COMPLEXITY MEASURES ON THE SYMMETRIC GROUP AND BEYOND

Neta Dafni[1], Yuval Filmus*[2], Noam Lifshitz[3], Nathan Lindzey[4], and Marc Vinyals[5]

[1] *Henry and Marylin Taub Faculty of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel*
netadafni@gmail.com
[2] *Henry and Marylin Taub Faculty of Computer Science and Faculty of Mathematics, Technion — Israel Institute of Technology, Haifa, Israel*
yuvalfi@cs.technion.ac.il
[3] *Einstein Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem, Israel*
noam.lifshitz@mail.huji.ac.il
[4] *Department of Mathematical Sciences, University of Memphis, Memphis, Tennessee, U.S.A.*
lindzeyn@memphis.edu
[5] *School of Computer Science, University of Auckland, Auckland, New Zealend*
marc.vinyals@auckland.ac.nz

Submitted: May 30, 2023; Accepted: Oct 3, 2025; Published: Aug 20, 2026

Abstract. We extend the definitions of complexity measures of functions to domains such as the symmetric group. The complexity measures we consider include degree, approximate degree, decision tree complexity, sensitivity, block sensitivity, and a few others. We show that these complexity measures are polynomially related for the symmetric group and for many other domains.

To show that all measures but sensitivity are polynomially related, we generalize classical arguments of Nisan and others. To add sensitivity to the mix, we reduce to Huang's sensitivity theorem using "pseudo-characters", which witness the degree of a function.

Using similar ideas, we extend the characterization of Boolean degree 1 functions on the symmetric group due to Ellis, Friedgut and Pilpel to the perfect matching scheme. As another application of our ideas, we simplify the characterization of maximum-size t-intersecting families in the symmetric group and the perfect matching scheme.

Keywords. Boolean functions, complexity measures

Mathematics Subject Classifications. 06E30, 94D10

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 802020-ERC-HARMONIC.

1. Introduction

A classical result in complexity theory states that a Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ of degree d can be computed using a decision tree of depth $\mathrm{poly}(d)$. Conversely, a Boolean function computed by a decision tree of depth d has degree at most d. Thus degree and decision tree complexity are polynomially related. Other complexity measures which are polynomially related to the degree include approximate degree, certificate complexity, and block sensitivity. Recently, Huang [Hua19] added sensitivity to the list.

Can we prove similar results for Boolean functions on other domains? Such domains have been introduced to complexity theory in recent years: for example, O'Donnell and Wimmer [OW13] used Boolean functions on the so-called "slice" to construct optimal nets for monotone functions; Barak et al. [BGH+15] used Boolean functions on the Reed–Muller code to construct and analyze the influential "short code"; and recently, Khot, Minzer and Safra [KMS18] proved the 2-to-2 conjecture using Boolean functions on the Grassmann scheme.

Although yet to see applications to complexity theory, perhaps the most appealing domain is the symmetric group. We say that a function $f\colon S_n \to \mathbb{R}$ has degree at most d if any of the following equivalent conditions hold:

1. $f(\pi)$ can be written as a linear combination of *d-juntas*, which are functions depending on $\pi(i_1), \dots, \pi(i_d)$ for some $i_1, \dots, i_d \in [n]$.

2. Representing the input as a permutation matrix, f can be written as a degree d polynomial in the entries of the matrix.

3. f has Fourier-degree d, that is, it is supported on isotypic components corresponding to partitions λ with $\lambda_1 \geqslant n - d$.

(The reader who is not familiar with representation theory can ignore the last definition.)

What is the correct generalization of decision tree? We take our inspiration from the work of Ellis, Friedgut and Pilpel [EFP11], which characterized the Boolean degree 1 functions on S_n. These are functions that depend on some $\pi(i)$ or on some $\pi^{-1}(j)$. This suggests the following definition: a decision tree for functions on S_n is a decision tree with queries of the form "$\pi(i) = ?$" and "$\pi^{-1}(j) = ?$". Essentially the same definition ("matching decision trees") is used in lower bounds on the pigeonhole principle [UF96].

We show that this is a good definition by proving that degree and decision tree complexity are polynomially related for the symmetric group. In fact, we are able to generalize many other complexity measures to the symmetric group, and show that all of them are polynomially related:

Theorem 1.1. *The following complexity measures (appropriately defined) are all polynomially related for Boolean functions over the symmetric group: degree, approximate degree, decision tree complexity, certificate complexity, unambiguous certificate complexity, sensitivity, block sensitivity, fractional block sensitivity, quantum query complexity.*

Our results hold for many other domains, such as the perfect matching scheme (the set of all perfect matchings in K_{2n}) and balanced slices (the balanced slice consists of all vectors in $\{0,1\}^{2n}$ with equally many 0s and 1s, and is also known as the Johnson scheme $J(2n, n)$).

We prove Theorem 1.1 and its generalizations in an abstract framework based on simplicial complexes. In this framework, every point in the domain is a set. For example:

1. Boolean cube: We identify each vector $x \in \{0,1\}^n$ with the set $\{(i, x_i) : i \in [n]\}$.
2. Symmetric group: We identify each permutation $\pi \in S_n$ with the set $\{(i, \pi(i)) : i \in [n]\}$.

A function has degree d if it can be written as a linear combination of functions of the form "the input set contains S", where $|S| \leqslant d$; this generalizes the usual notion of degree in these two domains. Our decision trees allow any queries of the form "which element of the set Q does the input set contain?", as long as there is a unique answer for every input.

With this setup in place, we are able to polynomially relate all complexity measures other than sensitivity by generalizing classical arguments, as presented by Buhrman and de Wolf [BdW02], for example. To add sensitivity to the mix, we reduce to Huang's sensitivity theorem [Hua19] using basic representation theory.

Generalizing ideas of Gopalan et al. [GNS+16], we also prove the following simple result:

Theorem 1.2. *If a function on the symmetric group has sensitivity s, then it can be recovered from its evaluation on a ball of radius $O(s)$ around an arbitrary permutation.*

Degree 1 functions Our results show that in a wide variety of domains, Boolean degree 1 functions can be computed by constant depth decision trees. Can we say more?

Boolean degree 1 functions on the Boolean cube are *dictators*, that is, depend on a single coordinate, and the same holds for functions on the balanced slice. Ellis, Friedgut and Pilpel [EFP11] showed that the same holds for the symmetric group, with the correct interpretation of "dictator": a function depending only on some $\pi(i)$, or only on some $\pi^{-1}(j)$. In contrast, Filmus and Ihringer [FI19b] showed that Boolean degree 1 functions on the Grassmann scheme (k-dimensional subspaces of an n-dimensional vector space over a finite field) could depend on two different "data points".

Among the domains we consider, in many cases Boolean degree 1 functions are trivially dictators. In some other cases, describing all Boolean degree 1 functions seems difficult. We identify one case in which the problem is feasible:

Theorem 1.3. *A Boolean function on the perfect matching scheme has degree at most 1 if and only if it is one of the following: a constant function; a function depending on the match of some vertex i; or a function depending on whether the perfect matching intersects some triangle.*

Incidentally, this is another example in which there are non-dictatorial degree 1 functions, namely those depending on intersections with a triangle.

We prove Theorem 1.3 using polyhedral techniques. As in the proof of the corresponding result for the symmetric group by Ellis, Friedgut and Pilpel (which we paraphrase using our methods), we first characterize all *nonnegative* degree 1 functions, using the classical characterization of supporting hyperplanes of the perfect matching polytope. To deduce the result for Boolean functions, we use a simple result from the theory of complexity measures: a degree 1 function has sensitivity at most 1.

The reader is perhaps wondering about nonnegative functions of higher degree. Can we say anything intelligent about them? It turns out that the answer is negative already for the Boolean cube: classical results on the Sherali–Adams hierarchy [GMT09] show that there exist nonnegative degree 2 functions which, if written as nonnegative linear combinations of monomials over literals (that is, products of factors of the form x_i and $1 - x_j$), require degree $\Omega(n)$.

Application to Erdős–Ko–Rado theory The work of Ellis, Friedgut and Pilpel, which has already been mentioned several times, is about intersecting families of permutations. A subset $\mathcal{F} \subseteq S_n$ is *t-intersecting* if any two $\pi_1, \pi_2 \in \mathcal{F}$ agree on the image of at least t points. In other words, if we think of π_1, π_2 as sets (as in our setup), then $|\pi_1 \cap \pi_2| \geqslant t$. How large can a t-intersecting family be? One construction is a *t-star*:

$$\mathcal{F} = \{\pi \in S_n : \pi(i_1) = j_1, \ldots, \pi(i_t) = j_t\}.$$

Ellis et al. show that for large enough n (depending on t), these families have the maximum possible size, and moreover uniquely so: every t-intersecting family of the maximum size $(n-t)!$ is a t-star. Unfortunately, their argument for the uniqueness claim is wrong, see [Fil17]. Uniqueness can be recovered from the work of Ellis [Ell11], which proves a much stronger result, and is quite complicated.

We give a much simpler proof of uniqueness, using the connection between degree and certificate complexity:

Theorem 1.4. *For every t, d, the following holds for large enough n. If f is the characteristic vector of a t-intersecting family and $\deg f \leqslant d$, then either f is contained in a t-star, or the corresponding family contains $O((n-t-1)!)$ permutations.*

Ellis et al. show that a t-intersecting family of size $(n-t)!$ must have degree t (for large enough n), and so Theorem 1.4 shows that for large enough n, such a family must be a t-star.

Theorem 1.4 generalizes to other domains for which similar intersection theorems are known, such as the perfect matching scheme [Lin18a, Lin18b], and to cross-t-intersecting families.

We see Theorem 1.4 as a contribution of theoretical computer science to extremal combinatorics. It illustrates the usefulness of the theory developed in this work.

Related work Ben-David [BD16] studied complexity measures on two types of domains: multislices (see Section 4.3 for a definition), and unions of multislices.

For multislices, he polynomially relates quantum query complexity and deterministic query complexity, using a proof which involves sensitivity, block sensitivity, and certificate complexity, and includes an analog of Lemma 3.5 (his Theorem 8). His polynomial relations hold for all multislices *uniformly*. In contrast, our results degrade with the unbalancedness of the multislice.

For unions of multislices, he polynomially relates quantum query complexity and randomized query complexity. His relation degrades with the number of multislices. He also gives a simple example, involving the union of only two multislices, which shows that the deterministic query complexity could be much larger than the randomized query complexity.

Paper organization We describe the basic setup in Section 2, using two running examples: the Boolean cube and the symmetric group. We prove our main theorem, polynomially relating all complexity measures other than sensitivity, in Section 3. After describing several domains to which our techniques apply in Section 4, we prove several sensitivity theorems (polynomially relating sensitivity to all other complexity measures) in Section 5. We discuss degree 1 functions in Section 6, and the application to intersecting families in Section 7. We close the paper with Section 8, in which we discuss some open questions.

Guide to the reader Our main contribution is the setup described in Section 2. A reader who is short on time can concentrate only on the running example of the symmetric group, and skip Section 4 altogether. The remaining sections are completely independent. Of these sections, Section 3 adapts known arguments, while Sections 5 to 7 contain novel arguments.

2. Basic setup

In this section we describe our general setup, using two running examples: the Boolean cube and the symmetric group.

2.1. Defining a domain

Domain A *domain* $\mathcal{X}$ is a collection of subsets of some universe $\mathcal{U}$, all of the same size n, known as the *dimension*.[1]

Boolean cube We think of the Boolean cube $\{0,1\}^n$ as the product set

$$\mathcal{X} = \bigtimes_{i=1}^{n} \{(i,0),(i,1)\}$$

over the universe $\mathcal{U} = \{(i,b) : i \in [n], b \in \{0,1\}\}$.

A vector $x \in \{0,1\}^n$ corresponds to the set $\{(i,x_i) : i \in [n]\}$.

Symmetric group We identify a permutation $\pi \in S_n$ with its graph $\{(i,\pi(i)) : i \in [n]\}$.

We can think of a permutation as a perfect matching in the bipartite graph $K_{n,n}$. In these terms, $\mathcal{U}$ is the set of edges of $K_{n,n}$, and $\mathcal{X}$ consists of all perfect matchings.

Query In order to define decision trees, we will also need to define the notion of query. A *query* Q is a subset of $\mathcal{U}$ which intersects each set in $\mathcal{X}$ in exactly one point. With each domain we will associate a set $\mathcal{Q}$ of allowed queries.

To avoid trivialities, we assume that $\mathcal{Q}$ satisfies the following property: $\bigcup \mathcal{Q} = \mathcal{U}$, that is, every element of $\mathcal{U}$ is an element of some query in $\mathcal{Q}$. This will ensure that any function can be represented as a decision tree.

[1] In the theory of simplicial complexes, $\mathcal{X}$ would actually have dimension $n-1$.

Boolean cube Decision trees on the Boolean cube use queries of the form "$x_i = ?$". In our formalism, such a query corresponds to the set $\{(i,0),(i,1)\}$, which is guaranteed to intersect each set in $\mathcal{X}$ at exactly one element. Therefore

$$\mathcal{Q} = \{\{(i,0),(i,1)\} : i \in [n]\}.$$

Symmetric group The correct notion of decision trees for the symmetric group is hinted at by the characterization of Boolean degree 1 functions due to Ellis, Friedgut and Pilpel [EFP11], and has also appeared in the proof complexity literature, in the context of lower bounds on the pigeonhole principle [UF96]. The allowed queries are "$\pi(i) =?$" and "$\pi^{-1}(j) =?$", which in our setup translate to:

$$\mathcal{Q} = \{\{(i,j) : j \in [n]\} : i \in [n]\} \cup \{\{(i,j) : i \in [n]\} : j \in [n]\}.$$

Note In order to completely define a domain, we need to specify both $\mathcal{X}$ and $\mathcal{Q}$, though in practice we will refer to a domain using $\mathcal{X}$ only, for brevity.

The set $\mathcal{Q}$ is not canonical: its choice determines the notion of decision tree used to define decision tree complexity, as well as the values of the parameters in Section 2.3, which in turn affect our main theorem, Theorem 3.1, quantitatively. As an example, the symmetric group can also be viewed as a multislice (see Section 4.3), in which case only queries of the form "$\pi(i) = ?$" are allowed.

Chunk size Looking ahead, the "blocks" in the definition of block sensitivity correspond to removing b elements from some $x \in \mathcal{X}$ and replacing them with b other elements. For example, on the Boolean cube, moving from 000 to 101 corresponds to replacing $(1,0),(3,0)$ with $(1,1),(3,1)$. The *size* of the block is b. In the definition of sensitivity, we require all blocks to have minimum size. Accordingly, we define the *chunk size* $\mathfrak{c}$ to be the minimal value of $|x \setminus y|$ for $x \neq y \in \mathcal{X}$.

Boolean cube For two points x, y on the Boolean cube, $|x \setminus y|$ is the Hamming distance between the vector representations of x and y. Consequently, $\mathfrak{c} = 1$.

Symmetric group The minimal number of changes required to move from one permutation to the other is 2 (corresponding to applying a transposition). Therefore, $\mathfrak{c} = 2$.

Note We could have made $\mathfrak{c}$ a free parameter, but for all domains for which we can prove a sensitivity theorem, we can prove it with respect to the current definition of $\mathfrak{c}$.

2.2. Complexity measures

With the setup in hand, we can define the various complexity measures we are interested in, for a given n-dimensional domain $\mathcal{X}$ over a universe $\mathcal{U}$ with queries $\mathcal{Q}$. While many of the complexity measures make sense for arbitrary functions, we will only define them for Boolean functions, that is, functions $\mathcal{X} \to \{0,1\}$. Our selection of complexity measures matches Aaronson, Ben David and Kothari [ABDK16, Table 2]. There are many other measures encountered in the literature,

for example tree sensitivity [GSW16] and quantum certificate complexity [Aar08], which we leave for future work.

All definitions that we give below coincide with the usual definitions in the case of the Boolean cube, as the reader can easily verify.

Degree and approximate degree A *polynomial* is a function $P\colon \mathcal{X} \to \mathbb{R}$ of the form

$$P(x) = \sum_{S \subseteq \mathcal{U}} c_S [\![x \supseteq S]\!],$$

where $[\![x \supseteq S]\!]$ is the Boolean function which equals 1 if $x \supseteq S$. The *degree* of a polynomial is the maximum size of a set S such that $c_S \neq 0$.

The *degree* of a Boolean function $f\colon \mathcal{X} \to \{0,1\}$, denoted $\deg(f)$, is the minimum degree of a polynomial P such that $f(x) = P(x)$ for all $x \in \mathcal{U}$.

The ϵ-*approximate degree* of a Boolean function $f\colon \mathcal{X} \to \{0,1\}$, denoted $\widetilde{\deg}_\epsilon(f)$, is the minimum degree of a polynomial P such that $|f(x) - P(x)| \leqslant \epsilon$ for all $x \in \mathcal{X}$. This definition is sensible for all $\epsilon \in (0, 1/2)$. The *approximate degree* is $\widetilde{\deg}(f) = \widetilde{\deg}_{1/3}(f)$. (The constant $1/3$ can be replaced by any value in $(0,1/2)$; the approximate degree will change by at most a constant factor.) Note that $\widetilde{\deg}_0(f) = \deg(f)$.

Symmetric group Recall that we represent a permutation $\pi \in S_n$ as a set of input-output pairs $\{(i, \pi(i)) : i \in [n]\}$. Equivalently, for every pair (i,j), there is a Boolean variable $x_{i,j}$ indicating whether $\pi(i) = j$. The variables $x_{i,j}$ together form the permutation matrix representation of the input permutation.

The degree of a function $f\colon S_n \to \{0,1\}$ is the minimum degree of a polynomial in the variables $x_{i,j}$ representing the function. This notion of degree coincides with a natural spectral notion, as described in Section 5.3. It also coincides with the following notion of *junta degree*: the degree of f is the minimal d such that f can be written as a linear combination of d-*juntas*, where a d-junta is a function depending on d data items of the form $\pi(i)$ or $\pi^{-1}(j)$.

Notes A given Boolean function can have several different polynomial representations. For example, if Q is any query then $P(X) = \sum_{e \in Q} [\![e \in x]\!] - 1$ is a representation of the zero function. In many cases it is possible to impose more constraints on the polynomial, thus enforcing the representation to be unique. For example, over the Boolean cube, if we only allow sets S with elements of the form $(i,1)$ then the representation will be unique. For the case of the slice, see [Fil16]; a similar unique representation should exist for the symmetric group.

In the definition of a polynomial, it suffices to consider sets S which are subsets of some $x \in \mathcal{X}$. Such sets are known as *faces* in the parlance of simplicial complexes (the sets in $\mathcal{X}$ are known as *facets*).

In the case of the Boolean cube, the maximum degree coincides with the dimension. In contrast, the maximum degree of a function on S_n is only $n-1$.

As mentioned in the introduction, this notion of degree ("spatial degree") coincides with algebraic notions of degree ("spectral degree") for all domains considered in this paper for which such notions exist. We expand on this in Section 4.4.

Certificate complexity Let $f\colon \mathcal{X} \to \{0,1\}$ be a Boolean function. A *certificate* for a point $x \in \mathcal{X}$ is a subset $c \subseteq x$ such that $f(y) = f(x)$ whenever $y \supseteq c$. In other words, a certificate for x is a set of elements of x which suffice to guarantee that f will attain the value $f(x)$.

The *certificate complexity* of f at x, denoted $C(f,x)$, is the minimum size of a certificate for x. For $b \in \{0,1\}$, the *b-certificate complexity* of f, denoted $C_b(f)$, is the maximum value of $C(f,x)$ over all points in $f^{-1}(b)$. The *certificate complexity* of f, denoted $C(f)$, is $\max_x C(f,x)$.

Symmetric group The certificate complexity of a function $f\colon S_n \to \{0,1\}$ at a permutation π is the minimum size of a *partial permutation* $\rho \subseteq \pi$ (that is, an injective function whose domain is a subset of $[n]$) that forces the value of f, in the sense that $f(\sigma) = f(\pi)$ whenever $\sigma \supseteq \rho$.

Note on negative certificates We can think of functions on $\mathcal{X}$ as partial functions on $\{0,1\}^{\mathcal{U}}$ with a fixed domain. From that perspective, it is natural to allow certificates which not only guarantee that some elements belong to the input, but also that some elements don't belong to the input. In other words, given a point $x \in \mathcal{X}$, we can consider pairs (c,d), where $c \subseteq x$ and $d \subseteq \overline{x}$. Such a pair is a certificate for x if $f(y) = f(x)$ whenever $y \supseteq c$ and $\overline{y} \supseteq d$. The size of such a certificate is $|c| + |d|$.

Such "negative" certificates do not reduce the certificate complexity. Indeed, consider a certificate (c,d) for some $x \in \mathcal{X}$, and an element $e \in d$. Our assumption that $\bigcup \mathcal{Q} = \mathcal{U}$ guarantees that e participates in some query $Q \in \mathcal{Q}$. Since Q is a query and $e \notin x$, we have $Q \cap x = \{e'\}$ for some $e' \neq e$. Conversely, if $e' \in y$ then $e \notin y$, since Q intersects y at a unique element. Therefore we can replace $e \in d$ with $e' \in c$, obtaining another certificate for x of the same size. Eliminating all elements of d in this way, we obtain a standard certificate for x of the same size.

Unambiguous certificate complexity Unambiguous certificate complexity is a notion closely related to decision tree complexity. A collection $\mathcal{C}$ of subsets of $\mathcal{U}$ is a *set of unambiguous certificates* for a Boolean function $f\colon \mathcal{X} \to \{0,1\}$ if for each $x \in \mathcal{X}$ there is a unique $y \in \mathcal{C}$ such that $y \subseteq x$, and this y is a certificate for x.

The *complexity* of a set $\mathcal{C}$ of unambiguous certificates is the size of the largest set in $\mathcal{C}$. The *unambiguous certificate complexity* of f, denoted $U(f)$, is the minimum complexity of a set of unambiguous certificates for f.

As in the case of certificate complexity, negative certificates do not reduce the unambiguous certificate complexity, for similar reasons. Indeed, consider a certificate with the negative guarantee $e \notin x$. Choosing some query Q containing x, the negative guarantee can be expanded *unambiguously* to positive guarantees $e' \in x$ for every $e' \in Q \setminus x$ which is consistent with the rest of the certificate.

Decision tree complexity A *decision tree* is a tree whose internal nodes are labeled by elements of $\mathcal{Q}$, whose edges are labeled by elements of $\mathcal{U}$, and whose leaves are labeled by values in $\{0,1\}$. For an internal node v labeled by $Q \in \mathcal{Q}$, let ℓ be the set of edge labels appearing

in the path from the root to v. For each $a \in Q$, say that a is *feasible* for v if some $x \in \mathcal{X}$ contains $\ell \cup \{a\}$. We require v to have exactly one child per feasible a, with the corresponding edge labeled a.

A decision tree computes a function in the natural way. The *decision tree complexity* of a function $f\colon \mathcal{X} \to \{0,1\}$, denoted $D(f)$, is the minimum depth (measured by edges) of a decision tree computing f.

The *ϵ-error randomized decision tree complexity* of f, denoted $R_\epsilon(f)$, is the minimum R such that there is a probability distribution $\mathcal{D}$ on decision trees of depth at most R such that

$$\Pr_{T\sim\mathcal{D}}[T(x) = f(x)] \geqslant 1 - \epsilon$$

for all $x \in \mathcal{X}$. We define $R(f) = R_{1/3}(f)$. As in the case of approximate degree, the constant $1/3$ only affects the measure up to a constant factor, as long as $\epsilon \in (0, 1/2)$.

A *zero-error decision tree* T for f is a decision tree whose leaves are labeled $\{0, 1, *\}$ (where $*$ has the interpretation "do not know") such that $T(x) \in \{f(x), *\}$ for all $x \in \mathcal{X}$. The *zero-error randomized decision tree complexity* of f, denoted $R_0(f)$, is the minimum R such that there is a probability distribution $\mathcal{D}$ on zero-error decision trees for f of depth at most R such that $\Pr_{T\sim\mathcal{D}}[T(x) = *] \leqslant 2/3$ form all $x \in \mathcal{X}$. Once again, the constant $2/3$ only affects the measure up to a constant factor, and can be replaced by any other constant in $(0, 1)$.

Symmetric group A decision tree for functions of S_n uses queries of the form "$\pi(i) =?$" and "$\pi^{-1}(j) =?$". This notion coincides with the *matching decision trees* appearing in [UF96].

Notes In the case of the Boolean cube, the number of possible answers is always the same: 2. In contrast, in the case of the symmetric group, a query at depth d only has $n - d$ answers (assuming no query is repeated along the way). This phenomenon is captured by the notion of feasibility.

Our assumption that $\bigcup \mathcal{Q} = \mathcal{U}$ implies that there is a query algorithm that determines the input using at most n queries, and in particular, the decision tree complexity of any function is at most n. The query algorithm proceeds in n rounds, each of which uncovers some element a_i belonging to the input. In the i'th round, we ask an arbitrary query not containing $a_1, \ldots, a_{i-1}$, and so the answer a_i differs from $a_1, \ldots, a_{i-1}$ by construction. To see that such a query must exist, let a be some element in the input other than $a_1, \ldots, a_{i-1}$. Since $\bigcup \mathcal{Q} = \mathcal{U}$, some query Q contains a. Since Q intersects the input at exactly one element, it cannot contain any of the elements $a_1, \ldots, a_{i-1}$.

Sensitivity and block sensitivity The definition of block sensitivity is less intuitive than the definitions we have seen so far.

Let us start by recalling the usual definition of block sensitivity. The block sensitivity of a function $f\colon \{0,1\}^n \to \{0,1\}$ at a point x is the maximum number of disjoint "blocks" $B_1, \ldots, B_s \subseteq [n]$ such that $f(x \oplus B_i) \neq f(x)$ for all i, where $x \oplus B_i$ is the result of flipping all the bits whose indices belong to B_i.

The operation $x \mapsto x \oplus B_i$ corresponds, in our formalism (viewing x as a set), to removing from x the points $\{(j, x_j) : j \in B_i\}$ and replacing them with $\{(j, 1 - x_j) : j \in B_i\}$. The fact that the blocks B_i are disjoint corresponds to the sets $\{(j, x_j) : j \in B_i\}$ being disjoint.

Accordingly, we define the *block sensitivity* of a function $f\colon \mathcal{X} \to \{0,1\}$ at a point $x \in \mathcal{X}$, denoted $bs(f,x)$, to be the maximum number of points $y_1, \ldots, y_s \in \mathcal{X}$ (corresponding to $x \oplus B_i$) such that (i) $f(y_i) \neq f(x)$ and (ii) the sets $x \setminus y_i$ are disjoint. The block sensitivity of f is $bs(f) = \max_{x\in\mathcal{X}} bs(f,x)$.

We define *sensitivity* at a point $s(f,x)$ and global sensitivity $s(f)$ in the same way, using one additional constraint: $|x \setminus y_i| = \mathfrak{c}$.

Symmetric group The block sensitivity of a function $f\colon S_n \to \{0,1\}$ at a permutation π is the maximal number of permutations τ_j such that $f(\tau_j\pi) \neq f(\pi)$ and the sets $D(\tau_j) = \{i \in [n] : \tau_j(i) \neq i\}$ are disjoint; we say that the permutations τ_j are disjoint.

This definition appears asymmetric, since we multiply π by τ_j on the left. Note, however, that $\tau_j\pi = \pi\sigma_j$, where $\sigma_j = \pi^{-1}\tau_j\pi$ is a conjugate of τ_j. Using the convention $(\alpha\beta)(i) = \alpha(\beta(i))$, we see that $D(\sigma_j) = \pi^{-1}(D(\tau_j))$, and so the permutations σ_j are also disjoint. Therefore the definition of block sensitivity is, in fact, symmetric.

If we require moreover that the τ_j be transpositions, we get the definition of sensitivity. This definition is also symmetric, since σ_j, as a conjugate of τ_j, is also a transposition.

Note on block sensitivity The definition of disjointness is asymmetric: we consider $x \setminus y_i$ but not $y_i \setminus x$. It turns out that if the sets $x \setminus y_i$ are disjoint, then so are the sets $y_i \setminus x$.

Lemma 2.1. *If $x, y_1, \ldots, y_s \in \mathcal{X}$ are such that $y_i \neq x$ for all i, and the sets $x \setminus y_i$ are disjoint, then so are the sets $y_i \setminus x$.*

Proof. Suppose that for some $i \neq j$, the sets $y_i \setminus x$ and $y_j \setminus x$ have a common element a. Since $\bigcup \mathcal{Q} = \mathcal{U}$, there is a query Q_a containing a. The query Q_a intersects x at a unique point $b \neq a$. Since Q_a intersects y_i at a unique point, necessarily $b \notin y_i$, and so $b \in x \setminus y_i$. The same argument shows that $b \in x \setminus y_j$, contradicting the assumption that these two sets are disjoint. □

Notes on sensitivity In the particular case of the symmetric group, another possible definition of sensitivity is as follows. Define the *edge sensitivity* of $f\colon S_n \to \{0,1\}$ at a point x, denoted $t(f,x)$, as the number of transpositions τ such that $f(x^\tau) \neq f(x)$. The edge sensitivity of f is then $t(f) = \max_x t(f,x)$.

This definition is natural from the point of view of Boolean function analysis. In the case of the Boolean cube, it is well-known that average sensitivity equals total influence, where total influence is defined via a graph structure imposed on $\{0,1\}^n$, namely the hypercube. We can define total influence in a similar way for the symmetric group, using the transposition graph. We then get that the average edge-sensitivity is the same as the total influence [Wim14].

On the Boolean cube, total influence is bounded by the degree. On the symmetric group, total influence is bounded by the degree times n. This is reflected in the inequality $t(f) = O(n \cdot s(f))$. Indeed, it is not hard to show that $t(f,x) \leqslant 2n \cdot s(f,x)$. Unfortunately, we don't have a matching bound in the other direction. To see this, let $f(x)$ be the function "the cycle decomposition of x contains at least one 2-cycle of the form $(2i-1\ 2i)$". Then $s(f) \geqslant s(f,\mathrm{id}) = n/2$ is maximal, but simple case analysis shows that $t(f) = O(n)$. On the other hand, the sign function maximizes both the sensitivity and the edge sensitivity.

Fractional block sensitivity Fractional block sensitivity, first defined by Tal [Tal13] and by Gilmer, Saks and Srinivasan [GSS16], is a relaxation of block sensitivity obtained by relaxing an integer program to a linear program.

We can express the block sensitivity of a function $f\colon \mathcal{X} \to \{0,1\}$ at a point $x \in \mathcal{X}$ as the following integer program. The variables are $c_y \in \{0,1\}$, for each $y \in \mathcal{X}$ such that $f(y) \neq f(x)$, which indicate a collection $y_1, \ldots, y_s$ of "blocks". We want to maximize $\sum_y c_y$ (the number of "blocks") under the constraints

$$\sum_{p \notin y} c_y \leqslant 1 \text{ for all } p \in x.$$

These constraints express the condition "for each $p \in x$, there can be at most one block y_i such that $p \in x \setminus y_i$." This is the same as asking for the sets $x \setminus y_i$ to be disjoint. Hence the solution to this integer program is $bs(f,x)$.

If we relax the constraint $c_y \in \{0,1\}$ to the linear constraint $0 \leqslant c_y \leqslant 1$ then we get a linear program whose solution we denote $fbs(f,x)$, the *fractional block sensitivity of f at x*. We also denote $fbs(f) = \max_{x \in \mathcal{X}} fbs(f,x)$.

Linear programming duality gives us another linear program for $fbs(f,x)$. The variables are d_p for each $p \in x$. The goal is to minimize $\sum_p d_p$ under the constraints $0 \leqslant d_p \leqslant 1$ and

$$\sum_{p \notin y} d_p \geqslant 1 \text{ for all } y \in \mathcal{X} \text{ such that } f(y) \neq f(x).$$

This measure is known as *fractional certificate complexity*, and is similar to *randomized certificate complexity* which had been defined by Aaronson [Aar08] (the two measures are identical up to constant factors, as shown by Tal [Tal13] and by Gilmer et al. [GSS16]). To understand the provenance of these terms, let us consider the corresponding integer program, in which the constraint $0 \leqslant d_p \leqslant 1$ is replaced with $d_p \in \{0,1\}$. The variables d_p define a subset $z \subseteq x$. For each y, the constraint above expresses the condition "if $f(y) \neq f(x)$ then y is not a superset of z". In other words, z is a certificate for x. The solution of this integer program is thus $C(f,x)$.

The foregoing shows how to obtain the strange definition of block sensitivity from the natural definition of certificate complexity: start with an integer program for $C(f,x)$; relax it to a linear program; dualize; tighten it up to an integer program for $bs(f,x)$. Hopefully this convinces the reader that our definition of block sensitivity is the correct one.

Quantum query complexity We assume that the reader is familiar with quantum query complexity in the case of the Boolean cube; see for example Buhrman and de Wolf [BdW02]. We introduce an alphabet $\mathcal{A}$ of answers to queries. For example, in the case of the Boolean cube we can choose $\mathcal{A} = \{0,1\}$, and in the case of the symmetric group we can choose $\mathcal{A} = [n]$.

A quantum query algorithm operates on a triplet of quantum registers: an input register of width $\lceil \log \mathcal{Q} \rceil$ qubits, an output register of width $\lceil \log \mathcal{A} \rceil$ qubits, and workspace of arbitrary width. The *quantum query operator* O is the unitary operator that maps $|q\rangle|y\rangle|z\rangle$ to $|q\rangle|y \oplus a\rangle|z\rangle$, where a is the answer to query q. (The exact encoding of queries and answers will not make a difference.)

A *quantum query algorithm* of complexity T consists of $T+1$ unitary transformations

$$U_0, U_1, \ldots, U_T.$$

To apply the algorithm on an input, we initialize the registers to $|0\rangle|0\rangle|0\rangle$ and apply the operations $U_0, O, U_1, O, \ldots, U_{T-1}, O, U_T$, in that order. Finally, we measure the first qubit (the choice of qubit to measure is arbitrary) and output the answer.

The *exact quantum query complexity* $Q_E(f)$ is the minimum complexity of a quantum query algorithm that always computes f correctly. The *bounded-error quantum query complexity* $Q_\epsilon(f)$ is the minimum complexity of a quantum query algorithm that on every input, computes f correctly with probability at least $1-\epsilon$. We define $Q(f) = Q_{1/3}(f)$. As in previous cases, changing ϵ only affects $Q_\epsilon(f)$ by at most a constant factor, as long as $\epsilon \in (0, 1/2)$. Also, $Q_E(f) = Q_0(f)$.

Index of notation We have defined quite a few complexity measures:

- $\deg(f), \widetilde{\deg}(f), \widetilde{\deg}_\epsilon(f)$: degree and approximate degree.
- $C(f), U(f)$: certificate complexity (general and unambiguous).
- $D(f), R(f), R_\epsilon(f), R_0(f)$: decision tree complexity (deterministic, randomized, and zero error).
- $s(f), bs(f), fbs(f)$: sensitivity, block sensitivity, fractional block sensitivity.
- $Q_E(f), Q(f), Q_\epsilon(f)$: quantum query complexity (exact and bounded-error).

Simple relations among the measures Some inequalities among the measures we have considered follow directly from the definitions.

Lemma 2.2. *The following hold for every function* $f\colon \mathcal{X} \to \{0,1\}$*:*

(a) $\widetilde{\deg}(f) \leqslant \deg(f)$.

(b) $C(f) \leqslant U(f)$.

(c) $R_0(f) \leqslant D(f)$.

(d) $s(f) \leqslant bs(f)$.

(e) $Q(f) \leqslant Q_E(f)$.

(f) $R(f) \leqslant R_0(f)$.

(g) $bs(f) \leqslant fbs(f) \leqslant C(f)$.

(h) $\deg(f) \leqslant D(f)$.

(i) $U(f) \leqslant D(f)$.

(j) $bs(f) \leqslant R_\epsilon(f)/(1-2\epsilon)$*, and so* $bs(f) \leqslant 3R(f)$ *and* $bs(f) \leqslant D(f)$.

(k) $Q_E(f) \leqslant D(f)$.

(l) $\widetilde{\deg}_\epsilon(f) \leqslant Q_\epsilon(f)$*, and so* $\widetilde{\deg}(f) \leqslant 2Q(f)$ *and* $\deg(f) \leqslant 2Q_E(f)$.

Proof. The first five relations are trivial.

We have $R(f) \leqslant R_0(f)$ since given a distribution of zero-error decision trees for f, if we convert every $*$-leaf into a random coin toss then we get a distribution of decision trees whose error probability is at most $\Pr[*] \cdot \frac{1}{2} \leqslant \frac{1}{3}$.

We have $bs(f) \leqslant fbs(f) \leqslant C(f)$ since fbs is a linear programming relaxation of a maximization integer program for $bs(f)$, and of a minimization integer program for $C(f)$.

We have $\deg(f) \leqslant D(f)$ since "reaching a leaf at depth d" is a degree d monomial, and f can be written as a sum of these monomials over all 1-leaves of a decision tree for f.

We have $U(f) \leqslant D(f)$ since the root-to-leaf paths in a decision tree form a set of unambiguous certificates.

Let us now show that $bs(f) \leqslant R_\epsilon(f)/(1-2\epsilon)$; substituting $\epsilon = 0$ gives $bs(f) \leqslant D(f)$. Let x be a point such that $bs(f,x) = bs(f)$, say as witnessed by $y_1, \ldots, y_{bs(f)}$. Let $\mathcal{D}$ be a distribution over decision trees of depth at most $R_\epsilon(f)$ such that $\Pr_{T\sim\mathcal{D}}[T(y) = f(y)] \geqslant 1-\epsilon$ for all $y \in \mathcal{X}$.

Denote by p_i the probability, over $T \sim \mathcal{D}$, that T asks a query which has a different answer on x and on y_i. When asked on x, such a query must return some element of $x \setminus y_i$. Since the sets $x \setminus y_i$ are disjoint, each query can differentiate x from at most one y_i. This implies that $\sum_i p_i \leqslant R_\epsilon(f)$.

On the other hand, the probability that $T(x) \neq T(y_i)$ (over $T \sim \mathcal{D}$) is at most p_i. Therefore

$$1-\epsilon \leqslant \Pr_{T\sim\mathcal{D}}[T(x) = f(x)] \leqslant \Pr_{T\sim\mathcal{D}}[T(x) \neq T(y_i)] + \Pr_{T\sim\mathcal{D}}[T(y_i) \neq f(y_i)] \leqslant p_i + \epsilon,$$

since $T(x) = T(y_i) = f(y_i)$ implies $T(x) \neq f(x)$. This shows that each p_i is at least $1-2\epsilon$. Since $\sum_i p_i \leqslant R_\epsilon(f)$ and there are $bs(f)$ many p_i, we conclude that $(1-2\epsilon)bs(f) \leqslant R_\epsilon(f)$.

We move on to the relations involving quantum query complexity. To see that $Q_E(f) \leqslant D(f)$, we show how to simulate a decision tree using a quantum query algorithm with complexity $D(f)$. We assume without loss of generality that all leaves are at depth $D(f)$. The workspace will contain the current node. The first unitary U_0 fixes the initial state to $|q(r)\rangle|0\rangle|r\rangle$, where $q(v)$ is the query at node v, and r is the root. The unitaries $U_1, \ldots, U_{D(f)-1}$ are defined to mimic the decision tree: U_i maps $|q(v)\rangle|a\rangle|v\rangle$ to $|q(v_a)\rangle|0\rangle|v_a\rangle$, where v is a node at depth $i-1$ and v_a is its child corresponding to answer a (such a unitary exists, since we can extend the constraints to a permutation of the basic states). Finally, $U_{D(f)}$ maps $|q(v)\rangle|a\rangle|v\rangle$ to $|q(v_a)\rangle|0\rangle|v_a\rangle$, where $q(v_a)$ is the label of the leaf v_a. By construction, just before the i'th query, the state of the algorithm is $|q(v_{i-1})\rangle|0\rangle|v_{i-1}\rangle$, where $v_0, \ldots, v_{D(f)}$ is the path in the decision tree corresponding to the input. In particular, at the end the state will be $|q(v_{D(f)})\rangle|0\rangle|v_{D(f)}\rangle = |f(x)\rangle|0\rangle|v_{D(f)}\rangle$, where x is the input, and so the algorithm outputs $f(x)$ correctly.

Finally, we show that $\widetilde{\deg}_\epsilon(f) \leqslant 2Q_\epsilon(f)$. Given a quantum query algorithm of complexity T, we prove inductively that the amplitudes of the state of the algorithms after k queries are polynomials of degree at most k. Since the initial state is constant and unitary operations do not

affect the degree, it suffices to show that the quantum query operator O increases the degree by at most 1. Indeed, denoting by $\alpha(\cdot)$ the magnitude before applying O and by $\beta(\cdot)$ the magnitude after applying O, we have

$$\beta(|q\rangle|y\rangle|z\rangle) = \sum_{a \in q} [\![a \in x]\!] \alpha(|q\rangle|y \oplus a\rangle|z\rangle).$$

The amplitudes of the final state thus have degree at most T. The probability that the quantum algorithm outputs 1 is a sum of squares of magnitudes, and so is a polynomial P of degree at most $2T$. If the quantum query algorithm outputs the correct answer with probability $1 - \epsilon$, then $|P(x) - f(x)| \leqslant \epsilon$ for all $x \in \mathcal{X}$, and so $\widetilde{\deg}_\epsilon(f) \leqslant 2Q_\epsilon(f)$. □

2.3. Admissible domains

Our arguments, which show that all complexity measures (except for sensitivity) are polynomially related, only work for domains satisfying an additional condition, *composability*. In addition, the big O constants involved depend on four parameters of the domain, which we define below.

2.3.1 Composability

Recall our definition of block sensitivity: the block sensitivity of a function $f\colon \mathcal{X} \to \{0,1\}$ at a point x is the maximum s such that there are points $y_1, \ldots, y_s \in \mathcal{X}$ satisfying $f(y_i) \neq f(x)$ and that the sets $x \setminus y_i$ are disjoint. When lower-bounding the approximate degree in terms of block sensitivity, we need the ability to compose these "blocks".

A domain is *composable* if whenever $x, y_1, \ldots, y_s \in \mathcal{X}$ are such that $y_i \neq x$ for all i and the sets $x \setminus y_i$ are disjoint, then

$$z := x \setminus \bigcup_{i=1}^{s} (x \setminus y_i) \cup \bigcup_{i=1}^{s} (y_i \setminus x) \in \mathcal{X}.$$

(Recall that by Lemma 2.1, the sets $y_i \setminus x$ are also disjoint.)

Composability implies that for each $s_1, \ldots, s_b \in \{0,1\}$, the following set is in $\mathcal{X}$:

$$x(s_1, \ldots, s_b) = x \setminus \bigcup_{i\colon s_i=1} (x \setminus y_i) \cup \bigcup_{i\colon s_i=1} (y_i \setminus x) = (x \cap y_1 \cap \cdots \cap y_b) \cup \bigcup_{i\colon s_i=0} (x \setminus y_i) \cup \bigcup_{i\colon s_i=1} (y_i \setminus x).$$

In other words, composability allows us to identify copies of the Boolean cube inside arbitrary domains. Using this, we can use results on the Boolean cube to deduce results on other domains. In particular, we will use this idea to obtain a lower bound on the approximate degree in terms of block sensitivity.

Criterion for composability For some domains, such as the Boolean cube, composability is easy to prove directly. For other domains such as the symmetric group, proving composability is less immediate. We will use the following simple criterion.

Lemma 2.3. *If $\mathcal{X}$, viewed as as a subset of $\{0,1\}^{\mathcal{U}}$, is the intersection of $\{0,1\}^{\mathcal{U}}$ and an affine subspace, then $\mathcal{X}$ is composable.*

Proof. Given sets $x, y_1, \ldots, y_s \in \mathcal{U}$ such that $y_i \neq x$ for all i and the sets $x \setminus y_i$ are disjoint, we have to show that

$$z := x \setminus \bigcup_{i=1}^{s} (x \setminus y_i) \cup \bigcup_{i=1}^{s} (y_i \setminus x) \in \mathcal{U}.$$

Lemma 2.1 shows that the sets $y_i \setminus x$ are also disjoint.

Since $\mathcal{X}$ is the intersection of $\{0,1\}^{\mathcal{U}}$ with an affine subspace, it suffices to show that for any linear form ℓ over $\mathbb{R}^{\mathcal{U}}$ such that $\ell(x) = \ell(y_1) = \cdots = \ell(y_s)$, we have $\ell(z) = \ell(x)$. By linearity,

$$\ell(x \setminus y_i) = \ell(x) - \ell(x \cap y_i) = \ell(y_i) - \ell(x \cap y_i) = \ell(y_i \setminus x).$$

Since the sets $x \setminus y_i$ and the sets $y_i \setminus x$ are disjoint, $\ell(z) = \ell(x)$ immediately follows. □

Boolean cube It is easy to prove directly that the Boolean cube is composable. Let S_i be the set of indices appearing in $x \setminus y_i$. Then y_i results from x by flipping the indices in S_i. The condition that the sets $x \setminus y_i$ are disjoint translates to the sets S_i being disjoint. The vector z results from x by flipping the indices in $S_1 \cup \cdots \cup S_s$, and in particular, it lies in the Boolean cube.

Symmetric group Using Lemma 2.3, it is easy to show that the symmetric group is composable. Indeed, it is the set of solutions to the following linear system:

$$\begin{aligned}
\sum_{j=1}^{n} x_{i,j} &= 1 && \text{for all } i \in [n] \\
\sum_{i=1}^{n} x_{i,j} &= 1 && \text{for all } j \in [n] \\
x_{i,j} &\in \{0,1\} && \text{for all } i,j \in [n]
\end{aligned}$$

This system states that the matrix formed by the elements $x_{i,j} \in \{0,1\}$ is bistochastic, and so a permutation matrix.

Note In both the Boolean cube and the symmetric group, the domain consists of all sets intersecting each query at exactly one point. But this is not the case for other domains, such as the slice.

2.3.2 Four parameters

We now introduce four parameters which control our results quantitatively.

Maximum degree The *degree* of an element a is the number of queries in $\mathcal{Q}$ mentioning it, that is, the number of $Q \in \mathcal{Q}$ such that $a \in Q$. Since $\bigcup \mathcal{Q} = \mathcal{U}$, each element has degree at least 1. We denote the maximum degree of an element by Δ.

Boolean cube Each element (i, j) is mentioned by exactly one query, "x_i =?". Therefore $\Delta = 1$.

Symmetric group Each element (i, j) is mentioned by two queries: "$\pi(i)$ =?" and "$\pi^{-1}(j)$ =?". Therefore $\Delta = 2$.

Conflict bound A *partial input* is a subset $c \subseteq \mathcal{U}$ which is a subset of some set $x \in \mathcal{X}$. Two partial inputs c_1, c_2 *conflict* if no set in $\mathcal{X}$ contains both. The *conflict bound* M is the maximal value such that if c_1, c_2 are two conflicting partial inputs of size at most M, then there is a query $Q \in \mathcal{Q}$ which "separates" them, that is, intersects c_1 and c_2 at different elements.

Boolean cube Two partial inputs are conflicting if one of them contains $(i, 0)$ and the other contains $(i, 1)$. Any two such partial inputs can be separated by the query "x_i =?". Therefore $M = n$.

Symmetric group Two partial inputs are conflicting if either one of them contains (i, j_1) and the other (i, j_2), in which case they are separated by "$\pi(i)$ =?", or one of them contains (i_1, j) and the other (i_2, j), in which case they are separated by "$\pi^{-1}(j)$ =?". Therefore again $M = n$.

To see this, suppose that c_1, c_2 are two partial inputs which are not separated by any query, and consider their union $c_1 \cup c_2$. By assumption, the union, considered as a set of edges of $K_{n,n}$, is a matching. Since every matching in $K_{n,n}$ can be completed to a perfect matching, we see that c_1, c_2 are not conflicting.

Note In both cases above $M = n$. However, on some domains M is smaller. As an example, consider the "slice" $\binom{[n]}{k}$, which is the set of all vectors in $\{0, 1\}^n$ with Hamming weight k. If $a + b > k$ then the two partial inputs $\{(1, 1), \ldots, (a, 1)\}$ and $\{(a + 1, 1), \ldots, (a + b, 1)\}$ conflict but are not separated by any query. One can check that $M = \lfloor \min(k, n - k)/2 \rfloor$.

Sensitivity ratio This double parameter is the most complicated to explain. When bounding the certificate complexity by the block sensitivity in the classical case, we need to show that block sensitivity is always witnessed by blocks whose size is at most the sensitivity. To do so, we show that a larger block can always be shortened by removing one of the elements. This is essentially because there are many ways of shortening a large block.

The corresponding property in our setup is a bit harder to state, and in fact we will have two different versions. The *block sensitivity ratio* β is the largest parameter such that for every distinct $x, y \in \mathcal{X}$ there exist distinct $z_1, \ldots, z_s \in \mathcal{X}$, with $s \geqslant \beta |x \setminus y|$, satisfying:

(a) $x \setminus z_i \subsetneq x \setminus y$.

(b) The sets $y \setminus z_i$ are disjoint.

Here is the idea behind this definition. We start with two sets $x, y \in \mathcal{X}$; one should think of y as one of the sets in the definition of $bs(f, x)$. We want to find many disjoint ways of bringing y closer to x: these are the sets $z_1, \ldots, z_s$. The first constraint states that z_i is closer to x than y: we obtain z_i from y by "fixing" some of the disagreements with x. The second constraint states

that the parts of y that had to be fixed are disjoint for different z_i; this will be useful when relating $|x \setminus y|$ to block sensitivity.

The *sensitivity ratio* $\beta_{\mathfrak{c}}$ is defined in the same way, but we also require $|y \setminus z_i| = \mathfrak{c}$. This will allow us to relate $|x \setminus y|$ to sensitivity rather than block sensitivity.

Boolean cube Given two vectors $x = \{(i, x_i) : i \in [n]\}$ and $y = \{(i, y_i) : i \in [n]\}$, let S be the set of coordinates on which they disagree; note that $|x \setminus y| = |S|$. For each $i \in S$, we define $z^i = (y \setminus \{(i, y_i)\}) \cup \{(i, x_i)\}$. That is, z^i is equal to y except at coordinate i, at which it agrees with x. Then $x \setminus z^i = (x \setminus y) \setminus \{(i, x_i)\}$, so the first property is satisfied, and $y \setminus z^i = \{(i, y_i)\}$, and so the second property is satisfied. This shows that $\beta = \beta_{\mathfrak{c}} = 1$.

Symmetric group This case is more complicated. Let us assume for simplicity that x is the identity permutation, and write y as a product of disjoint non-trivial cycles of lengths $\ell_1, \dots, \ell_m > 1$; note that $|x \setminus y| = \sum_i \ell_i$. Consider a specific cycle of y of length ℓ, without loss of generality $(1\ 2\ \cdots\ \ell)$. The relevant part of y is

$$(1, 2), (2, 3), \dots, (\ell - 1, \ell), (\ell, 1).$$

We form z_i by "shortcutting" over i, that is, replacing $(i-1, i), (i, i+1)$ with $(i-1, i+1), (i, i)$. The new set z_i is indeed a permutation. Moreover, $x \setminus z_i = (y \setminus z_i) \setminus \{(i, i)\}$, so the first property is satisfied. On the other hand, $y \setminus z_i = \{(i-1, i), (i, i+1)\}$, and so in order to satisfy the second property, we need to choose values of i which are at least 2 apart: z_1, z_3 and so on. The worst case is when $\ell = 3$, in which case we can only choose a single value of i. For general ℓ, we can always choose at least $\ell/3$ many i's which will satisfy the second property. Using the same construction for all cycles (the second property is automatically satisfied), we see that $\beta = \beta_{\mathfrak{c}} = 1/3$.

Note In both examples, the constructions bounded $\beta_{\mathfrak{c}}$ directly. However, in more complicated cases, similar arguments will result in changes which will be larger than the $\mathfrak{c}$ threshold.

3. Relating all measures except sensitivity

In this section, we prove our main theorem, relating the various complexity measures introduced in Section 2.

Theorem 3.1. *Let $(\mathcal{X}, \mathcal{U}, n)$ be a composable domain with parameters Δ, M, β.*

Every function $f : \mathcal{X} \to \{0, 1\}$ satisfies:

$$\begin{gathered}
\sqrt{bs(f)/6} \leqslant \widetilde{\deg}(f) \leqslant \deg(f) \leqslant D(f) \\
bs(f) \leqslant fbs(f) \leqslant C(f) \leqslant U(f) \leqslant D(f) \\
bs(f) \leqslant 3R(f) \leqslant 3R_0(f) \leqslant 3D(f) \\
\widetilde{\deg}(f) \leqslant 2Q(f) \leqslant 2Q_E(f) \leqslant 2D(f) \\
D(f) \leqslant \beta^{-2} \max\left(\Delta, \frac{n}{M}\right) bs(f)^4 \qquad (+)
\end{gathered}$$

In particular, if $\Delta = O(1)$, $M = \Omega(n)$, and $\beta = \Omega(1)$ then all complexity measures above are polynomially related.

Furthermore, we can strengthen (+) to

$$D(f) \leqslant \beta^{-2}\Delta bs(f)^4$$

if any of the following conditions holds:

(a) $C(f) \leqslant M$ *(implied by the same bound on* $U(f), D(f)$*).*

(b) $bs(f) \leqslant \sqrt{\beta M}$ *(implied by the same bound on* $fbs(f), R(f)/3, R_0(f)/3$*).*

(c) $\widetilde{\deg}(f) \leqslant \sqrt[4]{\beta M/36}$ *(implied by the same bound on* $\deg(f), Q(f)/2, Q_E(f)/2$*).*

In particular, if $\Delta = O(1)$, $\beta = \Omega(1)$, and one of the conditions above holds, then all complexity measures above are polynomially related.

In all cases we consider, we will have $\Delta = O(1)$ and $\beta = \Omega(1)$. In some cases, such as the unbalanced slice, M will be much smaller than n, and so the polynomial relation only holds for functions of low complexity. This is not just an artifact of our proof. As an extreme example, consider the subdomain of the Boolean cube consisting of all vectors with Hamming weight 1. Every function f has certificate complexity $C(f) = 1$, but balanced functions f have decision tree complexity $D(f) = \lfloor n/2 \rfloor$.

We also prove the following "ball property", generalizing a similar result in [GNS+16]. In this theorem, a ball of radius r around a point $x \in \mathcal{X}$ consists of all points $y \in \mathcal{X}$ such that $|x \setminus y| = |y \setminus x| \leqslant r$.

Theorem 3.2. *Let $(\mathcal{X}, \mathcal{U}, n)$ be a composable domain with chunk size $\mathfrak{c}$ and sensitivity ratio $\beta_{\mathfrak{c}}$.*

If $f\colon \mathcal{X} \to \{0,1\}$ has sensitivity $s = s(f)$, then f can be recovered from its evaluation on a ball of radius $r = \beta_{\mathfrak{c}}^{-1}(2s+1)$ around an arbitrary point.

(Recall that $s(f) \leqslant bs(f)$ by Lemma 2.2(d).)

Finally, let us briefly discuss one property of Boolean functions on the Boolean cube which fails on other domains: the junta property. Nisan and Szegedy [NS94] showed that Boolean degree d functions depend on at most $d2^{d-1}$ coordinates; this was later improved to $O(2^d)$ [CHS20, Wel22]. Similar results were proved for the slice [FI19a] and multislice [FOW19] (see also [Sal21]); the latter results hold unless the slice or multislice is extremely unbalanced.

In contrast, a similar junta property fails already for the symmetric group. Consider the function

$$f(x) = \sum_{i \neq 1} \sum_{j \neq 1,i} x_{1i} x_{ij} x_{j1},$$

which states that 1 participates in a 3-cycle. This function has degree 3, but cannot be computed given the answer to a constant number of queries. Contrast this with the case of Boolean degree 1 functions on the symmetric group: such functions are known to depend on the answer of a single query [EFP11].

3.1. Main theorem

Our proof of Theorem 3.1 closely follows the exposition in the survey paper of Buhrman and de Wolf [BdW02]. In some cases, we simply reduce to known results on the Boolean cube.

In all lemmas below, $f\colon \mathcal{X} \to \{0,1\}$ is a Boolean function on some composable domain. We will make use of the parameters $\Delta, M, \beta, \beta_{\mathfrak{c}}$ defined in Section 2.3, which we briefly recall:

- Δ, the *maximum degree* , is the maximum number of queries that can mention a point.
- M, the *conflict bound*, is the maximum value such that if c_1, c_2 are two conflicting partial inputs of size at most M then there is a query Q that separates them.
- β, the *block sensitivity ratio*, is the maximum value such that for every distinct $x, y \in \mathcal{X}$ there exist at least $\beta|x \setminus y|$ distinct $z_i \in \mathcal{X}$ such that $x \setminus z_i \subsetneq x \setminus y$ for all i, and the sets $y \setminus z_i$ are disjoint.
- $\beta_{\mathfrak{c}}$, the *sensitivity ratio*, is defined similarly to β, with the additional promise that $|y \setminus z_i| \leqslant \mathfrak{c}$ for all i.

We start by bounding the decision tree complexity in terms of the certificate complexity.

Lemma 3.3. *If $C(f) \leqslant M$ then $D(f) \leqslant \Delta C_0(f) C_1(f)$.*
For all f we have $D(f) \leqslant \max(\Delta, n/M) C_0(f) C_1(f)$.

Proof. Let us start by noting that the second statement follows immediately from the first: if $C(f) > M$ then $C_0(f)C_1(f) > M$, and so the inequality follows from $D(f) \leqslant n$. Also, the first statement trivially holds when f is constant. We prove the first statement when f is not constant by providing a query algorithm, Algorithm 1.

Input: $x \in \mathcal{X}$
Output: $f(x)$

1 initialize $z \leftarrow \varnothing$; `// z is a partial input representing our current knowledge of x`
2 **repeat** $C_0(f)$ *times* **:**
3 let c be a 1-certificate of f not conflicting with z ;
4 ask all queries containing elements in c, and add the answers to z ;
5 **if** *z is a b-certificate* **then return** b;

Algorithm 1: Query algorithm proving $D(f) \leqslant \Delta C_0(f) C_1(f)$ whenever $C(f) \leqslant M$.

It is easy to check that the algorithm makes at most $\Delta C_0(f) C_1(f)$ queries (since c contains at most $C_1(f)$ elements, and each of them appears in at most Δ queries), and that it outputs the correct value in line 5. To complete the proof, we will show that in line 3 a certificate c always exists, and that a value is always returned.

We start by showing that a certificate c always exists in line 3. This is the case in the first iteration, since we assume that f is not constant. In any subsequent iteration, if every 1-certificate

of f conflicts with z then this shows that z is a 0-certificate, and so the algorithm would have returned 0 in line 5 during the preceding iteration.

It remains to show that the algorithm always returns some value. Suppose first that $f(x) = 0$. We will show that each time that line 4 is executed, at least one new element of c_0 is added to z. Therefore, after at most $C_0(f)$ iterations, z will contain c_0, and so line 5 will return 0.

Consider, therefore, some iteration of the loop, and let c be the 1-certificate considered in line 3. Since $|c|, |c_0| \leqslant C(f) \leqslant M$, there exists a query $Q \in \mathcal{Q}$ which separates c and c_0, say $Q \cap c = \{a\}$ and $Q \cap c_0 = \{a_0\}$. The query Q gets asked in line 4 since $a \in c$. In the same line, a_0 gets added to z. To complete the proof, it suffices to show that a_0 did not belong to z. Indeed, if a_0 had belonged to z, then the certificate c couldn't have been chosen in line 3, since c and z would have conflicted: every set in $\mathcal{X}$ intersects Q exactly once, and so no such set can contain both a and a_0.

Finally, let us consider the case $f(x) = 1$. Above we have shown that if $f(x) = 0$ then after at most $C_0(f)$ iterations of the loop, z contains a 0-certificate. Therefore if execution reaches the last iteration and line 5 does not return 0, then necessarily $f(x) = 1$. In other words, at the conclusion of the final iteration, z is a 1-certificate, and so line 5 returns 1, completing the proof. □

In order to bound certificate complexity in terms of other measures, we first need to show that the size of every "minimal block" can be bounded in terms of sensitivity or block sensitivity.

Lemma 3.4. *Suppose that $f(x) \neq f(y)$. There exists $z \in \mathcal{X}$ satisfying the following properties:*

(a) $f(z) = f(y)$.

(b) $x \setminus z \subseteq x \setminus y$.

(c) $|x \setminus z| \leqslant \min(\beta_{\mathfrak{c}}^{-1} s(f), \beta^{-1} bs(f))$.

Proof. Among all $z \in \mathcal{X}$ satisfying (a) and (b), we choose one that minimizes $|x \setminus z|$. We will show that this z satisfies (c).

By definition of $\beta_{\mathfrak{c}}$, we can find distinct $w_1, \ldots, w_s \in \mathcal{X}$, where $s \geqslant \beta_{\mathfrak{c}} |x \setminus z|$, such that

(i) $x \setminus w_i \subsetneq x \setminus z$ for all i.

(ii) The sets $z \setminus w_i$ are disjoint.

(iii) $|z \setminus w_i| \leqslant \mathfrak{c}$.

Since $x \setminus w_i \subseteq x \setminus y$ and $|x \setminus w_i| < |x \setminus z|$, the minimality of $|x \setminus z|$ guarantees that $f(w_i) \neq f(z)$. This implies that $s \leqslant s(f)$ by definition of sensitivity, and so $|x \setminus z| \leqslant \beta_{\mathfrak{c}}^{-1} s \leqslant \beta_{\mathfrak{c}}^{-1} s(f)$.

The proof that $|x \setminus z| \leqslant \beta^{-1} bs(f)$ is very similar, and left to the reader. □

Armed with the preceding lemma, we can bound certificate complexity in terms of sensitivity and block sensitivity.

Lemma 3.5. *For all f we have $C(f) \leqslant \min(\beta_{\mathfrak{c}}^{-1} s(f) bs(f), \beta^{-1} bs(f)^2)$.*

Proof. We will bound $C(f,x)$ for every $x \in \mathcal{X}$. Let $b = bs(f,x) \leqslant bs(f)$, say witnessed by $y_1, \dots, y_b$. Thus $f(y_1), \dots, f(y_b) \neq f(x)$, and the sets $x \setminus y_i$ are disjoint. Lemma 3.4 shows that we can find $z_1, \dots, z_b$ such that $f(z_1), \dots, f(z_b) \neq f(x)$, the sets $x \setminus z_i$ are disjoint, and furthermore $|x \setminus z_1|, \dots, |x \setminus z_b| \leqslant \min(\beta_{\mathfrak{c}}^{-1} s(f), \beta^{-1} bs(f))$.

We claim that $c = (x \setminus z_1) \cup \cdots \cup (x \setminus z_b)$ is a certificate for x. Indeed, suppose that $f(w) \neq f(x)$ although $w \supseteq c$. For each i, since $w \supseteq c \supseteq x \setminus z_i$, we get that $x \setminus w$ is disjoint from $x \setminus z_i$. But this implies that $z_1, \dots, z_b, w$ also satisfy the conditions in the definition of $bs(f,x)$, and we reach a contradiction.

Clearly $|c| \leqslant bs(f) \cdot \min(\beta_{\mathfrak{c}}^{-1} s(f), \beta^{-1} bs(f))$, completing the proof. □

To complete the chain of inequalities, we bound block sensitivity in terms of approximate degree.

Lemma 3.6. *For all f we have $bs(f) \leqslant 6\widetilde{\deg}(f)^2$.*

Proof. The proof is by reduction to the classical case of the Boolean cube.

We will show that for every set $x \in \mathcal{X}$, the bound $bs(f,x) \leqslant 6\widetilde{\deg}(f)^2$ holds. Suppose that $b = bs(f,x)$, and let this be witnessed by $y_1, \dots, y_b$. Thus $f(y_i) \neq f(x)$ and the sets $x \setminus y_i$ are disjoint. Lemma 2.1 shows that the sets $y_i \setminus x$ are also disjoint, and so all of the following are disjoint:

$$x \cap y_1 \cap \cdots \cap y_b, \quad x \setminus y_1, \dots, x \setminus y_b, \quad y_1 \setminus x, \dots, y_b \setminus x.$$

Furthermore, by composability, for each $s_1, \dots, s_b \in \{0,1\}$, the following set is in $\mathcal{X}$:

$$x(s_1, \dots, s_b) = (x \cap y_1 \cap \cdots \cap y_b) \cup \bigcup_{i\colon s_i = 0} (x \setminus y_i) \cup \bigcup_{i\colon s_i = 1} (y_i \setminus x).$$

Here is another way to describe the sets $x(s_1, \dots, s_b)$. Let $s_1, \dots, s_b \in \{0,1\}$, and define an assignment to variables x_e, for $e \in \mathcal{U}$:

$$x_e = \begin{cases} 1 & \text{if } e \in x \cap y_1 \cdots \cap y_b, \\ \overline{s_i} & \text{if } e \in x \setminus y_i, \\ s_i & \text{if } e \in y_i \setminus x, \\ 0 & \text{otherwise.} \end{cases} \qquad (*)$$

The set $x(s_1, \dots, s_b)$ results from applying this assignment, and interpreting the result as the characteristic vector of a set.

Let us now define a function on the Boolean cube $\{0,1\}^b$:

$$g(s_1, \dots, s_b) = f(x(s_1, \dots, s_b)).$$

By construction, $bs(g) = b$: indeed, $g(0, \dots, 0) = f(x)$ while $g(0, \dots, 0, 1, 0, \dots, 0) = f(y_i)$, where the 1 is in the i'th coordinate. We will soon show that $\widetilde{\deg}_\epsilon(g) \leqslant \widetilde{\deg}_\epsilon(f)$ for all ϵ. The lemma then follows from the well-known bound $bs(g) \leqslant 6\widetilde{\deg}(g)^2$ [NS94, Lemma 3.5].

Recall that $\widetilde{\deg}_\epsilon(f)$ is the minimum degree of a polynomial P such that $|f(x) - P(x)| \leqslant \epsilon$ for all $x \in \mathcal{X}$. If we apply substitution $(*)$ to P, we get a polynomial Q such that $|g(s) - Q(s)| = |f(x(s)) - P(x(s))| \leqslant \epsilon$ for all $s \in \{0,1\}^b$. Since $(*)$ is an affine substitution, clearly $\deg Q \leqslant \deg P$. Therefore $\widetilde{\deg}_\epsilon(g) \leqslant \deg Q \leqslant \deg P$. □

We will need a strengthening of this result for the case $\deg(f) = 1$ in Section 6.2.

Lemma 3.7. *If* $\deg(f) \leqslant 1$ *then* $bs(f) \leqslant 1$.

Proof. Construct the function $g\colon \{0,1\}^{bs(f)} \to \{0,1\}$ as in the proof of Lemma 3.6. By construction, $\deg(g) \leqslant \deg(f) \leqslant 1$ and $bs(g) = bs(f)$. A Boolean degree 1 function on the Boolean cube is a dictator, and in particular, its block sensitivity is at most 1. Therefore $bs(f) = bs(g) \leqslant 1$. □

We can now prove Theorem 3.1.

Theorem 3.1. *Let* $(\mathcal{X}, \mathcal{U}, n)$ *be a composable domain with parameters* Δ, M, β.
Every function $f\colon \mathcal{X} \to \{0,1\}$ *satisfies:*

$$\begin{gathered}
\sqrt{bs(f)/6} \leqslant \widetilde{\deg}(f) \leqslant \deg(f) \leqslant D(f) \\
bs(f) \leqslant fbs(f) \leqslant C(f) \leqslant U(f) \leqslant D(f) \\
bs(f) \leqslant 3R(f) \leqslant 3R_0(f) \leqslant 3D(f) \\
\widetilde{\deg}(f) \leqslant 2Q(f) \leqslant 2Q_E(f) \leqslant 2D(f) \\
D(f) \leqslant \beta^{-2} \max\left(\Delta, \frac{n}{M}\right) bs(f)^4 \qquad (+)
\end{gathered}$$

In particular, if $\Delta = O(1)$, $M = \Omega(n)$, *and* $\beta = \Omega(1)$ *then all complexity measures above are polynomially related.*

Furthermore, we can strengthen (+) *to*

$$D(f) \leqslant \beta^{-2} \Delta bs(f)^4$$

if any of the following conditions holds:

(a) $C(f) \leqslant M$ *(implied by the same bound on* $U(f), D(f)$*).*

(b) $bs(f) \leqslant \sqrt{\beta M}$ *(implied by the same bound on* $fbs(f), R(f)/3, R_0(f)/3$*).*

(c) $\widetilde{\deg}(f) \leqslant \sqrt[4]{\beta M/36}$ *(implied by the same bound on* $\deg(f), Q(f)/2, Q_E(f)/2$*).*

In particular, if $\Delta = O(1)$, $\beta = \Omega(1)$, *and one of the conditions above holds, then all complexity measures above are polynomially related.*

Proof. The first four lines of inequalities follow from Lemma 2.2 and Lemma 3.6:

$$\begin{gathered}
\sqrt{bs(f)/6} \overset{3.6}{\leqslant} \widetilde{\deg}(f) \overset{2.2(a)}{\leqslant} \deg(f) \overset{2.2(h)}{\leqslant} D(f), \\
bs(f) \overset{2.2(g)}{\leqslant} fbs(f) \overset{2.2(g)}{\leqslant} C(f) \overset{2.2(b)}{\leqslant} U(f) \overset{2.2(i)}{\leqslant} D(f), \\
bs(f) \overset{2.2(j)}{\leqslant} 3R(f) \overset{2.2(f)}{\leqslant} 3R_0(f) \overset{2.2(c)}{\leqslant} 3D(f), \\
\widetilde{\deg}(f) \overset{2.2(l)}{\leqslant} 2Q(f) \overset{2.2(e)}{\leqslant} 2Q_E(f) \overset{2.2(k)}{\leqslant} 2D(f).
\end{gathered}$$

The fourth line (+) follows by combining Lemma 3.3 and Lemma 3.5:

$$D(f) \overset{3.3}{\leqslant} \max\left(\Delta, \frac{n}{M}\right) C(f)^2 \overset{3.5}{\leqslant} \beta^{-2} \max\left(\Delta, \frac{n}{M}\right) bs(f)^4.$$

If $C(f) \leqslant M$, then according to Lemma 3.3 we can replace $\max(\Delta, n/M)$ with Δ. This explains condition (a). Condition (b) follows from the inequality $C(f) \leqslant \beta^{-1} bs(f)^2$ of Lemma 3.5, and condition (c) then follows from the inequality $bs(f) \leqslant 6\widetilde{\deg}(f)^2$ of Lemma 3.6. □

3.2. Ball property

Our proof of Theorem 3.2 closely follows the argument of Gopalan et al. [GNS+16].

Theorem 3.2. *Let $(\mathcal{X}, \mathcal{U}, n)$ be a composable domain with chunk size $\mathfrak{c}$ and sensitivity ratio $\beta_{\mathfrak{c}}$.*

If $f\colon \mathcal{X} \to \{0,1\}$ has sensitivity $s = s(f)$, then f can be recovered from its evaluation on a ball of radius $r = \beta_{\mathfrak{c}}^{-1}(2s+1)$ around an arbitrary point.

Proof. Suppose that we are given the values of f at all points at distance at most r from some point $x \in \mathcal{X}$, where the distance between two points $x, y \in \mathcal{X}$ is $d(x,y) = |x \setminus y| = |y \setminus x|$. We will show that the value of $f(y)$ at a point at distance $d \geqslant r$ from x can be determined from the value of f at points at distance less than d from x, and so the entire function f can be constructed step by step from its values on the initial ball.

Let y be a point at distance $d \geqslant r$ from x. By definition of $\beta_{\mathfrak{c}}$, we can find distinct $z_1, \dots, z_t \in \mathcal{X}$, with $t \geqslant \beta_{\mathfrak{c}} d(x,y) \geqslant 2s+1$, such that $d(x, z_i) < d$, the sets $y \setminus z_i$ are disjoint, and $|y \setminus z_i| = \mathfrak{c}$. We will show that $f(y)$ is the majority value of $f(z_1), \dots, f(z_{2s+1})$, completing the proof.

Indeed, suppose that $f(y)$ were not the majority value of $f(z_1), \dots, f(z_{2s+1})$. This means that at least $s+1$ of the points $z_1, \dots, z_{2s+1}$ satisfy $f(z_i) \neq f(y)$. However, since the sets $y \setminus z_i$ are disjoint and $|y \setminus z_i| \leqslant \mathfrak{c}$, that would imply that $s(f, y) \geqslant s+1$, contradicting $s(f) = s$. □

4. Examples of domains

So far we have seen two examples of domains: the Boolean cube, and the symmetric group. In this section we describe three families of domains: product domains (generalizing the Boolean cube), perfect matching domains (generalizing the symmetric group), and multislices. We close the section by briefly discussing spectral notions of degree.

4.1. Product domains

The prototypical product domain is the Boolean cube $\{0,1\}^n$. Much of the theory of the Boolean cube extends to the so-called Hamming scheme $\{1, \dots, m\}^n$. Here we consider a slightly more general domain, in which the number of values in each coordinate could depend on the coordinate:

$$H(m_1, \dots, m_n) = \bigtimes_{i=1}^{n} \{1, \dots, m_i\}.$$

Formal definition and composability The universe is

$$\mathcal{U} = \{(i,j) : 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m_i\}.$$

The sets in the domain are

$$\mathcal{X} = \big\{\{(1,j_1),\ldots,(n,j_n)\} : 1 \leqslant j_i \leqslant m_i\big\}.$$

As in the case of the Boolean cube, the chunk size is $\mathfrak{c} = 1$.

Queries Thinking of the input as a vector $x_1,\ldots,x_n$, we allow queries of the form "$x_i =?$". Formally,

$$\mathcal{Q} = \big\{\{(i,1),\ldots,(i,m_i)\} : 1 \leqslant i \leqslant n\big\}.$$

Composability The domain $\mathcal{X}$ consists of all sets intersecting each query at exactly one point, hence is composable due to Lemma 2.3.

Parameters We now calculate the maximum degree, conflict bound, and sensitivity ratio.

Maximum degree Every element $(i,j) \in \mathcal{U}$ only participates in the query "$x_i =?$", so $\Delta = 1$.

Conflict bound Two partial inputs are conflicting if they specify different values for some coordinate i. Such inputs can be separated by the query "$x_i =?$". Therefore $M = n$.

Sensitivity ratio Given two vectors x and y at distance $d = |x \setminus y|$, we can form d vectors z^i, for each coordinate i for which $x_i \neq y_i$, defined by $z^i = (y \setminus \{(i,y_i)\}) \cup \{(i,x_i)\}$. These vectors satisfy $x \setminus z^i = (x \setminus y) \setminus \{(i,x_i)\}$, and the sets $y \setminus z^i = \{(i,y_i)\}$ are disjoint. Hence $\beta = \beta_{\mathfrak{c}} = 1$.

Main result Applying Theorem 3.1, we deduce the following corollary:

Corollary 4.1. *All complexity measures considered in the paper, except for sensitivity, are polynomially related for all functions on $H(m_1,\ldots,m_n)$. Furthermore, the polynomial relations do not depend on the values of $m_1,\ldots,m_n$.*

In fact, sensitivity is also polynomially related to the other measures, as we show in Section 5.2.

4.2. Perfect matching domains

Perfect matching domains generalize the symmetric group and the perfect matching scheme (the set of all perfect matchings in K_{2n}) to hypergraphs.

Let $\lambda = \lambda_1,\ldots,\lambda_m$ be a sequence of positive integers. The domain $P(n;\lambda)$ consists of all perfect hypermatchings in a $|\lambda|$-uniform hypergraph, where $|\lambda| = \lambda_1 + \cdots + \lambda_m$. The vertex set of the hypergraph is partitioned into m parts $P_1,\ldots,P_m$, where P_i contains $\lambda_i n$ vertices. The hyperedges consist of a choice of λ_i vertices from the the i'th part P_i, for each i. Every perfect hypermatching consists of n hyperedges.

Special cases include the symmetric group $P(n;1,1)$ (perfect matchings in $K_{n,n}$) and the perfect matching scheme $P(n;2)$ (perfect matchings in K_{2n}).

To avoid trivialities, we assume that $n, |\lambda| \geqslant 2$.

Formal definition and composability For $i \in \{1, \dots, m\}$, define

$$P_i = \{(i, j) : 1 \leqslant j \leqslant \lambda_i n\},$$

and let $P = P_1 \cup \cdots \cup P_m$. The universe is

$$\mathcal{U} = \{S \subseteq P : |S \cap P_i| = \lambda_i\}.$$

The sets in the domain are

$$\mathcal{X} = \big\{\{S_1, \dots, S_n\} \subseteq \mathcal{U} : S_1 \cup \cdots \cup S_n = \mathcal{U}\big\}.$$

The chunk size is $\mathfrak{c} = 2$. Indeed, given a set A, if we remove S_j, then the only way to complete it to a set in $\mathcal{X}$ is by adding S_j back. In contrast, we can switch the P_1-parts of S_1 and S_2 in order to get a point B at distance $|B \setminus A| = 2$.

Queries The queries we allow are of the form "which hyperedge does vertex (i, j) participate in?". This generalizes the queries we considered in the case of the symmetric group, and corresponds to the queries "which vertex is i connected to?" in the case of the perfect matching scheme. Formally,

$$\mathcal{Q} = \big\{\{S \in \mathcal{U} : S \ni v\} : v \in P\big\}.$$

Composability Although described slightly differently, $\mathcal{X}$ consists of all sets intersecting each query at exactly one vertex, hence is composable due to Lemma 2.3.

Parameters We now bound the maximum degree, conflict bound, and block sensitivity ratio.

Maximum degree A hyperedge $S \in \mathcal{U}$ appears in one query per each vertex $v \in S$. Hence $\Delta = |\lambda|$.

Conflict bound Suppose that C_1, C_2 are two partial inputs, which we think of as hypermatchings. If $C_1 \setminus C_2$ and $C_2 \setminus C_1$ mention different vertices (that is, different elements of P) then the union $C_1 \cup C_2$ is also a hypermatching. It is not hard to check that every hypermatching can be extended to a perfect hypermatching, and so C_1, C_2 do not conflict. This means that if C_1, C_2 do conflict, then there must exist a vertex $v \in P$ which belongs to different hyperedges in C_1 and C_2. Hence the two partial inputs can be separated using the question "which hyperedge does vertex v participate in?". This shows that $M = n$.

Block sensitivity ratio This is the only nontrivial part in the analysis of perfect matching domains. Our bound on the block sensitivity ratio takes inspiration from the case of the symmetric group, but the argument looks rather different.

Consider two perfect hypermatchings $A, B \in \mathcal{X}$. Recall that our goal is to find as many C_i as possible (in terms of $|A \setminus B|$) such that $A \setminus C_i \subsetneq A \setminus B$, and the sets $B \setminus C_i$ are disjoint.

The idea is to pick a hyperedge $S \in A \setminus B$, and form another perfect hypermatching C_S by modifying B so that it contains S. Later on we will show that we can choose many different hyperedges for which the corresponding perfect hypermatchings are disjoint.

Let $T_1, \dots, T_r$ be the hyperedges in $B \setminus A$ which share vertices with S; note that $r \geqslant 2$, since otherwise $S \in A \cap B$. To form C_S from B, we start by replacing T_1 with S. This means that we are no longer covering vertices in $T_1 \setminus S$, and vertices in $S \setminus T_1$ are covered twice; the two sets contain an equal number of vertices in each part. The vertices covered twice all appear in the sets $T_2, \dots, T_r$. We modify $T_2, \dots, T_r$ by replacing the vertices appearing twice with the vertices $T_1 \setminus S$, replacing each vertex by another vertex in the same part. Thus $B \setminus C_S = \{T_1, \dots, T_r\}$.

To obtain C_S from B, we have modified only hyperedges which conflict with S, and in particular do not belong to A. This shows that $A \setminus C_S \subseteq A \setminus B$. Since $S \in C_S$ by construction, in fact $A \setminus C_S \subsetneq A \setminus B$.

Let us illustrate this process using the case of the symmetric group. We start with two permutations A, B. We choose some edge $S = \{(1,i),(2,j)\} \in A \setminus B$ (encoding that $A(i) = j$). The vertices $(1,i),(2,j)$ appear in two edges of B, say $T_1 = \{(1,i),(2,J)\}$ and $T_2 = \{(1,I),(2,j)\}$. The permutation C_S is given by

$$C_S = B \setminus \big\{\{(1,i),(2,J)\},\{(1,I),(2,j)\}\big\} \cup \big\{\{(1,i),(2,j)\},\{(1,I),(2,J)\}\big\}.$$

The sets $B \setminus C_S$ are not necessarily disjoint. In order for two sets $B \setminus C_{S_1}$ and $B \setminus C_{S_2}$ to intersect, there needs to be a set $T \in B \setminus A$ which intersects both S_1 and S_2. Since each set $S \in A \setminus B$ intersects at most $|\lambda|$ sets $T \in B \setminus A$ and vice versa, we see that each set $B \setminus C_{S_1}$ intersects at most $|\lambda|(|\lambda|-1)$ sets $B \setminus C_{S_2}$. We immediately obtain $\beta \geqslant 1/(|\lambda|(|\lambda|-1)+1)$.

Sensitivity ratio When forming C_S from B, we remove $r \leqslant |\lambda|$ sets from B. Hence if $|\lambda| \leqslant 2$, then $|B \setminus C_S| \leqslant \mathfrak{c}$, implying that we get a bound on $\beta_{\mathfrak{c}}$, namely $\beta_{\mathfrak{c}} \geqslant 1/3$. This holds for both the symmetric group and the perfect matching scheme.

Main result Applying Theorem 3.1, we deduce the following corollary:

Corollary 4.2. *All complexity measures considered in the paper, except for sensitivity, are polynomially related for all functions on $P(n;\lambda)$. Furthermore, the polynomial relations do not depend on the value of n.*

When $|\lambda| = 2$, sensitivity is also polynomially related to the other measures, as we show in Section 5.3 and Section 5.4.

4.3. Multislices

The *slice*, or *Johnson scheme*, consists of all vectors in the Boolean cube $\{0,1\}^n$ with fixed Hamming weight. We consider a multicolored generalization of the slice, known as the *multislice*.

Let $\lambda = \lambda_1, \dots, \lambda_m$ be a sequence of positive integers summing to n; to avoid trivialities, we assume that $m \geqslant 2$. The multislice $M(\lambda)$ is the subset of $\{1, \dots, m\}^n$ consisting of all vectors having exactly λ_c coordinates labeled c. The set of all functions on the multislice is known in the representation theory of the symmetric group as the *permutation module* M^λ.

Formal definition and composability The universe is

$$\mathcal{U} = \{1, \ldots, n\} \times \{1, \ldots, m\}.$$

The sets in the domain are

$$\mathcal{X} = \{S \subseteq \mathcal{U} : \#\{c \colon (i, c) \in S\} = 1, \#\{i \colon (i, c) \in S\} = \lambda_c\}.$$

This definition shows that the domain is composable, by Lemma 2.3.

We sometimes think of the domain as the set of vectors $x = (x_1, \ldots, x_n)$, where $x_i \in \{1, \ldots, m\}$, with λ_c many elements "colored" c.

The chunk size is $\mathfrak{c} = 2$. Indeed, if we "uncolor" an element, there is only one way to color it, so $\mathfrak{c} > 1$. In contrast, we can switch the colors of two elements, and so $\mathfrak{c} = 2$.

Queries We allow queries of the form "$x_i =?$", that is, "what is the color of x_i?". Formally:

$$\mathcal{Q} = \{\{i\} \times \{1, \ldots, m\} : 1 \leqslant i \leqslant n\}.$$

Parameters We now bound the maximum degree, conflict bound, and block sensitivity ratio.

Maximum degree Every element (i, j) appears in exactly one query "$x_i = ?$", and so $\Delta = 1$.

Conflict bound In contrast to the previous domains we consider, the conflict bound in this case is smaller than n. If two conflicting partial inputs a, b disagree on the color of some coordinate i, then they can be separated by the query "$x_i =?$", but this need not be the case in general: another way for two partial inputs to conflict is if in total, they specify more than λ_c coordinates of color c; in all other cases, it is not hard to see that a, b do not conflict. The only way to guarantee that two partial inputs do not specify more than λ_c coordinates of some color c is to limit them to size $M = \lfloor \min(\lambda)/2 \rfloor$, where $\min(\lambda) = \min(\lambda_1, \ldots, \lambda_m)$.

Sensitivity ratio Let us start with the case of the slice, that is $m = 2$. In this case, we are able to get slightly better bounds.

Suppose that $x, y \in \mathcal{X}$ are at distance $d = |x \setminus y|$. This means that there are d indices whose color is different. Since the total number of indices of each color is the same in both vectors, there must be $d/2$ indices $i_1, \ldots, i_{d/2}$ such that $x_{i_t} = 0$ and $y_{i_t} = 1$, and $d/2$ indices $j_1, \ldots, j_{d/2}$ where the opposite happens.

We define z^t, for $t \in \{1, \ldots, d/2\}$, to result from y by switching the colors of indices i_t and j_t. Thus $x \setminus z^t = (y \setminus z^t) \setminus \{(i_t, 0), (j_t, 1)\}$ and $y \setminus z^t = \{(i_t, 1), (j_t, 0)\}$, showing that the sets $y \setminus z^t$ are disjoint. This shows that $\beta_{\mathfrak{c}} = 1/2$ when $m = 2$.

For general multislices, we will show that $\beta_{\mathfrak{c}} \geqslant 1/3$ by reduction to the sensitivity ratio of the symmetric group. Suppose that $x, y \in \mathcal{X}$ are at distance $d = |x \setminus y|$. Let $\pi \in S_n$ be a permutation such that $y_i = x_{\pi(i)}$, and furthermore $\pi(i) = i$ if $y_i = x_i$.

Let id be the identity permutation. By construction, $|\pi \setminus \mathsf{id}| = d$. We have shown in Section 4.2 that $\beta_{\mathfrak{c}} \geqslant 1/3$ for S_n, and so applying the sensitivity ratio property to the pair id, π we

obtain $\ell \geqslant d/3$ permutations $\sigma_1, \ldots, \sigma_\ell$ such that $\mathsf{id} \setminus \sigma_t \subsetneq \mathsf{id} \setminus \pi$, the sets $\pi \setminus \sigma_t$ are disjoint, and $|\pi \setminus \sigma_t| = 2$.

Define $z^t \in \mathcal{X}$ by $z^t_i = x_{\sigma_t(i)}$, so that

$$\begin{aligned} x \setminus y &= \{(i, x_i) : x_i \neq x_{\pi(i)}\}, \\ x \setminus z^t &= \{(i, x_i) : x_i \neq x_{\sigma_t(i)}\}, \\ y \setminus z^t &= \{(i, x_{\pi(i)}) : x_{\pi(i)} \neq x_{\sigma_t(i)}\}. \end{aligned}$$

We have to prove the following properties: $x \setminus z^t \subsetneq x \setminus y$, the sets $y \setminus z^t$ are disjoint, and $|y \setminus z^t| = 2$.

If $x_i \neq x_{\sigma_t(i)}$ then certainly $i \neq \sigma_t(i)$, and so $(i, i) \in \mathsf{id} \setminus \sigma_t$. Therefore $(i, i) \in \mathsf{id} \setminus \pi$, and so $\pi(i) \neq i$. By construction, this implies that $x_i \neq x_{\pi(i)}$. This shows that $x \setminus z^t \subset x \setminus y$.

By assumption, $\mathsf{id} \setminus \sigma_t \subsetneq \mathsf{id} \setminus \pi$, and so there exists i such that $\pi(i) \neq i$ but $\sigma_t(i) = i$. Since $\pi(i) \neq i$, by construction $x_i \neq x_{\pi(i)}$. On the other hand, clearly $x_i = x_{\sigma_t(i)}$, showing that $x \setminus z^t \subsetneq x \setminus y$, proving the first property.

If $(i, x_{\pi(i)}) \in y \setminus z^t$ then $\pi(i) \neq \sigma_t(i)$, and so $(i, \pi(i)) \in \pi \setminus \sigma_t$. Since the sets $\pi \setminus \sigma_t$ are disjoint, so are the sets $y \setminus z^t$, proving the second property. Since $|\pi \setminus \sigma_t| = 2$, also $|y \setminus z^t| = 2$ (since we already know that $y \neq z^t$ and 2 is the minimal distance), proving the third property.

Main result Applying Theorem 3.1, we deduce the following corollary:

Corollary 4.3. *Suppose that $\lambda = \lambda_1, \ldots, \lambda_m$ is a sequence of positive integers summing to n, and let $\lambda_{\min} = \min(\lambda_1, \ldots, \lambda_m)$.*

All complexity measures considered in the paper, except for sensitivity, are polynomially related for all functions on $M(\lambda)$ whose degree is at most $O_m(\lambda_{\min}^{1/4})$.

Furthermore, for any constant $c > 0$, if $\lambda_{\min} \geqslant cn$ then all complexity measures considered in the paper, except for sensitivity, are polynomially related for all functions on $M(\lambda)$ (without any constraint on the degree). Moreover, the polynomial relations do not depend on the value of n (but do depend on m and c).

In fact, sensitivity can also be added to the list of polynomially related measures, as we show in Section 5.5.

4.4. Spectral notions of degree

Analysis of Boolean functions [O'D14] studies functions on the Boolean cube $\{0,1\}^n$ from a spectral perspective. The starting point is the Fourier expansion of a function $f\colon \{0,1\}^n \to \mathbb{R}$:

$$f = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S, \text{ where } \chi_S(x_1, \ldots, x_n) = \prod_{i \in S} (-1)^{x_i}.$$

It is natural to partition the Fourier characters χ_S into levels according to their size, and to define

$$f^{=d} = \sum_{|S|=d} \hat{f}(S)\chi_S.$$

The degree of $f \neq 0$ is defined to be the maximal d such that $f^{=d} \neq 0$. This spectral notion of degree coincides with the spatial notion of degree appearing in Section 2.2.

Stated differently, we can decompose the space of real-valued functions on $\{0,1\}^n$ into $n+1$ subspaces $V_0, \ldots, V_n$, where V_d is spanned by the characters χ_S for $|S| = d$. The degree of a function f is the maximal d such that the projection of f to V_d is non-zero. Furthermore, the subspaces $V_0, \ldots, V_n$ are orthogonal.

Going in the other direction, given the spatial notion of degree appearing in Section 2.1, we can define the space $V_{\leqslant d}$ of all functions of degree at most d, and can then recover V_d as the orthogonal complement of $V_{\leqslant d-1}$ inside $V_{\leqslant d}$.

As we explain in Sections 5.3 to 5.5, there are natural spectral notions of degree for the symmetric group, the perfect matching scheme, and multislices. In all of these cases, the space of real-valued functions on the domain decomposes into isotypic components with respect to the action of the symmetric group. These isotypic components are indexed by partitions, and they can be grouped into levels according to the maximal part. These levels are the analogs of the subspaces V_d described above.

The decomposition $V_0, \ldots, V_n$ for real-valued functions on the Boolean cube can be obtained in a similar way by considering the action of the hyperoctahedral group, which is the symmetry group of the Boolean cube (as a graph or polytope); see for example [Bac09, §2.8.1].

The theory can likely be extended to other homogeneous product domains, that is, product domains of the form $[m]^n$, by considering the action of the wreath product $S_m \wr S_n$.

A different way of recovering the level decomposition in some cases is using the theory of distance regular graphs [BCN89] and cometric association schemes. For example, homogeneous product domains correspond to Hamming schemes, and slices correspond to Johnson schemes.

5. Sensitivity theorems

Theorem 3.1 shows that many different complexity measures are all polynomially related. Sensitivity is conspicuously missing. It had long been conjectured that sensitivity can also be added to the mix, but only recently Huang [Hua19] managed to prove this.

Theorem 5.1 (Huang's sensitivity theorem)**.** *If f is a function on the Boolean cube then $s(f) \geqslant \sqrt{\deg(f)}$.*

In this section we show that similar sensitivity theorems hold for many of the domains considered in Section 4: all product domains, the symmetric group, the perfect matching scheme, and multislices. We do so by reducing the sensitivity theorem on these domains to Huang's sensitivity theorem, using basic facts from representation theory. Our technique is unable to capture hypergraphical perfect matching domains, since the representation theory of these domains is not well-understood (this is related to the notorious Schur plethysm problem, see [Sta99, Theorem A2.8]; see also [Ker99, §7.8]).[2]

[2] In particular, it was recently shown that determining the multiplicities of the permutation representation of the symmetric group S_{mn} acting on the set of perfect hypermatchings of the complete m-uniform hypergraph on mn vertices is #P-hard for all $m > 2$ [FI20].

We introduce our method in Section 5.1, and prove the various sensitivity theorems in the subsequent subsections.

5.1. The method

The basic idea of our method is that if a function on some domain has degree d, then this is often witnessed by some *pseudo-character* $\chi\colon \mathcal{X} \to \{-1,0,1\}$, which decomposes the domain into parts which behave like a Boolean cube of some possibly smaller dimension D. (Formally, a pseudo-character is any function mapping $\mathcal{X}$ to $\{-1,0,1\}$.)

Given a domain $\mathcal{X}$, we will say that a pseudo-character $\chi\colon \mathcal{X} \to \{-1,0,1\}$ is $\{0,1\}^D$-*inducing* if the support of χ can be partitioned into subsets C_i of size 2^D, each of them accompanied with a bijection $\phi_i\colon \{0,1\}^D \to C_i$ satisfying:

(a) If $a \in \{0,1\}^D$ and $b_1,\dots,b_D$ are its neighbors (at Hamming distance 1) then $|\phi_i(a) \setminus \phi_i(b_j)| = \mathfrak{c}$, and furthermore the sets $\phi_i(a) \setminus \phi_i(b_j)$ are disjoint; we say that ϕ_i *maps neighbors to disjoint neighbors*.

(b) If $a,b \in \{0,1\}^D$ are neighbors then $\chi(\phi_i(a)) = -\chi(\phi_i(b))$, that is, the restriction of χ to C_i behaves like the parity character.

If a function $f\colon \mathcal{X} \to \{0,1\}$ has nontrivial correlation with a $\{0,1\}^D$-inducing pseudo-character, then we can apply Huang's sensitivity theorem to obtain a lower bound on the sensitivity of f.

Lemma 5.2. *Let $\mathcal{X}$ be a domain with chunk size $\mathfrak{c}$. Suppose that a function $f\colon \mathcal{X} \to \{0,1\}$ has nontrivial correlation with a $\{0,1\}^D$-inducing pseudo-character χ, that is,*

$$\sum_{x\in\mathcal{X}} f(x)\chi(x) \neq 0.$$

Then $s(f) \geqslant \sqrt{D}$.

Proof. Let C_i be the partition of the support of χ promised by the definition of $\{0,1\}^D$-inducing, and let ϕ_i be the corresponding embedding functions. Since

$$\sum_{x\in\mathcal{X}} f(x)\chi(x) = \sum_i \sum_{x\in C_i} f(x)\chi(x),$$

there must exist some part C_i such that

$$\sum_{x\in C_i} f(x)\chi(x) \neq 0.$$

Define a function $g\colon \{0,1\}^D \to \{0,1\}$ by $g = f \circ \phi_i$. Since $\chi \circ \phi_i$ is the Fourier character $\chi_{\{1,\dots,D\}}$ and ϕ_i is a bijection, it follows that $\deg(g) = D$. Huang's sensitivity theorem shows that $s(g) \geqslant \sqrt{D}$, and in particular, there is a point $y \in \{0,1\}^D$ which has at least $s(g)$ neighbors $z_1,\dots,z_{s(g)} \in \{0,1\}^D$ satisfying $g(y) \neq g(z_j)$. Since $|\phi_i(y) \setminus \phi_i(z_j)| = \mathfrak{c}$ and the sets $\phi_i(y) \setminus \phi_i(z_j)$ are disjoint, this shows that $s(f) \geqslant s(g) \geqslant \sqrt{D}$. □

In order to apply this lemma, we will need our domains to possess pseudo-characters witnessing the degrees of all functions on the domain. Specifically, we will say that a domain is *α-witnessing* if for every $d \leqslant n$, the space of functions of degree at most d is spanned by a collection X_d of pseudo-characters, and every degree d pseudo-character in X_d is $\{0,1\}^D$-inducing for some $D \geqslant \alpha d$.

Theorem 5.3. *If a domain $\mathcal{X}$ is α-witnessing, then every function $f\colon \mathcal{X} \to \{0,1\}$ satisfies*

$$s(f) \geqslant \sqrt{\alpha \deg(f)}.$$

Proof. Suppose that $\deg(f) = d$. If $f = 0$ then $d = 0$, and so there is nothing to prove. Otherwise, $\sum_{x\in\mathcal{X}} f(x)^2 \neq 0$, and so, since $f(x)$ is a linear combination of pseudo-characters in X_d and has degree d, there must be some degree d pseudo-character $\chi \in X_d$ such that $\sum_{x\in\mathcal{X}} f(x)\chi(x) \neq 0$. By assumption, χ is $\{0,1\}^D$-inducing for some $D \geqslant \alpha d$, and so Lemma 5.2 immediately shows that $s(f) \geqslant \sqrt{D} \geqslant \sqrt{\alpha d}$. □

5.2. Product domains

Let us recall the definition of product domains:

$$H(m_1,\dots,m_n) = \{1,\dots,m_1\} \times \cdots \times \{1,\dots,m_n\}.$$

We can assume without loss of generality that $m_1,\dots,m_n \geqslant 2$. We think of each element of $H(m_1,\dots,m_n)$ as an n-dimensional vector v whose i'th entry satisfies $v_i \in \{1,\dots,m_i\}$.

By definition, a function has degree d if it is a linear combination of functions of the form

$$[\![x_{i_1} = j_1]\!] \cdots [\![x_{i_e} = j_e]\!],$$

where $e \leqslant d$ and we can assume, without loss of generality, that the indices i_t are all different. (Recall that $[\![E]\!]$ is the indicator variable of the condition E.)

We will use the identity

$$[\![x_i = j]\!] = \frac{1}{m_i}\sum_{j'=1}^{m_i}([\![x_i = j]\!] - [\![x_i = j']\!]) + \frac{1}{m_i} = \frac{1}{m_i}\sum_{j'\neq j}([\![x_i = j]\!] - [\![x_i = j']\!]) + \frac{1}{m_i}.$$

This identity shows that every degree d function is a linear combination of functions of the form

$$([\![x_{i_1} = j_1]\!] - [\![x_{i_1} = j_1']\!]) \cdots ([\![x_{i_e} = j_e]\!] - [\![x_{i_e} = j_e']\!]),$$

where $e \leqslant d$, all indices i_t are different, and $j_t \neq j_t'$. All such functions are pseudo-characters, that is, are $\{-1,0,1\}$-valued, and they form the collection X_d. Such a pseudo-character has degree d if $e = d$ (this requires a short proof, and follows from standard facts about Fourier expansion in Abelian groups).

We claim that the pseudo-character

$$\chi = ([\![x_{i_1} = j_1]\!] - [\![x_{i_1} = j_1']\!]) \cdots ([\![x_{i_d} = j_d]\!] - [\![x_{i_d} = j_d']\!])$$

is $\{0,1\}^d$-inducing. First notice that the support of χ consists of all vectors in which $x_{i_t} \in \{j_t, j_t'\}$. The support naturally breaks into subcubes C_y, where the index y specifies the values on all coordinates other than $i_1, \dots, i_d$:

$$C_y = \{x \in \mathcal{X} : x_{i_t} \in \{j_t, j_t'\}, x_i = y_i \text{ for all other } i\}.$$

For each C_y, we have to construct a bijection $\phi_y \colon \{0,1\}^d \to C_y$ that maps neighbors to disjoint neighbors and such that $\chi \circ \phi_y$ is the sign character. The bijection is very simple: for $z \in \{0,1\}^d$, $\phi_y(z)$ is the vector given by

$$\phi_y(z)_i = \begin{cases} j_t & \text{if } i = i_t \text{ and } z_t = 0, \\ j_t' & \text{if } i = i_t \text{ and } z_t = 1, \\ y_i & \text{otherwise.} \end{cases}$$

It is easy to check that ϕ_y maps neighbors to disjoint neighbors. For the other property, notice that if z, w are neighbors differing in the t'th coordinate then $\chi(\phi_y(w))$ results from $\chi(\phi_y(z))$ by flipping the sign of the factor $[\![x_{i_t} = j_t]\!] - [\![x_{i_t} = j_t']\!]$, and so $\chi(\phi_z(w)) = -\chi(\phi_y(w))$, as needed.

The foregoing shows that $\mathcal{X}$ is 1-witnessing. Applying Theorem 5.3, we obtain a generalization of Huang's sensitivity theorem to all product domains:

Corollary 5.4. *If $\mathcal{X}$ is a product domain and $f \colon \mathcal{X} \to \{0,1\}$ then*

$$s(f) \geqslant \sqrt{\deg(f)}.$$

Together with Corollary 4.1, this shows that all complexity measures considered in this paper are polynomially related for functions on product domains.

5.3. Symmetric group

Our treatment of the symmetric group will involve basic representation theory, from a slightly unusual perspective. The representation theory of the symmetric group gives an orthogonal decomposition of the space of all real-valued functions on the symmetric group:

$$\mathbb{R}[S_n] = \bigoplus_{\lambda \vdash n} V^\lambda,$$

where $\lambda \vdash n$ means that λ is a partition of n, that is, a nonincreasing sequence of positive integers summing to n. The subspaces V^λ are known as *isotypic components*. Before describing how these isotypic components look like, let us mention a theorem of Ellis, Friedgut and Pilpel [EFP11, Theorem 7] relating degree and the isotypic decomposition:

Theorem 5.5. *The degree of a function $f \colon S_n \to \mathbb{R}$ is the maximum d such that f has a nonzero component in an isotypic component V^λ with $\lambda_1 = n - d$.*

For the sake of completeness, we prove Theorem 5.5 in Appendix A.1.

We now describe a spanning set for V^λ, which can be decoded, for example, from the treatment in Sagan [Sag01, Chapter 2]. Let $\lambda = (\lambda_1, \ldots, \lambda_m)$. A *tabloid of shape* λ consists of the numbers $1, \ldots, n$ arranged in m lines, the i'th line containing λ_i numbers. For example, here is a pair of tabloids of shape $(3, 2)$:

1	2	3
4	5	

4	5	3
1	2	

A pair of tabloids defines a Boolean function on S_n, which indicates that the each row in the first tabloid is mapped by the input permutation to the corresponding row in the second tabloid. For example, the function corresponding to the pair of tabloids above equals 1 whenever the input permutation π satisfies $\pi(\{1, 2, 3\}) = \{4, 5, 3\}$ and $\pi(\{4, 5\}) = \{1, 2\}$; we stress that the order of numbers in each row doesn't affect the definition of the function. Let us denote the function defined in this way for a pair of tabloids A, B by $e_{A,B}$.

For each pair of tabloids A, B of shape λ we define a pseudo-character $\chi_{A,B}$, as follows. We fix A, and consider all ways to permute the columns of B, forming various tabloids B^σ. We sum the functions e_{A,B^σ}, twisted by the sign of σ:

$$\chi_{A,B} = \sum_\sigma (-1)^\sigma e_{A,B^\sigma},$$

where the sum is over all ways to permute the entries of the columns of B, and $(-1)^\sigma$ is the sign of σ. As an illustration, in our running example, representing e_{A,B^σ} by the tabloid B^σ, the pseudo-character $\chi_{A,B}$ is equal to

$$\begin{array}{|c|c|c|}\hline 4&5&3\\\hline 1&2\\\cline{1-2}\end{array} - \begin{array}{|c|c|c|}\hline 1&5&3\\\hline 4&2\\\cline{1-2}\end{array} - \begin{array}{|c|c|c|}\hline 4&2&3\\\hline 1&5\\\cline{1-2}\end{array} + \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}$$

The punch line is that V^λ is spanned by the pseudo-characters $\chi_{A,B}$, where A, B go over all tabloids of shape λ. Moreover, if we restrict A, B to *standard* tabloids, in which the numbers increase in each row and each column, then we get a basis for V^λ; our argument does not use this fact.

Accordingly, we define X_d as follows:

$$X_d = \{\chi_{A,B} : A, B \text{ are of shape } \lambda, \text{ where } \lambda \vdash n \text{ and } \lambda_1 \geqslant n - d\}.$$

Theorem 5.5 shows that $\deg(\chi_{A,B}) = d$ if A, B have shape λ, where $\lambda_1 = n - d$. Let us consider such a pseudo-character $\chi = \chi_{A,B}$. We will show that χ is $\{0, 1\}^D$-inducing for some $D \geqslant d/2$.

Consider an empty tabloid of shape λ. Go over all columns, and partition each column into pairs and, if the column has odd length, a singleton. For a column of length $\ell + 1$, this results in $\lceil \ell/2 \rceil \geqslant \ell/2$ pairs. The total number of positions beyond the first row is d, so in total this gives $D \geqslant d/2$ pairs. We identify each pair with a transposition: if the corresponding entries in A are x, y, then the transposition switches $\pi(x)$ and $\pi(y)$, an operation that we denote by $\pi^{(x\ y)}$.

By construction, the transpositions are disjoint (affect disjoint sets of elements). We denote the set consisting of these transposition by T.

The main observation driving our approach is that if $\tau \in T$ then $\chi(\pi) = -\chi(\pi^\tau)$. This follows directly from the definition of $\chi_{A,B}$, since applying a transposition changes the sign of the column permutation. This suggests looking at the subgroup $\langle T \rangle$ generated by the transposition in T, and considering the cosets

$$C_\pi = \{\pi^\sigma : \sigma \in \langle T \rangle\}.$$

(Note that every coset is described by 2^D different permutations π.) Each such coset is either entirely in the support of χ, or entirely outside its support. The cosets C_π in the support thus partition the support of χ.

Let us consider some coset C_π in the support of T, and write $T = \{\tau_1, \ldots, \tau_D\}$. We define a bijection $\phi_\pi \colon \{0,1\}^D \to C_\pi$ as follows: $\phi_\pi(x) = \pi^\sigma$, where σ is the product of τ_i for all i such that $x_i = 1$. If $x, y \in \{0,1\}^D$ are neighbors then $\phi_\pi(x), \phi_\pi(y)$ differ by a transposition $\tau \in T$, so ϕ_π maps neighbors to neighbors. Since the transpositions in T are disjoint, ϕ_π actually maps neighbors to disjoint neighbors. As already noted above, if $x, y \in \{0,1\}^D$ are neighbors then also $\phi_\pi(x) = -\phi_\pi(y)$. This completes the proof that χ is $\{0,1\}^D$-inducing, for some $D \geqslant d/2$.

We have shown that the symmetric group is $1/2$-witnessing. Applying Theorem 5.3, we generalize Huang's sensitivity theorem to the symmetric group:

Corollary 5.6. *If $f \colon S_n \to \{0,1\}$ then*

$$s(f) \geqslant \sqrt{\deg(f)/2}.$$

Together with Corollary 4.2, this shows that all complexity measures considered in this paper are polynomially related for functions on the symmetric group.

The main insight behind the proof of the sensitivity theorem for the symmetric group is that the pseudo-characters $\chi_{A,B}$ witness the degree of a function. As another illustration of this point of view, we prove the following lemma in Appendix B. In this lemma, a *t-star* is a set of the form

$$\{\pi \in S_n : \pi(i_1) = j_1, \ldots, \pi(i_t) = j_t\}.$$

Lemma 5.7. *Suppose that $\mathcal{F}$ is a subset of S_n contained in a t-star, and its characteristic function has degree at most t. Then either $\mathcal{F}$ is empty, or it is equal to the t-star.*

5.4. Perfect matching scheme

Our treatment of the perfect matching scheme also involves representation theory. We will follow the exposition in Lindzey's Ph.D. thesis [Lin18b], using the notation $\mathcal{M}_{2n}$ for the collection of all perfect matchings in K_{2n}; see also Ceccherini-Silberstein et al. [CSST08, Chapter 11]. As in the case of the symmetric group, there is an orthogonal decomposition

$$\mathbb{R}[\mathcal{M}_{2n}] = \bigoplus_{\lambda \vdash n} V^{2\lambda},$$

where 2λ is the partition obtained by doubling each part in λ. (Note that λ is a partition of n rather than of $2n$.) Lindzey [Lin18b, Theorem 5.1.1] related degree to this decomposition:

Theorem 5.8. *The degree of a function $f\colon \mathcal{M}_{2n} \to \mathbb{R}$ is the maximum d such that f has a nonzero component in an isotypic component $V^{2\lambda}$ with $\lambda_1 = n - d$.*

For the sake of completeness, we prove Theorem 5.8 in Appendix A.2.

Lindzey [Lin18b, Theorem 5.2.6] gave a spanning set for $V^{2\lambda}$, in terms of tabloids of shape 2λ. Given a *single* tabloid A of shape 2λ, the function e_A is the indicator of all perfect matchings in which the two vertices in each edge lie in the same row. For example, here is a tabloid of shape $2(3,2)$:

1	2	3	4	5	6
7	8	9	10		

The corresponding function e_A equals 1 whenever the perfect matching consists of a perfect matching over the vertices $\{1,\dots,6\}$ together with a perfect matching over the vertices $\{7,\dots,10\}$.

For each tabloid A of shape 2λ we define a pseudo-character χ_A by considering all ways of permuting the columns of A, and summing them according to the sign of σ:

$$\chi_A = \sum_{\sigma} (-1)^{\sigma} e_{A^{\sigma}},$$

where σ goes over all ways of permuting the entries of the columns of A. We define

$$X_d = \{\chi_A : A \text{ is of shape } 2\lambda, \text{ where } \lambda \vdash n \text{ and } \lambda_1 \geqslant n - d\}.$$

Theorem 5.8 shows that χ_A has degree d if A is of shape 2λ, where $\lambda_1 = n - d$. We will show that each such pseudo-character is $\{0,1\}^D$-inducing for some $D \geqslant d$.

Go over the columns of A, and partition each column into pairs and, possible, a singleton. For a column of length $\ell+1$, this results in at least $\ell/2$ pairs. Since there are $2d$ elements beyond the first row, this results in $D \geqslant d$ pairs. We think of each pair $(a\ b)$ as a transposition, which acts on a perfect matching by switching the mates of vertices a and b (if a and b are an edge, this has no effect). We denote the set consisting of these transpositions by T. By construction, if m is a matching and $\tau \in T$ then $\chi_A(m^{\tau}) = -\chi_A(m)$.

We can decompose $\mathcal{M}_{2n}$ according to the subgroup $\langle T \rangle$ generated by T:

$$C_m = \{m^{\sigma} : \sigma \in \langle T \rangle\}.$$

Each C_m is either wholly contained in the support of χ, or wholly outside. For every C_m inside the support, we can construct a bijection $\phi_m\{0,1\}^D \to C_m$ just as in the case of the symmetric group, concluding that χ_A is $\{0,1\}^D$ inducing for some $D \geqslant d$.

We have shown that the perfect matching scheme is 1-witnessing. Applying Theorem 5.3, we generalize Huang's sensitivity theorem to the perfect matching scheme:

Corollary 5.9. *If $f\colon \mathcal{M}_{2n} \to \{0,1\}$ then*

$$s(f) \geqslant \sqrt{\deg(f)}.$$

Together with Corollary 4.2, this shows that all complexity measures considered in this paper are polynomially related for functions on the perfect matching scheme.

5.5. Multislices

The final domain we consider is the multislice $M(\mu)$, where μ is a sequence of m positive integers summing to n. This is the collection of all vectors in $\{1,\dots,m\}^n$ in which exactly μ_i coordinates have the value i.

When μ consists of n many 1s, then $M(\mu)$ is just the symmetric group. In other words, the symmetric group is a multislice. Conversely, a mild generalization of the argument for the symmetric group applies to general multislices.

The representation theory of the symmetric group refers to multislices as *permutation modules* M^μ. The isotypic decomposition generalizes:

$$\mathbb{R}[M^\mu] = \bigoplus_{\mu \trianglelefteq \lambda} V^\lambda,$$

where $\mu \trianglelefteq \lambda$ means that μ is dominated by λ, a partial order whose exact definition is immaterial here. We stress that the isotypic components V^λ depend on μ. The relation between degree and this decomposition was worked out by Filmus, O'Donnell and Wu [FOW19, Claim 27]:

Theorem 5.10. *The degree of a function $f\colon M^\mu \to \mathbb{R}$ is the maximum d such that f has a nonzero component in an isotypic component V^λ with $\lambda = n - d$.*

For the sake of completeness, we prove Theorem 5.10 in Appendix A.3.

The spanning set described in Sagan [Sag01, Chapter 2] and mentioned in Section 5.3 actually applies to arbitrary multislices. In the case of the symmetric group, we had a pair of tabloids, each filled with the numbers 1 to n. In this case, the first tabloid is the same, and the second one has *content* μ, that is, it contains μ_i copies of the number i. For example, if $\mu = (3,3)$ and $\lambda = (4,2)$, then one possible pair of tabloids is

1	2	3	4
5	6		

1	1	1	2
2	2		

The function $e_{A,B}$ corresponding to such a pair is the indicator of all vectors in which the multiset of values in the indices appearing in each row of the first tabloid is the contents of the same row in the second tabloid. For example, the pair indicated above corresponds to all vectors $x \in M(3,3)$ satisfying $\langle\!\langle x_1, x_2, x_3, x_4 \rangle\!\rangle = \langle\!\langle 1,1,1,2 \rangle\!\rangle$ and $\langle\!\langle x_5, x_6 \rangle\!\rangle = \langle\!\langle 2,2 \rangle\!\rangle$, where $\langle\!\langle \cdot \rangle\!\rangle$ defines a multiset.

We can now define $\chi_{A,B}$ as in the case of the symmetric group:

$$\chi_{A,B} = \sum_\sigma (-1)^\sigma e_{A^\sigma,B},$$

where σ goes over all ways to permute the columns. We can define X_d in exactly the same way:

$$X_d = \{\chi_{A,B} : A, B \text{ are of shape } \lambda, \text{ where } \lambda \vdash n \text{ and } \lambda_1 \geqslant n - d\}$$

Following the exact same argument as for the symmetric group, we deduce that the multislice is $1/2$-witnessing. Applying Theorem 5.3, we obtain a sensitivity theorem for multislices:

Corollary 5.11. *If* $f\colon M(\mu) \to \{0,1\}$ *then*

$$s(f) \geqslant \sqrt{\deg(f)/2}.$$

Together with Corollary 4.3, this shows that all complexity measures considered in this paper are polynomially related for functions on balanced multislices, and for low complexity functions on arbitrary multislices.

6. Degree 1 functions

Theorem 3.1 shows that in a composable domain with parameters Δ, M, β such that $M \geqslant 1$, a Boolean degree 1 function can be described by a decision tree of depth at most $\beta^{-2}\Delta$ (the constant 6 can be removed by replacing Lemma 3.6 with Lemma 3.7). In many cases, we can actually say more. For example, it is a classical result that a Boolean degree 1 function on the Boolean cube is a *dictator*, that is, depends on at most one coordinate. Ellis, Friedgut and Pilpel [EFP11, Corollary 2] extended this to the symmetric group:

Theorem 6.1. *If* $f\colon S_n \to \{0,1\}$ *has degree at most* 1 *then either* f *depends only on some* $\pi(i)$*, or it depends only on some* $\pi^{-1}(j)$*, where* π *denotes the input permutation.*

Similarly, Filmus and Ihringer [FI19b] extended the dictator result to many domains, including product domains and multislices.

Our main goal in this section is to prove a counterpart of Theorem 6.1 for the perfect matching scheme $\mathcal{M}_{2n}$:

Theorem 6.2. *If* $f\colon \mathcal{M}_{2n} \to \{0,1\}$ *has degree at most* 1 *then* f *either depends only on which vertex gets matched to some vertex* i*, or on whether the perfect matching intersects some triangle* $\{i,j\},\{j,k\},\{k,i\}$.

Since a perfect matching intersects a triangle at most once, Boolean functions depending on whether the input intersects a triangle indeed have degree 1.

One can capture both theorems using a single formulation: if f is a Boolean degree 1 function on S_n or on $\mathcal{M}_{2n}$, then there is a collection of mutually intersecting edges E in the corresponding graph ($K_{n,n}$ for S_n, K_{2n} for $\mathcal{M}_{2n}$) such that f depends only on whether the input perfect matching intersects E. Conversely, all such functions have degree 1.

Our general approach follows that of Ellis, Friedgut and Pilpel: we first describe all *nonnegative* degree 1 functions on the perfect matching scheme, and from this deduce the characterization of Boolean degree 1 function.

Describing all nonnegative degree 1 functions on a domain is essentially the same as determining the *H-representation* of the polytope whose vertices are the characteristic vectors of sets in the domain; an H-representation is simply a set of linear inequalities whose common solution is the polytope.

For the case of the symmetric group, the relevant polytope is the Birkhoff polytope, whose H-representation is implicitly used by Ellis et al. For the case of the perfect matching scheme,

the relevant polytope is the perfect matching polytope of the complete graph, first described by Edmonds [Edm65].

We illustrate our method by reproving Theorem 6.1 in Section 6.1. The more complicated proof of Theorem 6.2 appears in Section 6.2.

6.1. Symmetric group

We defined the symmetric group S_n in Section 4.2 as a collection of sets over the universe

$$\{\{(1,i),(2,j)\} : 1 \leqslant i,j \leqslant n\}.$$

It will be more convenient to change the universe to the set of all pairs (i,j), where $1 \leqslant i,j \leqslant n$. A permutation is then a collection of pairs which contains exactly one pair of the form $(i,\cdot)$ for each i, and exactly one pair of the form $(\cdot,j)$ for each j. We will identify such a set with its characteristic vector x, indexed by pairs i,j.

Suppose that $f\colon S_n \to \mathbb{R}$ has degree at most 1. By definition, this means that it is a linear combination of the functions $1, x_{i,j}$. Since $1 = x_{1,1} + \cdots + x_{1,n}$, we can eliminate 1, and deduce that f can be written in the form

$$f = \sum_{i,j=1}^{n} c_{i,j} x_{i,j}.$$

This representation is not unique: indeed, the space of all functions of degree at most 1 has dimension $(n-1)^2+1$, whereas here we have n^2 parameters. Ellis, Friedgut and Pilpel [EFP11, Theorem 28] show that if f is nonnegative, then we can choose a representation in which $c_{ij} \geqslant 0$.[3] We will show this using a (superficially) different argument, and deduce Theorem 6.1.

The *Birkhoff polytope* B_n is the convex hull of the characteristic vectors of all permutations in S_n. The Birkhoff–von Neumann theorem shows that the polytope has the following H-representation:

$$\begin{aligned}
x_{i,j} &\geqslant 0 && \text{for all } 1 \leqslant i,j \leqslant n \\
\sum_{j=1}^{n} x_{i,j} &= 1 && \text{for all } 1 \leqslant i \leqslant n \\
\sum_{i=1}^{n} x_{i,j} &= 1 && \text{for all } 1 \leqslant j \leqslant n
\end{aligned}$$

If $f \geqslant 0$ on S_n, then since f is linear, $f \geqslant 0$ on the entire polytope B_n. Therefore $f^{-1}(0)$ is a (possibly empty) face of B_n. In contrast, every face of B_n is specified by a set of tight inequalities. This allows us to describe all nonnegative functions on S_n.

Theorem 6.3. *If $f\colon S_n \to \mathbb{R}$ is nonnegative then f is a nonnegative linear combination of the functions x_{ij}.*

[3] Ellis, Friedgut and Pilpel also claim a similar result for larger degrees (their Theorem 27). However, the theorem is false, as explained in [Fil17].

Proof. We prove the theorem by induction on the support of f. The base case, $f = 0$, is trivial.

Let $\min(f) = \min_{x \in S_n} f(x)$ be the minimum value of f. If $\min(f) > 0$, then we can represent $f - \min(f)$ as a nonnegative linear combination of the functions x_{ij}. Since $1 = x_{11} + \cdots + x_{1n}$, it follows that f is also a nonnegative linear combination of the functions x_{ij}.

Now suppose that $\min(f) = 0$ but $f \neq 0$. Note first that since the minimum of f over B_n is attained at a vertex, f is nonnegative on the entire polytope B_n. Consider the set $F = \{x \in B_n : f(x) = 0\}$. Since F is the set of points at which the linear function f is minimized, F is a face of B_n, that is,

$$F = B_n \cap \{x_{i_1,j_1} = \cdots = x_{i_m,j_m} = 0\},$$

for some set of pairs $(i_1, j_1), \ldots, (i_m, j_m)$; note that $m \geqslant 1$ since $f \neq 0$.

Let $(i, j) = (i_1, j_1)$. The definition of F implies that if $x_{i,j} = 1$ then $f(x) \neq 0$. Let m be the minimal value of f among all points satisfying $x_{i,j} = 1$ (it doesn't matter if we take the minimum over S_n or over B_n, since the minimum is the same). Then $f - mx_{i,j}$ is nonnegative and has smaller support, and so we can represent it as a linear combination of the functions x_{ij}. It follows that f can be represented in the same way. □

We deduce Theorem 6.1 following the argument of Ellis, Friedgut and Pilpel.

Theorem 6.1. *If $f \colon S_n \to \{0, 1\}$ has degree at most 1 then either f depends only on some $\pi(i)$, or it depends only on some $\pi^{-1}(j)$, where π denotes the input permutation.*

Proof. Every Boolean function is a fortiori nonnegative, and so Theorem 6.3 shows that we can write

$$f = \sum_{i,j} c_{i,j} x_{i,j},$$

where $c_{i,j} \geqslant 0$. If $c_{i,j} > 0$ then $f(x) > 0$ whenever $x_{i,j} = 1$. Since f is Boolean, this shows that in fact $f(x) = 1$ whenever $x_{i,j} = 1$, and so the function $f - x_{i,j}$ is also a Boolean degree 1 function. A simple induction on the support thus shows that we can write

$$f = \sum_{(i,j) \in S} x_{i,j},$$

for some subset S of pairs. Any two pairs in S must conflict, that is, cannot belong to the same permutation. As edges of $K_{n,n}$, this means that they intersect at a vertex. It is easy to check that a set of edges are pairwise intersecting if and only if they are either all of the form $(i, \cdot)$ or all of the form $(\cdot, j)$. In the former case, f depends only on $\pi(i)$, and in the latter case, f depends only on $\pi^{-1}(j)$. □

In this argument, we were lucky that the functions x_{ij} are all Boolean. This won't be the case for the perfect matching scheme, a difficulty which will require an additional argument to address.

6.2. Perfect matching scheme

The perfect matching scheme $\mathcal{M}_{2n}$, as defined in Section 4.2, is the collection of all perfect matchings of K_{2n}, considered as sets of edges. We will identify a perfect matching with its characteristic vector x, indexed by *unordered* pairs i, j. As in the case of the symmetric group, using the identity $1 = x_{1,2}+\cdots+x_{1,2n}$ we can represent each degree 1 function f (non-uniquely) as a linear combination

$$\sum_{i,j} c_{i,j} x_{i,j}.$$

The *perfect matching polytope* P_n is the convex hull of the characteristic vectors of all perfect matchings in $\mathcal{M}_{2n}$. Edmonds [Edm65] determined the H-representation of P_n in terms of the cut functions

$$\delta_S(x) = \sum_{i\in S} \sum_{j\notin S} x_{i,j}.$$

Edmonds' theorem [Sch03, Chapter 25] states that the H-representation of P_n is:

$x_{i,j} \geqslant 0$	for all $1 \leqslant i \neq j \leqslant 2n$
$\delta_{\{i\}}(x) = 1$	for all $1 \leqslant i \leqslant 2n$
$\delta_S(x) \geqslant 1$	for all $S \subseteq \{1, \ldots, 2n\}$ such that $3 \leqslant \lvert S\rvert \leqslant n$ is odd

It will be more useful to replace the constraint $\delta_S(x) \geqslant 1$ with the equivalent constraint $d_S(x) \geqslant 0$, where

$$d_S(x) = \frac{\delta_S(x) - 1}{2}.$$

If x is a vertex of P_n then $d_S(x)$ is always an integer; indeed, $\lvert S\rvert - \delta_S(x)$ is the number of vertices in S matched inside S, which is always even.

We can repeat the argument of Theorem 6.3, extending it to the perfect matching scheme.

Theorem 6.4. *If $f\colon \mathcal{M}_{2n} \to \mathbb{R}$ is nonnegative then f is a nonnegative linear combinations of the functions $x_{i,j}$ and $d_S(x)$, where $3 \leqslant \lvert S\rvert \leqslant n$ is odd.*

Proof. The proof is by induction on the support of f. The base case, $f = 0$, is trivial. If $\min(f) > 0$, then the result follows by inducting on $f - \min(f)$, using the formula $1 = x_{1,2} + \cdots + x_{1,2n}$.

Suppose now that $\min(f) = 0$ and $f \neq 0$, and let $F = \{x \in P_n : f(x) = 0\}$. Since F is the set of points minimizing the linear function f over P_n, we see that F is a face of P_n. Therefore F is the intersection of P_n with equations of the form $x_{i,j} = 0$ and $d_S(x) = 0$, the latter for odd $3 \leqslant \lvert S\rvert \leqslant n$. Since $f \neq 0$, there must be at least one such equation.

If the defining equations of F include an equation $x_{i,j} = 0$, then this means that whenever $x_{i,j} = 1$, we must have $f(x) > 0$. Let m be the minimum value of f on all points at which $x_{i,j} = 1$. Then $f - mx_{i,j}$ is nonnegative and has smaller support, so can be represented inductively as a linear combination of the required form, and the same holds for f.

Similarly, if the defining equations of F include an equations $d_S(x) = 0$, then this means that whenever $d_S(x) \geqslant 1$, we must have $f(x) > 0$. Let m be the minimum value of $f(x)/d_S(x)$ over all points at which $d_S(x) \neq 0$. Then $f - md_S$ is nonnegative and has smaller support, so can be represented inductively in the required form, and the same holds for f. □

In order to deduce Theorem 6.2, we would like to use an argument similar to that of Theorem 6.1. However, direct replication of the argument fails, since the function $d_S(x)$ is only Boolean if $|S| = 3$. In order to rule out the appearance of $d_S(x)$ for larger S, we appeal to Lemma 3.7, which states that $\deg(f) \leqslant 1$ implies $bs(f) \leqslant 1$.

Theorem 6.2. *If $f\colon \mathcal{M}_{2n} \to \{0,1\}$ has degree at most 1 then f either depends only on which vertex gets matched to some vertex i, or on whether the perfect matching intersects some triangle $\{i,j\},\{j,k\},\{k,i\}$.*

Proof. Since the function f is nonnegative, it can be written as

$$f = \sum_{i,j} c_{i,j} x_{i,j} + \sum_S c_S d_S,$$

where $c_{i,j}, c_S \geqslant 0$, and the second sum is over all sets S whose size is odd and satisfies $3 \leqslant |S| \leqslant n$.

If $f = 1$ then the theorem is trivial. We will show that if $f \neq 1$ then $c_S = 0$ whenever $|S| \geqslant 5$.

Suppose, to the contrary, that $c_S > 0$ for $S = \{1,\dots,2k+1\}$, where $k \geqslant 2$. Since $f \neq 1$, there is some perfect matching y at which $f(y) = 0$, and so $d_S(y) = 0$. This means that y matches $2k$ of the vertices of S, say $\{1,2\},\dots,\{2k-1,2k\} \in Y$, where Y is the perfect matching whose characteristic vector is y. The vertex $2k+1$ is matched to some other vertex, say $2k+2$. Since $2n \geqslant 2|S| \geqslant 4k+2 \geqslant 2k+6$, the perfect matching Y must contain at least two more edges, say $\{2k+3,2k+4\}$ and $\{2k+5,2k+6\}$. So

$$\{1,2\},\dots,\{2k+5,2k+6\} \in Y.$$

We construct two new perfect matchings:

$$Z = Y \setminus \{\{1,2\},\{2k+3,2k+4\}\} \cup \{\{1,2k+3\},\{2,2k+4\}\},$$
$$W = Y \setminus \{\{3,4\},\{2k+5,2k+6\}\} \cup \{\{3,2k+5\},\{4,2k+6\}\}.$$

The characteristic vectors y, z of these perfect matchings satisfy $d_S(z) = d_S(w) = 1$, and so $f(z), f(w) > 0$, implying that $f(z) = f(w) = 1$. This shows that $s(f,y) \geqslant 2$, contradicting the bound $bs(f) \leqslant 1$ given by Lemma 3.7.

It follows that f is a nonnegative linear combination

$$f = \sum_{i,j} c_{i,j} x_{i,j} + \sum_{|S|=3} c_S d_S.$$

If $c_{i,j} > 0$ then $f(x) > 0$ and so $f(x) = 1$ whenever $x_{i,j} = 1$, implying that $f - x_{i,j}$ is still a Boolean degree 1 function. The same holds for c_S, since d_S is Boolean when $|S| = 3$. A simple induction shows that we can write f as a sum of $x_{i,j}$ and d_S. Since f is Boolean, in this sum no two terms can equal 1 at the same time.

The term $x_{i,j}$ equals 1 if the perfect matching contains the edge $\{i,j\}$; we will denote this "edge event" by $E_{i,j}$. The term $c_{i,j,k}$ equals 1 if the perfect matching contains none of the edges $\{i,j\},\{j,k\},\{k,i\}$; we will denote this "triangle event" by $\Delta_{i,j,k}$. To complete the proof, we need to consider when two such events, of either type, cannot co-occur. Note that triangle events are only relevant when $2n \geqslant 6$ (since $|S| \leqslant n$).

We start by considering two triangle events $\Delta_{S_1}, \Delta_{S_2}$. There are three cases to consider, depending on the size of the intersection $S_1 \cap S_2$:

S_1, S_2 **disjoint:** Say $S_1 = \{1,2,3\}$ and $S_2 = \{4,5,6\}$. There is a perfect matching containing the edges $\{1,4\},\{2,5\},\{3,6\}$, in which both events occur.

S_1, S_2 **share an edge:** Say $S_1 = \{1,3,4\}$ and $S_2 = \{2,3,4\}$. There is a perfect matching containing the edges $\{1,2\},\{3,5\},\{4,6\}$, in which both events occur.

S_1, S_2 **share a vertex:** Say $S_1 = \{1,2,5\}$ and $S_2 = \{3,4,5\}$. There is a perfect matching containing the edges $\{1,3\},\{2,4\},\{5,6\}$, in which both events occur.

This means that the sum representing f contains at most one triangle event.

We move on to consider a triangle event Δ_S and an edge event E_T. Again there are three cases to consider, depending on the size of the intersection $S \cap T$:

S, T **disjoint:** Say $S = \{1,2,3\}$ and $T = \{4,5\}$. Suppose first that $2n \geqslant 8$. In this case there is a perfect matching containing the edges $\{1,6\},\{2,7\},\{3,8\},\{4,5\}$, in which both events occur.

When $2n = 6$, the two events cannot co-occur. A simple case analysis shows that $\Delta_{1,2,3} \vee E_{4,5}$ is equivalent to $E_{1,6} \vee E_{2,6} \vee E_{3,6}$.

S, T **share a vertex:** Say $S = \{1,2,3\}$ and $T = \{3,4\}$. There is a perfect matching containing the edges $\{1,5\},\{2,6\},\{3,4\}$, in which both events occur.

S, T **share an edge:** Say $S = \{1,2,3\}$ and $T = \{1,2\}$. In this case the events are clearly mutually exclusive. If both of them *do not occur* then the perfect matching doesn't contain $\{1,2\}$ but contains either $\{1,3\}$ or $\{2,3\}$, and vice versa: if the perfect matching contains either $\{1,3\}$ or $\{2,3\}$ then both events do not occur. Therefore $\Delta_{1,2,3} \vee E_{1,2}$ is equivalent to $E_{3,4} \vee \cdots \vee E_{3,2n}$.

This means that the sum representing f either consists of a single triangle event, or otherwise is equivalent to another sum consisting only of edge events. In the latter case, the edges must pairwise intersect, and so are either all adjacent to a single vertex, or else form a triangle. □

7. Intersecting families

Consider some domain $\mathcal{X}$. A subset $\mathcal{F} \subseteq \mathcal{X}$ is *t-intersecting* if any two sets $S_1, S_2 \in \mathcal{F}$ have at least t elements in common: $|S_1 \cap S_2| \geqslant t$. If $t = 1$, then we call $\mathcal{F}$ an *intersecting* family. There is also a bipartite version: two subsets $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{X}$ are *cross-t-intersecting* (or *cross-intersecting* when $t = 1$) if any $S_1 \in \mathcal{F}_1$ and $S_2 \in \mathcal{F}_2$ have at least t elements in common.

Erdős, Ko and Rado [EKR61] determined the maximum size of an intersecting family for the domain $\binom{[n]}{k}$, which consists of all subsets of $\{1,\ldots,n\}$ of size k. This domain is similar to, but not identical with, the multislice $M(k, n-k)$ (the difference is that we consider $\overline{S_1 \triangle S_2}$ instead of $S_1 \cap S_2$). Ahlswede and Khachatrian [AK97, AK99] extended this to arbitrary t (for the same domain), and also considered [AK98] *t-agreeing* families in $\{1,\ldots,m\}^n$, which are just t-intersecting families in $H(m,\ldots,m)$ (n many copies).

Ellis, Friedgut and Pilpel [EFP11] studied t-intersecting families in the symmetric group, following earlier work on the case $t = 1$ [FD77, CK03, LM04]. They showed that for every fixed t and large enough n (large enough in terms of t), the maximum size of a t-intersecting family is $(n-t)!$, matching the size of *t-stars*, which are families of the form

$$\{\pi \in S_n : \pi(i_1) = j_1, \ldots, \pi(i_t) = j_t\}.$$

Furthermore, they claimed that for large enough n, the t-stars are the only t-intersecting families of size $(n-t)!$. Unfortunately, their proof is wrong when $t \geqslant 2$, as pointed out in [Fil17], although the result does follow from subsequent work of Ellis [Ell11]. The argument of Ellis actually proves a much stronger structural result, and is quite complicated. Our goal in this section is to present an alternative proof of this property, known as *uniqueness*, using the connection between degree and certificate complexity.

Ellis et al. also studied cross-t-intersecting families in the symmetric group, showing that for every fixed t and large enough n, the maximum product of sizes of two cross-t-intersecting families is $(n-t)!^2$. They claimed that the only extremal examples are when the two families are the same t-star. Again the argument is flawed, but the result follows from the work of Ellis. Our technique also applies to this result.

Our arguments apply to other domains as well, such as the perfect matching scheme, simplifying the characterization of maximum size t-intersecting families of perfect matchings due to Lindzey [Lin18a, Lin18b].

7.1. Main result

Setup Our result is stated in terms of two parameters of a domain $(\mathcal{X}, \mathcal{U}, n)$.

Maximum size of links For any $S \subseteq \mathcal{U}$, the *link* of S is

$$\mathcal{X}_S = \{T : T \in \mathcal{X}, T \supseteq S\}.$$

(We typically consider only non-empty links.) A link $\mathcal{X}_S$ is a *t-link* if $|S| = t$.

We define N_t to be the maximum size of a t-link.

Intersection bound Suppose that $x \in \mathcal{X}$ is a set that t-intersects all $y \in \mathcal{X}$ containing some partial input C (in our case, C will be a 1-certificate of a t-intersecting family). One way in which this can happen is if C itself t-intersects x, but this can fail if C is very large. For example, in the symmetric group, the identity permutation intersects the single permutation containing $\{1,2\}, \{2,3\}, \ldots, \{n-1,1\}$. Intersection bounds are bounds on the size of C which guarantee that this strange situation does not happen.

For any integer $t \geqslant 1$, the *intersection bound* I_t is the maximal value such that for any $x \in \mathcal{X}$ and any partial input C of size at most I_t, if x t-intersects all total inputs extending C then $|x \cap C| \geqslant t$.

We can now state the main result.

Theorem 7.1. *Consider a domain $(\mathcal{X}, \mathcal{U}, n)$ with parameters N_t, I_t.*

If $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{X}$ are cross-t-intersecting, and the characteristic function $f_1 \colon \mathcal{X} \to \{0,1\}$ of $\mathcal{F}_1$ satisfies $C(f_1) \leqslant I_t$, then either $\mathcal{F}_1$ is contained in a t-link, or the size of $\mathcal{F}_2$ can be bounded:

$$|\mathcal{F}_2| \leqslant \binom{C(f_1)}{t} C(f_1) N_{t+1}.$$

Before proving this result, let us indicate why it is useful in the case of the symmetric group (the case of the perfect matching scheme is similar), concentrating on the case $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{F}$. Using a spectral argument, Ellis et al. [EFP11] showed that if $|\mathcal{F}| = (n-t)!$ and n is large enough, then $\deg(f) \leqslant t$, which by Corollary 4.2 implies that $C(f) = O_t(1)$. Therefore either $\mathcal{F}$ is a t-star, or it has size $O((n-t-1)!)$, which for large enough n is smaller than $(n-t)!$.

Proof. Suppose that $\mathcal{F}_1$ is not contained in any t-link. In particular, it is not empty. Choose arbitrarily some $x \in \mathcal{F}_1$, and let C_x be a certificate for x of size at most $C(f_1)$. For any subset $S \subseteq C_x$ of size t, by assumption there is some $y_S \in \mathcal{F}_1$ which does not belong to $\mathcal{X}_S$. Let C_{y_S} be a certificate for y_S of size at most $C(f_1)$. By construction, $S \nsubseteq C_{y_S}$.

We will show that $\mathcal{F}_2$ can be covered by $\binom{C(f_1)}{t} C(f_1)$ many $(t+1)$-links, thus proving the theorem. To this end, let $p \in \mathcal{F}_2$ be arbitrary. Since p t-intersects all total inputs extending C_x and $|C_x| \leqslant C(f_1) \leqslant I_t$, necessarily $|p \cap C_x| \geqslant t$. Let S be an arbitrary subset of $p \cap C_x$ of size t. Since p also t-intersects all total inputs extending C_{y_S} and $|C_{y_S}| \leqslant C(f_1) \leqslant I_t$, necessarily $|p \cap C_{y_S}| \geqslant t$. Since C_{y_S} doesn't contain all of S, the intersection $p \cap C_{y_S}$ must contain some element $a \in C_{y_S} \setminus S$. In particular, $p \in \mathcal{X}_{S \cup \{a\}}$.

Since the element p was arbitrary, we conclude that

$$\mathcal{F}_2 \subseteq \bigcup_{S \in \binom{C_x}{t}} \bigcup_{a \in C_{y_S} \setminus S} \mathcal{X}_{S \cup \{a\}},$$

and so $|\mathcal{F}_2| \leqslant \binom{C(f_1)}{t} C(f_1) N_{t+1}$. □

In Section 7.2, we calculate the intersection bounds of the various domains considered in Section 4. In Section 7.3 we describe the spectral method used by Ellis, Friedgut and Pilpel [EFP11] and by Lindzey [Lin18a, Lin18b], and show how to apply Theorem 7.1 in this setting.

7.2. Intersection bounds

In this section, we compute or bound the intersection bounds for all domains considered in Section 4.

Product domains Recall that the product domain $H(m_1, \ldots, m_n)$ consists of all sets of the form

$$\{(1, j_1), \ldots, (n, j_n)\}, \text{ where } j_i \in \{1, \ldots, m_i\}.$$

To avoid trivialities, we assume that $m_1, \ldots, m_n \geqslant 2$.

Lemma 7.2. *The intersection bounds of $H(m_1, \ldots, m_n)$ are $I_t = \infty$ for all t.*

Proof. Suppose that C is a partial input, and that x is an input that t-intersects all extensions of C. We can think of x as a vector $x_1, \dots, x_n$, where $x_i \in \{1, \dots, m_i\}$. Similarly, C is a partial vector. We extend C to an input y according to the following rule: if C_i is undefined, we choose y_i to be some element different from x_i; this is possible since $m_i \geqslant 2$. This guarantees that $|x \cap C| = |x \cap y| \geqslant t$. □

Perfect matching domains Recall that a perfect matching domain $P(n; \lambda)$ is given by an integer $n \geqslant 2$ and a sequence $\lambda = \lambda_1, \dots, \lambda_m$ of positive integers summing to $|\lambda| \geqslant 2$. The domain consists of all $|\lambda|$-uniform hypermatchings in the complete "λ-partite" hypergraph on $|\lambda|n$ vertices. In more detail, the vertices are partitioned into m parts $P_1, \dots, P_m$, the i'th part containing $\lambda_i n$ vertices $(i, 1), \dots, (i, \lambda_i n)$, and the hyperedges consist of a choice of λ_i elements from P_i for each $i \in \{1, \dots, m\}$.

Lemma 7.3. *The intersection bounds of $P(n; \lambda)$ are $I_t = n - 2$ for all $t \leqslant n - 2$.*

Proof. We start by showing that $I_t \geqslant n - 2$. Let C be a partial input of size at most $n - 2$, and let x be an input which t-intersects all total inputs extending C.

Denote by V the set of vertices mentioned by hyperedges in C; thus C is a perfect hypermatching of V. Let $x|_{\overline{V}}$ be the set of hyperedges in x only involving vertices outside of V. We can complete $x|_{\overline{V}}$ to a perfect hypermatching z of $\overline{V}$.

Without loss of generality, suppose that $\overline{V} \cap P_i$ consists of the vertices $(i, 1), \dots, (i, \lambda_i r)$, where $r = n - |C| \geqslant 2$, and that z consists of the hyperedges

$$\{(i, 1), \dots, (i, \lambda_i) : 1 \leqslant i \leqslant m\}, \{(i, \lambda_i + 1), \dots, (i, 2\lambda_i) : 1 \leqslant i \leqslant m\}, \dots$$

Let w consist of the hyperedges

$$\{(i, \lambda_i r), (i, 1), \dots, (i, \lambda_i - 1) : 1 \leqslant i \leqslant m\}, \{(i, \lambda_i), \dots, (i, 2\lambda_i - 1) : 1 \leqslant i \leqslant m\}, \dots$$

Notice that z and w are disjoint; we can think of w as a generalized derangement.[4]

Since $C \cup w$ is a total input extending C, we have $|C \cap x| = |(C \cup w) \cap x| \geqslant t$.

To complete the proof, we show that $I_t \leqslant n - 2$. Let x, y be two inputs such that $|x \cap y| = t$; such inputs can be constructed using generalized derangements. Let C be obtained from y by removing one of the hyperedges in $x \cap y$. By construction, $|x \cap C| = t - 1$. On the other hand, y is the only total input extending C, and so x t-intersects all total inputs extending C. □

Multislices Recall that a multislice $M(\lambda)$ is specified by a sequence $\lambda_1, \dots, \lambda_m$ of positive integers summing to n, where $m \geqslant 2$. The multislice consists of all vectors $x \in \{1, \dots, m\}^n$ (encoded as sets $\{(i, x_i) : 1 \leqslant i \leqslant n\}$) having exactly λ_c coordinates equal to c.

Lemma 7.4. *The intersection bounds of $M(\lambda)$ satisfy $I_t \geqslant n - 2\max(\lambda_1, \dots, \lambda_m)$.*

[4] Recall that a *derangement* is a permutation in S_n without fixed points. In our setting, we can think of a derangement as a permutation in S_n which is disjoint from the identity permutation $\mathrm{id} \in S_n$. More generally, given a domain $\mathcal{X}$, we can define $a \in \mathcal{X}$ to be a derangement with respect to $b \in \mathcal{X}$ if $a \cap b = \varnothing$. Under this definition (and working inside an appropriate link of $P(n; \lambda)$), w is a derangement with respect to z.

Proof. Let C be a partial input of size at most $n - 2\max(\lambda_1, \dots, \lambda_m)$, and let x be an input which t-intersects all total inputs extending C.

The link of C is another multislice $M(\mu)$, where $\mu_1 + \cdots + \mu_m = n - |C|$, and

$$\max(\mu_1, \dots, \mu_m) \leqslant \max(\lambda_1, \dots, \lambda_m) \leqslant \frac{n - |C|}{2}.$$

Consider a bipartite graph with $n - |C|$ vertices on both sides. We color the first μ_1 vertices on each side by color 1, the following μ_2 vertices by color 2, and so on. We connect a vertex of color i on the left to all vertices on the right of colors different from i.

We claim that any set S on the left has at least $|S|$ neighbors on the right, and so the graph has a perfect matching by Hall's criterion. Indeed, if S contains vertices of more than one color, then its neighborhood consists of all vertices on the right. If all vertices in S are colored i then $|S| \leqslant \mu_i$ and S has exactly $n - |C| - \mu_i$ neighbors, which is at least $|S|$ since $\mu_i \leqslant (n - |C|)/2$.

Using the perfect matching whose existence is promised by Hall's criterion, we can extend C to a total input y which disagrees with x on all indices outside of those mentioned by C. It follows that $|x \cap C| = |x \cap y| \geqslant t$. □

When $\lambda_1 = \cdots = \lambda_m = 1$, this bound coincides with the bound for the symmetric group implied by Lemma 7.3.

7.3. Spectral technique

The results of Ellis, Friedgut and Pilpel [EFP11] and of Lindzey [Lin18a, Lin18b] are proved using a spectral technique known as the *weighted Hoffman bound*, pioneered in this context by Wilson [Wil84] and Frankl–Wilson [FW86], and later cast in a different form by Friedgut [Fri08]; see also the monograph of Godsil and Meagher [GM15].

In a nutshell, in order to prove a t-intersecting theorem for a domain $\mathcal{X}$, the idea is to construct a $\mathcal{X} \times \mathcal{X}$ matrix A, supported on pairs on non-t-intersecting elements, satisfying certain spectral properties.

Definition 7.5. A real $\mathcal{X} \times \mathcal{X}$ matrix A is *t-good* if the following properties hold:

1. A is symmetric.
2. If $|x \cap y| \geqslant t$ then $A(x, y) = 0$.
3. $A\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the constant 1 vector.
4. If $\deg(f) \leqslant t$ and $\mathbb{E}[f] = 0$, where $\mathbb{E}[f] = \sum_{x \in \mathcal{X}} f(x)/|\mathcal{X}|$, then $Af = -\omega f$, where
$$\omega = \frac{N_t}{|\mathcal{X}| - N_t}.$$
5. If $Af = \lambda f$ and $\deg(f) > t$ then $|\lambda| < \omega$.

We remark that if we replace $|\lambda| < \omega$ with $\lambda > -\omega$ in the final property, then we can still recover all results about t-intersecting families; the stronger guarantee $|\lambda| < \omega$ is only needed to handle cross-t-intersecting families.

Ellis, Friedgut and Pilpel [EFP11, Theorem 26] constructed a t-good matrix for $\mathcal{X} = S_n$ for large enough n (as a function of t). Similarly, Lindzey [Lin18a, Lin18b] constructed a t-good matrix for $\mathcal{X} = \mathcal{M}_{2n}$ for large enough n (as a function of t).

Wilson [Wil84] constructed a matrix satisfying a similar condition for $\mathcal{X} = M(n-k,k)$; however, his definition of t-intersecting is different from ours. Frankl and Wilson [FW86] constructed such a matrix when $\mathcal{X}$ is the Grassmann scheme, a domain which doesn't fall into our framework (see Section 8). Friedgut [Fri08] constructed such a matrix for the Boolean cube $\{0,1\}^n$ under a biased measure, also using a different definition of t-intersecting.

It is well-known that the existence of a t-good operator implies that a t-intersecting family contains at most N_t elements. For completeness, we reproduce the proof in Appendix C.

Proposition 7.6. *Suppose that there exists a t-good matrix for $\mathcal{X}$. Then a t-intersecting family contains at most N_t points, and furthermore, the characteristic function of a t-intersecting family of size N_t has degree at most t.*

Using Theorem 7.1, we can show that for large enough n, the bound N_t is achieved only by t-links.

Theorem 7.7. *Suppose that there exists a t-good matrix for $\mathcal{X}$, and every function of degree at most t has certificate complexity at most $C_t \leqslant I_t$. If $C_t\binom{C_t}{t}N_{t+1} < N_t$ then any t-intersecting family of size N_t is a t-link.*

Proof. Let $\mathcal{F}$ be a t-intersecting family of size N_t, and let f be its characteristic function. According to Proposition 7.6, $\deg(f) \leqslant t$. By assumption, $C(f) \leqslant I_t$, and so Theorem 7.1 shows that either $\mathcal{F}$ is contained in a t-link, or $|\mathcal{F}| \leqslant C_t\binom{C_t}{t}N_{t+1} < N_t$. The second case cannot happen by assumption. We conclude that $\mathcal{F}$ is contained in a t-link. Since $|\mathcal{F}| = N_t$, it has to be the complete t-link. □

We can extend this result to cross-t-intersecting families. We start with a cross-intersecting version of Proposition 7.6, which is also well-known. For completeness, we include the proof in Appendix C.

Proposition 7.8. *Suppose that there exists a t-good matrix for $\mathcal{X}$. If $\mathcal{F},\mathcal{G}$ are two cross-t-intersecting families then $\sqrt{|\mathcal{F}|\cdot|\mathcal{G}|} \leqslant N_t$. Furthermore, if equality holds then $\mathcal{F} = \mathcal{G}$, and the common characteristic function has degree at most t.*

As a simple corollary, we can generalize Theorem 7.7 to cross-t-intersecting families.

Theorem 7.9. *Suppose that there exists a t-good matrix for $\mathcal{X}$, and every function of degree at most t has certificate complexity at most $C_t \leqslant I_t$. Assume that $C_t\binom{C_t}{t}N_{t+1} < N_t$. If $\mathcal{F}$ and $\mathcal{G}$ are cross-t-intersecting families and $\sqrt{|\mathcal{F}|\cdot|\mathcal{G}|} = N_t$ then $\mathcal{F} = \mathcal{G}$ is a t-link.*

Proof. Proposition 7.8 implies that $\mathcal{F} = \mathcal{G}$, and so $\mathcal{F}$ is a t-intersecting family of size N_t. The result now follows from Theorem 7.7. □

If $\mathcal{X} = S_n$ or $\mathcal{X} = \mathcal{M}_{2n}$, then $C_t = O(t^8)$ by Theorem 3.1. Since $N_t/N_{t+1} = \Theta(n)$ and $I_t = n - 2$ in both cases, the premises of Theorem 7.7 and Theorem 7.9 other than the existence of a t-good matrix hold as long as $n \geqslant Kt^{8(t+1)}$, for some absolute constant K. Ellis, Friedgut and Pilpel [EFP11] and Lindzey [Lin18a, Lin18b] showed that a t-good matrix holds for large enough n, and so we conclude the following corollary.

Theorem 7.10. *Let $\mathcal{X} = S_n$ or $\mathcal{X} = \mathcal{M}_{2n}$. If n is large enough (as a function of t), then every t-intersecting family has size at most $(n-t)!$ (if $\mathcal{X} = S_n$) or $(2n-2t-1)!! = (2n-2t-1)(2n-2t-3)\cdots(1)$ (if $\mathcal{X} = \mathcal{M}_{2n}$). Furthermore, this bound is attained only by t-links.*

An identical bound holds on the geometric mean of the sizes of two cross-t-intersecting families. The bound is attained only if both families are the same t-link.

8. Open questions

Robust relations The most interesting open direction, in our view, is proving a robust version of Theorem 3.1. Concretely, Theorem 3.1 describes the structure of Boolean low degree functions: they correspond to shallow decision trees. What can we say about Boolean functions which are "almost" low degree, that is, are close to some (not necessarily Boolean) low-degree function?

To make this question precise, let us say that a function $f\colon \mathcal{X} \to \{0,1\}$ is *ϵ-close to degree d* if there exists a degree d function $g\colon \mathcal{X} \to \mathbb{R}$ such that $\mathbb{E}[(f-g)^2] \leqslant \epsilon$, where the expectation is with respect to the uniform distribution over $\mathcal{X}$. Does this imply that f is close to a function computed by a decision tree of depth $\mathrm{poly}(d)$? More ambitiously, is f close to a *Boolean* degree d function?

In the case of the Boolean cube, this has been answered in the affirmative by Kindler and Safra [KS02, Kin03], who showed that for every *fixed* d, if a Boolean function f is ϵ-close to degree d, then f is $O(\epsilon)$-close to a Boolean degree d function; this was extended to slices (multislices of the form $M(k, n-k)$) by Keller and Klein [KK20]. In ongoing work with Dor Minzer, the result of Kindler and Safra is extended to arbitrary d by making the proof of Theorem 3.1 robust.

Ellis, Filmus and Friedgut [EFF15a, EFF15b, EFF17] considered this question for the case of the symmetric group, showing (among else) that if a Boolean function is close to degree 1 then it is close to a Boolean degree 1 function. They also proved initial results for higher-degree functions, but these apply only to very sparse functions.

We conjecture that if a Boolean function on the symmetric group is close to degree d, then it is close to a function computable by a decision tree of depth $\mathrm{poly}(d)$; and perhaps even close to a Boolean degree d function. Furthermore, we expect the same to hold for the perfect matching scheme.

q-analogs Another interesting open direction is generalizing our framework to *q-analogs*, that is, domains such as the Grassmann scheme (consisting of all k-dimensional subspaces of an n-dimensional vector space over $\mathbb{F}_q$), the bilinear scheme (all $n \times m$ matrices over $\mathbb{F}_q$), and the

general linear group (all $n \times n$ invertible matrices over $\mathbb{F}_q$); all of these come with natural notions of degree.

What separates the domains considered in our paper and q-analogs in the natural symmetries they possess. All domains considered in this paper have a natural action of the symmetric group on them. In contrast, in the case of q-analogs the role of the symmetric group is played by the general linear group. This is related to the fact that the Grassmann scheme is not a simplicial complex.

Filmus and Ihringer [FI19b] have initiated the study of q-analogs from this perspective. They classified the Boolean degree 1 functions on the Grassmann scheme for $q = 2, 3, 4, 5$ (whenever $k, n - k$ are larger than a small constant), and proposed a conjectured classification of Boolean degree 1 functions on the Grassmann and bilinear schemes for all q.

At present we can neither extend Theorem 3.1 to q-analogs nor find nontrivial counterexamples.

More relations The literature contains many other relations between complexity measures. For example, Midrijānis [Mid04] proved that $D(f) \leqslant 2\deg(f)^3$, and Kulkarni and Tal [KT16] used the same method to prove that $R_0(f) = O(R(f)^2 \log R(f)^2)$. So far we have been unable to extend these arguments, which rely on *maxonomials*, to our setting.

Acknowledgements

We thank Nitin Saurabh for many helpful discussions.

References

[Aar08] Scott Aaronson. Quantum certificate complexity. *J. Comput. System Sci.*, 74(3):313–322, 2008. doi:10.1016/j.jcss.2007.06.020.

[ABDK16] Scott Aaronson, Shalev Ben-David, and Robin Kothari. Separations in query complexity using cheat sheets. In *STOC'16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 863–876. ACM, New York, 2016. doi:10.1145/2897518.2897644.

[AK97] Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *European J. Combin.*, 18(2):125–136, 1997. doi:10.1006/eujc.1995.0092.

[AK98] Rudolf Ahlswede and Levon H. Khachatrian. The diametric theorem in Hamming spaces—optimal anticodes. *Adv. in Appl. Math.*, 20(4):429–449, 1998. doi:10.1006/aama.1998.0588.

[AK99] Rudolf Ahlswede and Levon H. Khachatrian. A pushing-pulling method: new proofs of intersection theorems. *Combinatorica*, 19(1):1–15, 1999. doi:10.1007/s004930050042.

[Bac09] Christine Bachoc. Semidefinite programming, harmonic analysis and coding theory. lecture notes, 2009. arXiv:0909.4767.

[BCN89] A. E. Brouwer, A. M. Cohen, and A. Neumaier. *Distance-regular graphs*, volume 18 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1989. doi:10.1007/978-3-642-74341-2.

[BD16] Shalev Ben-David. The structure of promises in quantum speedups. In *11th Conference on the Theory of Quantum Computation, Communication and Cryptography*, volume 61 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 7, 14. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.

[BdW02] Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theoret. Comput. Sci.*, 288(1):21–43, 2002. Complexity and logic (Vienna, 1998). doi:10.1016/S0304-3975(01)00144-X.

[BGH+15] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. *SIAM J. Comput.*, 44(5):1287–1324, 2015. doi:10.1137/130929394.

[CHS20] John Chiarelli, Pooya Hatami, and Michael Saks. An asymptotically tight bound on the number of relevant variables in a bounded degree Boolean function. *Combinatorica*, 40(2):237–244, 2020. doi:10.1007/s00493-019-4136-7.

[CK03] Peter J. Cameron and C. Y. Ku. Intersecting families of permutations. *European J. Combin.*, 24(7):881–890, 2003. doi:10.1016/S0195-6698(03)00078-7.

[CSST08] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Harmonic analysis on finite groups*, volume 108 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008. Representation theory, Gelfand pairs and Markov chains. doi:10.1017/CBO9780511619823.

[Edm65] Jack Edmonds. Maximum matching and a polyhedron with $0,1$-vertices. *J. Res. Nat. Bur. Standards Sect. B*, 69B(1 and 2):125–130, 1965. doi:10.6028/jres.069b.013.

[EFF15a] David Ellis, Yuval Filmus, and Ehud Friedgut. A quasi-stability result for dictatorships in S_n. *Combinatorica*, 35(5):573–618, 2015. doi:10.1007/s00493-014-3027-1.

[EFF15b] David Ellis, Yuval Filmus, and Ehud Friedgut. A stability result for balanced dictatorships in $\mathrm{S_n}$. *Random Structures Algorithms*, 46(3):494–530, 2015. doi:10.1002/rsa.20515.

[EFF17] David Ellis, Yuval Filmus, and Ehud Friedgut. Low-degree Boolean functions on S_n, with an application to isoperimetry. *Forum Math. Sigma*, 5:e23, 46, 2017. doi:10.1017/fms.2017.24.

[EFP11] David Ellis, Ehud Friedgut, and Haran Pilpel. Intersecting families of permutations. *J. Amer. Math. Soc.*, 24(3):649–682, 2011. doi:10.1090/S0894-0347-2011-00690-5.

[EKR61] P. Erdős, Chao Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12(1):313–320, 1961. doi:10.1093/qmath/12.1.313.

[Ell11] David Ellis. Stability for t-intersecting families of permutations. *J. Combin. Theory Ser. A*, 118(1):208–227, 2011. doi:10.1016/j.jcta.2010.04.005.

[FD77] Péter Frankl and Mikhail Deza. On the maximum number of permutations with given maximal or minimal distance. *J. Combinatorial Theory Ser. A*, 22(3):352–360, 1977. doi:10.1016/0097-3165(77)90009-7.

[FI19a] Yuval Filmus and Ferdinand Ihringer. Boolean constant degree functions on the slice are juntas. *Discrete Math.*, 342(12):111614, 7, 2019. doi:10.1016/j.disc.2019.111614.

[FI19b] Yuval Filmus and Ferdinand Ihringer. Boolean degree 1 functions on some classical association schemes. *J. Combin. Theory Ser. A*, 162:241–270, 2019. doi:10.1016/j.jcta.2018.11.006.

[FI20] Nick Fischer and Christian Ikenmeyer. The computational complexity of plethysm coefficients. *computational complexity*, 29, 12 2020. doi:10.1007/s00037-020-00198-4.

[Fil16] Yuval Filmus. An orthogonal basis for functions over a slice of the Boolean hypercube. *Electron. J. Combin.*, 23(1):Paper 1.23, 27, 2016. doi:10.37236/4567.

[Fil17] Yuval Filmus. A comment on intersecting families of permutations. Unpublished Manuscript, 2017. arXiv:1706.10146.

[FOW19] Yuval Filmus, Ryan O'Donnell, and Xinyu Wu. A log-Sobolev inequality for the multislice, with applications. In *10th Innovations in Theoretical Computer Science*, volume 124 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 34, 12. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019.

[Fri08] Ehud Friedgut. On the measure of intersecting families, uniqueness and stability. *Combinatorica*, 28(5):503–528, 2008. doi:10.1007/s00493-008-2318-9.

[FW86] P. Frankl and R. M. Wilson. The Erdős-Ko-Rado theorem for vector spaces. *J. Combin. Theory Ser. A*, 43(2):228–236, 1986. doi:10.1016/0097-3165(86)90063-4.

[GM15] Chris Godsil and Karen Meagher. *Erdős-Ko-Rado theorems: algebraic approaches*, volume 149 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2015. doi:10.1017/CBO9781316414958.

[GMT09] Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. Optimal Sherali-Adams gaps from pairwise independence. In *Approximation, randomization, and combinatorial optimization*, volume 5687 of *Lecture Notes in Comput. Sci.*, pages 125–139. Springer, Berlin, 2009. doi:10.1007/978-3-642-03685-9_10.

[GNS+16] Parikshit Gopalan, Noam Nisan, Rocco A. Servedio, Kunal Talwar, and Avi Wigderson. Smooth Boolean functions are easy: efficient algorithms for low-sensitivity functions. In *ITCS'16—Proceedings of the 2016 ACM Conference on Innovations*

in Theoretical Computer Science, pages 59–70. ACM, New York, 2016. doi:10.1145/2840728.2840738.

[GSS16] Justin Gilmer, Michael Saks, and Srikanth Srinivasan. Composition limits and separating examples for some boolean function complexity measures. *Combinatorica*, 36(3):265–311, 2016. doi:10.1007/s00493-014-3189-x.

[GSW16] Parikshit Gopalan, Rocco A. Servedio, and Avi Wigderson. Degree and sensitivity: tails of two distributions. In *31st Conference on Computational Complexity*, volume 50 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 13, 23. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.

[Hua19] Hao Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Ann. of Math.*, 190(3):949–955, 2019. doi:10.4007/annals.2019.190.3.6.

[Ker99] Adalbert Kerber. *Applied finite group actions*, volume 19 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1999. doi:10.1007/978-3-662-11167-3.

[Kin03] Guy Kindler. *Property Testing, PCP and Juntas*. PhD thesis, Tel Aviv University, 2003.

[KK20] Nathan Keller and Ohad Klein. A structure theorem for almost low-degree functions on the slice. *Israel Journal of Mathematics*, 240:179–221, 2020. doi:10.1007/s11856-020-2062-4.

[KMS18] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in Grassmann graph have near-perfect expansion. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 592–601, 2018.

[KS02] Guy Kindler and Shmuel Safra. Noise-resistant Boolean functions are juntas. (Unpublished Manuscript), 2002.

[KT16] Raghav Kulkarni and Avishay Tal. On fractional block sensitivity. *Chic. J. Theoret. Comput. Sci.*, 2016(8):Art. 8, 16, 2016. doi:10.4086/cjtcs.2016.008.

[Lin18a] Nathan Lindzey. Intersecting families of perfect matchings. 2018. arXiv:1811.06160.

[Lin18b] Nathan Lindzey. *Matchings and representation theory*. PhD thesis, University of Waterloo, 2018.

[LM04] Benoit Larose and Claudia Malvenuto. Stable sets of maximal size in Kneser-type graphs. *European J. Combin.*, 25(5):657–673, 2004. doi:10.1016/j.ejc.2003.10.006.

[Mid04] Gatis Midrijānis. Exact quantum query complexity for total boolean functions, 2004. arXiv:quant-ph/0403168.

[NS94] Noam Nisan and Márió Szegedy. On the degree of Boolean functions as real polynomials. *Comput. Complexity*, 4(4):301–313, 1994. Special issue on circuit complexity (Barbados, 1992). doi:10.1007/BF01263419.

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014. doi:10.1017/CBO9781139814782.

[OW13] Ryan O'Donnell and Karl Wimmer. KKL, Kruskal-Katona, and monotone nets. *SIAM J. Comput.*, 42(6):2375–2399, 2013. doi:10.1137/100787325.

[Sag01] Bruce E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions. doi:10.1007/978-1-4757-6804-6.

[Sal21] Justin Salez. A sharp log-Sobolev inequality for the multislice. *Ann. H. Lebesgue*, 4:1143–1161, 2021. doi:10.5802/ahl.99.

[Sch03] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. A*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Paths, flows, matchings, Chapters 1–38.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. doi:10.1017/CBO9780511609589.

[Tal13] Avishay Tal. Properties and applications of Boolean function composition. In *ITCS'13—Proceedings of the 2013 ACM Conference on Innovations in Theoretical Computer Science*, pages 441–454. ACM, New York, 2013.

[UF96] Alasdair Urquhart and Xudong Fu. Simplified lower bounds for propositional proofs. *Notre Dame J. Formal Logic*, 37(4):523–544, 1996. doi:10.1305/ndjfl/1040046140.

[Wel22] Jake Wellens. Relationships between the number of inputs and other complexity measures of Boolean functions. *Discrete Anal.*, pages Paper No. 19, 21, 2022.

[Wil84] Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4(2-3):247–257, 1984. doi:10.1007/BF02579226.

[Wim14] Karl Wimmer. Low influence functions over slices of the Boolean hypercube depend on few coordinates. In *IEEE 29th Conference on Computational Complexity—CCC 2014*, pages 120–131. IEEE Computer Soc., Los Alamitos, CA, 2014. doi:10.1109/CCC.2014.20.

A. Spatial degree coincides with spectral degree

In this appendix we prove Theorems 5.5, 5.8 and 5.10.

A.1. Symmetric group

In this section we prove Theorem 5.5 using the explicit description of the isotypic decomposition in Section 5.3.

Recall that the space of all real-valued functions on the symmetric group decomposes orthogonally as

$$\mathbb{R}[S_n] = \bigoplus_{\lambda \vdash n} V^\lambda,$$

where V^λ is spanned by the pseudo-characters $\chi_{A,B}$, where A, B are a pair of tabloids of shape λ. Moreover,

$$\chi_{A,B} = \sum_\sigma (-1)^\sigma e_{A,B^\sigma},$$

where σ goes over all ways to permute the columns of B, $(-1)^\sigma$ is the sign of σ, and $e_{A,B^\sigma}(\pi)$ is the indicator of the condition that π maps the i'th row of A to the i'th row of B^σ (in any order).

Theorem 5.5 is equivalent to the following claim: a function $f\colon S_n \to \mathbb{R}$ has degree at most d iff it is supported on the isotypic components V^λ with $\lambda_1 \geqslant n-d$.

Suppose first that f is supported on the isotypic components V^λ with $\lambda_1 \geqslant n-d$. It is thus a linear combination of functions $e_{A,B}$, where A, B are tabloids of shape λ, where $\lambda_1 \geqslant n-d$. It suffices to show that $e_{A,B}$ has degree at most $n-\lambda_1$. We can write

$$e_{A,B} = \sum_\tau E_{A,B^\tau},$$

where τ goes over all ways to permute the rows of B other than the first one, and $E_{A,B}(\pi)$ is the indicator of the condition that π maps each element on each row of A other than the first to the corresponding element of B. The function $E_{A,B}$ is a monomial of degree $n-\lambda_1$, and so it has degree at most $n-\lambda_1$. Consequently, so does $e_{A,B}$.

Suppose next that f has degree d. We would like to show that the projection of f to V^λ vanishes if $\lambda_1 < n-d$. Since the isotypic decomposition is orthogonal and V^λ is spanned by the functions $\chi_{A,B}$ where A, B have shape λ, it suffices to show that f is orthogonal to all such functions $\chi_{A,B}$. Since f is a linear combination of monomials of degree at most d, it suffices to show that a monomial $\mathfrak{m} = x_{i_1 j_1} \cdots x_{i_e j_e}$ is orthogonal to all functions $\chi_{A,B}$, where A, B have shape λ satisfying $\lambda_1 < n-e$.

We claim that there is some column of B containing at least two elements different from $j_1, \ldots, j_e$. Otherwise, each column of B contains at most one element different from $j_1, \ldots, j_e$, and so B contains at most λ_1 elements different from $j_1, \ldots, j_e$ (since B has λ_1 columns); but this contradicts $\lambda_1 < n-e$. Choose two arbitrary elements different from $j_1, \ldots, j_e$ appearing in the same column of B, and let σ_0 be the permutation switching them.

Let σ be a permutation of the columns of B. By construction, $\sigma\sigma_0$ is another such permutation, and $e_{A,B^\sigma}(\pi) = e_{A,B^{\sigma\sigma_0}}(\pi^{\sigma_0})$, where π^{σ_0} is obtained from π by applying σ_0 to its image. Since σ_0 doesn't involve $j_1, \ldots, j_e$, $\mathfrak{m}(\pi) = \mathfrak{m}(\pi^{\sigma_0})$. It follows that

$$\mathfrak{m}(\pi) e_{A,B^\sigma}(\pi) = \mathfrak{m}(\pi^{\sigma_0}) e_{A,B^{\sigma\sigma_0}}(\pi^{\sigma_0}),$$

and so

$$\langle \mathfrak{m}, e_{A,B^\sigma} \rangle = \langle \mathfrak{m}, e_{A,B^{\sigma\sigma_0}} \rangle.$$

Since $(-1)^\sigma = (-1)^{\sigma\sigma_0}$, we conclude that $\langle \mathfrak{m}, \chi_{A,B} \rangle = 0$.

A.2. Perfect matching scheme

In this section we prove Theorem 5.8 using the explicit description of the isotypic decomposition in Section 5.4.

Recall that the space of all real-valued functions on the perfect matching scheme decomposes orthogonally as

$$\mathbb{R}[\mathcal{M}_{2n}] = \bigoplus_{\lambda \vdash n} V^{2\lambda},$$

where $V^{2\lambda}$ is spanned by the pseudo-characters χ_A, where A is a tabloid of shape 2λ. Moreover,

$$\chi_A = \sum_{\sigma} (-1)^{\sigma} e_{A^\sigma},$$

where σ goes over all ways to permute the columns of A, $(-1)^\sigma$ is the sign of σ, and $e_{A^\sigma}(m)$ is the indicator of the condition that all edges in m match elements in the same row of A^σ.

Theorem 5.8 is equivalent to the following claim: a function $f\colon \mathcal{M}_{2n} \to \mathbb{R}$ has degree at most d iff it is supported on the isotypic components $V^{2\lambda}$ with $\lambda_1 \geqslant n-d$.

Suppose first that f is supported on the isotypic components $V^{2\lambda}$ with $\lambda_1 \geqslant n-d$. It is thus a linear combination of functions e_A, where A is a tabloid of shape 2λ, where $\lambda_1 \geqslant n-d$. It suffices to show that e_A has degree at most $n-\lambda_1$. We can write

$$e_A = c_\lambda \sum_{\tau} E_{A^\tau},$$

where c_λ is a constant depending only on λ, τ goes over all ways to permute the rows of A other than the first one, and $E_{A^\tau}(m)$ is the indicator of the condition that m contains the edges specified by the rows of A^τ other than the first, where a row $a_1, \ldots, a_{2\ell}$ specifies the edges $\{a_1, a_2\}, \{a_3, a_4\}, \ldots, \{a_{2\ell-1}, a_{2\ell}\}$. The value of c_λ is thus

$$c_\lambda = \prod_{i=2}^{r} 2^{\lambda_i} \lambda_i!,$$

where r is the number of parts in λ.

The function E_A is a monomial of degree $n-\lambda_1$, and so it has degree at most $n-\lambda_1$. Consequently, so does e_A.

Suppose next that f has degree d. We would like to show that the projection of f to $V^{2\lambda}$ vanishes if $\lambda_1 < n-d$. Since the isotypic decomposition is orthogonal and $V^{2\lambda}$ is spanned by the functions χ_A where A has shape 2λ, it suffices to show that f is orthogonal to all such functions χ_A. Since f is a linear combination of monomials of degree at most d, it suffices to show that a monomial $\mathfrak{m} = x_{i_1 j_1} \cdots x_{i_e j_e}$ is orthogonal to all functions χ_A, where A has shape 2λ satisfying $\lambda_1 < n-e$.

We claim that there is some column of A containing at least two elements different from $i_1, \ldots, i_e, j_1, \ldots, j_e$. Otherwise, each column of A contains at most one element different from $i_1, \ldots, i_e, j_1, \ldots, j_e$, and so A contains at most $2\lambda_1$ elements different from $i_1, \ldots, i_e, j_1, \ldots, j_e$ (since A has $2\lambda_1$ columns); but this contradicts $\lambda_1 < n-e$. Choose

two arbitrary elements different from $i_1, \dots, i_e, j_1, \dots, j_e$ appearing in the same column of A, and let σ_0 be the permutation switching them.

Let σ be a permutation of the columns of A. By construction, $\sigma\sigma_0$ is another such permutation, and $e_{A^\sigma}(m) = e_{A^{\sigma\sigma_0}}(m^{\sigma_0})$, where m^{σ_0} is obtained from m by applying σ_0 to the vertices. Since σ_0 doesn't involve $i_1, \dots, i_e, j_1, \dots, j_e$, $\mathfrak{m}(m) = \mathfrak{m}(m^{\sigma_0})$. It follows that

$$\mathfrak{m}(m)e_{A^\sigma}(m) = \mathfrak{m}(m^{\sigma_0})e_{A^{\sigma\sigma_0}}(m^{\sigma_0}),$$

and so

$$\langle \mathfrak{m}, e_{A^\sigma} \rangle = \langle \mathfrak{m}, e_{A^{\sigma\sigma_0}} \rangle.$$

Since $(-1)^\sigma = (-1)^{\sigma\sigma_0}$, we conclude that $\langle \mathfrak{m}, \chi_A \rangle = 0$.

A.3. Multislices

In this section we prove Theorem 5.10 using the explicit description of the isotypic decomposition in Section 5.5, slightly modifying the proof in Appendix A.1.

The isotypic decomposition for the symmetric group generalizes to the multislice:

$$\mathbb{R}[M^\mu] = \bigoplus_{\mu \trianglelefteq \lambda} V^\lambda.$$

The isotypic component V^λ is spanned by the pseudo-characters $\chi_{A,B}$, where A, B are tabloids of shape λ, A contains the numbers $1, \dots, n$, and B has content μ. Moreover,

$$\chi_{A,B} = \sum_\sigma (-1)^\sigma e_{A^\sigma, B},$$

where $e_{A,B}$ is defined as in the case of the symmetric group, and σ goes over all ways to permute the columns of A.

Theorem 5.10 is equivalent to the claim that a function $f\colon M^\mu \to \mathbb{R}$ has degree at most d iff it is supported on the isotypic components V^λ with $\lambda_1 \geqslant n - d$.

In one direction, it suffices to show that the function $e_{A,B}$ has degree at most $n - \lambda_1$, where λ is the common shape of A, B. We can write

$$e_{A,B} = \sum_\tau E_{A,B^\tau},$$

where τ goes over all ways to permute the rows of B (going over each possible tabloid exactly once), and $E_{A,B^\tau}(\pi)$ indicates that the colors of the elements of A beyond the first line are the corresponding elements of B^τ. It is easy to check that E_{A,B^τ} has degree at most $n - \lambda_1$, hence so does $e_{A,B}$.

In the other direction, it suffices to show that if $\mathfrak{m}$ is a degree e monomial and $\chi_{A,B}$ is a pseudo-character corresponding to tabloids whose shape λ satisfies $\lambda_1 < n - e$, then $\langle \mathfrak{m}, \chi_{A,B} \rangle = 0$. The proof is the same as the one in Appendix A.1.

B. Application of pseudo-characters

In this section, we illustrate the utility of pseudo-characters by proving the following lemma. Before stating the lemma, let us recall that a t-star is a family of the form

$$\{\pi \in S_n : \pi(i_1) = j_1, \dots, \pi(i_t) = j_t\}.$$

(This coincides with the definition of *t-link* in Section 7.)

Lemma 5.7. *Suppose that $\mathcal{F}$ is a subset of S_n contained in a t-star, and its characteristic function has degree at most t. Then either $\mathcal{F}$ is empty, or it is equal to the t-star.*

Proof. For concreteness, suppose the t-star consists of all permutations containing $(1,1), \dots, (t,t)$. Let f be the characteristic function of $\mathcal{F}$. Denote by f_s the restriction of f to the s-star $\mathcal{S}_s$ consisting of all permutations containing $(1,1), \dots, (s,s)$.

We claim that if $f_1 \neq 0$ then $\deg(f_1) < \deg(f)$. Indeed, let $\deg(f_1) = d$. Our work in Section 5.3 shows that f_1 has nonzero correlation with some $\chi_{A,B}$ over $\mathcal{S}_1$, where A, B is a pair of tabloids of shape λ, with $\lambda_1 = n - 1 - d$. Form new tabloids A', B' by adding one more box filled with 1 at the very bottom. This is a pair of tabloids of shape $\mu = \lambda, 1$, which satisfies $\mu_1 = n - 1 - d = n - (d+1)$. Since f vanishes outside of $\mathcal{S}_1$, we have

$$\sum_{x \in S_n} f(x)\chi_{A',B'}(x) = \sum_{\sigma} (-1)^{\sigma} \sum_{x \in \mathcal{S}_1} f_1(x) e_{A',(B')^{\sigma}}(x),$$

where σ varies over all permutations of the columns of B'. If σ moves the bottom box of B' then $e_{A',(B')^{\sigma}}(x) = 0$ for all $x \in \mathcal{S}_1$, so we need not consider such σ. All remaining σ permute the columns of B and satisfy $e_{A',(B')^{\sigma}} = e_{A,B}$ over $\mathcal{S}_1$, and so

$$\sum_{x \in S_n} f(x)\chi_{A',B'}(x) = \sum_{x \in \mathcal{S}_1} f(x)\chi_{A,B}(x) \neq 0.$$

This shows that $\deg(f) \geqslant d + 1$.

The same argument shows that $\deg(f_t) < \cdots < \deg(f)$ unless $f_t = 0$, and so f_t is constant. Since f_t is Boolean, either $f_t = 0$ or $f_t = 1$. □

We comment that the result can be proven, for large enough n (as a function of t), using the methods outlined in Section 7.3. Indeed, suppose that a t-good matrix A exists for S_n (Ellis, Friedgut and Pilpel [EFP11] showed that this holds for large enough n, as a function of t). Let $\mathcal{F}$ be a subset of a t-star whose characteristic function f has degree at most t. Since $\mathcal{F}$ is a subset of a t-star, it is automatically t-intersecting. The proof of Proposition 7.6 shows that

$$0 = \langle f, Af \rangle = \mu^2 - \omega(\mu - \mu^2),$$

where $\omega = (n-t)!/(n! - (n-t)!)$. Thus either $\mu = 0$ (and so $\mathcal{F} = \varnothing$) or $\mu = \omega/(1+\omega) = (n-t)!/n!$ (and so $\mathcal{F}$ is a t-star).

C. More on the spectral technique

Section 7.3 outlines a spectral approach for bounding the size of t-intersecting families and for characterizing t-intersecting families of maximum size. The input to the spectral approach is a matrix satisfying the properties given in the following definition.

Definition 7.5. A real $\mathcal{X} \times \mathcal{X}$ matrix A is *t-good* if the following properties hold:

1. A is symmetric.
2. If $|x \cap y| \geqslant t$ then $A(x, y) = 0$.
3. $A\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the constant 1 vector.
4. If $\deg(f) \leqslant t$ and $\mathbb{E}[f] = 0$, where $\mathbb{E}[f] = \sum_{x \in \mathcal{X}} f(x)/|\mathcal{X}|$, then $Af = -\omega f$, where
$$\omega = \frac{N_t}{|\mathcal{X}| - N_t}.$$
5. If $Af = \lambda f$ and $\deg(f) > t$ then $|\lambda| < \omega$.

(In this definition, N_t is the maximum size of a *t-link*, as defined in Section 7.1.)

In this brief section, we explain how the existence of a t-good matrix implies an upper bound on the size of t-intersecting and cross-t-intersecting families. The proofs below are standard and appear in many papers, and are reproduced here for the sake of completeness.

We start by showing that a t-good matrix implies a bound on the size of t-intersecting families. The argument is known as the *weighted Hoffman bound*, which is closely related to the Lovász θ function.

Proposition 7.6. *Suppose that there exists a t-good matrix for $\mathcal{X}$. Then a t-intersecting family contains at most N_t points, and furthermore, the characteristic function of a t-intersecting family of size N_t has degree at most t.*

Proof. Let A be a t-good matrix, let $\mathcal{F}$ be a t-intersecting family, and let f be its characteristic vector. Define $\mu = \mathbb{E}[f] = |\mathcal{F}|/|\mathcal{X}|$.

Let V_0 consist of all constant functions, and let V_1 consist of all functions of degree at most t orthogonal to V_0. Both of these are eigenspaces of A. Denote by $V_2, \ldots, V_m$ its remaining eigenspaces. Since A is symmetric, these eigenspaces are orthogonal. Denote the corresponding eigenvalues by $\lambda_0, \ldots, \lambda_m$. Thus $\lambda_0 = 1$, $\lambda_1 = -\omega$, and $|\lambda_i| < \omega$ for $i \geqslant 2$.

We will use the following notations, for functions g, h on $\mathcal{X}$: $\langle g, h\rangle = \mathbb{E}[gh]$, and $\|g\|^2 = \langle g, g\rangle$.

Let $f_0, \ldots, f_m$ be the projections of f into $V_0, \ldots, V_m$, respectively. Since
$$\langle \mu\mathbf{1}, f - \mu\mathbf{1}\rangle = \mu\mathbb{E}[f] - \mu^2\mathbb{E}[\mathbf{1}] = 0,$$
we see that $f_0 = \mu\mathbf{1}$. Also, orthogonality guarantees that
$$\|f_0\|^2 + \cdots + \|f_m\|^2 = \|f\|^2 = \mathbb{E}[f^2] = \mathbb{E}[f] = \mu.$$

Since $\mathcal{F}$ is t-intersecting,

$$\langle f, Af\rangle = \frac{1}{|\mathcal{X}|}\sum_{x,y\in\mathcal{X}} A(x,y)f(x)f(y) = 0.$$

On the other hand, since $Af_i = \lambda_i f_i$ and the eigenspaces are orthogonal,

$$0 = \langle f, Af\rangle = \sum_{i=0}^{m}\langle f_i, \lambda_i f_i\rangle = \sum_{i=0}^{m}\lambda_i\|f_i\|^2 \geqslant \mu^2 - \omega\sum_{i=1}^{m}\|f_i\|^2 = \mu^2 - \omega(\mu-\mu^2),$$

with equality if and only if $f_i = 0$ for all $i \geqslant 2$, that is, $\deg f \leqslant t$.

We thus obtain the inequality $(1+\omega)\mu \leqslant \omega$, hence

$$\mu \leqslant \frac{\omega}{1+\omega} = \frac{N_t}{|\mathcal{X}|}.$$

In other words, $|\mathcal{F}| \leqslant N_t$. Furthermore, if equality holds then $\deg f \leqslant t$. □

The argument extends to the cross-t-intersecting case, by throwing in several applications of the Cauchy–Schwartz inequality.

Proposition 7.8. *Suppose that there exists a t-good matrix for $\mathcal{X}$. If $\mathcal{F}, \mathcal{G}$ are two cross-t-intersecting families then $\sqrt{|\mathcal{F}|\cdot|\mathcal{G}|} \leqslant N_t$. Furthermore, if equality holds then $\mathcal{F} = \mathcal{G}$, and the common characteristic function has degree at most t.*

Proof. We will use the same notation as in the proof of Proposition 7.6. Let $\mu = \mathbb{E}[f]$ and $\nu = \mathbb{E}[g]$. Since $\mathcal{F}$ and $\mathcal{G}$ are cross-t-intersecting, we have

$$0 = \langle f, Ag\rangle = \sum_{i=0}^{m}\lambda_i\langle f_i, g_i\rangle = \mu\nu + \sum_{i=1}^{m}\lambda_i\langle f_i, g_i\rangle.$$

The Cauchy–Schwartz inequality shows that $|\langle f_i, g_i\rangle| \leqslant \|f_i\|\cdot\|g_i\|$, and so

$$0 = \mu\nu + \sum_{i=1}^{m}\lambda_i\langle f_i, g_i\rangle \geqslant \mu\nu - \omega\sum_{i=1}^{m}\|f_i\|\cdot\|g_i\|. \tag{$*$}$$

Applying the Cauchy–Schwartz inequality to the sum gives

$$0 \geqslant \mu\nu - \omega\sqrt{\sum_{i=1}^{m}\|f_i\|^2}\sqrt{\sum_{i=1}^{m}\|g_i\|^2} = \mu\nu - \omega\sqrt{\mu-\mu^2}\sqrt{\nu-\nu^2} \geqslant \mu\nu - \omega(\sqrt{\mu\nu} - \mu\nu),$$

using the following consequence of the AM-GM inequality (applied twice):

$$\sqrt{(\mu-\mu^2)(\nu-\nu^2)} = \sqrt{\mu\nu}\sqrt{(1-\mu)(1-\nu)} \leqslant \sqrt{\mu\nu}\left(1 - \frac{\mu+\nu}{2}\right) \leqslant \sqrt{\mu\nu}(1-\sqrt{\mu\nu}).$$

As in the proof of Proposition 7.6, this shows that $\sqrt{\mu\nu} \leqslant N_t/|\mathcal{X}|$, and so $\sqrt{|\mathcal{F}| \cdot |\mathcal{G}|} \leqslant N_t$.

If $\sqrt{|\mathcal{F}| \cdot |\mathcal{G}|} = N_t$ then the AM-GM inequality is tight, implying that $\mu = \nu$. Furthermore, the inequalities in $(*)$ are tight, implying that $\deg(f), \deg(g) \leqslant t$ (since $|\lambda_i| < \omega$ for $i \geqslant 2$) and that f_1, g_1 are parallel (since the Cauchy–Schwartz inequality $\langle f_1, g_1 \rangle \leqslant \|f_1\| \cdot \|g_1\|$ is tight). Since $f_1 = f - \mu$ and $g_1 = g - \nu = g - \mu$ are both $\{-\mu, 1-\mu\}$-valued, necessarily $f_1 = g_1$, and so $f = g$. (Note that $f_1 = -g_1$ is impossible even when $\mu = 1/2$, since then $\langle f_1, g_1 \rangle = -\|f_1\| \cdot \|g_1\|$.) □