CrossMark

Connection between a class of polynomial optimization problems and maximum cliques of non-uniform hypergraphs

Yanming Chang · Yuejian Peng · Yuping Yao

Abstract In 1965, Motzkin and Straus provided a connection between the order of a maximum clique in a graph G and a global solution of a quadratic optimization problem determined by G which is called the Lagrangian function of G. This connection and its extensions have been useful in both combinatorics and optimization. In 2009, Rota Bulò and Pelillo extended the Motzkin–Straus result to r-uniform hypergraphs by establishing a one-to-one correspondence between local (global) minimizers of a family of homogeneous polynomial functions of degree r (different from Lagrangian function) and the maximal (maximum) cliques of an r-uniform hypergraph. In this paper, we study similar optimization problems related to non-uniform hypergraphs and obtain some extensions of their results to non-uniform hypergraphs. In particular, we provide a one-to-one correspondence between local (global) minimizers of a family of non-homogeneous polynomial functions and the maximal (maximum) cliques of $\{1, r\}$-hypergraphs. An application of a main result gives an upper bound on the Turán density of complete $\{1, r\}$-hypergraphs. The approach applied in the proof follows from the approach in Rota Bulò and Pelillo (2009).

Keywords Hypergraph · Maximum clique · Polynomial optimization

Y. Chang · Y. Peng (✉) · Y.Yao
College of Mathematics, Hunan University, Changsha 410082,
People's Republic of China
e-mail: ypeng1@hnu.edu.cn

Y. Chang
e-mail: mingmingmonkey@163.com

Y. Yao
e-mail: yupingYao1989@163.com

1 Introduction and main results

A hypergraph $H = (V, E)$ consists of a vertex set V and an edge set $E \subset 2^V$, where 2^V is the family of all subset of V. For a hypergraph H, let $V(H)$ denote the vertex set of H and $E(H)$ denote the edge set of H. If all edges of H have the same cardinality r, then H is an *r-uniform hypergraph* or *r-graph*. The set $T(H) = \{|e| : e \in E(H)\}$ is called the set of *edge types* of hypergraph H. Hypergraph H is non-uniform if it has at least two edge types. If $T(H) = \{r_1, r_2, \ldots, r_l\}(r_1 < r_2 < \ldots < r_l)$, then we say that H is an $\{r_1, r_2, \ldots, r_l\} - hypergraph$. For any $r \in T(H)$, the *r-th level hypergraph H^r* is the r-uniform hypergraph consisting of all edges of H with r vertices. For a positive integer r, let $V^{(r)}$ be the family of all r-subsets of V. We write $\overline{H}^r = V^{(r)} \backslash H^r$, then the complement of H is $\overline{H} = (V, \overline{H}^{r_1} \cup \overline{H}^{r_2} \cup \ldots \cup \overline{H}^{r_l})$. For any integer $n \in \mathbb{N}$, denote the set $\{1, 2, \ldots, n\}$ by $[n]$. Let $[t]^{\{r_1, r_2, \ldots, r_l\}}$ denote the complete $\{r_1, r_2, \ldots, r_l\}$-hypergraph of order t, that is, the $\{r_1, r_2, \ldots, r_l\}$-hypergraph of order t containing all r_i-subsets of the vertex set V for $1 \leq i \leq l$. A hypergraph H is a *subhypergraph* of a hypergraph G, denoted by $H \subseteq G$ if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. A complete subhypergraph of a hypergraph H with the same edge type as $T(H)$ is called a clique of H. If $U \subseteq V(H)$, then the subhypergraph of H induced by U is denoted by $H[U]$. We assume that all hypergraphs or graphs have the vertex set $[n]$ throughout the paper if it is not specified. The *characteristic vector* of a set U, denoted by $\vec{x}^U = (x_1^U, x_2^U, \ldots, x_n^U)$, is the vector in S defined as:

$$x_i^U = \frac{1_{i \in U}}{|U|}$$

where $|U|$ denotes the cardinality of U and 1_P is the indicator function returning 1 if property P is satisfied and 0 otherwise.

Definition 1.1 Let H be an r-uniform hypergraph. Let $S = \{\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, x_i \geq 0$ for $i = 1, 2, \ldots, n\}$ and let $\vec{x} = (x_1, x_2, \ldots, x_n) \in S$. The Lagrange function of H, denoted by $\lambda(H, \vec{x})$, is

$$\lambda(H, \vec{x}) = \sum_{e \in E(H)} \prod_{i \in e} x_i.$$

The Lagrangian of H, denoted by $\lambda(H)$, is

$$\lambda(H) = \max\{\lambda(H, \vec{x}) : \vec{x} \in S\}.$$

We call $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ a feasible weighting for H if and only if $\vec{x} \in S$. A vector $\vec{y} \in S$ is called an optimal weighting for H if and only if $\lambda(H, \vec{y}) = \lambda(H)$.

Lagrangians were introduced for 2-graphs by Motzkin and Straus (1965). They determined the following expression for the Lagrangian of a 2-graph.

Theorem 1.1 (Motzkin and Straus (1965)) *If G is a 2-graph and the order of its maximum cliques is t, then*

$$\lambda(G) = \lambda\left([t]^{(2)}\right) = \frac{1}{2}\left(1 - \frac{1}{t}\right).$$

Moreover, the characteristic vector of a maximum clique of G is an optimal weighting for G.

This result and its extensions have applications in both combinatorics and optimization. The Lagrangian of a hypergraph has been a useful tool in hypergraph extremal problems. For example, Sidorenko (1987) and Frankl and Füredi (1989) applied graph-Lagrangians of hypergraphs in finding Turán densities of hypergraphs. Frankl and Rödl (1984) applied it in disproving Erdös long standing jumping constant conjecture. The Motzkin–Straus result and its extension were successfully employed in optimization to provide heuristics for the maximum clique problem Bomze (1997); Budinich (2003); Busygin (2006); Gibbons et al. (1997); Pardalos and Phillips (1990). However, the obvious generalization of Motzkin and Straus' result to hypergraphs is false, i.e., the Lagrangian of a hypergraph is not always the same as the Lagrangian of its maximum cliques. In fact, there are many examples of hypergraphs that do not achieve their Lagrangian on any proper subhypergraph. A generalization of Theorem 1.1 to r-uniform hypergraphs was given by Rota Bulò and Pelillo (2009) by associating the edge set of an r-uniform hypergraph H to a polynomial function of degree r different from the Lagrange function. In Rota Bulò and Pelillo (2009) considered the following non-linear program:

$$minimize \ \ h_H(\vec{x}) = \lambda(\overline{H}, \vec{x}) + \tau \sum_{i=1}^{n} x_i^r = \sum_{e \in \overline{H}} \prod_{i \in e} x_i + \tau \sum_{i=1}^{n} x_i^r$$

$$subject \ to \ \ \vec{x} \in S, \tag{1}$$

where $\tau \in \mathbb{R}$ and $\lambda(\overline{H}, \vec{x}) = \sum_{e \in \overline{H}} \prod_{i \in e} x_i$ is the Lagrangian of $\overline{H}$. They obtained the following generalization of Motzkin–Straus Theorem.

Theorem 1.2 (Rota Bulò and Pelillo (2009)) *Let H be an r-uniform hypergraph and $0 < \tau \leq \frac{1}{r(r-1)}$ (with strict inequality for r = 2). A vector $\vec{x} \in S$ is a local (global) solution of (1) if and only if it is the characteristic vector of a maximal (maximum) clique of H. If H has a maximum clique of order t, then h attains its minimum over S at τt^{1-r} and the characteristic vector of a maximum clique is a global solution of (1) (This is true for r = 2 and $\tau = \frac{1}{2}$ as well).*

If H is a 2-graph on n vertices, then

$$1 = (x_1 + x_2 + \ldots + x_n)^2$$

$$= \sum_{i=1}^{n} x_i^2 + 2\left(\sum_{\{i,j\} \in E(H)} x_i x_j + \sum_{\{i,j\} \in E(\overline{H})} x_i x_j\right)$$

$$= \sum_{i=1}^{n} x_i^2 + 2\left(\lambda(H, \vec{x}) + \lambda(\overline{H}, \vec{x})\right).$$

So

$$\lambda(H, \vec{x}) = \frac{1}{2} - \left(\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \lambda(\overline{H}, \vec{x})\right). \tag{2}$$

By (2), to obtain the maximal (maximum) value of $\lambda(H, \vec{x})$ is equivalent to find the minimal (minimum) value of $\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \lambda(\overline{H}, \vec{x})$. So Theorem 1.2 generalizes the Motzkin–Straus result.

Very recently, the study of extremal problems of non-uniform hypergraphs have been motivated by extremal poset problems. We will study a similar problem for non-uniform hypergraphs and explore its applications in determining Turán densities of non-uniform hypergraphs. Given an $\{r_1, r_2, \ldots, r_l\}$-hypergraph H, consider the following non-linear programming.

$$minimize \ g_H(\vec{x}) = \sum_{r \in T(H)} \alpha_r \sum_{e \in \overline{H}^r} \prod_{i \in e} x_i + \sum_{r \in T(H)} \beta_r \sum_{i=1}^{n} x_i^r$$

$$subject \ to \ \vec{x} \in S, \tag{3}$$

where $\alpha_r, \beta_r \in \mathbb{R}$, for all $r \in T(H)$. In this paper, we study problem (3) for $\{r_1, r_2\}$-hypergraphs and $\{1, r\}$-hypergraphs.

When $r_3 = r_4 = \ldots = r_l = 0$, that is, H is an $\{r_1, r_2\}$-hypergraph, (3) can be written as

$$minimize \ p_H(\vec{x}) = \alpha \sum_{e \in \overline{H}^{r_1}} \prod_{i \in e} x_i + \beta \sum_{e \in \overline{H}^{r_2}} \prod_{i \in e} x_i + \gamma \sum_{i=1}^{n} x_i^{r_1} + \tau \sum_{i=1}^{n} x_i^{r_2}$$

$$subject \ to \ \vec{x} \in S, \tag{4}$$

where $\alpha, \beta, \gamma, \tau \in \mathbb{R}$. Let $p_H = \min_{\vec{x} \in S} p_H(\vec{x})$.

Furthermore, when $r_1 = 1$ and $r_2 = r$, it is clear that $\gamma \sum_{i=1}^{n} x_i = \gamma$ is a constant, so we can write $f_H(\vec{x}) = p_H(\vec{x}) - \gamma$ for simplification,

$$minimize \ f_H(\vec{x}) = \alpha \sum_{i \in \overline{H}^1} x_i + \beta \sum_{e \in \overline{H}^r} \prod_{i \in e} x_i + \tau \sum_{i=1}^{n} x_i^r$$

$$subject \ to \ \vec{x} \in S. \tag{5}$$

In order to simplify the notation we write $f_H(\vec{x})$ as $f(\vec{x})$ where the context is non ambiguous. Let $f_H = \min_{\vec{x} \in S} f(\vec{x})$.

We call $\vec{x} = (x_1, x_2, \ldots, x_n)$ a feasible weighting for a non-uniform hypergraph H with n vertices if $\vec{x} \in S$. A *local solution* of (5) is a vector $\vec{x} \in S$ for which there exists

a neighborhood $\Gamma(\vec{x})$ of $\vec{x}$ such that $f(\vec{y}) \geq f(\vec{x})$ for all $\vec{y} \in \Gamma(\vec{x})$. A *global solution* is a vector $\vec{x} \in S$ such that $f(\vec{y}) \geq f(\vec{x})$, for all $\vec{y} \in S$. We say that $\vec{x}$ is a strict local (global) solution if the inequalities are strict for $\vec{y} \neq \vec{x}$. Applying an approach similar in Rota Bulò and Pelillo (2009), we obtain the following results for $\{1, r\}$-hypergraphs and $\{r_1, r_2\}$-hypergraphs.

Theorem 1.3 *Let $\tau > 0$, $\alpha \geq \tau r$ and $\beta \geq \tau r(r-1)$ (with strict inequality for $r = 2$) be constants, let H be a $\{1, r\}$-hypergraph. A feasible weighting $\vec{x}$ is a local (global) solution of (5) if and only if it is the characteristic vector of a maximal (maximum) clique of H. In particular, if H has a maximum clique of order t, then $f_H = \tau t^{1-r}$ and the characteristic vector of a maximum clique is a global solution of (5).*

Theorem 1.4 *Let $\beta, \gamma, \tau > 0$, $r_2 > r_1$ aand $\alpha \geq \gamma r_1(r_1 - 1) + \tau r_2(r_2 - 1)$ be constants. Let H be an $\{r_1, r_2\}$-hypergraph and U be the vertex set of a maximum clique of H^{r_1}, then $p_H = p_{H[U]}$.*

The concept of Turán density of a non-uniform hypergraph F was given in [7]. For a non-uniform hypergraph H on n vertices, the Lubell function of H is defined to be

$$h_n(H) = \sum_{r \in T(H)} \frac{|E(H^r)|}{\binom{n}{r}}.$$

Given a hypergraph F with edge-types T, the Turán density of F is defined to be

$$\pi(F) = \lim_{n \to \infty} \max\{h_n(H) : |v(H)| = n, H \subseteq K_n^T, \text{ and}$$
$$H \text{ does not contain } F \text{ as a subhypergraph.}\}.$$

The proof of the existence of this limit can be found in [7]. Determining the Turán density of a hypergraph in general has been a very challenging problem. Very few results are known and a survey on this topic can be found in [9]. Applying Theorem 1.3, we give an upper bound of the Turán densities of complete $\{1, r\}-$hypergraphs.

Corollary 1.5 *(a) Let $\alpha \geq \tau r$, $\beta \geq \tau r(r - 1)$, $\tau > 0$ be constants, let H be a $\{1, r\}$-hypergraph on n vertices. If H doesn't contain a maximum complete $\{1, r\}$-subhypergraph of order t, then*

$$\frac{\alpha|E(H^1)|}{n} + \frac{\beta|E(H^r)|}{n^r} \leq \alpha + \binom{n}{r}\frac{\beta}{n^r} - \tau\left[(t - 1)^{1-r} - n^{1-r}\right].$$

(b) The Turán density of $[t]^{\{1,r\}}$ satisfies $\pi([t]^{\{1,r\}}) \leq 2 - \frac{1}{r}(t - 1)^{1-r}$.

2 Proofs of the main results

The *support* of a vector $\vec{x} \in S$, denoted by $\sigma(\vec{x})$, is the set of indices corresponding to positive components of $\vec{x}$, i.e.,

$$\sigma(\vec{x}) = \{i : x_i > 0, 1 \leq i \leq n\}.$$

For $\vec{x} \in S$ and function $g(\vec{x})$, let $\partial_j g(\vec{x})$ denote the partial derivative of $g(\vec{x})$ with respect to x_j and $\partial_{jl} g(\vec{x})$ will denote the partial derivative with respect to x_j and x_l. In particular,

$$\partial_j p_H(\vec{x}) = \alpha \sum_{e \in \overline{H}^{r_1}} 1_{j \in e} \prod_{i \in e \setminus \{j\}} x_i + \beta \sum_{e \in \overline{H}^{r_2}} 1_{j \in e} \prod_{i \in e \setminus \{j\}} x_i + \gamma r_1 x_j^{r_1 - 1} + \tau r_2 x_j^{r_2 - 1}, \quad (6)$$

$$\partial_{jl} p_H(\vec{x}) = 1_{j \neq l} \left(\alpha \sum_{e \in \overline{H}^{r_1}} 1_{j,l \in e} \prod_{i \in e \setminus \{j,l\}} x_i + \beta \sum_{e \in \overline{H}^{r_2}} 1_{j,l \in e} \prod_{i \in e \setminus \{j,l\}} x_i \right)$$
$$+ 1_{j=l} \left[\gamma r_1 (r_1 - 1) x_j^{r_1 - 2} + \tau r_2 (r_2 - 1) x_j^{r_2 - 2} \right]. \quad (7)$$

$$\partial_j f(\vec{x}) = \alpha 1_{j \in \overline{H}^1} + \beta \sum_{e \in \overline{H}^r} 1_{j \in e} \prod_{i \in e \setminus \{j\}} x_i + \tau r x_j^{r-1}, \quad (8)$$

$$\partial_{jl} f(\vec{x}) = \beta 1_{j \neq l} \sum_{e \in \overline{H}^r} 1_{j,l \in e} \prod_{i \in e \setminus \{j,l\}} x_i + 1_{j=l} \tau r (r - 1) x_j^{r-2}. \quad (9)$$

Lemma 2.1 (KKT necessary condition, Luenberger (1984)) *If a feasible weighting $\vec{x} = (x_1, x_2, \ldots, x_n)$ is a local solution of (3), then there exists $\theta \in \mathbb{R}$ such that for all $j \in [n]$,*

$$\partial_j g_H(\vec{x}) \begin{cases} = \theta, \ j \in \sigma(\vec{x}), \\ \geq \theta, \ j \notin \sigma(\vec{x}). \end{cases} \quad (10)$$

Lemma 2.2 *A sufficient condition for a feasible weighting $\vec{x} \in S$ to be a local solution of (3) is to be a KKT point and to have the Hessian matrix of $g_H(\vec{x})$ in $\vec{x}$ positive definite on the subspace $M(\vec{x})$ defined as*

$$M(\vec{x}) = \{\vec{\varepsilon} \in \mathbb{R}^n : \sum_{i=1}^n \varepsilon_i = 0, \ \text{and} \ \varepsilon_j = 0 \ \text{for all } j \text{ such that } \partial_j g_H(\vec{x}) > \theta\},$$

where the Hessian matrix of $g_H(\vec{x})$ in $\vec{x}$ is defined as

$$H(\vec{x}) = \left[\partial_{jl} g_H(\vec{x}) \right]_{j,l \in [n]}.$$

In other words, if $\vec{x}$ is a KKT point and for all $\vec{\varepsilon} \in M(\vec{x}) \setminus \{\vec{0}\}$, $\vec{\varepsilon}' H(\vec{x}) \vec{\varepsilon} > 0$, then $\vec{x}$ is a (strict) local solution of (3).

Certainly, Lemmas 2.1 and 2.2 are also suitable for (4) and (5).

2.1 Proofs of Theorem 1.3 and Corollary 1.5

Lemma 2.3 *Let $\alpha, \beta, \tau > 0$ be constants. Let H be a $\{1, r\}$-hypergraph and let a feasible weighting $\vec{x}$ be a local (global) solution of (5). If $H[\sigma(\vec{x})]$ is a clique of H, then it is a maximal (maximum) clique of H and $\vec{x}$ is the characteristic vector of $\sigma(\vec{x})$.*

Proof If $\vec{x}$ is a local solution of (5), then it satisfies the KKT necessary condition in Lemma 2.1. Therefore for every $j \in \sigma(\vec{x})$, we have that $\theta = \partial_j f(\vec{x}) = \tau r x_j^{r-1} > 0$, and it follows that $\vec{x}$ is the characteristic vector of $\sigma(\vec{x})$. Moreover if there exists a set C that contains $\sigma(\vec{x})$, such that $H[C]$ is a clique of H, then for any $j \in C \backslash \sigma(\vec{x})$, $j \in H^1$ and $x_j = 0$, then $\partial_j f(\vec{x}) = 0 < \theta$. This contradicts the KKT necessary condition in Lemma 2.1. Hence, $H[C]$ is a maximal clique of H.

If $\vec{x}$ is the characteristic vector of a maximal clique of order s, then by a direct calculation, $f_H(\vec{x}) = \tau s^{1-r}$ and $f_H(\vec{x}) = \tau s^{1-r}$ decreases as s increases. So $f_H = \tau t^{1-r}$, where t is the maximum clique of H. Combining with the conclusion obtained for a local solution in the previous paragraph, we obtain the conclusion for a global solution.

Hence Lemma 2.3 holds.   □

Lemma 2.4 *Let $\vec{x}$ be a local (global) solution of (5). If both the following conditions hold:*

1. *$\alpha > 0, \tau > 0$ and $\beta \geq \tau r (r-1)$,*
2. *For $r = 2$, $\alpha > 0, \tau > 0$, and $\beta = 2\tau$, the support size of $\vec{x}$ is minimum among all feasible weighting $\vec{y}$ such that $f(\vec{y}) = f(\vec{x})$, then $H^r[\sigma(\vec{x})]$ is a clique in H^r.*

Proof Suppose that there exists $\tilde{e} \in \sigma(\vec{x})^{(r)}$ such that $\tilde{e} \notin H^r$. We define a new feasible weighting $\vec{y}$ for H as follows. Let $j, l \in \tilde{e}$ such that $x_j \leq x_l \leq x_i$ for all $i \in \tilde{e} \backslash \{j, l\}$ and take $y_i = x_i$ for $i \neq j, l$, $y_j = x_j + \varepsilon$ and $y_l = x_l - \varepsilon$, where $0 < \varepsilon \leq x_l$. Then $\vec{y}$ is clearly a feasible weighting for H.

We study the sign of $f(\vec{y}) - f(\vec{x})$ in a neighbourhood of $\vec{x}$ as $\varepsilon \to 0$ by means of the Taylor expansion of f. By Lemma 2.1, $\partial_j f(\vec{x}) = \partial_l f(\vec{x})$ for $j, l \in \sigma(x)$.

$$\begin{aligned}
f(\vec{y}) - f(\vec{x}) &= \varepsilon \partial_j f(\vec{x}) - \varepsilon \partial_l f(\vec{x}) + \frac{\varepsilon^2}{2!} \partial_{jj} f(\vec{x}) - \frac{2\varepsilon^2}{2!} \partial_{jl} f(\vec{x}) + \frac{\varepsilon^2}{2!} \partial_{ll} f(\vec{x}) + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \left[\partial_{jj} f(\vec{x}) + \partial_{ll} f(\vec{x}) - 2 \partial_{jl} f(\vec{x}) \right] + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \left[\tau r (r-1)(x_j^{r-2} + x_l^{r-2}) - 2\beta \sum_{e \in \overline{H}^r} 1_{j,l \in e} \prod_{i \in e \backslash \{j,l\}} x_i \right] + \mathcal{O}(\varepsilon^3).
\end{aligned}$$
(11)

We will distinguish 2 cases, each of which yields a contradiction, hence proving that $H^r[\sigma(\vec{x})]$ is a clique in H^r.

Case a: $\beta > \tau r (r-1)$.

In this case, since at least $\tilde{e} \in \overline{H}^r$ and $x_j \leq x_l \leq x_i$ for all $i \in \tilde{e} \backslash \{j, l\}$, (11) can be written as

$$\begin{aligned}
f(\vec{y}) - f(\vec{x}) &\leq \frac{\varepsilon^2}{2} [\tau r (r-1)(x_j^{r-2} + x_l^{r-2}) - 2\beta x_l^{r-2}] + \mathcal{O}(\varepsilon^3) \\
&\leq \varepsilon^2 [\tau r (r-1) - \beta] x_l^{r-2} + \mathcal{O}(\varepsilon^3).
\end{aligned}$$
(12)

So $f(\vec{y}) - f(\vec{x}) < 0$ for sufficiently small value of ε. This contradicts the assumption that $\vec{x}$ is a local solution of f.

Case b: $\beta = \tau r(r-1)$.

Then

$$
\begin{aligned}
f(\vec{y}) - f(\vec{x}) &= \frac{\varepsilon^2}{2} \left[\tau r(r-1) \left(x_j^{r-2} + x_l^{r-2} \right) - 2\beta \sum_{e \in \overline{H}^r} 1_{j,l \in e} \prod_{i \in e \setminus \{j,l\}} x_i \right] + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \tau r(r-1) \left[\left(x_j^{r-2} + x_l^{r-2} \right) - 2 \sum_{e \in \overline{H}^r} 1_{j,l \in e} \prod_{i \in e \setminus \{j,l\}} x_i \right] + \mathcal{O}(\varepsilon^3).
\end{aligned}
\tag{13}
$$

Let $\mu = 2 \sum_{e \in \overline{H}^r} 1_{j,l \in e} \prod_{i \in e \setminus \{j,l\}} x_i - (x_j^{r-2} + x_l^{r-2})$. It is obvious that $\mu \geq 0$ since at least $\tilde{e} \in \overline{H}^r$ and $x_j \leq x_l \leq x_i$ for all $i \in \tilde{e} \setminus \{j,l\}$.

Case b1: $\mu > 0$.

In this case, $f(\vec{y}) - f(\vec{x}) < 0$ for sufficiently small value of ε. But $\vec{x}$ is a local solution of f, this is a contradiction.

Case b2: $\mu = 0$ and $r = 2$.

Since $r = 2$, $\mathcal{O}(\varepsilon^3)$ is missing in (13), then $f(\vec{y}) - f(\vec{x}) = 0$. Taking $\varepsilon = x_l$, we get another local solution $\vec{y}$ with smaller support size. This contradicts to the minimality of the support size of $\vec{x}$.

Case b3: $\mu = 0$ and $r \geq 3$.

If $\mu = 0$, then $\tilde{e}$ is the only edge in $\overline{H}^r$ with vertices in $\sigma(x)^{(r)}$ that contains both j and l, and $x_i = x_j$ for all $i, k \in \tilde{e}$. Setting $x_i = \xi$ for all $i \in \tilde{e}$.

We define a new feasible weighting $\vec{z}$ as follows. Let $m \in \tilde{e} \setminus \{j,l\}$ and take $z_i = x_i$ for $i \neq j, l, m$, $z_j = x_j + \frac{\varepsilon}{2}$, $z_l = x_l + \frac{\varepsilon}{2}$ and $z_m = x_m - \varepsilon$ where $0 < \varepsilon \leq x_m$. We study the sign of $f(\vec{z}) - f(\vec{x})$ in a neighbourhood of $\vec{x}$ as $\varepsilon \to 0$ by means of the Taylor expansion of f. By Lemma 2.1 $\partial_j f(\vec{x}) = \partial_l f(\vec{x}) = \partial_m f(\vec{x})$ for $j, l, m \in \sigma(x)$.

$$
\begin{aligned}
f(\vec{z}) - f(\vec{x}) &= \frac{\varepsilon^2}{2} \left[\frac{\partial_{jj} f(\vec{x}) + \partial_{ll} f(\vec{x})}{4} + \partial_{mm} f(\vec{x}) - \partial_{jm} f(\vec{x}) - \partial_{lm} f(\vec{x}) + \frac{\partial_{jl} f(\vec{x})}{2} \right] \\
&\quad + \frac{\varepsilon^3}{6} \left[\frac{\partial_{jjj} f(\vec{x}) + \partial_{lll} f(\vec{x})}{8} - \partial_{mmm} f(\vec{x}) - \frac{3\partial_{jlm} f(\vec{x})}{2} \right] + \mathcal{O}(\varepsilon^4),
\end{aligned}
$$

where ∂_{jlm} denotes the partial derivative of f with respect to x_j, x_l, and x_m, i.e.

$$
\partial_{jlm} f(\vec{x}) = 1_{j \neq l} 1_{l \neq m} 1_{j \neq m} \beta \sum_{e \in \overline{H}^r} 1_{j,l,m \in e} \prod_{i \in e \setminus \{j,l,m\}} x_i + 1_{j=l=m} \beta(r-2) x_j^{r-3}.
$$

Recall that $x_j = \xi$ for all $j \in \tilde{e}$, then $\forall j, l \in \tilde{e}$, $\partial_{jj} f(\vec{x}) = \tau r(r-1)\xi^{r-2} = \beta \xi^{r-2}$, $\partial_{jl} f(\vec{x}) = \beta \xi^{r-2}$, and $\forall j, l, m \in \tilde{e}$, $\partial_{jjj} f(\vec{x}) = \tau r(r-1)(r-2)\xi^{r-3} = \beta(r-2)\xi^{r-3}$ and $\partial_{jlm} f(\vec{x}) = \beta \xi^{r-3}$. Hence, the sign of $f(\vec{z}) - f(\vec{x})$ for sufficiently

small values of ε is given by the sign of $-\frac{r-1}{4}\beta\varepsilon^3\xi^{r-3}$ which is clearly negative and this contradicts the local minimality of $\vec{x}$.

Hence Lemma 2.4 holds.   □

Lemma 2.5 *Let $\beta > 0$, $\tau > 0$ and $\alpha \geq \tau r$ be constants. If $\vec{x}$ is a local (global) solution of (5), then $\sigma(\vec{x})^{(1)} \subseteq H^1$.*

Proof If $\sigma(\vec{x})^{(1)} \nsubseteq H^1$, then there are two possible cases to consider.

Case 1: $\exists j, l \in \sigma(\vec{x})$ such that $j \in H^1$ but $l \notin H^1$.

Since $\vec{x}$ is a local solution of (5), it satisfies the KKT necessary condition in Lemma 2.1. Therefore $\forall j, l \in \sigma(\vec{x})$, we have $\theta = \partial_j f(\vec{x}) = \partial_l f(\vec{x})$. But $\partial_j f(\vec{x}) = \tau r x_j^{r-1}$, $\partial_l f(\vec{x}) = \alpha + \tau r x_l^{r-1}$, so $\tau r x_j^{r-1} = \alpha + \tau r x_l^{r-1}$ i.e. $\tau r(x_j^{r-1} - x_l^{r-1}) = \alpha$, then $x_j^{r-1} = \frac{\alpha}{\tau r} + x_l^{r-1} > 1$ since $\alpha \geq \tau r$ and $0 < x_i < 1$ for all $i \in \sigma(x)$. It is a contradiction.

Case 2: $\forall j \in \sigma(x)$, $j \notin H^1$.

For $j \in \sigma(x)$ but $j \in \overline{H}^1$, $\theta = \partial_j f(\vec{x}) = \alpha + \tau r x_j^{r-1}$. Let $l \notin \sigma(x)$, then $\partial_l f(\vec{x}) = \alpha < \theta$, this contradicts the KKT condition.

Hence Lemma 2.5 holds.   □

Applying Lemmas 2.4 and 2.5, we will obtain the following claim:

Claim 2.6 *(a) Let $\tau > 0$, $\alpha \geq \tau r$ and $\beta \geq \tau r(r-1)$ (with strict inequality if $r = 2$) be constants. If $\vec{x}$ is a local (global) solution of (5), then $H[\sigma(x)]$ is a clique of H.*

(b) For $r = 2$, $\alpha > 0$, $\tau > 0$, and $\beta = 2\tau$, if $\vec{x}$ is a local (global) solution of (5) such that the support size of $\vec{x}$ is minimum among all feasible weighting $\vec{y}$ such that $f(\vec{y}) = f(\vec{x})$, then $H[\sigma(x)]$ is a clique of H.

By Lemma 2.3 and Claim 2.6, we can get the following claim:

Claim 2.7 *(a) Let $\tau > 0$, $\alpha \geq \tau r$ and $\beta \geq \tau r(r-1)$ (with strict inequality if $r = 2$) be constants. If $\vec{x}$ is a local (global) solution of (5), then $H[\sigma(x)]$ is a maximal (maximum) clique of H and $\vec{x}$ is the characteristic vector of $\sigma(x)$.*

(b) For $r = 2$, $\alpha > 0$, $\tau > 0$, and $\beta = 2\tau$, if $\vec{x}$ is a local (global) solution of (5) such that the support size of $\vec{x}$ is minimum among all feasible weighting $\vec{y}$ such that $f(\vec{y}) = f(\vec{x})$, then $H[\sigma(x)]$ is a maximal (maximum) clique of H and $\vec{x}$ is the characteristic vector of $\sigma(x)$.

Lemma 2.8 *Let $\tau > 0$, $\alpha \geq \tau r$ and $\beta \geq \tau r(r-1)$ be constants. Let H be a $\{1, r\}$-hypergraph. If $\vec{x}$ is the characteristic vector of a maximal (maximum) clique C of H, then $\vec{x}$ is a strict local (global) solution of (5).*

Proof We will show that $\vec{x}$ is a strict local solution of (5), by showing that it satisfies the sufficient conditions in Lemma 2.2. First we prove that $\vec{x}$ satisfies the KKT necessary condition in Lemma 2.1, for all $j \in \sigma(x)$ we have $\theta = \partial_j f(\vec{x}) = \tau r x_j^{r-1}$ $= \tau r |V(C)|^{1-r}$, since $\vec{x}$ is the characteristic vector of C. For all $j \notin \sigma(x)$, since C is a maximal clique of H, then there exists at least one edge in $\overline{H}^r$ joining j and $r-1$ vertices in C, so $\partial_j f(\vec{x}) \geq \beta |V(C)|^{1-r} \geq \tau r(r-1)|V(C)|^{1-r} \geq (r-1)\theta \geq \theta$. Hence $\vec{x}$ is a KKT point.

Next we show that $H(\vec{x})$ is positive definite on the subspace $M(\vec{x})$. Since $H(\vec{x})|_{\sigma(x)}$ is a diagonal matrix with positive diagonal entries,

$$H(\vec{x})|_{\sigma(x)} = \tau r(r-1)|V(C)|^{r-2}I,$$

where I is the identity matrix. So all eigenvalues of $H(\vec{x})|_{\sigma(x)}$ are positive. This implies that $H(\vec{x})$ is positive definite on the subspace $M(\vec{x})$.

By a direct calculation, $f(\vec{x}) = \tau|V(C)|^{1-r}$ attains its global minimum when $|C|$ is as large as possible, i.e., a maximum clique. □

Proof of Theorem 1.3 Following from Claim 2.7 and Lemma 2.8, we can obtain the first part. Let $\vec{x}$ be global solution of (5) with the minimum support size. Then by Claim 2.7, $\vec{x}$ is the characteristic vector of a maximum clique of H. By a direct calculation, $f_H = f(\vec{x}) = \tau t^{1-r}$. □

Proof of Corollary 1.5 (a) Since H doesn't contain a clique of order t, then the order of its maximum cliques is at most $t-1$. By Theorem 1.3, $f(\vec{x}) \geq \tau(t-1)^{1-r}$. In particular if we take $\vec{x} = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$, then $f((\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})) \geq \tau(t-1)^{1-r}$. By a direct calculation, we obtain that

$$\frac{\alpha|E(H^1)|}{n} + \frac{\beta|E(H^r)|}{n^r} \leq \alpha + \binom{n}{r}\frac{\beta}{n^r} - \tau\left[(t-1)^{1-r} - n^{1-r}\right].$$

(b) Taking $\alpha = 1$, $\beta = r!$, and $\tau = \frac{1}{r}$, we obtain that

$$\begin{aligned}
\pi([t]^{\{1,r\}}) &= \lim_{n\to\infty}\left(\frac{|E(H^1)|}{\binom{n}{1}} + \frac{|E(H^r)|}{\binom{n}{r}}\right) \\
&= \lim_{n\to\infty}\left(\frac{|E(H^1)|}{n} + \frac{r!|E(H^r)|}{n^r}\right) \\
&\leq \lim_{n\to\infty}\left\{1 + \binom{n}{r}\frac{r!}{n^r} - \frac{1}{r}\left[(t-1)^{1-r} - n^{1-r}\right]\right\} \\
&= 2 - \frac{1}{r}(t-1)^{1-r}.
\end{aligned}$$

□

2.2 Proof of Theorem 1.4

Lemma 2.9 *Let H be an $\{r_1, r_2\}$-hypergraph. If $\vec{x}$ is a local (global) solution of (4), then $H^{r_1}[\sigma(x)]$ is a clique of H^{r_1} provided that $\beta, \gamma, \tau > 0$ and $\alpha \geq \gamma r_1(r_1-1) + \tau r_2(r_2-1)$.*

Proof Suppose that there exists $\tilde{e} \in \sigma(x)^{(r_1)}$ such that $\tilde{e} \notin H^{r_1}$. We define a new feasible weighting $\vec{y}$ for H as follows. Let $j, l \in \tilde{e}$ such that $x_j \leq x_l \leq x_i$ for all

$i \in \tilde{e} \backslash \{j, l\}$ and take $y_i = x_i$ for $i \neq j, l$, $y_j = x_j + \varepsilon$ and $y_l = x_l - \varepsilon$, where $0 < \varepsilon \leq x_l$. Then $\vec{y}$ is clearly a feasible weighting for H.

We study the sign of $p_H(\vec{y}) - p_H(\vec{x})$ in a neighbourhood of $\vec{x}$ as $\varepsilon \to 0$ by means of the Taylor expansion of p_H. By Lemma 2.1, $\partial_j p_H(\vec{x}) = \partial_l p_H(\vec{x})$ for $j, l \in \sigma(x)$.

$$
\begin{aligned}
p_H(\vec{y}) - p_H(\vec{x}) &= \varepsilon \partial_j p_H(\vec{x}) - \varepsilon \partial_l p_H(\vec{x}) + \frac{\varepsilon^2}{2!} \partial_{jj} p_H(\vec{x}) \\
&\quad - \frac{2\varepsilon^2}{2!} \partial_{jl} p_H(\vec{x}) + \frac{\varepsilon^2}{2!} \partial_{ll} p_H(\vec{x}) + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \left[\partial_{jj} p_H(\vec{x}) + \partial_{ll} p_H(\vec{x}) - 2\partial_{jl} p_H(\vec{x}) \right] + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \Bigg[\gamma r_1(r_1-1) \left(x_j^{r_1-2} + x_l^{r_1-2} \right) + \tau r_2(r_2-1) \left(x_j^{r_2-2} + x_l^{r_2-2} \right) \\
&\quad - 2 \left(\alpha \sum_{e \in \overline{H}^{r_1}} 1_{j,l \in e} \prod_{i \in e \backslash \{j,l\}} x_i + \beta \sum_{e \in \overline{H}^{r_2}} 1_{j,l \in e} \prod_{i \in e \backslash \{j,l\}} x_i \right) \Bigg] + \mathcal{O}(\varepsilon^3).
\end{aligned}
\tag{14}
$$

Since $x_j \leq x_l \leq x_i$ for all $i \in \tilde{e} \backslash \{j, l\}$ and $r_1 < r_2$, we can estimate (14) as

$$
\begin{aligned}
p_H(\vec{y}) - p_H(\vec{x}) &< \frac{\varepsilon^2}{2} \left[2\gamma r_1(r_1-1)x_l^{r_1-2} + 2\tau r_2(r_2-1)x_l^{r_2-2} - 2\alpha x_l^{r_1-2} \right] + \mathcal{O}(\varepsilon^3) \\
&= \varepsilon^2 \left[\gamma r_1(r_1-1)x_l^{r_1-2} + \tau r_2(r_2-1)x_l^{r_1-2} - \alpha x_l^{r_1-2} \right] \\
&\quad + \varepsilon^2 \tau r_2(r_2-1) \left(x_l^{r_2-2} - x_l^{r_1-2} \right) + \mathcal{O}(\varepsilon^3) \\
&= \varepsilon^2 \left[\gamma r_1(r_1-1) + \tau r_2(r_2-1) - \alpha \right] x_l^{r_1-2} \\
&\quad + \varepsilon^2 \tau r_2(r_2-1) \left(x_l^{r_2-2} - x_l^{r_1-2} \right) + \mathcal{O}(\varepsilon^3).
\end{aligned}
\tag{15}
$$

The value of $p_H(\vec{y}) - p_H(\vec{x})$ is less than zero for small enough ε since $r_2 > r_1$ and $\alpha \geq \gamma r_1(r_1 - 1) + \tau r_2(r_2 - 1)$. This contradicts to that $\vec{x}$ is a local solution of $p_H(\vec{x})$. Hence Lemma 2.9 holds.   □

Proof of Theorem 1.4 Clearly, $p_H \leq p_{H[U]}$. By Lemma 2.9, $\sigma(x)^{(r_1)} \subseteq H^{r_1}$, then $\sigma(x) \subseteq U$ since U is the vertex set of a maximum complete r_1-subhypergraph in H. So we obtain that $p_H = p_{H[\sigma(x)]} \geq p_{H[U]}$. Hence $p_H = p_{H[U]}$.   □

3 Remarks

For $\{r_1, r_2\}$-hypergraphs, we are not able to obtain a result to similar to Theorem 1.3. The obstruction is that we are not able to verify that the vertices corresponding to the support of a solution (with minimum number of positive weights) induce a complete r_2-subgraph in the r_2-level hypergraph. There is no evidence that this is not true, but

we are not able to confirm it by the current method. A result similar to Theorem 1.4 can be obtained for $\{r_1, r_2, \ldots, r_l\}$-hypergraphs.

References

Bomze IM (1997) Evolution towards the maximum clique. J Glob Optim 10:143–164

Budinich M (2003) Exact bounds on the order of the maximum clique of a graph. Discrete Appl Math 127:535–543

Busygin S (2006) A new trust region technique for the maximum weight clique problem. Discrete Appl Math 154:2080–2096

Frankl P, Füredi Z (1989) Extremal problems whose solutions are the blow-ups of the small Witt-designs. J Comb Theory (A) 52:129–147

Frankl P, Rödl V (1984) Hypergraphs do not jump. Combinatorica 4:149–159

Gibbons LE, Hearn DW, Pardalos PM, Ramana MV (1997) Continuous characterizations of the maximum clique problem. Math Oper Res 22:754–768

Johston T, Lu L Turán problems on non-uniform hypergraphs, submitted

Katona G, Nemetz T, Simonovits M (1964) On a graph problem of Turán. Mat Lapok 15:228–238

Keevash P (2011) Surveys in combinatorics. Cambridge University Press, Cambridge, pp 83–140

Luenberger DG (1984) Linear and nonlinear programming. Addison Wesley, Reading

Motzkin TS, Straus EG (1965) Maxima for graphs and a new proof of a theorem of Turán. Can J Math 17:533–540

Pardalos PM, Phillips AT (1990) A global optimization approach for solving the maximum clique problem. Int J Comput Math 33:209–216

Peng Y, Peng H, Tang Q, Zhao C An extension of Motzkin–Straus theorem to non-uniform hypergraphs and its applications, preprint

Rota Bulò S, Pelillo M (2009) A generalization of the Motzkin–Straus theorem to hypergraph. Optim Lett 3:287–295

Sidorenko AF (1987) The maximal number of edges in a homogeneous hypergraph containing no prohibited subgraphs. Math Notes 41: 247–259. Translated from Mat Zametki

Sidorenko AF (1987) Solution of a problem of Bollobas on 4-graphs. Mat Zametki 41:433–455

Sós V, Straus EG (1982) Extremal of functions on graphs with applications to graphs and hypergraphs. J Comb Theory Ser B 63:189–207

Turán P (1941) On an extremal problem in graph theory (in Hungarian). Mat Fiz Lapok 48:436–452