Minimizing the weight of the union-closure of families of two-sets

Leck, Uwe; Roberts, Ian; Simpson, Jamie

Published in:
Australasian Journal of Combinatorics

Published: 01/01/2012

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Minimizing the weight of the union-closure of families of two-sets

UWE LECK

Department of Mathematics and Computer Science
University of Wisconsin-Superior
Belknap & Catlin POB 2000, Superior, WI 54880
U.S.A.
uleck@uwsuper.edu

IAN T. ROBERTS

School of Engineering and Information Technology
Charles Darwin University
Darwin 0909
Australia
ian.roberts@cdu.edu.au

JAMIE SIMPSON

Centre for Stringology and Applications, DEBII
Department of Mathematics and Statistics
Curtin University of Technology
GPO Box U1987, Perth, WA 6845
Australia
simpson@maths.curtin.edu.au

Abstract

It is proved that, for any positive integer \( m \), the weight of the union-closure of any \( m \) distinct 2-sets is at least as large as the weight of the union-closure of the first \( m \) 2-sets in squashed (antilexicographic) order, where all \( i \)-sets have the same non-negative weight \( w_i \) with \( w_i \leq w_{i+1} \) for all \( i \), and the weight of a family of sets is the sum of the weights of its members. As special cases, solutions are obtained for the problems of minimising size and volume of the union-closure of a given number of distinct 2-sets.
1 Introduction

In recent years there has been a number of papers published on Frankl's Union-Closed Sets Conjecture, which is that for any non-empty union-closed collection of sets there is an element appearing in at least half the sets (see [1] and its list of references, see also [3]). There has, however, been little investigation of such collections independent of this conjecture.

This paper is a step towards solving the following problem: Given positive integers $m, i$, find a family of $m$ $i$-sets such that its union-closure is of smallest possible size. The only known results on this problem of minimising union-closure appear in [2], where it is shown that the family of the first $m$ 2-sets in squashed (antilexicographic) order (see definition below) solves the above problem for $i = 2$, and under the assumption of a ground set of smallest possible size. This paper removes the assumption of a minimum size ground set, and a more general weighted version of the result is given.

Let $[n] := \{1, 2, \ldots, n\}$, and let $2^{[n]}$ denote the power set of $[n]$. All sets and families of sets in this paper are subsets of $[n]$ and $2^{[n]}$, respectively. If $|F| = i$, then $F$ is called an $i$-set. The family of $i$-subsets of $[n]$ is denoted by $\binom{[n]}{i}$. Given a family $\mathcal{F}$, the following notations are used:

$$\mathcal{F}_x := \{F \in \mathcal{F} : x \in F\},$$

$$\mathcal{F}_x := \{F \in \mathcal{F} : x \notin F\}.$$ 

$\mathbb{R}^+$ denotes the set of non-negative real numbers. A function $w : 2^{[n]} \to \mathbb{R}^+$ is called a weight function on $2^{[n]}$, and the weight of a family $\mathcal{F}$ of sets is defined as

$$w(\mathcal{F}) := \sum_{F \in \mathcal{F}} w(F)$$

with $w(\mathcal{F}) = 0$ if $\mathcal{F}$ is empty.

**Property 1.** A weight function $w$ on $2^{[n]}$ is said to have Property 1 if there are $w_0, w_1, \ldots, w_n \in \mathbb{R}^+$ such that $w_0 \leq w_1 \leq \cdots \leq w_n$ and $w(F) = w_i$ whenever $|F| = i$.

Two particular weight functions with Property 1 are of special interest. If $w(F) = 1$ for all $F \subseteq [n]$, then the weight $w(\mathcal{F})$ of a family $\mathcal{F}$ is equal to its size $|\mathcal{F}|$. If $w_i = i$ for all $i$, then $w(\mathcal{F})$ is equal to the volume

$$V(\mathcal{F}) := \sum_{F \in \mathcal{F}} |F|$$

defined of $\mathcal{F}$.

The union-closure of a family $\mathcal{F}$ is

$$\text{UC}(\mathcal{F}) := \{\bigcup_{G \in \mathcal{G}} G : \emptyset \neq G \subseteq \mathcal{F}\} \cup (\mathcal{F} \cap \{\emptyset\}).$$

$\mathcal{F}$ is said to be union-closed if $\text{UC}(\mathcal{F}) = \mathcal{F}$. The generating set of a union-closed family $\mathcal{F}$ is the family $\mathcal{B}_F$ of all members of $\mathcal{F}$ that are not the union of two distinct sets in $\mathcal{F}$, and $\mathcal{F}$ is said to be generated by $\mathcal{B}_F$.

For two distinct $F, G \subseteq [n]$, $F$ occurs before $G$ in squashed order, denoted by $F \prec_S G$, whenever

$$\max(F \Delta G) \in G,$$
where $F \Delta G$ denotes the symmetric difference of $F$ and $G$.

The compressed family $C(m, i)$ is the family of the first $m$ $i$-sets in squashed order. Throughout this paper, $C(m, 2)$ is abbreviated by $C(m)$, and $U(m)$ denotes the union-closure of $C(m)$.

Furthermore, integers $m$ with $0 \leq m \leq \binom{n}{2}$ will be represented in the form

$$m = \left(\frac{a}{2}\right) + b$$

with $0 \leq b \leq a$.  

Proposition 2. Let $m$ and $U(m)$ be as above, and let $w : 2^{[n]} \to \mathbb{R}^+$ be a weight function having Property 1, then

$$w(U(m)) = \sum_{i=2}^{a+1} \left(\frac{a+1}{i} - \frac{a-b}{i-1}\right) w_i.$$  

In particular, $w(U(0)) = 0$.

Proof. By the definition of the squashed order,

$$C(m) = \binom{\left[\frac{a}{2}\right]}{2} \cup \{h, a+1 : h \in [b]\} = \left[\frac{a+1}{2}\right] \setminus \{j, a+1 : j \in [a] \setminus [b]\}.$$  

This implies that

$$U(m) \cap \binom{[n]}{i} = \left[\frac{a+1}{i}\right] \setminus \{S \cup \{a+1\} : S \subseteq [a] \setminus [b], |S| = i-1\}$$

for $i \geq 2$. Hence, the number of $i$-sets in $U(m)$ is $\binom{a+1}{i} - \binom{a-b}{i-1}$ for every $i \geq 2$, which implies the claim.  

In the next section we prove a lemma which will be used to prove our main result, namely Theorem 3. Its proof is given in Section 3 where it is followed by a discussion of the theorem’s consequences and by two conjectures.

Theorem 3. Let $m \leq \binom{n}{2}$ be a non-negative integer. If $G$ is a family of $m$ 2-sets and $w : 2^n \to \mathbb{R}^+$ is a weight function having Property 1, then

$$w(UC(G)) \geq w(U(m)).$$

2 Preparations

Lemma 4. Let $G$ be a non-empty subset of $\binom{[n]}{2}$ and $F := UC(G)$. Furthermore, let $w$ and $w'$ be weight functions on $2^{[n]}$ such that $w$ has Property 1 and $w'(F) = w_{i+1}$ whenever $|F| = i < n$. Then there exists an $x \in [n]$ with $|G_x| \geq 1$ such that

$$w(F_x) \geq w'(F_x) + |G_x| \cdot w_2.$$  

(3)
Proof. The proof is based on the simple argument that was used by Sarvate and Renaud to show that the Union Closed Sets Conjecture is true for families containing a 2-set (see [4]).

Let \( \{x, y\} \in \mathcal{G} \), and define
\[
\mathcal{X} := \{ F \in \mathcal{F} : \{x, y\} \subseteq F, F \setminus \{x\} \notin \mathcal{F}, F \setminus \{y\} \in \mathcal{F}, F \setminus \{x, y\} \in \mathcal{F} \},
\]
\[
\mathcal{Y} := \{ F \in \mathcal{F} : \{x, y\} \subseteq F, F \setminus \{y\} \notin \mathcal{F}, F \setminus \{x\} \in \mathcal{F}, F \setminus \{x, y\} \in \mathcal{F} \}.
\]

Without loss of generality, we can assume
\[
w(\mathcal{X}) \geq w(\mathcal{Y}). \tag{4}
\]

We will show that we then have
\[
w(\mathcal{F}_x \setminus \mathcal{G}_x) \geq w'(\mathcal{F}_x'), \tag{5}
\]
which clearly implies (3).

For \( F \in \mathcal{Y} \), we define \( f(F) := F \setminus \{x, y\} \) and \( \mathcal{Y}' := \{ f(F) : F \in \mathcal{Y} \} \). By the definitions of \( \mathcal{X} \) and \( \mathcal{Y} \), we have \( \mathcal{X} \subseteq \mathcal{F}_x \setminus \mathcal{G}_x \) and \( \mathcal{Y}' \subseteq \mathcal{F}_x' \). Furthermore, \( f : \mathcal{Y} \mapsto \mathcal{Y}' \) is a bijection and \( |F| = |f(F)| + 2 \) for all \( F \in \mathcal{Y} \). This implies that \( w(\mathcal{Y}) \geq w'(\mathcal{Y}') \), and by (4) we obtain
\[
w(\mathcal{X}) \geq w'(\mathcal{Y}'). \tag{6}
\]

For \( F \in \mathcal{F}_x \setminus \mathcal{Y}' \), we define
\[
g(F) := \begin{cases} F \cup \{x\} & \text{if } F \cup \{x\} \in \mathcal{F}, \\ F \cup \{x, y\} & \text{otherwise}. \end{cases}
\]

Let \( F \in \mathcal{F}_x \setminus \mathcal{Y}' \). We will show that \( g(F) \in (\mathcal{F}_x \setminus \mathcal{G}_x) \setminus \mathcal{X} \). As \( \{x, y\} \in \mathcal{G} \), we have \( F \cup \{x, y\} \in \mathcal{F}_x \), which implies \( g(F) \in \mathcal{F}_x \). As \( |g(F)| \geq |F| + 1 \geq 3 \), it follows that \( g(F) \notin \mathcal{G}_x \). Finally, assume for a contradiction that \( g(F) \in \mathcal{X} \). Then \( \{x, y\} \subseteq g(F) \) and \( g(F) \setminus \{x\} \notin \mathcal{F} \), which implies that \( F = g(F) \setminus \{x, y\} \). As \( g(F) \in \mathcal{X} \), we have \( g(F) \setminus \{y\} \in \mathcal{F} \). The definition of \( g \) and \( g(F) \setminus \{y\} \in \mathcal{F} \) imply that \( g(F) = F \cup \{x\} \neq g(F) \), which is a contradiction.

We next show that \( g : \mathcal{F}_x \setminus \mathcal{Y}' \mapsto (\mathcal{F}_x \setminus \mathcal{G}_x) \setminus \mathcal{X} \) is injective. Assume for a contradiction that there are sets \( F, F' \in \mathcal{F}_x \setminus \mathcal{Y}' \), \( F \neq F' \), and \( H \), such that \( g(F) = g(F') = H \). By the definition of \( g \), we have \( x \in H \) and
\[
F, F' \in \{ H \setminus \{x\}, H \setminus \{x, y\} \}.
\]

As \( F \neq F' \), without loss of generality we can assume that \( F = H \setminus \{x\} \) and \( F' = H \setminus \{x, y\} \), where \( y \in H \). As \( F' \notin \mathcal{Y}' \), we have \( H \notin \mathcal{Y} \). As \( F = H \setminus \{x\} \) and \( F' = H \setminus \{x, y\} \) are in \( \mathcal{F} \), \( H \setminus \{y\} \in \mathcal{F} \). On the other hand, \( H \setminus \{y\} = F' \cup \{x\} \), and by the definition of \( g \), we obtain \( g(F') = H \setminus \{y\} \neq H \), which is a contradiction.

As \( g : \mathcal{F}_x \setminus \mathcal{Y}' \mapsto (\mathcal{F}_x \setminus \mathcal{G}_x) \setminus \mathcal{X} \) is injective, we have
\[
w((\mathcal{F}_x \setminus \mathcal{G}_x) \setminus \mathcal{X}) \geq w'(\mathcal{F}_x \setminus \mathcal{Y}'). \tag{7}
\]

Adding (6) and (7), we obtain (5) which completes the proof. \( \blacksquare \)
3 Proof of the main result

Proof of Theorem 3. Let $\mathcal{G}$ be a family of $m$ sets in $\binom{[n]}{2}$, and let $w : 2^n \rightarrow \mathbb{R}^+$ have Property 1.

To show (2), we proceed by induction on $m$. If $m = 0$, then trivially $w(\text{UC}(\mathcal{G})) = w(\mathcal{U}(0)) = 0$. Assume that $m \geq 1$ is represented as in (1), and that the assertion is true for any integer $m'$ with $0 \leq m' < m$.

Let $\mathcal{F}$, $w'$ and $x$ be as in Lemma 4, and define $g_x := |\mathcal{G}_x|$. Without loss of generality, we can assume that $w'(\emptyset) = 0$. By the assumption in Lemma 4, the weight function $w'$ also has Property 1.

Obviously, we have

$$w(\mathcal{F}) = w(\mathcal{F}_x) + w(\mathcal{F}_x). \quad (8)$$

Case 1. Assume that $g_x \leq a - 1$.

Note that if $a = b$, then with $a' = a + 1$ we have $m = \binom{a'}{2}$ and $g_x < a' - 1$. Hence, without loss of generality, we can assume $b \leq a - 1$.

By (3), (8), and the induction hypothesis, we obtain

$$w(\mathcal{F}) \geq w(\mathcal{U}(m - g_x)) + w'(\mathcal{U}(m - g_x)) + g_x w_2.$$ It is clear that the size of $\mathcal{U}(m)$ is strictly increasing with $m$, so the right hand side decreases as $g_x$ increases, and thus it attains its minimum when $g_x = a - 1$. Hence,

$$w(\mathcal{F}) \geq w\left(\mathcal{U}\left(\binom{a-1}{2} + b\right)\right) + w'\left(\mathcal{U}\left(\binom{a-1}{2} + b\right)\right) + (a - 1)w_2$$

and, by Proposition 2,

$$w(\mathcal{F}) \geq \sum_{i=2}^{a} \left(\binom{a}{i} - \binom{a-b-1}{i-1}\right) w_i + w_2$$

Case 2. Assume that $g_x = a \geq b \geq 1$.

In this case, (8) and the induction hypothesis imply

$$w(\mathcal{F}) \geq w(\mathcal{U}\left(\binom{a-1}{2} + (b - 1)\right)) + w(\mathcal{F}_x)$$

$$\geq \sum_{i=2}^{a} \left(\binom{a}{i} - \binom{a-b}{i-1}\right) w_i$$

Case 3. Assume that $g_x = a$, $b = 0$ or $g_x \geq a + 1$.

Note that if $g_x = a$ and $b = 0$, then with $a' = a - 1$ we have $m = \binom{a'}{2} + a'$ and $g_x = a' + 1$. Hence, without loss of generality, we can assume that $g_x \geq a + 1$ and obtain

$$w(\mathcal{F}) \geq w(\mathcal{F}_x) \geq \sum_{i=2}^{a+2} \binom{a+1}{i-1} w_i \geq w(\mathcal{U}(m)).$$

This completes the proof of (2).
4 Cases of special interest

The next corollary provides formulas for size and volume of \( U(m) \). It follows immediately from Theorem 3, as size and volume correspond to weight functions that have Property 1 (see Section 1). The right hand sides of the inequalities in the corollary are equal to \( |U(m)| \) and \( V(U(m)) \), respectively, which is easily derived from Proposition 2 and consistent with formulas given in [2].

**Corollary 5.** Let \( m \leq \binom{n}{2} \) be a non-negative integer represented in the form (1). If \( G \) is a family of \( m \)-2-subsets of \([n]\), then

\[
|UC(G)| \geq 2^{a+1} - 2^{a-b} - a - 1
\]

and

\[
V(UC(G)) \geq (a+1) \cdot 2^a - (a-b+2) \cdot 2^{a-b-1} - a.
\]

These bounds are best possible.

As an example consider the case \( m = 4 \). Since \( m = \binom{3}{2} + 1 \) we have \( a = 3 \) and \( b = 1 \). The first four 2-sets in squashed order are 12, 13, 23, and 14. (For brevity, braces and commas are omitted.) The union-closure of these is

\[
\{12, 13, 123, 23, 14, 124, 134, 1234\}.
\]

So the size of the union-closure is 8 and its volume is 21, in agreement with the formulas above.

5 Concluding remarks

For two distinct sets \( F, G \subseteq [n] \), \( F \) occurs before \( G \) in order \( U \), denoted by \( F <_U G \), whenever

\[
\max F < \max G, \text{ or } \max F = \max G \text{ and } \min(F \Delta G) \in F.
\]

Order \( U \) and squashed order coincide on \( \binom{n}{i} \) for \( i \leq 2 \) but not for larger \( i \). For example, the family of the first seven 3-sets in squashed order is (again, using abbreviated notation without braces and commas)

\[
\mathcal{F} = \{123, 124, 134, 234, 125, 135, 235\}
\]

but the family of the first seven 3-sets in Order \( U \) is

\[
\mathcal{G} = \{123, 124, 134, 234, 125, 135, 145\}.
\]

It is easily checked that both the size and volume of the union-closed family generated by \( \mathcal{G} \) are smaller than the corresponding values for the union-closed family generated by \( \mathcal{F} \).

It is conjectured in [2] that choosing the first \( m \) \( i \)-sets in Order \( U \) simultaneously minimises both size and volume of the union-closure over all choices of generating families of \( m \) \( i \)-sets. That conjecture is generalised here.
Conjecture 6. Let $i \leq n$ and $m \leq \binom{n}{i}$ be non-negative integers, and let $w : 2^{[n]} \rightarrow \mathbb{R}^+$ be a weight function having Property 1. If $\mathcal{B}$ and $\mathcal{B}_U$ are a family of $m$ $i$-subsets of $[n]$ and the family of the first $m$ $i$-subsets of $[n]$ in order $U$, respectively, then

$$w(\text{UC}(\mathcal{B})) \geq w(\text{UC}(\mathcal{B}_U)).$$

In principle, the proof of Theorem 3 can easily be generalised to prove Conjecture 6. However, it would rely upon a generalised version of Lemma 4, the formulation of which is straight-forward. In the case $w \equiv 1$, it becomes the following conjecture.

Conjecture 7. If $\mathcal{B}$ is a collection of $m$ $i$-sets and $\mathcal{U} = \text{UC}(\mathcal{B})$, then there is an $x \in [n]$ with $\mathcal{B}_x \neq \emptyset$ and such that $x$ is contained in at least half of the sets in $\mathcal{U} \setminus \mathcal{B}_x$.

In view of the Union Closed Sets Conjecture, a proof of this seems to be out of reach at this point.

References


(Received 21 Dec 2010; revised 9 Sep 2011)