Abstract

Given a family $\mathcal{A}$ of subsets of $[n]$, $n \in \mathbb{N}$, finding another family of optimal size satisfying a certain relationship with sets in $\mathcal{A}$ constitute a class of problems studied in extremal combinatorics. This class includes the set cover problem, the problem of separating families and its variants [59, 38, 71], and the test cover problem [50, 32, 21]. Another class of well studied problems is "Covering the $\{0,1\}^n$ Hamming cube with the minimum number of affine hyperplanes" - a point $x \in \{0,1\}^n$ is said to be *covered* by a hyperplane $H(a,b)$ if $\langle a, x \rangle = b$ (see [2, 43, 66, 65]). Viewing the elements of a family $\mathcal{A}$ as the points of the $\{0,1\}^n$ Hamming cube and putting restrictions on the covering hyperplanes $H(a,b)$, the covering hyperplanes correspond to a family $\mathcal{B}$ with interesting combinatorial connections with $\mathcal{A}$. In this thesis, we study few such connections, their variants and the underlying minimization problems.

Let $\mathcal{A}$ be a family of subsets of $[n]$, where $[n] = \{1, \ldots, n\}$. For any set $A \subseteq [n]$, let $\overline{A}$ denote the complement set of A, i.e. $\overline{A} = [n] \setminus A$. Given a $D \subseteq \{-n, -n+1, \ldots, 0, \ldots, n\}$, we say a family $\mathcal{B}$ is *D-secting* for $\mathcal{A}$ if for each subset $A \in \mathcal{A}$, there exists a subset $B \in \mathcal{B}$ such that $|A \cap B| - |A \cap \overline{B}| = i$, where $i \in D$. A *D-secting* family $\mathcal{B}$ of $\mathcal{A}$, where $D = \{-1, 0, 1\}$, is a *bisecting* family ensuring the existence of a subset $B \in \mathcal{B}$ such that $|A \cap B| \in \{\lceil \frac{|A|}{2} \rceil, \lfloor \frac{|A|}{2} \rfloor\}$, for each $A \in \mathcal{A}$. Let $\beta_D(\mathcal{A})$ denote the minimum cardinality of a D-secting family for $\mathcal{A}$. We define $\beta_D(n)$ as

$$\beta_D(n) = \max_{\mathcal{A}} \beta_D(\mathcal{A}).$$

Let Y denote a ± 1 bicoloring of elements of $[n]$, i.e. $Y : [n] \to \{+1, -1\}$. We abuse the notation to denote the subset of $[n]$ colored with +1 (-1) with respect to bicoloring Y as $Y(+1)$ (respectively, $Y(-1)$). Note that to describe a bicoloring of $[n]$, it suffices to specify either $Y(+1)$ or $Y(-1)$. Allowing $B = Y(+1)$, for any $A \subseteq [n]$, $|A \cap B| - |A \cap \overline{B}|$ is equivalent to $|A \cap Y(+1)| - |A \cap Y(-1)|$ (this is called as *discrepancy* of A with

respect to bicoloring Y). Therefore, $|A \cap B| - |A \cap \overline{B}|$ can represent the difference of +1 colored points and -1 colored points in any A with respect to a bicoloring Y, where $Y(+1) = B$. This connection to discrepancy also leads to the following reformulation in terms of covering the $\{0,1\}^n$ Hamming cube.

For any subset $A \subseteq [n]$, let (i) $X_A = (x_1, \ldots, x_n) \in \{0,1\}^n$ be the incidence vector such that $x_i = 1$ if and only if $i \in A$; and, (ii)$Y_A = (y_1, \ldots, y_n) \in \{-1,1\}^n$ be the incidence vector such that $y_i = 1$ if and only if $i \in A$. Observe that for any two subsets A and B of $[n]$, the dot product of $X_A = (x_1, \ldots, x_n)$ with $Y_B = (y_1, \ldots, y_n)$, denoted by $\langle X_A, Y_B \rangle$, is equivalent to $|A \cap B| - |A \cap \overline{B}|$. Therefore, $\beta_D(n)$ may be alternatively defined as the minimum cardinality of a family $\mathcal{H}$ of structures $H(a, D)$ such that

1. $a \in \{-1,+1\}^n$ for each $H(a, D)$;
2. each $H(a, D)$ is a collection of $|D|$ hyperplanes $H(a, i)$, $i \in D$;
3. a point $x \in \{0,1\}^n$ is said to be *covered* by some $H(a, D)$ if $\langle a, x \rangle \in D$;
4. for each $x \in \{0,1\}^n$, there exists some $H(a, D) \in \mathcal{H}$ that covers x.

Varying the domain of a, D, and x leads to various kinds of combinatorial question (see Table 1). When a is restricted to the domain of $\{-1,0,+1\}^n$ with exactly d nonzero coordinates (combinatorially, this can be viewed as a partial bicoloring of d out of n points), and $D = \{0\}$, the covering problem of the $\{0,1\}^n$ Hamming cube translates into a special kind of D-secting family problem - "the Induced bisection problem". When the $\{0,1\}^n$ Hamming cube is replaced with the $\{-1,+1\}^n$ Hamming cube, $a \in \{0,1\}^n$, and $D = \{0\}$, the covering problem reduces to an inverse-D-secting family problem - " the System of unbiased representatives (SUR) problem". In the thesis, we study each of these notions in detail and establish bounds on cardinalities of such families.

Let $\mathcal{A}_e$ be a family of subsets of $[n]$ where each $A \in \mathcal{A}_e$ has an even cardinality. Recall that when D is restricted to the set $\{0\}$, any D-*secting* family $\mathcal{B}$ for $\mathcal{A}$ becomes a *bisecting* family for $\mathcal{A}_e$: for each subset $A \in \mathcal{A}$, there exists a subset $B \in \mathcal{B}$ such that $|A \cap B| = \frac{|A|}{2}$. We have the following one family extension of the bisecting family

x	a	$D \subseteq \{-n, \dots, n\}$	combinatorial property
$\{0,1\}^n$	$\{-1,1\}^n$	D	D-section
$\{0,1\}^n$	$\{-1,0,1\}^n$ with d non-zeros	$\{0\}$	induced bisection
$\{-1,1\}^n$	$\{0,1\}^n$	$\{0\}$	SUR

Table 1: Various combinatorial notions corresponding to alterations in x, a and D.

notion. A family $\mathcal{F}$ of subsets of $[n]$ is called *bisection closed* if for each pair $A, B \in \mathcal{F}$, either A bisects B or B bisects A. We study extremal question regarding bisection closed families in detail and establish bounds on cardinalities of such families. We also study the problem of computation of bisecting families for products of families $\mathcal{A}$ for which $\beta(\mathcal{A})$ is known.

Keywords: discrepancy, hypergraphs, separating family, bisecting families, Hamming cube, covering, hyperplane, hypergraph bicoloring, hitting set, test cover, unbiased representatives