**DIMACS Technical Report 2000-17**
**June 2000**

# Generating Weighted Transversals of a Hypergraph

by

Endre Boros[1]
RUTCOR, Rutgers University
640 Bartholomew Road
Piscataway NJ 08854
(boros@rutcor.rutgers.edu)

Vladimir Gurvich
Russian Academy of Sciences
Moscow, Russia
(gurvich@rutcor.rutgers.edu)

Leonid Khachiyan[1]
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(leonid@cs.rutgers.edu)

Kazuhisa Makino
Department of Systems and Human Science
Graduate School of Engineering Science
Osaka University
Toyonaka, Osaka, 560-8531, Japan
(makino@sys.es.osaka-u.ac.jp)

[1]Permanent Members of DIMACS

[2]The research of the first two authors was supported in part by the Office of Naval Research (Grant N00014-92-J-1375), the National Science Foundation (Grant DMS 98-06389). The second author's visit to Rutgers University was supported in part by the DIMACS special year on *Large Scale Discrete Optimization*. The research of the third author was supported in part by the National Science Foundation (Grant CCR-9618796). The research of the fourth author was supported in part by a Scientific Grant-in-Aid by the Ministry of Education, Science, Sports, and Culture of Japan.

---

## ABSTRACT

We consider a generalization of the notion of transversal to a finite hypergraph, so called *weighted transversals*. Given a non-negative weight vector assigned to each hyperedge of the input hypergraph, we define a weighted transversal as a minimal vertex set which intersects a collection of hyperedges of sufficiently large total weight. We show that the hypergraph of all weighted transversals is dual-bounded, i.e, the size of its dual hypergraph is polynomial in the number of weighted transversals and the size of the input hypergraph. Our bounds are based on new inequalities of extremal set theory and threshold Boolean logic, which may be of independent interest. For instance, we show that for any threshold frequency, the number of maximal frequent sets of columns in a binary matrix is bounded by the number of minimal infrequent sets of columns in the same matrix multiplied by the number of its rows. We also prove that the problem of generating all weighted transversals for a given hypergraph is polynomial-time reducible to the generation of all ordinary transversals for another hypergraph, i.e., to the well-known hypergraph dualization problem. As a corollary, we obtain an incremental quasi-polynomial-time algorithm for generating all weighted transversals for a given hypergraph. This result includes as special cases the generation of all the minimal Boolean solutions to a given system of non-negative linear inequalities and the generation of all minimal infrequent sets of columns for a given binary matrix.

# 1 Introduction

In this paper we consider some problems involving the generation of all subsets of a finite set satisfying certain conditions. The most well-known problem of this type, the generation of all minimal *transversals*, has applications in combinatorics [33], graph theory [21, 23, 34, 36], artificial intelligence [14], game theory [16, 17, 32], reliability theory [11, 32], database theory [1, 28, 37] and learning theory [2].

Given a finite set $V$ of $n = |V|$ points, and a hypergraph (set family) $\mathcal{A} \subseteq 2^V$, a subset $B \subseteq V$ is called a *transversal* of the family $\mathcal{A}$ if $B \cap A \neq \emptyset$ for all sets $A \in \mathcal{A}$; it is called a *minimal transversal* if no proper subset of $B$ is a transversal of $\mathcal{A}$. The hypergraph $\mathcal{A}^d$ consisting of all minimal transversals of $\mathcal{A}$ is called the *dual* (or *transversal*) hypergraph of $\mathcal{A}$. It is easy to see that if $A \in \mathcal{A}$ is not minimal in $\mathcal{A}$, i.e. if $A' \subset A$ for some $A' \in \mathcal{A}$, then $(\mathcal{A} \setminus \{A\})^d = \mathcal{A}^d$. We can assume therefore that all sets in $\mathcal{A}$ are minimal, i.e. that the hypergraph $\mathcal{A}$ is Sperner. (The dual hypergraph $\mathcal{A}^d$ is Sperner by definition.) It is then easy to verify that $(\mathcal{A}^d)^d = \mathcal{A}$ and $\bigcup_{A \in \mathcal{A}} A = \bigcup_{B \in \mathcal{A}^d} B$.

## 1.1 Dualization

Given a Sperner hypergraph $\mathcal{A}$, a frequently arising task is the generation of the transversal hypergraph $\mathcal{A}^d$. This problem, known as *dualization*, can be stated as follows:

**DUAL($\mathcal{A}, \mathcal{B}$):** Given a complete list of all hyperedges of $\mathcal{A}$ and a set of minimal transversals $\mathcal{B} \subseteq \mathcal{A}^d$, either prove that $\mathcal{B} = \mathcal{A}^d$, or find a new transversal $X \in \mathcal{A}^d \setminus \mathcal{B}$.

Clearly, we can generate all of the hyperedges of $\mathcal{A}^d$ by initializing $\mathcal{B} = \emptyset$ and iteratively solving the above problem $|\mathcal{A}^d|+1$ times. Note also that in general, $|\mathcal{A}^d|$ can be exponentially large both in $|\mathcal{A}|$ and $|V|$. For this reason, the complexity of dualization is customarily measured in the input and output sizes. In particular, we say that $\mathcal{A}^d$ can be generated in *incremental polynomial time* if problem DUAL($\mathcal{A}, \mathcal{B}$) can be solved in time polynomial in $|V|, |\mathcal{A}|$ and $|\mathcal{B}|$.

The dualization problem can be efficiently solved for many classes of hypergraphs. For example, if the sizes of all the hyperedges of $\mathcal{A}$ are limited by a constant $c$, then problem DUAL($\mathcal{A}, \mathcal{B}$) can be solved in polynomial time (see e.g. [7, 14]). In addition, for $c = 2$ there are dualization algorithms that run with polynomial delay, i.e. in $poly(|V|, |\mathcal{A}|)$ time for a specific sequence $\emptyset \subset \mathcal{B}_1 \subset \mathcal{B}_2 \subset \ldots \subset \mathcal{A}^d$, see e.g. [21, 23, 36]). Efficient algorithms exist also for the dualization of 2-monotonic, threshold, matroid, read-bounded, acyclic and some other classes of hypergraphs (see e.g. [4, 10, 12, 13, 25, 26, 30, 31]).

Even though no incremental polynomial time algorithm for the dualization of arbitrary hypergraphs is known, an incremental *quasi-polynomial time* one exists (see [15]). This algorithm solves the dualization problem in $O(nm) + m^{o(\log m)}$ time, where $n = |V|$ and $m = |\mathcal{A}| + |\mathcal{B}|$ (see also [19] for more detail). In fact, the algorithm of [15] can solve problem DUAL($\mathcal{A}, \mathcal{B}$) in $poly(n, m)$ time on $m^{o(\log m)}$ parallel processors. In particular, this means that problem DUAL($\mathcal{A}, \mathcal{B}$) cannot be NP-hard unless any NP-complete problem of size $s$ can be solved in $poly(s)$ time on $s^{o(\log s)}$ processors.

## 1.2 Weighted transversals

In [8, 9] we considered *partial* and *multiple* transversals of a hypergraph (see Sections 1.3 and 1.4), and showed that the generation of those are equivalent with dualization. Improving on a combinatorial inequality instrumental in those proofs, we are further generalizing those results in this paper.

We shall consider the following natural generalization of minimal transversals, the so called *weighted transversals*: Given a (not necessarily Sperner) hypergraph $\mathcal{A} \subseteq 2^V$, a non-negative $r$-dimensional weight vector $w(A) \in \mathbb{R}^r_+$ associated with every hyperedge $A \in \mathcal{A}$, and a threshold $r$-vector $t$, a vertex set $X$ is called a *$w, t$-transversal* if $X$ intersects all the hyperedges of $\mathcal{A}$, except for a sub-family of total weight at most $t$:

$$\sum \{w(A) \mid A \in \mathcal{A},\ A \cap X = \emptyset\} \leq t. \tag{1}$$

We call minimal $w, t$-transversals *weighted* transversals and let $\mathcal{A}_{w,t}$ denote the set of all weighted transversals for $\mathcal{A}$. Note that for the special case

$$r = 1, \quad t = 0, \text{ and } \ w(A) = 1 \ \text{ for all } A \in \mathcal{A}, \tag{2}$$

we have $\mathcal{A}_{w,t} = \mathcal{A}^d$, i.e., the set of weighted transversals turns into the transversal hypergraph for $\mathcal{A}$.

Consider the problem of generating all weighted transversals for a given weighted hypergraph:

**GEN($\mathcal{A}_{w,t}, \mathcal{B}$):** Given a complete list of all hyperedges of $\mathcal{A}$ along with their weights $w(A) \in \mathbb{R}^r_+$, a threshold vector $t \in \mathbb{R}^r_+$ and a set $\mathcal{B}$ of weighted transversals $\mathcal{B} \subseteq \mathcal{A}_{w,t}$, either prove that $\mathcal{B} = \mathcal{A}_{w,t}$, or find a new transversal $X \in \mathcal{A}_{w,t} \setminus \mathcal{B}$.

As mentioned above, problem GEN($\mathcal{A}_{w,t}, \mathcal{B}$) includes DUAL($\mathcal{A}, \mathcal{B}$) as a special case. We show in this paper that in fact these two problems are polynomially related.

**Theorem 1** *Problem GEN($\mathcal{A}_{w,t}, \mathcal{B}$) is polytime reducible to the hypergraph dualization problem.*

In particular, problem GEN($\mathcal{A}_{w,t}, \mathcal{B}$) can be solved in quasi-polynomial time.

The notion of weighted transversals generalize both partial and multiple transversals introduced in [8, 9]. Similarly to those cases, the above equivalence of GEN($\mathcal{A}_{w,t}, \mathcal{B}$) and DUAL($\mathcal{A}, \mathcal{B}$) is based on the fact that the hypergraph of weighted transversals is *uniformly dual-bounded* in the sense that the cardinality of the dual hypergraph of any sub-family $\mathcal{H} \subseteq \mathcal{A}_{w,t}$ can be bounded by a polynomial in the cardinality of $\mathcal{H}$ and the length of the input description of $\mathcal{A}_{w,t}$.

**Theorem 2** *Let $\mathcal{H} \subseteq \mathcal{A}_{w,t}$ be an arbitrary non-empty sub-hypergraph of the hypergraph of all weighted transversals, then*

$$|(\mathcal{A}_{w,t})^d \cap \mathcal{H}^d| \leq r \sum_{H \in \mathcal{H}} |\{A \in \mathcal{A} \mid A \cap H \neq \emptyset\}|, \tag{3}$$

*where $r$ is the dimension of the edge weights of $\mathcal{A}$. In particular,*

$$|(\mathcal{A}_{w,t})^d \cap \mathcal{H}^d| \leq rm|\mathcal{H}|, \tag{4}$$

*where $m$ is the number of hyperedges of $\mathcal{A}$. For $\mathcal{H} = \mathcal{A}_{w,t} \neq \emptyset$ we thus obtain*

$$|(\mathcal{A}_{w,t})^d| \leq rm|\mathcal{A}_{w,t}|. \tag{5}$$

Before proceeding further, in Sections 1.3 and 1.4 we discuss two special cases of the above results.

## 1.3 Minimal feasible and maximal infeasible binary solutions for systems of monotone linear inequalities

Consider a system of $r$ linear inequalities in $n$ binary variables

$$Ax \geq b, \quad x \in \{0,1\}^n \tag{6}$$

where $A$ is a given non-negative $r \times n$ -matrix and $b$ is a given $r$-vector. Let $V = \{1, \ldots, n\}$ be the (index set of the) columns of $A$, and let $\mathcal{A}$ be the hypergraph consisting of $m = n$ singletons $A_1 = \{1\}, \ldots, A_n = \{n\}$. Each column of the $r \times n$ matrix $A$ can be interpreted as a non-negative $r$-dimensional weight vector associated with the corresponding hyperedge of $\mathcal{A} = \{\{1\}, \ldots, \{n\}\}$. Let also $t = Ae - b$, where $e \in \mathbb{R}^n$ is the vector of all ones. Then the characteristic vector of each minimal $w,t$-transversal is a minimal binary solution to (6) and vice versa, the supporting set of any minimal binary solution to (6) is a minimal $w,t$-transversal. It is also easy to see that under this interpretation, the anti-characteristic vector of any set in $(\mathcal{A}_{w,t})^d$ is a maximal infeasible binary vector for (6) and conversely, the complement to the supporting set of any maximal infeasible binary vector for (6) yields a hyperedge of $(\mathcal{A}_{w,t})^d$. From (3) we now conclude that for any feasible system (6),

$$\#\ maximal\ infeasible\ x \quad \leq r \sum \{\, p(x) \mid x\ minimal\ feasible\,\}, \tag{7}$$

where $p(x)$ is the number of positive components in $x$.

**Corollary 1** *Let (6) be a system of $r$ non-negative linear inequalities in $n$ binary variables. Suppose that the set of feasible solutions for (6) is non-empty, i.e., $t = Ae - b \geq 0$. Then*

$$\#\ \textit{maximal infeasible vectors for (6)} \leq\ rn[\#\ \textit{minimal feasible vectors for (6)}].$$

Note that the bound of Corollary 1 is sharp for $r = 1$, e.g., for the inequality $x_1+\ldots+x_n \geq n$. For large $r$, this bound is accurate up to a factor polylogarithmic in $r$. To see this, let $n = 2k$ and consider the system of $r = 2^k$ inequalities of the form

$$x_{i_1} + x_{i_2} + \cdots + x_{i_k} \geq 1,\ i_1 \in \{1,2\},\ i_2 \in \{3,4\}, \ldots, i_k \in \{2k-1, 2k\}. \tag{8}$$

This system has $2^k$ maximal infeasible binary vectors and only $k$ minimal feasible binary vectors, i.e.,

$$\#\ max\ infeasible\ vectors\ for\ (8) = \frac{rn}{2(\log r)^2}[\#\ min\ feasible\ vectors\ for\ (8)].$$

It is also worth mentioning that in general, the number of minimal feasible vectors for (6) cannot be bounded by a polynomial in $r$, $n$, and the number of maximal infeasible vectors for (6). For instance, for $n = 2k$, the system of $k$ inequalities $x_1 + x_2 \geq 1, x_3 + x_4 \geq 1, \ldots, x_{2k-1} + x_{2k} \geq 1$ has $2^k$ minimal feasible binary vectors and only $k$ maximal infeasible binary vectors.

As shown in [24], the problem of generating all maximal infeasible binary vectors for (6) is NP-hard already for binary matrices $A$:

**Theorem 3** ([24]) *Given a system $Ax \geq b$ with 0,1-matrix $A$ and a collection $\mathcal{X} \subseteq \{0,1\}^n$ of maximal infeasible binary points, it is NP-hard to decide whether $\mathcal{X}$ can be extended by adding a new maximal infeasible binary vector for $Ax \geq b$.*

In contrast to Theorem 3, from Theorem 1 we obtain the following.

**Corollary 2** *Consider the problem of generating all minimal feasible binary solutions to a system of non-negative linear inequalities, i.e. the following problem:*

> *Given a system of non-negative Boolean inequalities $Ax \geq b$ and a collection $\mathcal{X}$ of minimal feasible binary vectors for $Ax \geq b$, either show that*
>
> $$\{x \in \{0,1\}^n \mid Ax \geq b\} = \{y \in \{0,1\}^n \mid y \geq x\ for\ some\ x \in \mathcal{X}\},$$
>
> *or find a new minimal binary solution for $Ax \geq b$.*

*The above problem can be reduced in polynomial time to the hypergraph dualization problem. In particular, all minimal feasible binary solutions to $Ax \geq b$ can be enumerated in incremental quasi-polynomial time.*

Let us note that if the number of non-zero entries in each row of $A$ is bounded, then the above problem is obviously solvable in polynomial time.

Let us also add that inequality (7) and Corollaries 1, 2 can in fact be shown to hold for any *monotone* system of linear inequalities in binary variables, i.e. whenever $Ax \geq b$, $y \geq x$, $x, y \in \{0,1\}^n$ implies $Ay \geq b$. (See [9].) Clearly, $A \geq 0$ implies that $Ax \geq b$ is monotone.

## 1.4 Partial transversals, unions, maximal frequent and minimal infrequent sets

We now turn to the special case of weighted transversals for

$$r = 1, \quad t \in \{0, 1, \ldots, |\mathcal{A}| - 1\}, \text{ and } \; w(A) = 1 \; \text{ for all } A \in \mathcal{A}. \tag{9}$$

Following [9], we call weighted transversals for (9) *partial* $t$-transversals. Thus, $X \subseteq V$ is a partial $t$-transversal to $\mathcal{A}$ if $X$ is a minimal set which intersects all but at most $t$ of the hyperedges of $\mathcal{A}$, i.e., $|\{A \in \mathcal{A} \mid A \cap X = \emptyset\}| \leq t$. Let $\mathcal{A}_{e,t}$ be the family of all partial $t$-transversals of $\mathcal{A}$. Define a $t$-*union* from $\mathcal{A}$ as the union of some $t$ hyperedges of $\mathcal{A}$, and let $\mathcal{A}_{u_t}$ denote the family of all minimal $t$-unions of $\mathcal{A}$. In other words, $\mathcal{A}_{u_t}$ is the family of all the minimal subsets of $V$ which contain at least $t$ hyperedges of $\mathcal{A}$. By the above definitions, $t$-union and $t$-transversal families both are Sperner (even if the input hypergraph $\mathcal{A}$ is not). It is also easy to see that the families of all minimal $t$-transversals and $(t+1)$-unions are in fact dual, i.e.,

$$\mathcal{A}_{u_{t+1}} = (\mathcal{A}_{e,t})^d, \quad t = 0, 1, \ldots, m-1.$$

The notion of frequent sets in the data-mining literature (see e.g. [1, 28, 37, 35]) can be related naturally to the families considered above. More precisely, given a binary $m \times n$-matrix $A$ and a threshold $t$, a subset $X$ of the columns of $A$ is called *frequent* if there are more than $t$ rows having a 1 entry in each of the corresponding positions: $\#\{i \mid i \in \{1, \ldots, m\},\ a_{ij} = 1 \ for\ j \in X\} \geq t$. The problems of generating all *maximal frequent* sets and their duals, the so called *minimal infrequent* sets (for a given binary matrix) were proposed, and the complexity of the corresponding decision problems were asked in [35].

Let $V = \{1, \ldots, n\}$ be the column set of $A$ and let $\mathcal{A} = \{A_1, \ldots, A_m\}$ be the hypergraph whose incidence matrix is $A$. Let also $\mathcal{A}^c = \{A_1^c, \ldots, A_m^c\}$ be the complementarity hypergraph for $\mathcal{A}$, where $X^c = V \setminus X$ denotes the complement of a vertex set $X$. Since the family $(\mathcal{A}^c)_{u_{t+1}}$ consists of all the minimal sets $X$ which contain at least $t+1$ hyperedges of $\mathcal{A}^c$, we have

$$X \in (\mathcal{A}^c)_{u_{t+1}} \iff X^c \text{ is a maximal frequent set for } A. \tag{10}$$

Analogously, $(\mathcal{A}^c)_{e,t}$ consists of all the minimal $t$-transversals to $\mathcal{A}^c$, i.e. subsets of $V$ which are contained in at most $t$ hyperedges of $\mathcal{A}$. It is easy to recognize that

$$X \in (\mathcal{A}^c)_{e,t} \iff X \text{ is a minimal infrequent set for } A. \tag{11}$$

If $\mathcal{A}$ is a graph, i.e. all hyperedges of $\mathcal{A}$ are of size 2, the complementary family to $\mathcal{A}_{e,t}$ is also known as the family of the so called *fairly independent sets* of the graph $\mathcal{A}$, i.e. all the vertex subsets which induce at most $t$ edges (see [35]).

Results of [27] imply that the task of generating all maximal frequent sets or, equivalently, all minimal $t$-unions is NP-hard.

**Theorem 4** ([27]) *Given a binary matrix $A$ and a collection of maximal frequent sets of columns for $A$, it is NP-hard to tell whether there are other maximal frequent sets. Equivalently, given a collection of minimal $t$-unions for a hypergraph, it is NP-hard to determine whether the given collection is complete.*

On the other hand, Theorem 1 implies that the tasks of generating partial and ordinary transversals are polynomially equivalent. In view of (11), we thus obtain the following corollary.

**Corollary 3** *The problem of generating all minimal infrequent sets (for a given binary matrix $A$) polynomially reduces to dualization. In particular, all minimal infrequent sets for a given $A$ can be generated in quasi-polynomial incremental time.*

Theorem 2 also yields the following inequality.

**Corollary 4** *Let $A$ be a binary $m \times n$ matrix. Then for any threshold frequency,*

$$\#\ \textit{maximal frequents sets for } A \ \leq\ m\,[\#\ \textit{minimal infrequent sets for } A],$$

*provided that the right-hand side of the above inequality is positive.*

For large $m$, the inequality of Corollary 4 is again accurate up to a factor of $\log m$. For instance, let $A$ be the constraint matrix of (8), i.e., each of the $m = 2^k$ rows of $A$ contains exactly one 0 and 1 in each pair of adjacent columns $\{1,2\},\{3,4\},\ldots,\{2k-1,2k\}$. It is easy to see that for small $t$, this matrix has $2^k$ maximal frequent columns sets and $k$ minimal infrequent sets, i.e.,

$$\#\ maximal\ frequents\ sets\ for\ A = \frac{m}{\log m}[\#\ minimal\ infrequents\ sets\ for\ A].$$

## 1.5 Inequalities

Our proof of Theorems 2 makes use of two lemmas which we state in this section. The first lemma is an intersection inequality for two set families $\mathcal{S} = \{S_1,\ldots,S_\alpha\}$ and $\mathcal{T} = \{T_1,\ldots,T_\beta\}$ on a common ground set $U$, say $U = \{1,\ldots,m\}$. We say that $\mathcal{T}$ *covers all pairwise intersections of* $\mathcal{S}$ if $\alpha \geq 2$ and for any $1 \leq i < j \leq \alpha$ there is an index $k \in \{1,\ldots,\beta\}$ such that $S_i \cap S_j \subseteq T_k$.

We also say that $\mathcal{S}$ and $\mathcal{T}$ are *threshold separable* if there is a non-negative weight-function $w : U \to \mathbb{R}_+$ and a real $t$ such that

$$w(T_k) \leq t\ for\ all\ k \in \{1,\ldots,\beta\}\ \ and\ \ w(S_i) > t\ for\ all\ i \in \{1,\ldots,\alpha\}, \tag{12}$$

where the weight of a set $X \subseteq U$ is defined in the usual way: $w(\emptyset) = 0$ and $w(X) = \sum\{w_u \mid u \in X\}$ for $X \neq \emptyset$. Let us note that $t > 0$ can always be assumed without any loss of generality, since $w(S_i) > 0$ is implied by the above definition, and thus $t = \frac{1}{2}[\min\{w(S_1),\ldots,w(S_\alpha\} + \max\{w(T_1),\ldots,w(T_\beta)\}] > 0$ is also a separating threshold for these families.

**Lemma 1** *Suppose that $\mathcal{S} = \{S_1, \ldots, S_\alpha\}$ and $\mathcal{T} = \{T_1, \ldots, T_\beta\}$ are threshold separable families of subsets of $U = \{1, \ldots, m\}$ such that $\mathcal{T}$ covers all pairwise intersections of $\mathcal{S}$. Then*

$$\alpha \leq \sum_{k=1}^{\beta} |U \setminus T_k|. \tag{13}$$

*In particular, $\alpha \leq m\beta$.*

We note that the threshold separability of $\mathcal{S}$ and $\mathcal{T}$ is essential for the validity of the lemma. In particular, there are Sperner set families $\mathcal{S} = \{S_1, \ldots, S_\alpha\}$ and $\mathcal{T} = \{T_1, \ldots, T_\beta\}$ such that $\mathcal{T}$ covers all pairwise intersections of $\mathcal{S}$ but $\alpha$ is exponentially larger than $\beta$. For instance, let $U$ consist of two special vertices $a, b$ and $s$ disjoint sets $U_1, \ldots, U_s$ of size 3 each, $m = 2 + 3s$. Consider the hypergraph $\mathcal{S}$ with $\alpha = 2s + 3^s$ hyperedges of the following three types:

(a) $U \setminus (\{a\} \cup U_i), \quad i = 1, \ldots, s,$

(b) $U \setminus (\{b\} \cup U_i), \quad i = 1, \ldots, s,.$

(c) $3^s$ hyperedges obtained by selecting exactly one vertex from each $U_i$.

Let $\mathcal{T}$ be the family of all maximal pairwise intersections of two distinct sets from $\mathcal{S}$. It is easy to check that $\mathcal{S}$ has a total of $\beta = s^2$ maximal intersections. Thus, we have $\alpha = 3^{(m-2)/3} + 2(m-2)/3$ and $\beta = (m-2)^2/9$.

Let us remark that Lemma 1 improves on the quadratic inequality shown as Lemma 3 in [8]. Let us add that the interested reader can find several other inequalities on hypergraphs with restricted intersections in Chapter 4 of [3].

To state our second lemma, we need to introduce some notation. Given a monotone Boolean function $h : 2^V \to \{0, 1\}$, let us denote by

$$\max[h] \stackrel{\text{def}}{=} \{X \mid X \subseteq V \textit{ maximal set such that } h(X) = 0\}$$

its family of all maximal false sets, and by

$$\min[h] \stackrel{\text{def}}{=} \{X \mid X \subseteq V \textit{ minimal set such that } h(X) = 1\}$$

its family of all minimal true sets.

Given a family $\mathcal{A} \subseteq 2^V$, a weight function $w : \mathcal{A} \to \mathbb{R}^r_+$, and a threshold $t \in \mathbb{R}^r_+$, let us associate a Boolean function $g_{\mathcal{A},w,t} : 2^V \to \{0, 1\}$ to the family of $w, t$-transversals of $\mathcal{A}$ by defining $g_{\mathcal{A},w,t}(X) = 1$ if and only if $X$ is a $w, t$-transversal of $\mathcal{A}$, i.e. if $X \supseteq A$ for some $A \in \mathcal{A}_{w,t}$. It is clear from the definition of $w, t$-transversals that if $X \subseteq V$ is a $w, t$-transversal of $\mathcal{A}$, and $Y \supseteq X$, then $Y$ is also a $w, t$-transversal of $\mathcal{A}$. Hence, $g_{\mathcal{A},w,t}$ is a monotone Boolean function.

**Lemma 2** *Let $\mathcal{A}$ be a hypergraph on $|V| = n$ vertices, $w(A)$ be non-negative scalar weights associated to each hyperedges $A \in \mathcal{A}$, and $t \geq 0$ be a given real. Let us consider an arbitrary monotone Boolean function $h : 2^V \to \{0,1\}$ such that $h \not\equiv 0$ and $g_{\mathcal{A},w,t}(X) \geq h(X)$ for all $X \subseteq V$. Then, we have*

$$|\max[h] \ \cap \ \{X \mid g_{\mathcal{A},w,t}(X) = 0\}| \ \leq \sum_{X \in \min[h]} |\{A \in \mathcal{A} \mid A \cap X \neq \emptyset\}|. \tag{14}$$

Note that if $\mathcal{A}$ consists of $n$ singletons $\{1\}, \ldots, \{n\}$ then $g_{\mathcal{A},w,t}$ is a threshold function and we obtain the following threshold inequality.

**Corollary 5** ([9]) *Let $h : 2^V \to \{0,1\}$ be a monotone Boolean function such that $w(X) = \sum_{i \in X} w_i > t$, whenever $h(X) = 1$, where $w_1, \ldots, w_n$ and $t$ are given nonnegative reals. If $h \not\equiv 0$, then*

$$|\max[h] \cap \{X \mid w(X) \leq t\}| \ \ \leq \sum_{X \in \min[h]} |X|.$$

*In particular,* $|\max[h] \cap \{X \mid w(X) \leq t\}| \ \leq \ n|\min[h]|$.

If the function $h$ is also threshold and $h \equiv g$, then $|\max[h]| \leq n|\min[h]|$ and, by symmetry, $|\min[h]| \leq n|\max[h]|$, well-known inequalities (see [4, 12, 30, 31]). Corollary 5 thus extends these two threshold inequalities to arbitrary monotone functions $h$.

The remainder of the paper is organized as follows. In Section 2 we discuss the complexity of jointly generating a pair of dual hypergraphs defined via a superset oracle. For a polynomial-time superset oracle the above problem reduces to dualization. This reduction along with the bounds stated in Theorem 2 yields Theorem 1. In Section 3 we prove Lemmas 1, 2, and Theorem 2. In Section 4 we consider various generalizations, and show that essentially all are NP-hard. Finally, Section 5 contains some closing remarks.

# 2 Joint generation of dual hypergraphs

For completeness, we shall recall in this section several useful notions and an algorithm for generating jointly all minimal true and maximal false sets of a monotone Boolean function (see [8, 18]).

## 2.1 Superset oracles

Let $\mathcal{G} \subseteq 2^V$ be a Sperner hypergraph on $V = \{1, \ldots, n\}$. In many applications, $\mathcal{G}$ is represented by a *superset oracle*, i.e., an algorithm which, given an input description $\mathfrak{D}$ of $\mathcal{G}$ and a vertex set $X \subseteq V$, can decide whether or not $X$ contains a hyperedge of $\mathcal{G}$. We assume that the length $|\mathfrak{D}|$ of the input description of $\mathcal{G}$ is at least $n$ and denote by $T_s = T_s(|\mathfrak{D}|)$ the worst-case running time of the oracle on any superset query "Does $X$ contains a hyperedge of $\mathcal{G}$?". In particular, $T_s$ is polynomial-time if $T_s \leq poly(|\mathfrak{D}|)$. In what follows, we do not distinguish the superset oracle and the input description $\mathfrak{D}$ of $\mathcal{G}$. Note that a vertex set $X$ contains a hyperedge of $\mathcal{G}$ if and only if the complement $X^c \stackrel{\text{def}}{=} V \setminus X$ does not contain a hyperedge of $\mathcal{G}^d$. For this reason, any superset oracle $\mathfrak{D}$ for $\mathcal{G}$ can also be regarded as a superset oracle for the dual hypergraph $\mathcal{G}^d$. We list below a few examples.

*1) Weighted transversals.* Let $\mathcal{G}$ be the hypergraph $\mathcal{A}_{w,t}$ of weighted transversal introduced in Section 1.2. Then the input description $\mathfrak{D}$ of $\mathcal{A}_{w,t}$ consists of a complete list of all hyperedges of $\mathcal{A}$, their weights $w(A) \in \mathbb{R}^r_+$ and the threshold vector $t \in \mathbb{R}^r$. Clearly, given a vertex set $X \subseteq V$, we can easily determine whether or not $X$ contains a hyperedge of $\mathcal{A}_{w,t}$ by checking the condition $\sum\{w(A) \mid A \cap X = \emptyset,\ A \in \mathcal{A}\} \leq t$.

*2) Monotone Boolean formulae.* Let $f$ be a $(\vee, \wedge)$-formula with $n$ variables and let $\mathcal{G} = \mathcal{A}_f$ be the supporting sets of all the minimal true vectors for $f$. Then $\mathfrak{D} \sim f$ and the superset oracle checks if (the characteristic vector of) $X \subseteq V$ satisfies $f$. The dual hypergraph $\mathcal{G}^d$ is the set of all the (complements to the support sets of) maximal false vectors of $f$.

*3) Two terminal connectivity.* Consider a digraph $\Gamma$ with a source $s$ and a sink $t$, each arc $a$ of which is assigned a *relay* $x(a) \in V$ (two or more distinct arcs may be assigned identical relays). Let $\mathcal{G}$ be the set of *relay s-t paths*, i.e., minimal subsets of relays that connect $s$ and $t$. Then $\mathfrak{D} \sim \Gamma$, and for a given relay set $X \subseteq V$, the superset oracle can use breadth-first search to check the reachability of $t$ from $s$ via a path consisting of relays in $X$. Note that the dual hypergraph $\mathcal{G}^d$ is the set of all relay *s-t cuts*, i.e., minimal subsets of relays that disconnect $s$ and $t$.

*4) Helly's systems of polyhedra.* Consider a family of $n$ convex polyhedra $P_i \subseteq \mathbb{R}^r,\ i \in V$, and let $\mathcal{G}$ denote the minimal sub-families with no point in common. Then (the complementary hypergraph to) $\mathcal{G}^d$ is the family of all maximal sub-families with a nonempty intersection. (In particular, if $P_1, \ldots, P_n$ are the facets of a convex polytope $Q$, then $\mathcal{G}^d$ corresponds to the set of vertices of $Q$.) We have $\mathfrak{D} \sim (P_1, \ldots, P_n)$ and, given a subsets of polytopes $X \subseteq V$, the superset oracle can use linear programming to check whether $\cap_{i \in X} P_i \neq \emptyset$.

## 2.2 Generation of dual pairs of hypergraphs

In all of the above examples, we have pairs of dual Sperner hypergraphs given by polynomial-time superset oracles. Let $\mathcal{G}, \mathcal{G}^d \subseteq 2^V$ be a pair of dual Sperner hypergraphs given by a superset oracle $\mathfrak{D}$. Consider the following problem GEN$(\mathcal{G}, \mathcal{G}^d, \mathcal{A}, \mathcal{B})$ of *generating jointly* all the hyperedges of $\mathcal{G}$ and $\mathcal{G}^d$:

> Given two explicitly listed set families $\mathcal{A} \subseteq \mathcal{G}$ and $\mathcal{B} \subseteq \mathcal{G}^d$, either prove that these families are complete, $(\mathcal{A}, \mathcal{B}) = (\mathcal{G}, \mathcal{G}^d)$, or find a new set in $(\mathcal{G} \setminus \mathcal{A}) \cup (\mathcal{G}^d \setminus \mathcal{B})$.

For the special case when $\mathcal{A} = \mathcal{G}$ and $\mathfrak{D}$ is a complete list of all the sets in $\mathcal{G} = \mathcal{A}$, we obtain problem $DUAL(\mathcal{A}, \mathcal{B})$ as stated in Section 1.1. In fact, as observed in [5, 18], for any polynomial-time superset oracle $\mathfrak{D}$ problem GEN$(\mathcal{G}, \mathcal{G}^d, \mathcal{A}, \mathcal{B})$ can be reduced in polynomial time to problem $DUAL(\mathcal{A}, \mathcal{B})$. This can be done via the following Algorithm $\mathcal{J}$:

*Step 1.* Check whether each element of $\mathcal{B}$ is a minimal transversal to $\mathcal{A}$, i.e., $\mathcal{B} \subseteq \mathcal{A}^d$. (Recall that $\mathcal{A}$ and $\mathcal{B}$ are given explicitly.) Note that each set $X \in \mathcal{B}$ is a transversal to $\mathcal{A}$ because $\mathcal{A} \subseteq \mathcal{G}$ and $\mathcal{B} \subseteq \mathcal{G}^d$. If some transversal $X \in \mathcal{B}$ is not *minimal* for $\mathcal{A}$ then we can easily find a proper subset $Y$ of $X$ such that $Y$ is also a transversal to $\mathcal{A}$. Since $Y$ is a proper subset of $X$, and $X$ is a minimal transversal to $\mathcal{G}$, $Y$ must miss some hyperedges of $\mathcal{G}$. Hence $Y^c$, the complement of $Y$, contains a hyperedge of $\mathcal{G}$. By querying the superset oracle $\mathfrak{O}$ at most $|Y^c|$ times we can find such a hyperedge $Z \in \mathcal{G}$. Note that $Z \cap Y = \emptyset$ whereas $A \cap Y \neq \emptyset$ for all hyperedges $A \in \mathcal{A}$. This means that $Z$ is a new hyperedge of $\mathcal{G}$. Thus, if the inclusion $\mathcal{B} \subseteq \mathcal{A}^d$ is not satisfied, we can obtain an element in $\mathcal{G} \setminus \mathcal{A}$ and halt.

*Step 2.* Check whether $\mathcal{A} \subseteq \mathcal{B}^d$. If $\mathcal{A}$ contains a non-minimal transversal to $\mathcal{B}$, find a new element in $\mathcal{G}^d \setminus \mathcal{B}$ and halt.

*Step 3.* Suppose that $\mathcal{B} \subseteq \mathcal{A}^d$ and $\mathcal{A} \subseteq \mathcal{B}^d$. Then $\mathcal{B} = \mathcal{A}^d \Rightarrow (\mathcal{A}, \mathcal{B}) = (\mathcal{G}, \mathcal{G}^d)$. (This is because any hyperedge $X \in \mathcal{G} \setminus \mathcal{A}$ is a transversal to $\mathcal{B}$. If $\mathcal{A} = \mathcal{B}^d$, then $X$ must contain some hyperedge of $\mathcal{A}$, a contradiction. By symmetry, the duality of $\mathcal{A}$ and $\mathcal{B}$ also implies the emptiness of $\mathcal{G}^d \setminus \mathcal{B}$.) Hence $(\mathcal{A}, \mathcal{B}) = (\mathcal{G}, \mathcal{G}^d) \Leftrightarrow \mathcal{B} = \mathcal{A}^d$. The condition $\mathcal{B} = \mathcal{A}^d$ can be checked by solving problem $DUAL(\mathcal{A}, \mathcal{B})$. If $\mathcal{B} \neq \mathcal{A}^d$, we obtain a new minimal transversal $X \in \mathcal{A}^d \setminus \mathcal{B}$, see Section 1.1. By definition, $X$ contains no hyperedge in $\mathcal{B}$ and $X^c$ contains no hyperedge in $\mathcal{A}$. Due to the duality of $\mathcal{G}$ and $\mathcal{G}^d$, either (i) $X^c$ contains a hyperedge of $\mathcal{G}$, or (ii) $X$ contains a hyperedge of $\mathcal{G}^d$ , but not both. We can call the superset oracle to decide which of the two cases holds. In case (i) we obtain a new hyperedge in $\mathcal{G} \setminus \mathcal{A}$ by querying the superset oracle at most $|X^c|$ times. Similarly, in case (ii) we get a new hyperedge in $\mathcal{G}^d \setminus \mathcal{B}$ in at most $|X|$ calls to the oracle.

Algorithm $\mathcal{J}$ readily implies the following result.

**Proposition 1** ([5, 18]) *Problem GEN*$(\mathcal{G}, \mathcal{G}^d, \mathcal{A}, \mathcal{B})$ *can be solved in time*

$$n[poly(|\mathcal{A}|, |\mathcal{B}|) + T_s(|\mathfrak{O}|)] + T_{dual},$$

*where* $T_{dual}$ *denotes the time required to solve problem* $DUAL(\mathcal{A}, \mathcal{B})$.

In particular, for any polynomial-time oracle $\mathfrak{O}$, problem GEN$(\mathcal{G}, \mathcal{G}^d, \mathcal{A}, \mathcal{B})$ can be solved in quasi-polynomial time. Thus, for each of the four examples above we can jointly generate all the hyperedges of $(\mathcal{G}, \mathcal{G}^d)$ in incremental quasi-polynomial time. Note, however, that separately generating all the hyperedges of $\mathcal{G}$ or all the hyperedges of $\mathcal{G}^d$ may be substantially harder. For instance, as shown in [18], both problems GEN$(\mathcal{G}, \mathcal{A})$ and GEN$(\mathcal{G}^d, \mathcal{B})$ are NP-hard for examples 2-4. In fact, in example 2 these problems are NP-hard already for $\vee, \wedge$-formulae of depth 3; if the depth is 2 then the formula is either CNF or DNF and we get exactly dualization.

### 2.3 Dual-bounded hypergraphs

Algorithm $\mathcal{J}$ may not be efficient for solving either of the problems GEN$(\mathcal{G}, \mathcal{A})$ or GEN$(\mathcal{G}^d, \mathcal{B})$ separately for the simple reason that we do not control which of the families $\mathcal{G} \setminus \mathcal{A}$ or $\mathcal{G}^d \setminus \mathcal{B}$ contains each new hyperedge produced by the algorithm. Suppose, we want to generate $\mathcal{G}$, and the family $\mathcal{G}^d$ is exponentially larger than $\mathcal{G}$. Then, if we are unlucky, we can get hyperedges of $\mathcal{G}$ with exponential delay, while getting large sub-families of $\mathcal{G}^d$ (which are not needed at all) in between.

Such a problem will not arise and simultaneous generation of $(\mathcal{G}, \mathcal{G}^d)$ can be used to produce $\mathcal{G}$ efficiently, in some sense, if the size of $\mathcal{G}^d$ is polynomially bounded in the size of $\mathcal{G}$ and in the input size $|\mathfrak{O}|$, i.e. when there exists a polynomial $p$ such that

$$|\mathcal{G}^d| \leq p(|\mathfrak{O}|, |\mathcal{G}|). \tag{15}$$

We call such Sperner hypergraphs $\mathcal{G}$ *dual-bounded.*

If $\mathcal{G}$ is dual-bounded, we can generate both $\mathcal{G}$ and $\mathcal{G}^d$ in $|\mathcal{G}^d| + |\mathcal{G}| \leq poly(|\mathfrak{O}|, |\mathcal{G}|)$ rounds of Algorithm $\mathcal{J}$, and hence obtain all the hyperedges of $\mathcal{G}$ in *total quasi-polynomial* time.

This approach, however, may still be inefficient incrementally, i.e., for obtaining a single hyperedge of $\mathcal{G}$ as required in problem GEN$(\mathcal{G}, \mathcal{A})$. If $\mathcal{A}$ is much smaller than $\mathcal{G}$, getting a new hyperedge in $\mathcal{G} \setminus \mathcal{A}$ may require exponentially many (in $|\mathcal{A}|$) rounds of $\mathcal{J}$.

### 2.4 Uniformly dual-bounded hypergraphs

Let us call a Sperner hypergraph $\mathcal{G}$ *uniformly dual-bounded* if

$$|\mathcal{G}^d \cap \mathcal{H}^d| \leq p(|\mathfrak{O}|, |\mathcal{H}|) \text{ for any nonempty hypergraph } \mathcal{H} \subseteq \mathcal{G}.$$

Note that for $\mathcal{H} = \mathcal{G}$ the above condition gives (15).

**Proposition 2** *Problem GEN($\mathcal{G}$) is polytime reducible to dualization for any uniformly dual-bounded hypergraph $\mathcal{G}$ defined by a polynomial-time superset oracle.*

**Proof**. Given a proper sub-family $\mathcal{A}$ of $\mathcal{G}$, we wish to find a new hyperedge $\mathcal{G} \setminus \mathcal{A}$. Start with the given $\mathcal{A}$ and $\mathcal{B} = \emptyset$ and run Algorithm $\mathcal{J}$ repeatedly until it outputs a hyperedge in $\mathcal{G} \setminus \mathcal{A}$. Suppose this takes $k$ rounds of $\mathcal{J}$, then $\mathcal{J}$ produces $k-1$ hyperedges of $\mathcal{G}^d$, each of which is also a hyperedge of $\mathcal{A}^d$, see Step 1 of the algorithm. Since $\mathcal{G}$ is uniformly dual-bounded, we have $k - 1 \leq |\mathcal{A}^d \cap \mathcal{G}^d| \leq p(|\mathfrak{O}|, |\mathcal{A}|)$. □

Theorem 2 states that the hypergraph $\mathcal{G} = \mathcal{A}_{w,t}$ of weighted transversals is uniformly dual-bounded. In view of Proposition 2, this means that Theorem 1 follows from Theorem 2.

# 3 Proofs

As mentioned above, our proof of Theorem 2 relies on Lemmas 1 and 2.

### 3.1 Proof of Lemma 1

First of all, since $\mathcal{T}$ covers all pairwise intersection of $\mathcal{S}$, any inclusion of the form $S_i \subseteq S_j,\ \ i \neq j$ implies $S_i = S_i \cap S_j \subseteq T_k$ for some $k \in \{1, \dots, \beta\}$. However, $S_i \subseteq T_k$ contradicts the threshold separability of $\mathcal{T}$ and $\mathcal{S}$, see (12). We can therefore assume without loss of generality that $\mathcal{S} = \{S_1, \dots, S_\alpha\}$ is a Sperner hypergraph. In particular, all the sets in $\mathcal{S}$ are distinct. Without loss of generality we can also assume that $\mathcal{T}$ is Sperner as well, for otherwise we can replace $\mathcal{T}$ by the family of all maximal subsets of $\mathcal{T}$.

We first prove the lemma for $\alpha = 2$. Assume without loss of generality that $\beta = 1$ and $S_1 \cap S_2 \subseteq T_1$. If $|T_1| \geq |U| - 1$, then (12) implies $S_1 = S_2 = U$, a contradiction. If $|T_1| \leq |U| - 2$, we have $\alpha = 2 \leq |U \setminus T_1|$ and (13) follows.

We assume henceforth that $\alpha \geq 3$ and prove the lemma by induction on $|U| = m$. Clearly, the lemma holds for $m = 1$ because there does not exist a Sperner hypergraph with one vertex and $\alpha \geq 2$ hyperedges. Let $m \geq 2$.

For $u \in U$, let $\alpha_u$ (respectively $\beta_u$) denote the number of hyperedges in $\mathcal{S}$ (respectively $\mathcal{T}$) containing $u$. We split the inductive proof of the lemma into three cases.

*Case 1:* $\beta_u = \beta$ for some $u \in U$. In this case let us define $U' = U \setminus \{u\}$, $\mathcal{T}' = \{T_1 \setminus \{u\}, \dots, T_\beta \setminus \{u\}\}$ and $\mathcal{S}' = \{S_1 \setminus \{u\}, \dots, S_\alpha \setminus \{u\}\}$. Then, the hypergraphs $\mathcal{S}', \mathcal{T}' \subseteq 2^{U'}$ can be separated by the original weight function restricted to $U'$ if we use the threshold value $t' = t - w_u$. By the inductive hypothesis, this gives

$$\alpha \leq \sum_{k=1}^{\beta} |U' \setminus T_k'| = \sum_{k=1}^{\beta} |U \setminus T_k|.$$

*Case 2:* $\beta_u < \beta$ for all $u \in U$, and $\alpha_v \leq 1$ for some $v \in U$, i.e. there is at most one hyperedge in $\mathcal{S}$ which contains vertex $v$, and no vertex of $U$ is contained in all hyperedges in $\mathcal{T}$. If $\alpha_v = 1$, assume without loss of generality that vertex $v$ belongs to $S_\alpha$. Regardless of whether $\alpha_v = 1$ or $\alpha_v = 0$, define

$$U'' = U \setminus \{v\},\ \mathcal{S}'' = \{S_1, \dots, S_{\alpha-1}\},\ \mathcal{T}'' = \{T_1 \setminus \{v\}, \dots, T_\beta \setminus \{v\}\}.$$

Since $\alpha - 1 \geq 2$, hypergraphs $\mathcal{S}''$ and $\mathcal{T}''$ satisfy the assumptions of the lemma with the original weight function $w$ restricted to $U''$ and with the original threshold value $t$. By induction, we have

$$\alpha - 1 \leq \sum_{k=1}^{\beta} |U'' \setminus T_k''| = \sum_{k=1}^{\beta} |U \setminus T_k| - (\beta - \beta_v) \leq \sum_{k=1}^{\beta} |U \setminus T_k| - 1,$$

since $\beta > \beta_v$ holds in this case, and thus the lemma follows.

*Case 3:* $\alpha_u \geq 2$ for all $u \in U$, i.e. each vertex $u \in U$ is covered by at least two sets in $\mathcal{S}$. Let $U^{[u]} = U \setminus \{u\}$, $\mathcal{S}^{[u]} = \{S_i \setminus \{u\} \mid u \in S_i,\ i \in \{1, \dots, \alpha\}\}$, and $\mathcal{T}^{[u]} = \{T_k \setminus \{u\} \mid u \in T_k,\ k \in$

$\{1,\ldots,\beta\}\}$. Since $|\mathcal{S}^{[u]}| = \alpha_u \geq 2$ and hypergraphs $\mathcal{S}^{[u]}$, $\mathcal{T}^{[u]} \subseteq 2^{U^{[u]}}$ satisfy the assumptions of the lemma with the restriction of $w$ to $U^{[u]}$ and $t^{[u]} = t - w_u$, we can apply the inductive hypothesis to get

$$\alpha_u \leq \sum_{k:\ u \in T_k} |U^{[u]} \setminus T_k^{[u]}| = \sum_{k:\ u \in T_k} |U \setminus T_k|, \qquad u = 1, \ldots, m.$$

By multiplying the above inequalities by the non-negative weights $w_u$ and summing up the resulting bounds for all $u$, we obtain:

$$\sum_{u=1}^{m} w_u \alpha_u \leq \sum_{u=1}^{m} w_u \sum_{k:\ u \in T_k} |U \setminus T_k|.$$

From (12) it follows that

$$t\alpha < \sum_{i=1}^{\alpha} w(S_i) = \sum_{u=1}^{m} w_u \alpha_u$$

and

$$\sum_{u=1}^{m} w_u \sum_{k:\ u \in T_k} |U \setminus T_k| = \sum_{k=1}^{\beta} w(T_k)|U \setminus T_k| \leq t \sum_{k=1}^{\beta} |U \setminus T_k|.$$

Hence

$$t\alpha < t \sum_{k=1}^{\beta} |U \setminus T_k|,$$

from which (13) follows, since $t > 0$ can be assumed by (12). This completes the proof of the lemma. $\square$

## 3.2 Proof of Lemma 2

Denote $\max[h] \cap \{X \mid g_{\mathcal{A},w,t}(X) = 0\}$ by $\mathcal{X}$ and let $\mathcal{X} = \{X_1, \ldots, X_\alpha\}$. Let also $\min[h] = \mathcal{Y} = \{Y_1, \ldots, Y_\beta\}$. Since $h \not\equiv 0$, there is at least one minimal true set for $h$, i.e, $\beta \geq 1$.

Consider an arbitrary set $X_i$ in $\mathcal{X}$. Since $g_{\mathcal{A},w,t}(X_i) = 0$, this set cannot contain a $w,t$-transversal to $\mathcal{A}$, i.e., $\sum\{w(A) \mid X_i \cap A = \emptyset,\ A \in \mathcal{A}\} > t$. Equivalently, we have

$$\sum\{w(A) \mid A \subseteq X_i^c,\ A \in \mathcal{A}\} > t, \quad i = 1, \ldots, \alpha, \tag{16}$$

where $X^c = V \setminus X$ is the complements of $X$.

On the other hand, $h(Y_k) = 1$ for any set $Y_k \in \min[h]$. Since $g_{\mathcal{A},w,t}(X) \geq h(X)$ for all $X$, we conclude that $g_{\mathcal{A},w,t}(Y_k) = 1$ for any $k \in \{1, \ldots, \beta\}$. By the definition of $g_{\mathcal{A},w,t}$, this means that each set $Y_k$ contains a $w,t$-transversal to $\mathcal{A}$, i.e, $\sum\{w(A) \mid Y_k \cap A = \emptyset,\ A \in \mathcal{A}\} \leq t$. Equivalently,

$$\sum\{w(A) \mid A \subseteq Y_k^c,\ \ A \in \mathcal{A}\} \leq t, \quad k = 1, \ldots, \beta. \tag{17}$$

Given a set $X \subseteq V$, let $\phi(X) = \{A \mid A \subseteq X,\ A \in \mathcal{A}\}$ denote the set of those hyperedges of $\mathcal{A}$ which are contained in $X$. Clearly, $\phi$ is a monotonic mapping

$$X \subseteq X' \subseteq V \quad \Longrightarrow \phi(X) \subseteq \phi(X'),$$

and for any sets $X, X' \subseteq V$ we have the identity

$$\phi(X) \cap \phi(X') \equiv \phi(X \cap X'). \tag{18}$$

Denoting the number of hyperedges in $\mathcal{A}$ by $m$, we can view any set $\phi(X)$ as a subset of $U = \{1, \ldots, m\}$.

Consider the set families

$$\mathcal{S} = \{\phi(X_1^c), \ldots, \phi(X_\alpha^c)\}, \quad \mathcal{T} = \{\phi(Y_1^c), \ldots, \phi(Y_\beta^c)\}.$$

Inequalities (16) and (17) imply that $\mathcal{S}$ and $\mathcal{T}$ are threshold separable:

$$w(\phi(X_k^c)) > t, \quad k = 1, \ldots, \alpha; \qquad w(\phi(Y_k^c)) \leq t, \quad k = 1, \ldots, \beta.$$

We now split the proof into two cases.

*Case 1:* $\alpha \leq 1$. As mentioned above, $\min[h]$ contains at least one set, say $Y_1$. If $Y_1$ intersects at least one hyperedge of $\mathcal{A}$, we obtain (14) and the lemma follows. Otherwise $Y_1$ is disjoint from all hyperedges of $\mathcal{A}$ and (17) implies

$$\sum\{w(A) \mid A \subseteq Y_1^c,\ A \in \mathcal{A}\} = \sum\{w(A) \mid A \in A\} \leq t.$$

This contradicts (16) unless $\mathcal{S}$ is an empty family, i.e. $\alpha = 0$.

*Case 2:* $\alpha \geq 2$. Let us show that $\mathcal{T}$ covers all pairwise intersections of $\mathcal{S}$. Let $X_i, X_j,\ \ 1 \leq i < j \leq \alpha$, be two distinct sets in $\mathcal{X}$. Since $X_i, X_j \in \max[h]$ are *maximal* false sets for $h$, we have $h(X_i \cup X_j) = 1$. Consequently, there is a minimal true point $Y_k \in \min[h]$ such that $Y_k \subseteq X_i \cup X_k$. Equivalently, we can write $X_i^c \cap X_j^c \subseteq Y_k^c$. Hence $\phi(X_i^c \cap X_j^c) \subseteq \phi(Y_k^c)$ by the monotonicity of $\phi$. In view of (18) we now obtain $\phi(X_i^c) \cap \phi(X_j^c) \subseteq \phi(Y_k^c)$, i.e., the intersection of any two distinct sets in $\mathcal{S}$ can be covered by a set in $\mathcal{T}$.

We have thus shown that $\mathcal{S}$ and $\mathcal{T}$ satisfy the assumptions of Lemma 1. Hence

$$\alpha \leq \sum_{k=1}^{\beta}(|\mathcal{A}| - |\phi(Y_k^c)|).$$

However, $|\mathcal{A}| - |\phi(Y^c)|$ is exactly the number of hyperedges in $\mathcal{A}$ which have a non-empty intersection with $Y$ and (14) follows.

$\square$

### 3.3 Proof of Theorem 2

Let $\mathcal{H}$ be an arbitrary non-empty sub-hypergraph of $\mathcal{A}_{w,t}$, and let $h : 2^V \to \{0,1\}$ denote the monotone Boolean function defined by the condition

$$h(X) = 1 \iff H \subseteq X \text{ for some } H \in \mathcal{H}.$$

Note that $h \not\equiv 0$ because $\mathcal{H} \neq \emptyset$. Furthermore, $\mathcal{H} = \min[h]$ because $\mathcal{H}$ and $\mathcal{A}_{w,t}$ are Sperner hypergraphs.

Next, for each component $t_\rho$ of the threshold vector $t = (t_1, \ldots, t_r)$ and for each component $w_\rho$ of the weight function $w = (w_1, \ldots, w_r) : \mathcal{A} \to \mathbb{R}^r_+$ let $\mathcal{A}_{w_\rho, t_\rho}$ denote the hypergraph of all minimal $w_\rho, t_\rho$-transversals of $\mathcal{A}$, and let $g_\rho = g_{\mathcal{A}, w_\rho, t_\rho}$ be the associated monotone Boolean function, i.e.

$$g_\rho(X) = 1 \iff X \text{ is a } w_\rho, t_\rho\text{- transversal of } \mathcal{A}.$$

Consider an arbitrary set $X \subseteq V$ such that $h(X) = 1$. Since $X$ contains a $w, t$-transversal, we have

$$\sum \{w_\rho(A) \mid A \in \mathcal{A},\ A \cap X = \emptyset\} \leq t_\rho, \quad \rho = 1, \ldots, r.$$

Consequently, $X$ contains a $w_\rho, t_\rho$-transversal for each $\rho$. This shows that

$$g_\rho(X) \geq h(X) \text{ for all } X \subseteq V \text{ and } \rho = 1, \ldots, r.$$

Let us now consider an arbitrary set $X \in \mathcal{H}^d$. Clearly, we have $h(X^c) = 0$ because none of the sets $H \in \mathcal{H}$ can be contained in $X^c$. Thus, $X^c$ is a false set for $h$. In fact, it is easy to see that $X^c$ is a maximal false set for $h$, i.e., $h(X^c \cup \{u\}) = 1$ for any vertex $u \notin X^c$. This is because $h(X^c \cup \{u\}) = 0$ would imply that $X \setminus \{u\}$ intersects each hyperedge $H \in \mathcal{H}$ in contradiction with the fact that $X \in \mathcal{H}^d$ is a minimal transversal to $\mathcal{H}$. We have thus shown that

$$X \in \mathcal{H}^d \implies X^c \in max[h].$$

Suppose that we also have $X \in (\mathcal{A}_{w,t})^d$. Then $X^c$ contains no $w, t$-transversal, and consequently

$$\sum \{w_\rho(A) \mid A \in \mathcal{A},\ A \cap X^c = \emptyset\} > t_\rho$$

for some $\rho \in \{1, \ldots, r\}$. This means that $X^c$ cannot contain a $w_\rho, t_\rho$ transversal and therefore $g_\rho(X^c) = 0$. Hence

$$X \in (\mathcal{A}_{w,t})^d \cap \mathcal{H}^d \implies X^c \in \bigcup_{\rho=1}^{r} \mathcal{X}_\rho,$$

where

$$\mathcal{X}_\rho = \max[h] \cap \{X \mid g_\rho(X) = 0\}, \quad \rho = 1, \ldots, r.$$

In particular, we have

$$|(\mathcal{A}_{w,t})^d \cap \mathcal{H}^d| \leq \sum_{\rho=1}^{r} |\max[h] \cap \{X \mid g_\rho(X) = 0\}|. \tag{19}$$

It remains to apply Lemma 2 to each pair of functions $h$ and $g_\rho$ to obtain

$$|\max[h] \cap \{X \mid g_\rho(X) = 0\}| \leq \sum_{X \in \min[h]} |\{A \in \mathcal{A} \mid A \cap X \neq \emptyset\}|. \qquad (20)$$

Since $\min[h] = \mathcal{H}$, we have

$$\sum_{X \in \min[h]} |\{A \in \mathcal{A} \mid A \cap X \neq \emptyset\}| = \sum_{H \in \mathcal{H}} |\{A \in \mathcal{A} \mid A \cap H \neq \emptyset\}|.$$

Now (19) and (20) yield (3). □

# 4 NP-hard generalizations

As we have shown, generating weighted transversals is polytime reducible to dualization and thus can be executed in incremental quasi-polynomial time. However, even minor modifications in the definition of weighted transversals may lead to NP-hard generating problems. In this section we shall consider several possible variants, and will show that essentially all generalizations are NP-hard.

## 4.1 Partially feasible binary solutions to monotone systems of linear inequalities

Given a system of $m$ linear inequalities $Ax \geq b$ and an integral threshold $t$ ($0 \leq t < m$), we call vector $x$ a *partial* solutions or $t$-solution to this system of inequalities, if $x$ violates at most $t$ out of the $m$ inequalities of $Ax \geq b$.

**Proposition 3** *Generating all minimal partial binary solutions to a system $Ax \geq b$ of $m$ linear inequalities is NP-hard, even if the matrix $A$ is binary and $b = (2, ..., 2)$.*

**Proof**. To show the claim, we shall use arguments analogous to [24, 27]. Let us consider the well-known NP-hard problem of determining whether a given graph $G = (V, E)$ contains an independent vertex set of size $\alpha$, where $\alpha \geq 2$ is a given threshold. Introduce $|V| + 1$ binary variables $x_0$ and $x_v$, $v \in V$, and write $\alpha$ identical inequalities $x_u + x_v \geq 2$ for each edge $e = (u, v) \in E$, followed by one inequality of the form $x_0 + x_v \geq 2$ for each vertex $v \in V$. This gives a system of $m = \alpha|E| + |V|$ inequalities in $n = |V| + 1$ variables. It is easy to see that the characteristic vector of any edge $e = (u, v)$ is a minimal binary solution satisfying at least $\alpha$ inequalities of this system. On the other hand, deciding whether there are other minimal binary solutions satisfying at least $\alpha$ inequalities of the system is equivalent with determining whether the graph $G$ has an independent set of size at least $\alpha$. □

Our definition of weighted transversals requires that $X$ satisfy all of the $r$ linear inequalities in (1). Proposition 3 shows that we obtain an NP-hard problem if we weaken (1) by requiring that $X$ satisfy at least $t$ of these inequalities, where $t$ is a given threshold. Another possible generalization could be obtained by relaxing the conditions $A \cap X = \emptyset$ in (1) to $|A \cap X| \leq l$, where $l$ is a given integer. Again, Proposition 3 shows that this generalization is NP-hard already for scalar unit weights and $l = 1$.

## 4.2 Generating fairly independent sets and their kernels

Consider the hypergraph of partial transversals $\mathcal{A}_{e,t}$ introduced in Section 1.4. By definition, each hyperedge of $\mathcal{A}_{e,t}$ is a minimal vertex set $X$ which intersects at least $|\mathcal{A}| - t$ hyperedges of $\mathcal{A}$. Hence any hyperedge of the complementarity hypergraph $\mathcal{A}^c_{e,t} \stackrel{\text{def}}{=} \{V \setminus X \mid X \in \mathcal{A}_{e,t}\}$ contains at most $t$ hyperedges of $\mathcal{A}$. The hyperedges of $\mathcal{A}^c_{e,t}$ are known as *fairly independent sets.* In particular, if $t = 0$ then $\mathcal{A}^c_{e,t}$ consists of all maximal independent sets for $\mathcal{A}$.

Given a fairly independent set $Y \in \mathcal{A}^c_{e,t}$, denote by $\phi(Y) = \{A \mid A \subseteq Y,\ A \in \mathcal{A}\}$ the set of hyperedges of $\mathcal{A}$ contained in $Y$, and let $\Phi(Y) = \cup_{A \in \phi(Y)} A$. We call $\phi(Y) \subseteq \mathcal{A}$ and $\Phi(Y) \subseteq V$ the *kernel family* and the *kernel* of $Y$, respectively, and denote by $\mathcal{A}_{\phi,t} \subseteq 2^{\mathcal{A}}$ and $\mathcal{A}_{\Phi,t} \subseteq 2^V$ the corresponding hypergraphs. Note that $|\mathcal{A}_{\phi,t}| = |\mathcal{A}_{\Phi,t}|$ because there is an obvious one-to-one correspondence between the kernel families and the kernels of $\mathcal{A}$. Note also that any fairly independent set contains exactly one kernel, but exponentially many fairly independent sets may contain the same kernel. Generating all fairly independent sets $Y$ which map to a given kernel $C = \Phi(Y)$ reduces to the problem of generating all independent sets for the hypergraph $\{A \setminus C \mid A \in \mathcal{A} \setminus \phi(C)\}$, i.e. to dualization. In view of this fact, it is natural to ask whether there is an efficient algorithm for generating the kernel hypergraph $\mathcal{A}_{\Phi,t}$ (equivalently, $\mathcal{A}_{\phi,t}$). Unfortunately, generating all kernels for a given hypergraph turns out to be NP-hard. To show this fact, it is convenient to replace the problem of generating $\mathcal{A}_{\Phi,t}$ by the identical problem for another hypergraph.

Let $\mathcal{A}_{\Psi,t}$ be the family of all minimal vertex sets $Y \subseteq V$ of size $\geq t$ such that $Y$ is the union of some hyperedges of $\mathcal{A}$. Let us also denote by $\mathcal{A}^T \subseteq 2^{\mathcal{A}}$ the *transposed* hypergraph for $\mathcal{A}$. In other words, each hyperedge of $\mathcal{A}^T$ is a *bunch* of the form $\mathcal{V}_v \stackrel{\text{def}}{=} \{A \in \mathcal{A} \mid A \ni v\}$, where $v \in V$.

**Proposition 4** *The problems of generating $\mathcal{A}_{\Phi,t}$ and $(\mathcal{A}^T)_{\Psi,|\mathcal{A}|-t}$ are equivalent.*

**Proof**. Given a set $X \subseteq V$, define $\mathcal{T}_X = \bigcup_{v \in V \setminus X}$. It is easy to see that $X$ is a $t$-kernel if and only if $\mathcal{T}_X \in (\mathcal{A}^T)_{\Psi,|\mathcal{A}|-t}$, i.e., $\mathcal{T}_X$ is a minimal sub-family of $\mathcal{A}$ of cardinality at least $|\mathcal{A}| - t$, which is a union of some edges in $\mathcal{A}^T$. This equivalence shows that generating all $t$-kernels of $\mathcal{A}$ is equivalent with generating all minimal sub-families $\mathcal{Y} \subseteq \mathcal{A}$ of size at least $|\mathcal{A}| - t$ which are the unions of some bunches, i.e. generating $(\mathcal{A}^T)_{\Psi,|\mathcal{A}|-t}$. $\square$

We are now ready to show the NP-hardness of generating $\mathcal{A}_{\Phi,t}$ or, equivalently, $\mathcal{A}_{\Psi,t}$.

**Proposition 5** *Given a hypergraph $\mathcal{A}$, a threshold $t$ and a sub-family $\mathcal{B} \subseteq \mathcal{A}_{\Psi,t}$, it is NP-hard to decide if $\mathcal{B} = \mathcal{A}_{\Psi,t}$.*

**Proof**. We reduce our problem from the following NP-complete problem (see [6, 18]).

**Problem $Q$:** Given a set $U = \{x_1, y_1, \ldots, x_k, y_k\}$ and a family $\mathcal{S}$ of subsets of $U$, check if there is a set $\{z_1, \ldots, z_k\}$ which contains no $S \in \mathcal{S}$, where $z_i \in \{x_i, y_i\}$ for any $i = 1, \ldots, k$.

It is easy to see that without any loss of generality we can assume the family $S$ to be Sperner, and $|\{x_i, y_i\} \cap S| \leq 1$ to hold for all $S \in \mathcal{S}$ and $i = 1, \ldots, k$.

Let us associate to an instance of problem $Q$ a hypergraph $\mathcal{A}$, defined as follows: Let $Z_j$, $j = 1, \ldots, k$ be pairwise disjoint subsets such that $|Z_1| = \cdots = |Z_k| = s > 2k$. Assume that $Z = \bigcup_{j=1}^{k} Z_j$ is disjoint from $U$. Let us then define the hypergraph $\mathcal{A}$ on the set of vertices $V = U \cup Z$, consisting of the hyperedges of the form $\{x_i\} \cup Z_i$ and $\{y_i\} \cup Z_i$ for $i = 1, \ldots, k$, and the hyperedges in family $\mathcal{B} \overset{\text{def}}{=} \{S \cup Z \mid S \in \mathcal{S}\}$. Let us further set the threshold to $t = ks = |Z|$.

Then, by definition, the hypergraph $\mathcal{A}_{\Psi,t}$ contains all the hyperedges of $\mathcal{B}$. Additionally, a set of the form $\bigcup_{i=1}^{k}(\{z_i\} \cup Z_i)$, $z_i \in \{x_i, y_i\}$, $i = 1, \ldots, k$ belongs to $\mathcal{A}_{\Psi,t}$ if and only if the set $\{z_1, \ldots, z_k\}$ does not contain a hyperedge $S \in \mathcal{S}$. Hence, deciding if $\mathcal{A}_{\Psi,t} \setminus \mathcal{B} \neq \emptyset$ amounts to solving problem $Q$. □

## 4.3 NP-hard and quasi-polynomial generating problems for hypergraphs

Our results enable us now to clarify the complexity of many related generation problems. We give a short listing of those here below.

Let $\mathcal{A} \subseteq 2^V$ be a hypergraph, $|V| = n$, $|\mathcal{A}| = m$, and let $t \in \{0, 1, \ldots, m-1\}$ be a given threshold.

**Problem** $\min \mathcal{A}V$**:** *Generate all minimal families of hyperedges $\mathcal{Y} \subseteq \mathcal{A}$ whose union contains at least $t$ vertices.*

This problem is equivalent with generating $(m-t)$-transversals of the transposed hypergraph $\mathcal{A}^T$. Hence it reduces to dualization and can be solved in quasi-polynomial time, see Theorem 1.

**Problem** $\min V\mathcal{A}$**:** *Generate all minimal vertex sets $X \subseteq V$ containing at least $t$ hyperedges of $\mathcal{A}$. In other words, generate all minimal $t$-unions.*

NP-hard; see [27] and Theorem 4.

**Problem** $\min \mathcal{A}\mathcal{A}$**:** *Generate all minimal families of hyperedges $\mathcal{Y} \subseteq \mathcal{A}$ whose union contains at least $t$ hyperedges of $\mathcal{A}$.*

NP-hard; similar to problem $\min V\mathcal{A}$

**Problem** $\min V_U V$**:** *Generate all minimal vertex sets $X \subseteq V$, such that $|X| \geq t$, and $X$ is the union of some hyperedges of $\mathcal{A}$.*

NP-hard; see Propositions 4 and 5.

**Problem** $\max \mathcal{A}V$**:** *Generate all maximal families of hyperedges $\mathcal{Y} \subseteq \mathcal{A}$ whose union contains at most $t$ vertices of $V$.*

NP-hard; the complementary families of hyperedges are minimal $(m-t)$-unions for the transposed hypergraph $\mathcal{A}^T$.

**Problem** $\max V\mathcal{A}$**:** *Generate all maximal subsets $X \subseteq V$ containing at most $t$ hyperedges of $\mathcal{A}$. In other words, generate all maximal $t$-independent sets of $\mathcal{A}$.*

Since $t$-independent sets are complementary to $t$-transversals, this problem reduces to dualization and hence can be solved in incremental quasi-polynomial time.

**Problem** $\max V_U\mathcal{A}$**:** *Generate all maximal subsets $X \subseteq V$, such that $X$ contains at most $t$ hyperedges of $\mathcal{A}$, and $X$ is the union of those hyperedges. In other words, generate all $t$-kernels of $\mathcal{A}$.*

NP-hard, equivalent to $minV_UV$ for the transposed hypergraph $\mathcal{A}^T$ with threshold $m-t$.

**Problem** $\max \mathcal{A}\mathcal{A}$**:** *Generate all maximal sub-families of hyperedges $\mathcal{Y} \subseteq \mathcal{A}$ whose union contains at most $t$ hyperedges of $\mathcal{A}$. In other words, generate all $t$-kernel families of $\mathcal{A}$.*

NP-hard; equivalent to $\max V_U\mathcal{A}$, because generating kernel families and kernels are equivalent.

**Problem** $\max V_U V$**:** *Generate all maximal subsets $X \subseteq V$, such that $|X| \leq t$, and $X$ is the union of some hyperedges of $\mathcal{A}$.*

NP-hard; equivalent to $\min V\mathcal{A}$ and $\min V_U\mathcal{A}$, i.e. to generating $(n-t)$-unions, for the transposed hypergraph.

Finally let us consider the following dual pair of generating problems.

**Problem** $V \max \mathcal{A}$**:** *Generate all maximal subsets $X \subseteq V$, which are contained in the union of (at most) $t$ hyperedges of $\mathcal{A}$.*

**Problem** $V \min \mathcal{A}$**:** *Generate all minimal subsets $X \subseteq V$, which are* not *contained in the union of (at most) $t$ hyperedges of $\mathcal{A}$.*

Both of these problems are NP-hard; "minimum cover" is polytime reducible to any one of them, hence it is already NP-hard to decide, whether the set $V$ itself can be covered by $t$ hyperedges of $\mathcal{A}$. Let us add that this dual pair may be dual-bounded (an open problem), and this does not lead to a conflict with Proposition 2 because none of the two problems has a polynomial-time superset oracle.

Naturally, we can double the above list by complementation, i.e. by substituting *union* with *intersection* in all of the above problems.

### 4.4 Transversals to families of hypergraphs

Finally, let $\aleph = \{\mathcal{A}_1, ..., \mathcal{A}_m\}$ be a family of hypergraphs defined on a common vertex-set $V$, and suppose that we wish to generate incrementally all $t$-transversals to $\aleph$, i.e., all minimal subsets $X \subseteq V$ transversal to at least $m - t$ hypergraphs of $\aleph$, where $t \in \{0, \ldots, m-1\}$ is a given threshold. If each hypergraph $\mathcal{A} \in \aleph$ consists of a single hyperedge, we obtain the hypergraph of partial transversals defined in Section 1.4. In general, however, generating $t$-transversals to a given hypergraph family $\aleph$ is NP-hard. Consider a arbitrary family $\mathcal{S} = \{S_1, ..., S_m\}$ of subsets of $V$, and assign to each $S_i$ the hypergraph $\mathcal{A}_i = \mathcal{A}(S_i)$ consisting of the $|S_i|$ singletons of $S_i$. Then the minimal $t$-transversals to $\aleph$ are exactly the minimal $(m-t)$-unions for $\mathcal{S}$, which are NP-hard to generate by Theorem 4.

## 5 Concluding remarks

**1)** Even though Theorem 3 states that generating all *maximal* infeasible binary points for a non-negative system of Boolean inequalities is NP-hard, there is a polynomial randomized scheme for nearly uniform sampling from the set of *all* binary infeasible points for any (not necessarily monotone) system $Ax \geq b$. Such a scheme can be obtained by combining the algorithm [22] for approximating the size of set-unions with the rapidly mixing random walk [29] on the set of all binary vectors $\{0,1\}^n$ truncated by a single linear inequality. On the other hand, a similar randomized scheme for nearly uniform sampling from within the set of all binary (or all minimal binary) solutions to a given non-negative system $Ax \geq b$ would imply that any NP-complete problem can be solved in polynomial time by a randomized algorithm with arbitrarily small one-sided failure probability. By using the amplification technique of [20], this can be shown already for systems $Ax \geq b$ with binary matrices $A$ containing only two 1's in each row, see e.g. [18] for more detail.

**2)** It can be shown that inequality (7) and Corollary 1 hold for any monotone system of $r$ linear inequalities in $n$ *integer* variables

$$Ax \geq b, \quad 0 \leq x \leq c, \tag{21}$$

where $A$ is an $r \times n$ matrix, $b$ is an $r$-vector, and $c$ is an $n$-vector some or all of whose components may be infinite. For $c = (1, \ldots, 1)$ we obtain the binary case discussed in Section 1.3. Corollary 2 can also be extended to monotone integer systems (21).

**3)** The inequality of Corollary 4 can be slightly improved. More precisely, for any threshold value $t$,

$$\#\ maximal\ frequents\ sets\ for\ A \leq\ (m-t)[\#\ minimal\ infrequents\ sets\ for\ A],$$

provided that the right-hand side is positive. This can be shown by using the following intersection inequality:

*Let $\mathcal{S} = \{S_1, \ldots, S_\alpha\}$ and $\mathcal{T} = \{T_1, \ldots, T_\beta\}$ be two families of subsets of $U = \{1, \ldots, m\}$, and let non-negative weights $w_1, \ldots, w_m$ and thresholds $t_1 < t_2$ be such that $w(T_k) \leq t_1$ for $k = 1, \ldots, \beta$ and $w(S_i) \geq t_2$ for $i = 1, \ldots, \alpha$. Suppose that any intersection of $r$ sets in $\mathcal{S}$ with $r$ indices $1 \leq i_1 < i_2 < \ldots < i_r \leq \alpha$ is contained in some set $T_k$. Then*

$$\alpha \leq \frac{\omega - t_1}{t_2 - t_1}(r-1)\beta,$$

*where $\omega = w_1 + \ldots + w_m$.*

In particular, for $w_1 = ... = w_m = 1$, $r = 2$ and $t + 1 = t_2 > t_1 = t$ we obtain

$$\alpha \leq (m-t)\beta.$$

Finally, we mention that if $w_1 = ... = w_m = 1$, $t+1 = t_2 > t_1 = t$, and, in addition, every $s$ points of $V$ are contained in at least $r$ sets of $\mathcal{S}$ then the stronger inequality

$$\alpha \leq \big(1 - \frac{s}{t+1}\big)(m-t)(r-1)\beta$$

holds, as well.

# References

[1] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen and A. I. Verkamo, Fast discovery of association rules, In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth and R. Uthurusamy eds., *Advances in Knowledge Discovery and Data Mining*, 307-328, AAAI Press, Menlo Park, California, 1996.

[2] M. Anthony and N. Biggs, *Computational Learning Theory*, Cambridge University Press, 1992.

[3] L. Babai and P. Frankl, Linear Algebra Methods in Combinatorics, Manuscript, The University of Chicago, September 1992.

[4] P. Bertolazzi and A. Sassano, An $O(mn)$ time algorithm for regular set-covering problems, *Theoretical Computer Science* 54 (1987) 237-247.

[5] J. C. Bioch and T. Ibaraki, Complexity of identification and dualization of positive Boolean functions, *Information and Computation* 123 (1995) 50-63.

[6] E. Boros and V. Gurvich, Stable effectivity functions and perfect graphs, *Mathematical Social Sciences*, 39 (2000) 175-194. (also see RUTCOR Research Report 23-95).

[7] E. Boros, V. Gurvich, and P.L. Hammer, Dual subimplicants of positive Boolean functions, *Optimization Methods and Software*, 10 (1998) 147-156. (also see RUTCOR Research Report 11-93).

[8] E. Boros, V. Gurvich, L. Khachiyan and K. Makino, Dual-Bounded Hypergraphs: Generating Partial and Multiple Transversals, DIMACS Technical Report 99-62, Rutgers University, 1999. (http://dimacs.rutgers.edu/TechnicalReports/1999.html)

[9] E. Boros, V. Gurvich, L. Khachiyan and K.Makino, Generating partial and multiple transversals of a hypergraph, ICALP, July 9-15, 2000.

[10] E. Boros, P. L. Hammer, T. Ibaraki and K. Kawakami, Polynomial time recognition of 2-monotonic positive Boolean functions given by an oracle, *SIAM Journal on Computing* 26 (1997) 93-109.

[11] C. J. Colbourn, *The combinatorics of network reliability*, Oxford University Press, 1987.

[12] Y. Crama, Dualization of regular Boolean functions, *Discrete Applied Mathematics* 16 (1987) 79-85.

[13] C. Domingo, N. Mishra and L. Pitt, Efficient read-restricted monotone CNF/DNF dualization by learning with membership queries, *Machine learning*, 37 (1999) 89-110.

[14] T. Eiter and G. Gottlob, Identifying the minimal transversals of a hypergraph and related problems, *SIAM Journal on Computing*, 24 (1995) 1278-1304.

[15] M. L. Fredman and L. Khachiyan, On the complexity of dualization of monotone disjunctive normal forms. *J. of Algorithms*, 21 (1996) 618-628.

[16] V. Gurvich, To theory of multistep games, *USSR Comput. Math. and Math Phys.*, 13-6 (1973) 1485-1500.

[17] V. Gurvich, Nash-solvability of games in pure strategies, *USSR Comput. Math and Math. Phys.*, 15-2 (1975) 357-371.

[18] V. Gurvich and L. Khachiyan, On generating the irredundant conjunctive and disjunctive normal forms of monotone Boolean functions, *Discrete Applied Mathematics*, 96-97 (1999) 363-373.

[19] V. Gurvich and L. Khachiyan On the frequency of the most frequently occurring variable in dual DNFs, *Discrete Mathematics* 169 (1997) 245-248.

[20] M. R. Jerrum, L. G. Valiant and V. V. Vazirani, Random generation of combinatorial structures from a uniform distribution *Theoretical Computer Science* 43 (1986) 169-188.

[21] D. S. Johnson, M. Yannakakis and C. H. Papadimitriou, On generating all maximal independent sets, *Information Processing Letters*, 27 (1988) 119-123.

[22] R. Karp and M, Luby, Monte-Carlo algorithms for enumeration and reliability problems, in *Proc. 24th IEEE Symp. on Foundations of Computer Science* (1983) 56-64.

[23] E. Lawler, J. K. Lenstra and A. H. G. Rinnooy Kan, Generating all maximal independent sets: NP-hardness and polynomial-time algorithms, *SIAM Journal on Computing*, 9 (1980) 558-565.

[24] K. Makino and T. Ibaraki, Interor and exterior functions of Boolean functions, *Discrete Applied Mathematics*, 69 (1996) 209-231.

[25] K. Makino and T. Ibaraki, The maximum latency and identification of positive Boolean functions, *SIAM Journal on Computing*, 26 (1997) 1363-1383.

[26] K. Makino and T. Ibaraki, A fast and simple algorithm for identifying 2-monotonic positive Boolean functions, *Journal of Algorithms*, 26 (1998) 291-305.

[27] K. Makino and T. Ibaraki, Inner-core and outer-core functions of partially defined Boolean functions, *Discrete Applied Mathematics*, 96-97 (1999) 307-326.

[28] H. Mannila and K. J. Räihä, Design by example: An application of Armstrong relations, *Journal of Computer and System Science* 22 (1986) 126-141.

[29] B. Morris and A. Sinclair, Random walks on truncated cubes and sampling 0-1 knapsack problem, in *Proc. 40th IEEE Symp. on Foundations of Computer Science* (1999) 230-240.

[30] U. N. Peled and B. Simeone, Polynomial-time algorithm for regular set-covering and threshold synthesis, *Discrete Applied Mathematics* 12 (1985) 57-69.

[31] U. N. Peled and B. Simeone, An $O(nm)$-time algorithm for computing the dual of a regular Boolean function, *Discrete Applied Mathematics* 49 (1994) 309-323.

[32] K. G. Ramamurthy, *Coherent Structures and Simple Games*, Kluwer Academic Publishers, 1990.

[33] R. C. Read, Every one a winner, or how to avoid isomorphism when cataloging combinatorial configurations, *Annals of Discrete Mathematics* 2 (1978) 107-120.

[34] R. C. Read and R. E. Tarjan, Bounds on backtrack algorithms for listing cycles, paths, and spanning trees, *Networks* 5 (1975) 237-252.

[35] R. H. Sloan, K. Takata, G. Turan, On frequent sets of Boolean matrices, *Annals of Mathematics and Artificial Intelligence* 24 (1998) 1-4.

[36] S. Tsukiyama, M. Ide, H. Ariyoshi and I. Shirakawa, A new algorithm for generating all maximal independent sets, *SIAM Journal on Computing*, 6 (1977) 505-517.

[37] J. D. Ullman, *Principles of Database and Knowledge Base Systems*, Vols. 1 and 2, Computer Science Press, 1988.