)$,
- every element from $Z$ is a product of two distinct elements from $p \in \mathbb{P}(\sqrt[n]{n}, \sqrt{n})$.

**Proof.** We start with an arbitrary maximum $(n, k)$-shelf $A$. Then we apply the following steps, where we only move to the next step if the current one cannot be applied anymore.

- If there exist multiples of elements $p$ from $\mathbb{P}(1, \sqrt[n]{n})$, which are not of the desired form $pq$ with $q \in \mathbb{P}(\sqrt[n]{n}, n)$, they are exchanged using Lemma DWa and Lemma DWb.
- If there is some $p \in \mathbb{P}(\sqrt[n]{n}, \sqrt{n})$, for which $|\{p, p^2\} \cap A| < 2$ holds, we exchange one of its multiples by one of its first two powers using Lemma DWa.
- If there exists some $p \in \mathbb{P}(\sqrt[n]{n}, n)$, which is not contained in $A$, we replace a suitable element of $A$ by $p$, again by using Lemma DWa. Note that it can never happen that some $pq$ with $q \in \mathbb{P}(1, \sqrt[n]{n})$ can be exchanged by $p$ (otherwise $A$ would not have maximal cardinality).

The resulting set has the desired form. □

We only have to maximize $|Y| + |Z|$ to get a maximum $(n, k)$-shelf, since the elements of $\mathbb{P}(1, \sqrt[n]{n})$ can be paired with “free” primes from $p \in \mathbb{P}(\sqrt[n]{n}, n)$ using Lemma DWb without changing given $Y$ and $Z$.

A promising attempt would be to use the max-flow-min-cut theorem for networks (see for example [2]). The following graph $G_k(n) = (V, E)$ for $k \in \mathbb{N}_{\geq 4}$ should be considered:
\( V = \{s\} \cup \mathcal{P}(\sqrt[3]{n}, \sqrt{n}) \cup \mathcal{P}(\sqrt{n}, n) \cup \{t\}, \) where \( s \) is the source and \( t \) is the sink.

- The graph has edges between \( s \) and every element of \( \mathcal{P}(\sqrt[3]{n}, \sqrt{n}) \), between \( t \) and every element of \( \mathcal{P}(\sqrt{n}, n) \), and between \( p \in \mathcal{P}(\sqrt[3]{n}, \sqrt{n}) \) and \( q \in \mathcal{P}(\sqrt{n}, n) \) iff \( pq \leq n \).

- The capacity function \( c : E \to \mathbb{R} \) has values

\[
\begin{align*}
    c(\{s, \cdot\}) &= k - 2, \\
    c(\{t, \cdot\}) &= k - 1, \\
    c(\{p, q\}) &= 1.
\end{align*}
\]

The given choice of the capacity function ensures that a flow in \( G_k(n) \) cannot use elements from \( \mathcal{P}(\sqrt[3]{n}, \sqrt{n}) \) more than \( k - 2 \) times, elements from \( \mathcal{P}(\sqrt{n}, n) \) more than \( k - 1 \) times, and products of the form \( pq \) with \( p \in \mathcal{P}(\sqrt[3]{n}, \sqrt{n}) \) and \( q \in \mathcal{P}(\sqrt{n}, n) \) more than once.

We now have to find a maximum flow \( f \) in \( G_k(n) \) while also guaranteeing that the differences of \( f \) and the capacity on the edges starting at \( s \) yield a degree sequence of a graph or are at least close to one. Then the edges of the form \( \{p, q\} \) used by \( f \) (meaning where \( f \) has value 1) produce the set \( Y \) in Theorem 2, and from the degree sequence we can construct \( Z \) (by the same procedure as in the lower bound in [3]). The author does not know yet how to solve the “close to degree sequence” problem.

References

[1] C. Berge, Two theorems in graph theory, Proceedings of the National Academy of Sciences of the United States of America 43 (1957), no. 9, 842–844.

[2] A. Bondy and U. S. R. Murty, Graph Theory, Springer, London, 2008.

[3] J.-H. de Wiljes, Complete Subgraphs of the Coprime Hypergraph of Integers I: Introduction and Bounds, European Journal of Mathematics 3 (2017), no. 2, 379–386.

[4] J.-H. de Wiljes, Complete Subgraphs of the Coprime Hypergraph of Integers II: Structural Properties, to appear in European Journal of Mathematics.