

# Dual-Bounded Generating Problems: Weighted Transversals of a Hypergraph\*

E. Boros<sup>†</sup>   V. Gurvich<sup>†</sup>   L. Khachiyan<sup>‡</sup>   K. Makino<sup>§</sup>

January 15, 2003

## 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 an input hypergraph  $\mathcal{A}$  and a non-negative threshold vector, we define a weighted transversal as a minimal vertex set which intersects all the hyperedges of  $\mathcal{A}$  except for a subfamily of total weight not exceeding the given threshold vector. Weighted transversals generalize partial and multiple transversals introduced in [9] and also include minimal binary solutions to non-negative systems of linear inequalities and minimal weighted infrequent sets in databases. We show that the hypergraph of all weighted transversals is dual-bounded, i.e., the size of its transversal 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 row-weighted  $m \times n$  binary matrix and any threshold weight  $t$ , the number of maximal sets of columns whose row support has weight above  $t$  is at most  $m$  times the number of minimal sets of columns with row support of total weight below  $t$ . 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

---

\*The research of the first three authors was supported in part by the National Science Foundation (Grant IIS-0118636). The research of the first two authors was also supported in part by the Office of Naval Research (Grant N00014-92-J-1375). Visits of the second author to Rutgers University were supported by DIMACS, the National Science Foundation's Center for Discrete Mathematics and Theoretical Computer Science. The research of the fourth author was supported in part by a Scientific Grant-in-Aid by the Ministry of Education, Culture, Sports, Science, and Technology of Japan. A preliminary version of this paper appeared in [11]

<sup>†</sup>RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway NJ 08854-8003; (`{boros,gurvich}@rutcor.rutgers.edu`).

<sup>‡</sup>Department of Computer Science, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854; (`leonid@cs.rutgers.edu`)

<sup>§</sup>Division of Systems Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka, 560-8531, Japan; (`makino@sys.es.osaka-u.ac.jp`)

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 weighted infrequent sets of columns for a given binary matrix.

**Key Words:** Boolean programming, data mining, dualization, dual hypergraph, incremental polynomial time, intersection inequalities, minimal infrequent set, maximal frequent set, multiple transversal, partial transversal, threshold Boolean function.

## 1 Introduction

In this paper we consider some problems involving the generation of all sets in certain implicitly given set families. The most well-known problem of this type is the *hypergraph dualization problem* which calls for generating all minimal transversals for an explicitly given hypergraph. The hypergraph dualization problem has applications in combinatorics [36], graph theory [25, 26, 37, 40], artificial intelligence [16], game theory [21, 22, 35], convex programming [23], reliability theory [13, 35], database theory [1, 32, 41] 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  $\bigcup_{A \in \mathcal{A}} A = \bigcup_{B \in \mathcal{A}^d} B$  (see e.g., [3]).

### 1.1 Dualization

The dualization problem for a Sperner hypergraph  $\mathcal{A}$  is to incrementally generate all hyperedges of the dual hypergraph  $\mathcal{A}^d$ . This problem can thus 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 find a new transversal  $X \in \mathcal{A}^d \setminus \mathcal{B}$ , or conclude that  $\mathcal{B} = \mathcal{A}^d$ .

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  $\text{DUAL}(\mathcal{A}, \mathcal{B})$  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  $\text{DUAL}(\mathcal{A}, \mathcal{B})$  can be solved in polynomial time (see [8, 16, 17]); moreover, it can be efficiently solved in parallel (see [5]). In addition, for  $c = 2$  there are dualization algorithms that run with polynomial delay, i.e. in  $\text{poly}(|V|, |\mathcal{A}|)$  time for a specific sequence  $\emptyset \subset \mathcal{B}_1 \subset \mathcal{B}_2 \subset \dots \subset \mathcal{A}^d$ , see e.g. [25, 26, 40]). Efficient algorithms exist also for dualization of 2-monotonic, threshold, matroid, read-bounded, acyclic and some other classes of hypergraphs (see e.g. [4, 12, 14, 17, 20, 15, 27, 29, 30, 33, 34]).

Even though no incremental polynomial time algorithm for the dualization of arbitrary hypergraphs is known, an incremental *quasi-polynomial time* one exists (see [18]). 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 [24] for more detail). The other kinds of incremental quasi-polynomial time algorithms can be found in [19, 39]. Furthermore, the algorithm of [18] can solve problem  $\text{DUAL}(\mathcal{A}, \mathcal{B})$  in  $\text{poly}(n, m)$  time on  $m^{o(\log m)}$  parallel processors. In particular, this means that problem  $\text{DUAL}(\mathcal{A}, \mathcal{B})$  cannot be NP-hard unless any NP-complete problem of size  $s$  can be solved in  $\text{poly}(s)$  time on  $s^{o(\log s)}$  processors. In addition, it is also known that the duality of two hypergraphs  $\mathcal{A}$  and  $\mathcal{B}$  can be disproved in polynomial time with only  $o(\log^2(|\mathcal{A}| + |\mathcal{B}|))$  suitably guessed bits, i.e. with limited non-determinism [17].

## 1.2 Weighted transversals

In [9] we considered *partial* and *multiple* transversals of a hypergraph (see Sections 1.3 and 1.4 below), 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  $t$ :

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

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

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

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

in integer programming and data mining are discussed in Sections 1.3 and 1.4 below.

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 of weighted transversals  $\mathcal{B} \subseteq \mathcal{A}_{w,t}$ , either find a new weighted transversal  $X \in \mathcal{A}_{w,t} \setminus \mathcal{B}$ , or conclude that  $\mathcal{B} = \mathcal{A}_{w,t}$ .

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 reducible to the hypergraph dualization problem in polynomial time.*

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

Our proof of Theorem 1 is based on the fact that the hypergraph of weighted transversals is *dual-bounded* [9]: the cardinality of the dual hypergraph for  $\mathcal{A}_{w,t}$  can be bounded by a polynomial in the cardinality 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*

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

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

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

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

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

Let us remark that inequalities (3) and (4) are instrumental in the proof of Theorem 1 (see Proposition 2). Furthermore, the complicated looking intersection with  $\mathcal{A}_{w,t}^d$  on the left hand side is not only helpful in proving our algorithmic claims, but also essential for the validity of these inequalities. Without such an intersection the same inequalities would not hold, even if we replace the right hand sides with an arbitrary polynomial in  $r$ ,  $n$ ,  $|\mathcal{A}|$  and  $|\mathcal{H}|$ . Consider for example the hypergraph  $\mathcal{A} = \{\{1, 2, \dots, k\}, \{1', 2', \dots, k'\}\}$  consisting of two disjoint hyperedges of size  $k$  each, and let  $\mathcal{H} = \{\{i, i'\} \mid i = 1, 2, \dots, k\}$ . Then for  $r = 1$ ,  $w \equiv 1$  and  $t = 0$  we have  $\mathcal{H} \subseteq \mathcal{A}_{w,t}$ , however  $|\mathcal{H}^d| = 2^k$  cannot be bounded from above by any polynomial of  $r = 1$ ,  $|\mathcal{A}| = 2$ ,  $n = 2k$  and  $|\mathcal{H}| = k$ .

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

### 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

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

where  $A$  is a given non-negative  $r \times n$ -matrix and  $b$  is a given  $r$ -vector. Let  $V = \{1, \dots, n\}$  be the (index set of the) columns of  $A$ , and let  $\mathcal{A}$  be the hypergraph consisting of  $m = n$  singletons  $A_1 = \{1\}, \dots, A_n = \{n\}$ . 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  $\mathcal{A} = \{\{1\}, \dots, \{n\}\}$ . 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),

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

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*

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

Note that the bound of Corollary 1 is sharp for  $r = 1$ , e.g., for the inequality  $x_1 + \dots + x_n \geq n$ . 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

$$x_{i_1} + x_{i_2} + \dots + x_{i_k} \geq 1, \quad i_1 \in \{1, 2\}, \quad i_2 \in \{3, 4\}, \dots, \quad i_k \in \{2k-1, 2k\}. \quad (8)$$

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

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

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  $x_1 + x_2 \geq 1, x_3 + x_4 \geq 1, \dots, x_{2k-1} + x_{2k} \geq 1$  has  $2^k$  minimal feasible binary vectors and only  $k$  maximal infeasible binary vectors.

Results of [28] imply that the problem of generating all maximal infeasible binary vectors for (6) is NP-hard already for binary matrices  $A$ :

**Theorem 3** *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 ( $\mathcal{P}$ ):*

*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*

$$\{x \in \{0,1\}^n \mid Ax \geq b\} = \{y \in \{0,1\}^n \mid y \geq x \text{ for some } x \in \mathcal{X}\},$$

*or find a new minimal binary solution for  $Ax \geq b$ .*

*Problem ( $\mathcal{P}$ ) 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.*

It was conjectured in [26] that problem ( $\mathcal{P}$ ) cannot be solved in polynomial time unless  $P=NP$ .

Let us note that if the number of non-zero entries in each row of  $A$  is bounded, then problem ( $\mathcal{P}$ ) is solvable in polynomial time. To see this, suppose that each inequality  $i = 1, \dots, r$  in the input system  $Ax \geq b$  depends on at most  $k$  variables  $x_{i_1}, \dots, x_{i_k}$ . Then the satisfiability of this inequality can be expressed by a monotone conjunctive normal form  $\phi_i(x_{i_1}, \dots, x_{i_k})$ , which can easily be constructed by enumeration (provided that  $k$  is fixed). Now all minimal solutions to  $Ax \geq b$  can be generated by dualizing the monotone conjunctive normal form  $\phi = \phi_1 \phi_2 \cdots \phi_r$ , i.e., by converting  $\phi$  into an equivalent disjunctive normal form with prime implicants. Since each clause of  $\phi$  depends on at most  $k$  variables,  $\phi$  can be dualized in polynomial time [8, 16, 17].

Problem ( $\mathcal{P}$ ) can also be solved in polynomial time when the number of inequalities  $r$  is bounded [6]. 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. Furthermore, inequality (7) and Corollaries 1, 2 also hold for monotone systems of linear inequalities in *integer* variables [6].

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

We now turn to the special case of weighted transversals for

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

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.,

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

The notion of frequent sets in the data-mining literature (see e.g. [1, 32, 41, 38]) can be related naturally to the families considered above. More precisely, given a binary  $m \times n$ -matrix  $A$  and an integer 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:  $\#\{i \mid i \in \{1, \dots, m\}, a_{ij} = 1 \text{ for all } j \in X\} > t$ . 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 [38].

Let  $V = \{1, \dots, n\}$  be the column set of  $A$  and let  $\mathcal{A} = \{A_1, \dots, A_m\}$  be the hypergraph whose incidence matrix is  $A$ . Let also  $\mathcal{A}^c = \{A_1^c, \dots, A_m^c\}$  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

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

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

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

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 [38]).

As shown in [10], it can be derived from the results of [31] that the task of generating all maximal frequent sets or, equivalently, all minimal  $t$ -unions is NP-hard:

**Theorem 4** ([10, 31]) *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 by (10), given a collection of minimal  $t$ -unions for a hypergraph, it is NP-hard to determine whether the given collection is complete.*

**Proof.** Given an arbitrary graph  $G = (V, E)$ , let us assign a hyperedge  $\{v\}$  to each vertex of  $v \in V$  and  $t - 2$  identical hyperedges  $\{v', v''\}$  to each edge  $e = (v'v'') \in E$ , and denote the obtained hypergraph by  $\mathcal{A}$ . Then the  $t$ -unions of  $\mathcal{A}$  are the edges of  $G$  and also the independent sets of  $G$  of size  $\geq t$ . Thus the family  $\mathcal{E} = \{\{v', v''\} \mid (v', v'') \in E\} \subseteq \mathcal{A}_{u_t}$  is not complete if and only if  $G$  has an independent set of size  $\geq t$ , which is NP-hard to recognize.  $\square$

On the other hand, Theorem 1 implies that the tasks of generating partial and ordinary transversals are polynomially equivalent. In particular, given a binary matrix  $A$  and a threshold  $t$ , all minimal  $t$ -infrequent columns sets of  $A$  can be generated in incremental quasi-polynomial time. By Theorem 1 this is also true for the weighted variant of the problem, i.e., if row  $i$  of  $A$  is counted with a prescribed non-negative weight  $w_i$ ,  $i = 1, \dots, m$ . We thus obtain the following corollary.

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

Theorem 2 also yields the following inequality.

**Corollary 4** *Let  $A$  be a binary  $m \times n$  matrix. Then, regardless of the weights assigned to the rows of  $A$  and the threshold weight  $t$ ,*

$$\# \text{ maximal } t\text{-frequent sets for } A \leq m [\# \text{ minimal } t\text{-infrequent sets for } A],$$

*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  $\{1, 2\}, \{3, 4\}, \dots, \{2k - 1, 2k\}$ . Then, assigning unit weight to each row of  $A$ , it can be seen that for small  $t$  this matrix has  $2^k$  maximal frequent columns sets and  $k$  minimal infrequent sets, i.e.,

$$\# \text{ maximal frequent sets for } A = \frac{m}{\log m} [\# \text{ minimal infrequent sets for } A].$$

## 1.5 Inequalities

Our proof of Theorem 2 makes use of two lemmas which we state in this section. The first lemma is an intersection inequality for two set families  $\mathcal{S} = \{S_1, \dots, S_\alpha\}$  and  $\mathcal{T} = \{T_1, \dots, T_\beta\}$  on a common ground set  $U$ , say  $U = \{1, \dots, m\}$ . 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  $k \in \{1, \dots, \beta\}$  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  $w : U \rightarrow \mathbb{R}_+$  and a real  $t$  such that

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

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  $t = \frac{1}{2}[\min\{w(S_1), \dots, w(S_\alpha)\} + \max\{w(T_1), \dots, w(T_\beta)\}] > 0$  is also a separating threshold for these families.

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

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

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

We prove Lemma 1 in Section 3.1. A variant of the lemma for the case where  $\mathcal{S}$  and  $\mathcal{T}$  can be separated by the unit weight function  $w(X) \equiv 1$  can be found in [9]. We also mention that the threshold separability of  $\mathcal{S}$  and  $\mathcal{T}$  is essential for the validity of the lemma: there are exponentially large Sperner set families  $\mathcal{S} = \{S_1, \dots, S_\alpha\}$  that only have polynomially many maximal pairwise intersections [9].

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

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

its family of all maximal false sets, and by

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

its family of all minimal true sets.

Given a family  $\mathcal{A} \subseteq 2^V$ , a weight function  $w : \mathcal{A} \rightarrow \mathbb{R}_+^r$ , and a threshold  $t \in \mathbb{R}_+^r$ , let us associate a Boolean function  $g_{\mathcal{A}, w, t} : 2^V \rightarrow \{0, 1\}$  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  $h : 2^V \rightarrow \{0, 1\}$  such that  $h \not\equiv 0$  and  $g_{\mathcal{A}, w, t}(X) \geq h(X)$  for all  $X \subseteq V$ . Then, we have

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

Lemma 2 will be shown in Section 3.2. Note that if  $\mathcal{A}$  consists of  $n$  singletons  $\{1\}, \dots, \{n\}$  then  $g_{\mathcal{A}, w, t}$  is a threshold function and we obtain the following threshold inequality.

**Corollary 5** ([9]) *Let  $h : 2^V \rightarrow \{0, 1\}$  be a monotone Boolean function such that  $w(X) = \sum_{i \in X} w_i > t$ , whenever  $h(X) = 1$ , where  $w_1, \dots, w_n$  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, 14, 33, 34]). Lemma 2 and Corollary 5 thus extend these two threshold inequalities to arbitrary monotone functions  $h$  and arbitrary weighted hypergraphs  $\mathcal{A}$ .

The remainder of the paper is organized as follows. In Section 2 we show that Theorem 1 follows from Theorem 2. In Section 3 we prove Lemmas 1, 2, and Theorem 2. Finally, in Section 4 we consider various generalizations, and show that essentially all of them are NP-hard.

## 2 Generating dual-bounded hypergraphs and the proof of Theorem 1

For completeness, we shall recall in this section several useful notions and an algorithm for generating all hyperedges of a dual-bounded hypergraph [9].

### 2.1 Superset oracles

Let  $\mathcal{G}$  be the hypergraph  $\mathcal{A}_{w, t}$  of weighted transversals introduced in Section 1.2. The input description  $\mathcal{D}$  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  $t \in \mathbb{R}^r$ . The following straightforward observation relates the hypergraph of weighted transversals to the more general class of hypergraphs defined via polynomial-time superset oracles, which we discuss in this section.

**Proposition 1** *Given a vertex set  $X \subseteq V$ , we can determine in polynomial time whether or not  $X$  contains a hyperedge of  $\mathcal{A}_{w, t}$ .*

Proposition 1 follows from the obvious fact that  $X$  contains a hyperedge of  $\mathcal{A}_{w,t}$  if and only if  $\sum\{w(A) \mid A \cap X = \emptyset, A \in \mathcal{A}\} \leq t$ , and that the latter condition can be verified in polynomial time.

More generally, let  $\mathcal{G} \subseteq 2^V$  be a Sperner hypergraph on  $V = \{1, \dots, n\}$  represented by a *superset oracle*, i.e., an algorithm  $\mathcal{D} : \mathcal{D} \times 2^V \mapsto \{yes, no\}$  which, given an input description  $\mathcal{D}$  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  $|\mathcal{D}|$  of the input description of  $\mathcal{G}$  is at least  $n$  and call the oracle polynomial-time if  $T_s = T_s(|\mathcal{D}|) \leq poly(|\mathcal{D}|)$ , where  $T_s$  is the worst-case running time of the oracle on any superset query “Does  $X$  contains a hyperedge 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 for  $\mathcal{G}$  can also be regarded as a superset oracle for the dual hypergraph  $\mathcal{G}^d$ .

## 2.2 Dual-bounded hypergraphs

Consider a family of Sperner hypergraphs  $\mathcal{G} \subseteq 2^V$  defined by a polynomial-time superset oracle  $\mathcal{D} : \mathcal{D} \times 2^V \mapsto \{yes, no\}$ . Following [9], we say that the family of hypergraphs  $\mathcal{G}$  is *dual-bounded* if the number of hyperedges of the transversal hypergraph  $\mathcal{G}^d$  can be bounded by a polynomial in the number of hyperedges and the length of description of  $\mathcal{G}$ :

$$|\mathcal{G}^d| \leq poly(|\mathcal{D}|, |\mathcal{G}|). \quad (15)$$

Furthermore,  $\mathcal{G}$  is called *uniformly dual-bounded* if

$$|\mathcal{G}^d \cap \mathcal{H}^d| \leq poly(|\mathcal{D}|, |\mathcal{H}|) \text{ for any nonempty hypergraph } \mathcal{H} \subseteq \mathcal{G}. \quad (16)$$

Note that uniform dual-boundedness implies dual-boundedness because (15) follows from (16) for  $\mathcal{H} = \mathcal{G}$ .

The following result deals with the complexity of generating all hyperedges of an implicitly given Sperner hypergraph, i.e. with the following enumeration problem:

**GEN**( $\mathcal{G}, \mathcal{H}$ ): Given a polynomial-time superset oracle for some Sperner hypergraph  $\mathcal{G} \subseteq 2^V$ , and a set  $\mathcal{H}$  of hyperedges of  $\mathcal{H}$ , either find a new hyperedge  $X \in \mathcal{G} \setminus \mathcal{H}$  of  $\mathcal{G}$ , or conclude that  $\mathcal{H} = \mathcal{G}$ .

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

**Proof.** Given a list of hyperedges  $\mathcal{H} \subseteq \mathcal{G}$ , we wish to find a new hyperedge  $\mathcal{G} \setminus \mathcal{H}$  or show that  $\mathcal{H} = \mathcal{G}$ . Start generating minimal transversals  $X$  for  $\mathcal{H}$  and check, for each generated set  $X \in \mathcal{H}^d$ , whether or not  $V \setminus X$  contains a hyperedge of  $\mathcal{G}$ . Note that since  $X$  is a minimal transversal to  $\mathcal{H}$ , the complement  $V \setminus X$  is a maximal independent set for  $\mathcal{H}$ , i.e., a maximal vertex set that contains no

hyperedge of  $\mathcal{H}$ . If  $V \setminus X$  contains a hyperedge of  $\mathcal{G}$ , then  $V \setminus X$  contains a new hyperedge in  $\mathcal{G} \setminus \mathcal{H}$  which can be found by querying the superset oracle at most  $|X| + 1$  times. If  $V \setminus X$  contains no hyperedge of  $\mathcal{G}$  then  $V \setminus X$  is a maximal independent set for  $\mathcal{G}$  and hence  $X \in \mathcal{G}^d \cap \mathcal{H}^d$ . By (16), the number of such sets  $X \in \mathcal{G}^d \cap \mathcal{H}^d$  is polynomially bounded.  $\square$

**Proof of Theorem 1.** 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 Proof of Theorem 2

As mentioned earlier, our proof of Theorem 2 relies on Lemmas 1 and 2.  $\square$

#### 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

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

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  $\mathcal{T}^{[u]} = \{T_k \setminus \{u\} \mid u \in T_k, k \in \{1, \dots, \beta\}\}$ . 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

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

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  $\mathcal{X} = \{X_1, \dots, X_\alpha\}$ . Let also  $\min[h] = \mathcal{Y} = \{Y_1, \dots, Y_\beta\}$ . Since  $h \neq 0$ , there is at least one minimal true set for  $h$ , that is  $\beta \geq 1$ .

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

$$\sum\{w(A) \mid A \subseteq X_i^c, A \in \mathcal{A}\} > t, \quad i = 1, \dots, \alpha, \quad (17)$$

where  $X^c = V \setminus X$  is the complement 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  $k \in \{1, \dots, \beta\}$ . 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,

$$\sum\{w(A) \mid A \subseteq Y_k^c, A \in \mathcal{A}\} \leq t, \quad k = 1, \dots, \beta. \quad (18)$$

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

$$X \subseteq X' \subseteq V \implies \phi(X) \subseteq \phi(X'),$$

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

$$\phi(X) \cap \phi(X') = \phi(X \cap X'). \quad (19)$$

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

Consider the set families

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

Inequalities (17) and (18) imply that  $\mathcal{S}$  and  $\mathcal{T}$  are threshold separable:

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

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 (18) implies

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

This contradicts (17) 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  $Y_k \subseteq X_i \cup X_j$ . 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 (19) 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  $h : 2^V \rightarrow \{0, 1\}$  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  $h \neq 0$  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  $t = (t_1, \dots, t_r)$  and for each component  $w_\rho$  of the weight function  $w = (w_1, \dots, w_r) : \mathcal{A} \rightarrow \mathbb{R}_+^r$  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

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

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

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

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  $\rho \in \{1, \dots, r\}$ . 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

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

In particular, we have

$$|(\mathcal{A}_{w,t})^d \cap \mathcal{H}^d| \leq \sum_{\rho=1}^r |\max[h] \cap \{X \mid g_\rho(X) = 0\}|. \quad (20)$$

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

$$|\max[h] \cap \{X \mid g_\rho(X) = 0\}| \leq \sum_{X \in \min[h]} |\{A \in \mathcal{A} \mid A \cap X \neq \emptyset\}|. \quad (21)$$

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 (20) and (21) yield (3).  $\square$

## 4 NP-hard generalizations

As we have shown, generating weighted transversals is reducible to dualization in polynomial time 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 these generalizations are NP-hard.

## 4.1 Transversals to families of hypergraphs

Let  $\aleph = \{\mathcal{A}_1, \dots, \mathcal{A}_m\}$  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  $t \in \{0, \dots, m - 1\}$  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 an arbitrary family  $\mathcal{S} = \{S_1, \dots, S_m\}$  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.

## 4.2 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* solution 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** ([9]) *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  $b = (2, \dots, 2)$ .*

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  $l = 1$ .

## 4.3 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  $\mathcal{A}_{e,t}^c \stackrel{\text{def}}{=} \{V \setminus X \mid X \in \mathcal{A}_{e,t}\}$  contains at most  $t$  hyperedges of  $\mathcal{A}$ . The hyperedges of  $\mathcal{A}_{e,t}^c$  are known as *fairly independent* or  *$t$ -independent sets*. In particular, if  $t = 0$  then  $\mathcal{A}_{e,t}^c$  consists of all maximal independent sets for  $\mathcal{A}$ .

Let  $Y \in \mathcal{A}_{e,t}^c$  be a  $t$ -independent set. As before, we 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  $\phi(Y) \subseteq \mathcal{A}$  the *kernel family* and  $\Phi(Y) \subseteq V$  the *kernel* of  $Y$ . Let us denote by  $\mathcal{A}_{\phi,t} \subseteq 2^{\mathcal{A}}$  and  $\mathcal{A}_{\Phi,t} \subseteq 2^V$  the hypergraphs of all kernel families and all kernels of  $\mathcal{A}$ , respectively. 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 each  $t$ -independent set contains exactly one kernel, but exponentially many  $t$ -independent sets may contain the same kernel.

Generating all  $t$ -independent sets  $Y$  which map to a given kernel  $C = \Phi(Y)$  reduces to the problem of generating all maximal independent sets for the hypergraph  $\{A \setminus C \mid A \in \mathcal{A} \setminus \phi(C)\}$ , that is to a dualization problem. In view of this fact, it is natural to ask whether there is an efficient algorithm for generating the kernel hypergraph  $\mathcal{A}_{\Phi,t}$ . 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.

Given a hypergraph  $\mathcal{B}$  on a vertex set  $W$  and a threshold  $s$ , let  $\mathcal{B}_{\Psi,s}$  be the family of all minimal vertex sets  $Y \subseteq W$  of size  $|Y| \geq s$ , such that  $Y$  is the union of some hyperedges of  $\mathcal{B}$ .

Let us also denote by  $\mathcal{A}^T \subseteq 2^{\mathcal{A}}$  the *transposed* hypergraph for  $\mathcal{A}$ . Its hyperedges are in one-to-one correspondence with the vertices of  $V$ . A hyperedge  $\mathcal{V}_v$  of  $\mathcal{A}^T$  is a *bunch* of all hyperedges of  $\mathcal{A}$  which contain  $v$ , that is  $\mathcal{V}_v \stackrel{\text{def}}{=} \{A \in \mathcal{A} \mid v \in A\}$ . The above definitions imply the following claim.

**Lemma 3** *A subset  $X \subseteq V$  is a kernel of the hypergraph  $\mathcal{A}$  if and only if the set  $\mathcal{T}_X = \bigcup_{v \in V \setminus X} \mathcal{V}_v$  is a hyperedge of the hypergraph  $(\mathcal{A}^T)_{\Psi,|A|-t}$ .*

□

**Corollary 6** *The problems of generating  $\mathcal{A}_{\Phi,t}$  and  $(\mathcal{A}^T)_{\Psi,|A|-t}$  are equivalent.*

We are now ready to show that generating  $\mathcal{A}_{\Phi,t}$  (or equivalently,  $\mathcal{A}_{\Psi,t}$ ) is NP-hard.

**Proposition 4** *Given a hypergraph  $\mathcal{A}$ , a threshold  $t$  and a subfamily  $\mathcal{B} \subseteq \mathcal{A}_{\Psi,t}$ , it is NP-hard to decide if  $\mathcal{B} = \mathcal{A}_{\Psi,t}$ .*

**Proof.** We reduce our problem from the well-known NP-complete *satisfiability* problem, which we state here in the following equivalent form (see [7, 23]).

**Problem Q:** Given a set  $U = \{x_1, y_1, \dots, x_k, y_k\}$  and a family  $\mathcal{S}$  of subsets of  $U$ , determine whether there is a set  $\{z_1, \dots, z_k\}$  such that it contains no set  $S \in \mathcal{S}$  and  $z_i \in \{x_i, y_i\}$  for all  $i = 1, \dots, k$ .

Without loss of generality we may also assume that  $\mathcal{S}$  is a Sperner family and  $|\{x_i, y_i\} \cap S| \leq 1$  for all  $S \in \mathcal{S}$  and  $i = 1, \dots, k$ .

Let us associate to an instance of problem  $Q$  a hypergraph  $\mathcal{A}$ , defined as follows. Let  $Z_j$ ,  $j = 1, \dots, k$  be pairwise disjoint sets of size  $s > 2k$  each. Assume that  $Z = \bigcup_{j=1}^k Z_j$  is disjoint from  $U$ , and let  $V = U \cup Z$  be the vertex set of  $\mathcal{A}$ . The hyperedges of  $\mathcal{A}$  include  $2k$  sets  $\{x_i\} \cup Z_i$  and  $\{y_i\} \cup Z_i$  for  $i = 1, \dots, k$ , and  $|S|$  sets in the family  $\mathcal{B} \stackrel{\text{def}}{=} \{S \cup Z \mid S \in \mathcal{S}\}$ . 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  $\bigcup_{i=1}^k (\{z_i\} \cup Z_i)$  with  $z_i \in \{x_i, y_i\}$ ,  $i = 1, \dots, k$  belongs to  $\mathcal{A}_{\Psi,t}$  if and only if the set  $\{z_1, \dots, z_k\}$  does not contain a set  $S \in \mathcal{S}$ . Hence, deciding if  $\mathcal{A}_{\Psi,t} \setminus \mathcal{B} \neq \emptyset$  amounts to solving problem  $Q$ .  $\square$

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

Our results enable us now to clarify the complexity of many generation problems related to hypergraphs. 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  $t \in \{0, 1, \dots, m-1\}$  be a given threshold. The notation for these problems is self-explanatory: the first term *min* or *max* indicates whether we generate minimal or maximal subsets or subfamilies; the second term ( $V$ ,  $\mathcal{A}$  or  $V_U$ ) indicates whether we generate subsets of the vertices, subfamilies of hyperedges, or subsets of the vertices which should also be unions of some hyperedges; while the last term ( $V$  or  $\mathcal{A}$ ) indicates whether the cardinality restriction applies to the number of vertices, or to the number of hyperedges.

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

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 VA:** *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 [10, 31] and the proof of Theorem 4.

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

NP-hard; see the proof of Theorem 4.

**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 Proposition 4.

**Problem min  $V_U \mathcal{A}$ :** *Generate all minimal subsets  $X \subseteq V$  such that  $X$  contains at least  $t$  hyperedges of  $\mathcal{A}$ , and it is the union of those hyperedges.*

NP-hard; trivially equivalent to problem min VA.

**Problem max AV:** *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 VA:** *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 AA:** *Generate all maximal subfamilies 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  $\min V_U \mathcal{A}$ , that is to generating  $(n - t)$ -unions, for the transposed hypergraph.

**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  $\min V_U V$  for the transposed hypergraph  $\mathcal{A}^T$  with threshold  $m - t$ ; see Lemma 3.

Note that problems  $\max V \mathcal{A}$  and  $\max V_U \mathcal{A}$  differ a lot, while  $\min V \mathcal{A}$  and  $\min V_U \mathcal{A}$  are equivalent.

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 poly-time 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, that is by substituting *union* with *intersection*. Let us also note that if we “delete the words” MAXimal and MINimal in the formulations of the above problems, then all these problems become solvable in incremental polynomial time or even with polynomial delay because the set  $\mathcal{T}$  of all true (or all false) vectors of any polynomial-time computable monotone Boolean function  $f : \{0, 1\}^n \mapsto \{0, 1\}$  can be enumerated efficiently by performing a traversal of  $\mathcal{T}$  with the natural adjacency relation on  $\{0, 1\}^n$ .

**Acknowledgement** We thank the referees for their careful review of the manuscript and helpful comments and suggestions.

## 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] C. Berge, *Hypergraphs*, North-Holland, Amsterdam, 1989.
- [4] P. Bertolazzi and A. Sassano, An  $O(mn)$  time algorithm for regular set-covering problems, *Theoretical Computer Science* 54 (1987) 237-247.
- [5] E. Boros, K. Elbassioni, V. Gurvich and L. Khachiyan, An efficient incremental algorithm for generating all maximal independent sets for hypergraphs of bounded dimension, *Parallel Processing Letters* 10, No 4 (2000) 253-256.
- [6] E. Boros, K. Elbassioni, V. Gurvich, L. Khachiyan, and K. Makino, Dual-bounded generating problems: minimal integral solutions of monotone linear inequalities, *SIAM Journal on Computing* 31(5), (2002) 1624 - 1643.
- [7] E. Boros and V. Gurvich, Stable effectivity functions and perfect graphs, *Mathematical Social Sciences*, 39 (2000) 175-194. (Also see RUTCOR Research Report 23-1995).

- [8] 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 11-1993).
- [9] E. Boros, V. Gurvich, L. Khachiyan and K. Makino, Dual-bounded generating problem: partial and multiple transversals of a hypergraph, *SIAM Journal on Computing* 30 (6) (2001), 2036-2050.
- [10] E. Boros, V. Gurvich, L. Khachiyan and K. Makino, On the complexity of generating maximal frequent and minimal infrequent sets, in: *Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science, STACS 2002*, 133-142.
- [11] E. Boros, V. Gurvich, L. Khachiyan and K. Makino, Generating weighted transversals of a hypergraph, in: *Proceedings of the 2nd Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications*, A. Frank, G. Y. Katona, A. Recski, eds., pp. 13-23, Budapest, April 20-23, 2001.
- [12] 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.
- [13] C. J. Colbourn, *The combinatorics of network reliability*, Oxford University Press, 1987.
- [14] Y. Crama, Dualization of regular Boolean functions, *Discrete Applied Mathematics* 16 (1987) 79-85.
- [15] 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.
- [16] T. Eiter and G. Gottlob, Identifying the minimal transversals of a hypergraph and related problems, *SIAM Journal on Computing*, 24 (1995) 1278-1304.
- [17] T. Eiter, G. Gottlob, and K. Makino, New results on monotone dualization and generating hypergraph transversals. STOC 2002, 14-22
- [18] M. L. Fredman and L. Khachiyan, On the complexity of dualization of monotone disjunctive normal forms. *Journal of Algorithms*, 21 (1996) 618-628.
- [19] D. R. Gaur, *Satisfiability and self-duality of monotone Boolean functions*. Ph.D. thesis, School of Computing Science, Simon Fraser University, January 1999.

- [20] D. R. Gaur, R. Krishnamurti, Self-duality of bounded monotone Boolean functions and related problems, In: *Proc. 11th International Conference on Algorithmic Learning Theory (ALT)*, pp. 209-223, Springer LNCS 1968, 2000.
- [21] V. Gurvich, To theory of multistep games, *USSR Comput. Math. and Math Phys.*, 13-6 (1973) 1485-1500.
- [22] V. Gurvich, Nash-solvability of games in pure strategies, *USSR Comput. Math and Math. Phys.*, 15-2 (1975) 357-371.
- [23] 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.
- [24] V. Gurvich and L. Khachiyan On the frequency of the most frequently occurring variable in dual DNFs, *Discrete Mathematics* 169 (1997) 245-248.
- [25] D. S. Johnson, M. Yannakakis and C. H. Papadimitriou, On generating all maximal independent sets, *Information Processing Letters*, 27 (1988) 119-123.
- [26] 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.
- [27] K. Makino, Efficient dualization of  $O(\log n)$ -term monotone disjunctive normal forms, *Discrete Applied Mathematics*, 126 (2003) 305-312.
- [28] K. Makino and T. Ibaraki, Interior and exterior functions of Boolean functions, *Discrete Applied Mathematics*, 69 (1996) 209-231.
- [29] K. Makino and T. Ibaraki, The maximum latency and identification of positive Boolean functions, *SIAM Journal on Computing*, 26 (1997) 1363-1383.
- [30] K. Makino and T. Ibaraki, A fast and simple algorithm for identifying 2-monotonic positive Boolean functions, *Journal of Algorithms*, 26 (1998) 291-305.
- [31] K. Makino and T. Ibaraki, Inner-core and outer-core functions of partially defined Boolean functions, *Discrete Applied Mathematics*, 96-97 (1999) 307-326.
- [32] H. Mannila and K. J. Räihä, Design by example: An application of Armstrong relations, *Journal of Computer and System Science* 22 (1986) 126-141.

- [33] U. N. Peled and B. Simeone, Polynomial-time algorithm for regular set-covering and threshold synthesis, *Discrete Applied Mathematics* 12 (1985) 57-69.
- [34] 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.
- [35] K. G. Ramamurthy, *Coherent Structures and Simple Games*, Kluwer Academic Publishers, 1990.
- [36] R. C. Read, Every one a winner, or how to avoid isomorphism when cataloging combinatorial configurations, *Annals of Discrete Mathematics* 2 (1978) 107-120.
- [37] R. C. Read and R. E. Tarjan, Bounds on backtrack algorithms for listing cycles, paths, and spanning trees, *Networks* 5 (1975) 237-252.
- [38] R. H. Sloan, K. Takata, G. Turan, On frequent sets of Boolean matrices, *Annals of Mathematics and Artificial Intelligence* 24 (1998) 1-4.
- [39] H. Tamaki. Space-efficient enumeration of minimal transversals of a hypergraph. *IPSJ-AL* 75 (2000) 29-36.
- [40] 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.
- [41] J. D. Ullman, *Principles of Database and Knowledge Base Systems*, Vols. 1 and 2, Computer Science Press, 1988.