Acta Math. Univ. Comenianae
Vol. LXXXVIII, 3 (2019), pp. 1043–1048

EXTREMAL FAMILIES FOR KRUSKAL-KATONA THEOREM

O. SERRA and L. VENA

ABSTRACT. Given a set of size n and a positive integer $k < n$, Kruskal-Katona theorem gives the minimum size of the shadow of a family S of k-sets of $[n]$ in terms of the cardinality of S. We give a characterization of the families of k-sets satisfying equality in Kruskal-Katona theorem. This answers a question of Füredi and Griggs.

1. INTRODUCTION

The well-known Kruskal-Katona Theorem [**7, 9**] on the minimum shadow of a family of k-subsets of an n-set is a central result in Extremal Combinatorics with multiple applications, see e.g. [**5**]. The *shadow* of a family $S \subset \binom{[n]}{k}$ is the family $\Delta(S) \subset \binom{[n]}{k-1}$ of $(k-1)$-subsets which are contained in some set in S. The Shadow Minimization Problem asks for the minimum cardinality of $\Delta(S)$ of sets S with a given cardinality $m = |S|$. The answer given by the Kruskal-Katona theorem can be stated in terms of k-binomial decompositions. The *k-binomial decomposition* of a positive integer m is

$$m = \binom{a_0}{k} + \binom{a_1}{k-1} + \cdots + \binom{a_t}{k-t}, \quad a_0 > a_1 > \cdots > a_t \geq k-t \geq 1,$$

where the coefficients $a_0 > a_1 > \cdots > a_t \geq k-t \geq 1$ are well defined and uniquely determined by m.

Theorem 1 (Kruskal-Katona [**7, 9**]). *Let $S \subset \binom{[n]}{k}$ be a family of k-subsets of $[n]$ and let*

$$|S| = \binom{a_0}{k} + \binom{a_1}{k-1} + \cdots + \binom{a_t}{k-t}, \quad a_0 > a_1 > \cdots > a_t \geq k-t \geq 1,$$

be the k-binomial decomposition of $|S|$. Then

$$|\Delta S| \geq \binom{a_0}{k-1} + \binom{a_1}{k-2} + \cdots + \binom{a_t}{k-t-1}.$$

Received June 7, 2019.
2010 *Mathematics Subject Classification.* Primary 05D05; Secondary 06A07.
Key words and phrases. Extremal combinatorics; Kruskal Katona theorem; compressed sets.
O. Serra is supported by the Spanish Research Agency under project MTM2017-82166-P.
L. Vena is supported by the project 18-13685Y of the Czech Science Foundation (GAČR).

In this context, we say that a family S is *extremal* if the cardinality of its shadow achieves the lower bound in Theorem 1. For every m, the initial segment of length m in the colex order is an extremal family. We recall that the colex order on the k-subsets of $[n]$ is defined by $x \leq y$ if and only if $\max(x\Delta y) \in y$, where here Δ denotes the symmetric difference of the two k-subsets x and y.

Füredi and Griggs [**6**] (see also Mörs [**10**]) proved that, for cardinalities m for which the k-binomial decomposition has length $t+1<k$, these initial segments in the colex order are in fact the unique extremal families. They also gave some examples which show that this is not the case when $t+1=k$. This prompted the authors to ask about the characterization of the extremal families. The aim of this paper is to give an answer to this question.

For fixed k the set of integers for which the k-binomial decomposition has length k (so $t=k-1$) has asymptotic density one. This suggests that there may be a large number of extremal families besides the initial segments of the colex order. A consequence of our results is that this is indeed the case.

Our approach is based on a variation on the k-binomial decomposition of a positive integer, which will be directly related to the family S of k-subsets. We recall that the i-th shadow of a family S is defined recursively by $\Delta^i(S)=\Delta(\Delta^{i-1}(S))$. Our main result is the following.

Theorem 2. *Let $S\subset\binom{[n]}{k}$ be a family k-subsets of $[n]$. There is a sequence of (non necessarily positive) integers $b_0>b_1>\cdots>b_{k-2}\geq b_{k-1}$ such that, for each $0\leq i\leq k-1$,*

$$|\Delta^i(S)|=\binom{b_0}{k-i}+\binom{b_1}{k-i-1}+\cdots+\binom{b_{k-1}}{-i+1}.$$

Moreover, S is an extremal family if and only if $b_{k-1}\geq 1$.

We call the sequence $b(S)=(b_0,b_1,\ldots,b_{k-1})$ the *shadow k-binomial sequence* of S. For $k\geq 2$, the shadow sequence of a family S of k-subsets can be simply defined recursively by

$$b_i=\begin{cases}|\Delta^{k-1}(S)|-1, & i=0\\ |\Delta^{k-i-1}(S)|-\left(\binom{b_0}{i+1}+\cdots+\binom{b_{i-1}}{2}\right)-1, & 1\leq i<k-1,\\ |S|-\left(\binom{b_0}{k}+\cdots+\binom{b_{k-2}}{2}\right) & i=k-1\end{cases}.$$

The core of Theorem 2 is to show that the terms of the sequence satisfy the inequalities $b_0>b_1>\cdots>b_k\geq b_{k-1}$, and also to obtain some of its properties. Once this is proved, the characterization of extremal sets in the Theorem follows using Theorem 1 and the results in [**6**, Theorem 2.1] and [**10**].

An application of Theorem 2 is to provide several explicit examples of extremal families which are not initial segments in the colex order. An example of the ubiquity of extremal families is the following result.

Theorem 3. *For each family $S \subset \binom{[n]}{k}$ there is a positive integer $n_0 = n_0(S)$ such that, for every $m \geq n_0$, the family S' of k-subsets of $[n+m]$ defined as*

$$S' = \left\{ y \cup z \;\middle|\; y \subset s, s \in S, z \in \binom{[n+1, n+m]}{k - |y|} \right\}, \quad \textit{is extremal.}$$

Example 4. The family of 3-sets in [6] formed by $S = \{\{1,2,3\},\{4,5,6\}\}$ is not extremal as its shadow 3-binomial decomposition is $\binom{5}{3} + \binom{-5}{2} + \binom{-18}{1}$. The family of 3-sets in $[6+19]$ formed by

$$\{1,2,3\}, \{4,5,6\}, \{s \cup i\}_{s \in \binom{6}{2}, i \in [7,25]}, \quad \text{and} \quad \{s \cup \{i,j\}\}_{s \in [6], \{i,j\} \in \binom{[7,25]}{2}}$$

have shadow 3-binomial $\binom{24}{3} + \binom{14}{2} + \binom{1}{1}$ and is hence extremal.

Theorem 3 has Theorem 9 below as a complementary result which describes a structural property of extremal families.

The shadow sequences also provide a tool to analyze the structure of sets which are close to be extremal and derive stability results for the Kruskal-Katona theorem as the ones obtained by Keevash [**8**]. Another application, which was one of the motivation of the present paper, is the analysis of the isoperimetric function of the Johnson graphs (see e.g. [**2, 1**]), that we develop in a forthcoming paper.

2. Downsets

We denote a k-subset $x = \{x_1, \ldots, x_k\} \in \binom{[n]}{k}$ with $x_1 < x_2 < \cdots < x_k$ by the vector $(x_1, \ldots, x_k)$. We consider the partial order in $\binom{[n]}{k}$ defined as

$$x = (x_1, \ldots, x_k) \leq y = (y_1, \ldots, y_k) \Leftrightarrow x_i \leq y_i, \qquad \text{for all } i,\ 1 \leq i \leq k.$$

By a downset we mean a downset in the former order, namely a set S satisfying that $x \in S$ and $y \leq x$ implies $y \in x$. Downsets are equivalent to compressed sets defined in terms of the shift operator used by Erdős, Ko and Rado [**3**] in their proof of the Erdős-Ko-Rado theorem and also used by Frankl [**4**] in his proof of the Kruskal-Katona theorem. The computation of the shadow sequence of a family S is simpler for downsets than for general sets. We need some notation.

For $x \in \binom{[n]}{k}$ the denote the downset generated by x as

$$\langle x \rangle = \left\{ y \in \binom{[n]}{k} \;\middle|\; y \leq x \right\}.$$

For a family $R = \{r_1, \ldots, r_t\} \subset \binom{[n]}{k}$ the downset generated by R is $\langle R \rangle = \langle x_1, \ldots, x_t \rangle = \langle x_1 \rangle \cup \cdots \cup \langle x_t \rangle$. A generating set R is minimal if no proper subset of R generates $\langle R \rangle$. Every downset is generated by the family of all its members and has a uniquely defined minimum generating set.

For every $x = (x_1, \ldots, x_k) \in \binom{[n]}{k}$, the subset of the i largest elements in x is denoted by $x^{(i)} = (x_{k-i+1}, \ldots, x_k) \in \binom{[n]}{i}$. The notation is extended to a family S of k-sets $S \subset \binom{[n]}{k}$ by $S^{(i)} = \{x^{(i)} \mid x \in S\}$. The k-extension of a set $y = (y_1, \ldots, y_i) \in \binom{[n]}{i}$ such that $y_1 > k - i$ and $i \leq k$ is the k-subset obtained from y by adding the largest elements smaller than y_1 and it is denoted by

$$(y)_k = (y_1 - (k-i), y_1 - (k-i-1), \dots, y_1 - 1, y_1, y_2, \dots, y_i) \in \binom{[n]}{k},$$

The *conditional cone* of y is the set of all elements in $\langle (y)_k \rangle$ whose i largest coordinates coincide with y, denoted by

$$K_y = \{x \in \langle (y)_k \rangle : x^{(i)} = y\}.$$

We note that, by definition, for distinct subsets y, y' of $[n]$ with $|y| \le |y'|$

$$\text{either } K_y \subseteq K_{y'} \text{ (when } y^{(|y'|)} = y') \quad \text{or} \quad K_y \cap K_{y'} = \emptyset.$$

Lower shadows of downsets are also downsets and their generating sets can be easily obtained.

Proposition 5 (Generation of lower shadows)**.** *Let $S = \langle R \rangle \subset \binom{[n]}{k}$ be a downset generated by the family R. Then, for each $1 \le i < k$,*

$$\Delta^i(S) = \langle R^{(k-i)} \rangle.$$

The following Proposition gives an expression of a downset $S \subset \binom{[n]}{k}$ in terms of a principal downset and a family of cones.

Proposition 6. *Let $R = \{r_1, \dots, r_t\}$ be the minimal generating set of the downset S. Let $b = \max(\cup_{i=1}^t r_i)$. There is a set $M = M(R) \subset 2^{[n]}$ univocally determined by R such that*

$$S = \langle (b)_k \rangle \setminus \left(\cup_{y \in M} K_y \right).$$

The set $M = M(R)$ given in Proposition 6 can be explicitly described in terms of the elements in R and provides the means to obtain the shadow sequence of S described in Theorem 2, when S is compressed. Let $M_i = \{y \in M : |y| = i\}$, $2 \le i \le k$, and denote by $\lambda_{i,j} = |\{y \in M_j : y_1 = i+1\}|$. The following Theorem describes the shadow sequence of a downset.

Theorem 7. *Let $S = \langle R \rangle$ be the downset of $\binom{[n]}{k}$ generated by R. For $2 \le i \le k$ let $\mu_{i,2} = \lambda_{i,2}$ and let b be the maximum element in $\cup_{x \in S} x$*

$$\mu_{i,j} = \lambda_{i,j} + \max\left\{ \min\{i - (a_{i-1} - 1), 0\} + \sum_{t>i} \mu_{t,j-1}, 0 \right\}, \qquad 3 \le j \le k.$$

The shadow sequence of S is given by

$$a_i = \begin{cases} b - 1 & \text{if } i = 0 \\ a_{i-1} - 1 - \sum_{t \in [1,b]} \mu_{t,i+1} & \text{if } k-1 > i \ge 1, \\ a_{k-2} - \sum_{t \in [1,b]} \mu_{t,k} & i = k-1. \end{cases}$$

3. Shadow sequences

Theorem 2 involves the treatment of any family of sets, not only compressed ones. In the general case the explicit form of the shadow sequence heavily depends on the elements in S. Nevertheless sufficient information can be obtained from the shadow sequence for the applications mentioned in the Introduction.

The approach for the general case is parallel to the one for downsets. We first prove an analogous to Proposition 6.

Proposition 8. *Let S be a family of k-subsets of $[n]$. Let $b = |\cup_{x\in S} x|$ be the size of the support of S. There is a set $M = M(S)$ univocally defined by S such that $S = \langle (b)_k \rangle \setminus \left(\cup_{y\in M}\overline{K}_y\right)$.*

When S is not compressed, $\overline{K}_y$ is a cone analogous to K_y. A labeled tree is associated to the set $M(S)$ which provides parameters $\lambda_{i,j} = \lambda_{i,j}(S)$ analogous to the downset case. Once these are defined, the exact same description from Theorem 7 follows also for non-compressed sets, exchanging the $\lambda_{i,j}$ in Theorem 7 by these analogous $\lambda_{i,j}(S)$, ommiting the reference to the generating set R, and where b is the size of the support of S (as in Proposition 8).

4. Extremal families

Given S, a family of k-sets of $[n]$, we can define the hypergraph $\mathcal{H}(S)$ as follows. The vertex set of $\mathcal{H}(S)$ is $[n]$ and $\{h_1,\dots,h_s\}$ is an edge in $\mathcal{H}(S)$ if and only if, for each $x \in \binom{[n]\setminus\{h_1,\dots,h_s\}}{k-s}$ then $x \cup \{h_1,\dots,h_s\} \notin S$, and there is no subset of $\{h_1,\dots,h_s\}$ with the same extension property. Note that the edge set of $\mathcal{H}(S)$ is precisely the set M from Proposition 8. Any family of k-sets on $[n]$ determines a unique hypergraph $\mathcal{H}(S)$ on $[n]$ whose edges have at most k elements (complexity at most k), no edge is strictly contained in another (simple), and such that for each $s \in [k]$ and for each s-set X of $V(\mathcal{H}(S))$ there exists an $x \in V(\mathcal{H}(S))\setminus X$ such that $x \cup X \notin E(\mathcal{H}(S))$ (no s-set has the complete $s+1$ neighbourhood). Moreover, any hypergraph with the mentioned properties induce a family of k-sets. The hypergraph $\mathcal{H}(S)$ is said to be the *hypergraph of the family S*.

Theorem 9. *Let S be an extremal family of k-sets on $[n]$. Let $\mathcal{H}(S) = (V,E)$ be the hypergraph of the family S.*

Then any hypergraph $\mathcal{H}' = (V,E')$ with $E' \subset E$ is the hypergraph of S', a family of k-sets on $[n]$, which is extremal and contains S.

In particular, the families S' containing S are not trivial as long as $\mathcal{H}' = (V,E')$ is non-empty.

References

1. Christofides D., Ellis D. and Keevash P., *An approximate isoperimetric inequality for r-sets*, Electron. J. Combin. **20** (2013), #15.
2. Diego V., Serra O. and Vena L., *On a problem by Shapozenko on Johnson graphs*, Graphs Combin. **34** (2018), 947–964.

3. Erdős P., Ko C. and Rado R., *Intersection theorems for systems of finite sets*, Q. J. Math. **12** (1961), 313–320.
4. Frankl P., *A new short proof for the Kruskal-Katona theorem*, Discrete Math. **48** (1984), 327–329.
5. Frankl P. and Tokushige N., *The Kruskal-Katona theorem, some of its analogues and applications*, in: Extremal problems for finite sets (Visegrád, 1991), Bolyai Soc. Math. Stud. vol. 3, János Bolyai Math. Soc., Budapest, 1994, 229–250.
6. Füredi Z. and Griggs J. R., *Families of finite sets with minimum shadows*, Combinatorica **6** (1986), 355–363.
7. Katona G., *A Theorem of Finite Sets*, in: Classic Papers in Combinatorics, Springer, 2009.
8. Keevash P., *Shadows and intersections: stability and new proofs*, Adv. Math. **218** (2008), 1685–1703.
9. Kruskal J. B., *The number of simplices in a complex*, Mathematical optimization techniques **10** (1963), 251–278.
10. Mörs M., *A generalization of a theorem of Kruskal*, Graphs Combin. **1** (1985), 167–183.

O. Serra, Department of Mathematics, Universitat Politècnica de Catalunya, Barcelona, Spain,
e-mail: `oriol.serra@upc.edu`

L. Vena, Department of Applied Mathematics, Charles University, Prague, Czech Republic,
e-mail: `lluis.vena@gmail.com`