

Agreement tests on graphs and hypergraphs ^{*}

Irit Dinur[†]

Yuval Filmus[‡]

Prahladh Harsha[§]

December 11, 2020

Abstract

Agreement tests are a generalization of low degree tests that capture a local-to-global phenomenon, which forms the combinatorial backbone of most PCP constructions. In an agreement test, a function is given by an ensemble of local restrictions. The agreement test checks that the restrictions agree when they overlap, and the main question is whether average agreement of the local pieces implies that there exists a global function that agrees with most local restrictions.

There are very few structures that support agreement tests, essentially either coming from algebraic low degree tests or from direct product tests (and recently also from high-dimensional expanders). In this work, we prove a new agreement theorem which extends direct product tests to higher dimensions, analogous to how low degree tests extend linearity testing. As a corollary of our main theorem, it follows that an ensemble of small graphs on overlapping sets of vertices can be glued together to one global graph assuming they agree with each other on average.

We prove the agreement theorem by (re)proving the agreement theorem for dimension 1, and then generalizing it to higher dimensions (with the dimension 1 case being the direct product test, and dimension 2 being the graph case). A key technical step in our proof is the reverse union bound, which allows us to treat dependent events as if they are disjoint, and may be of independent interest. An added benefit of the reverse union bound is that it can be used to show that the “majority decoded” function also serves as a global function that explains the local consistency of the agreement theorem, a fact that was not known even in the direct product setting (dimension 1) prior to our work.

Beyond the motivation to understand fundamental local-to-global structures, our main theorem allows us to lift structure theorems from $\mu_{1/2}$ to μ_p . As a simple demonstration of this paradigm, we show how the low degree testing result of Alon *et al.* [*IEEE Trans. Inform. Theory*, 2005] and Bhattacharyya *et al.* [*Proc. 51st FOCS*, 2010], originally proved for $\mu_{1/2}$, can be extended to the biased hypercube μ_p , even for very small sub-constant p .

^{*}A preliminary version of this paper and the parallel work [DFH17] appeared in *Proc. 30th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)* [DFH19].

[†]Weizmann Institute of Science, ISRAEL. email: irit.dinur@weizmann.ac.il. Research supported by ERC-CoG grant 772839.

[‡]Technion Israel Institute of Technology, ISRAEL. email: yuvalfi@cs.technion.ac.il. Taub Fellow — supported by the Taub Foundations. The research was funded by ISF grant 1337/16.

[§]Tata Institute of Fundamental Research, INDIA. email: prahladh@tifr.res.in. Research supported by the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.01-0500 and in part by the UGC-ISF grant and Swarnajayanti fellowship. Part of this work was done when the author was visiting the Weizmann Institute of Science.

1 Introduction

Agreement tests are a type of PCP tests and capture a fundamental local-to-global phenomenon. In this paper, we study an agreement testing question that is a new extension of direct product testing to higher dimensions.

It is a basic fact of computation that any global computation can be broken down into a sequence of local steps. The PCP theorem [AS98, ALMSS98] says that moreover, this can be done in a robust fashion, so that as long as *most* steps are correct, the entire computation checks out. At the heart of this is a *local-to-global* argument that allows deducing a global property from local pieces that fit together only approximately.

A key example is the line vs. line [GLRSW91, RS96] low degree test in the proof of the PCP theorem. In the PCP construction, a function on a large vector space is replaced by an ensemble of (supposed) restrictions to all possible affine lines. These restrictions are supplied by a prover and are not a priori guaranteed to agree with any single global function. This is taken care of by the “low degree test”, which checks that restrictions on intersecting lines agree with each other, i.e. they give the same value to the point of intersection. The crux of the argument is the fact that the local agreement checks *imply* agreement with a single global function. Thus, the low degree test captures a local-to-global phenomenon.

In what other scenarios does such a local-to-global theorem hold? This question was first asked by Goldreich and Safra [GS00], who studied a combinatorial analog of the low degree test. Let us describe the basic framework of agreement testing in which we will study this question. In agreement testing, a global function is given by an ensemble of local functions. There are two key aspects of agreement testing scenarios:

- **Combinatorial structure:** for a given ground set V of size n , the combinatorial structure is a collection H of subsets $S \subset V$ such that for each $S \in H$ we get a local function. For example, if V is the points of a vector space then H can be the collection of affine lines.
- **Allowed functions:** for each subset $S \in H$, we can specify a space \mathcal{F}_S of functions on S that are allowed. The input to the agreement test is an ensemble of functions $\{f_S\}$ such that for every $S \in H$, $f_S \in \mathcal{F}_S$. For example, in the line vs. line low degree test we only allow local functions on each line that have low degree.

Given the ensemble $\{f_S\}$, the intention is that f_S is the restriction to S of a global function $F: V \rightarrow \Sigma$. Indeed, a local ensemble is called *global* if there is a global function $F: V \rightarrow \Sigma$ such that

$$\forall S \in H, \quad f_S = F|_S.$$

An *agreement check* for a pair of subsets S_1, S_2 is specified by a triple of sets (T, S_1, S_2) such that $T \subset S_1 \cap S_2$ and is denoted by $f_{S_1} \sim_T f_{S_2}$. It checks whether the local functions f_{S_1} and f_{S_2} agree on the intersection T . Formally,

$$f_{S_1} \sim_T f_{S_2} \iff \forall x \in T, \quad f_{S_1}(x) = f_{S_2}(x).$$

A local ensemble which is global passes all agreement checks. The converse is also true: a local ensemble that passes *all* agreement checks must be global.

An *agreement test* is specified by giving a distribution \mathcal{D} over triples of subsets (T, S_1, S_2) such tha

$T \subset S_1 \cap S_2$. We define the agreement of a local ensemble to be the probability of agreement:

$$\text{agree}_{\mathcal{D}}(\{f_S\}) := \Pr_{(T, S_1, S_2) \sim \mathcal{D}} [f_{S_1}|_T = f_{S_2}|_T].$$

An agreement theorem shows that if $\{f_S\}_S$ is a local ensemble with $\text{agree}_{\mathcal{D}}(\{f_S\}) > 1 - \varepsilon$ then it is close to being global.

Example: direct product tests Perhaps the simplest agreement test to describe is the direct product test, in which H contains all possible k -element subsets of V . For each S , we let \mathcal{F}_S be all possible functions on S , that is $\mathcal{F}_S = \{f: S \rightarrow \Sigma\}$. The input to the test is an ensemble of local functions $\{f_S\}$, and a natural testing distribution is to choose a random t -element set T and two k -element subsets S_1, S_2 so that they intersect on $S_1 \cap S_2 = T$, a distribution we denote $\mu_n(k, t)$. Suppose $\text{agree}(\{f_S\}) \geq 1 - \varepsilon$. Is there a global function $F: V \rightarrow \Sigma$ such that $F|_S = f_S$ for most subsets S , where most is measured with respect to the uniform distribution $\nu_n(k)$ over k -elements sets? This is the content of the direct product testing theorem of Dinur and Steurer [DS14].

Theorem 1.1 (agreement theorem, dimension 1). *For all positive integers n, k, t satisfying $1 \leq t < k \leq n$ and alphabet Σ , the following holds. Let $\{f_S: S \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \mu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon.$$

Then there exists a global function $G: [n] \rightarrow \Sigma$ that satisfies

$$\Pr_{S \sim \nu_n(k)} [f_S \neq G|_S] = O_\alpha(\varepsilon),$$

where $\alpha = t/k$.

The qualitatively strong aspect of this theorem is that in the conclusion, the global function agrees *perfectly* with $1 - O(\varepsilon)$ of the local functions. Achieving a weaker result where perfect agreement $f_S = F|_S$ is replaced by approximate one $f_S \approx F|_S$ would be significantly easier but also less useful. Quantitatively, this is manifested in that the fraction of local functions that end up disagreeing with the global function F is at most $O(\varepsilon)$ and is *independent of n and k* . It would be significantly easier to prove a weaker result where the closeness is $O(k\varepsilon)$ (via a union bound on the event that $F(i) = f_S(i)$). This theorem is proven by Dinur and Steurer [DS14] by imitating the proof of the parallel repetition theorem [Raz98]. This theorem is also used as a component in the recent work on agreement testing on high dimensional expanders [DK17].

Our Results

In order to motivate our extension of Theorem 1.1, let us describe it in a slightly different form. The global function F can be viewed as specifying the coefficients of a linear form $\sum_{i=1}^n F(i)x_i$ over variables x_1, \dots, x_n . For each S , the local function f_S specifies the partial linear form only over the variables in S . This f_S is supposed to be equal to F on the part of the domain where $x_i = 0$ for all $i \notin S$. Given an ensemble $\{f_S\}$ whose elements are promised to agree with each other on average, the agreement theorem allows us to conclude the existence of a global linear function that agrees with most of the local pieces.

This description naturally leads to the question of extending this to higher degree polynomials. Now, the global function is a degree d polynomial with coefficients in Σ , namely $F = \sum_T F(T)x_T$, where we sum over subsets $T \subset [n]$, $|T| \leq d$. The local functions f_S will be polynomials of degree $\leq d$, supposedly obtained by zeroing out all variables outside S . Two local functions f_{S_1}, f_{S_2} are said to agree, denoted $f_{S_1} \sim f_{S_2}$, if every monomial that is induced by $S_1 \cap S_2$ has the same coefficient in both polynomials. Our new agreement theorem says that in this setting as well, local agreement implies global agreement:

Theorem 1.2 (agreement theorem for high dimensions). *For all positive integers n, k, t, d satisfying $d \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: \binom{S}{d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \mu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon,$$

then the majority decoded function $G: \binom{[n]}{d} \rightarrow \Sigma$ satisfies

$$\Pr_{S \sim \nu_n(k)} [f_S \neq G|_S] \leq O_{d, \alpha}(\varepsilon),$$

where $\alpha = t/k$ and the majority-decoded function G is the one given by “popular vote”, namely for each $A \in \binom{[n]}{d}$ set $G(A)$ to be the most frequently occurring value among $\{f_S(A) \mid S \supset A\}$ (breaking ties arbitrarily).

For $d = 1$, this theorem is precisely [Theorem 1.1](#) (but for the “furthermore” clause). The additional “furthermore” clause strengthens our theorem by naming the popular vote function as a candidate global function that explains most of the local functions. This addendum strengthens also [Theorem 1.1](#).

Let us spell out how this theorem fits into the framework described above. The ground set is $V = \binom{[n]}{\leq d}$, and the collection of subsets H is the collection of all induced hypergraphs on k elements. In particular, if we focus on $\Sigma = \{0, 1\}$, we can view the local function of a subset $S \subset [n]$, $|S| = k$, as specifying a hypergraph on the vertices of S with hyperedges of size up to d . The theorem says that if these small hypergraphs agree with each other most of the time, then there is a global hypergraph that they nearly all agree with.

For the special case of $d = 2$ and $\Sigma = \{0, 1\}$, we get an interesting statement about combining small pieces of a graph into a global one.

Corollary 1.3 (agreement test for graphs). *For all positive integers n, k, t satisfying $2 \leq t < k \leq n$ the following holds.*

Let $\{G_S\}$ be an ensemble of graphs, where S is a k element subset of $[n]$ and G_S is a graph on vertex set S . Suppose that

$$\Pr_{\substack{S_1, S_2 \in \binom{[n]}{k} \\ |S_1 \cap S_2| = t}} [G_{S_1}|_{S_1 \cap S_2} = G_{S_2}|_{S_1 \cap S_2}] \geq 1 - \varepsilon.$$

Then there exists a single global graph $G = ([n], E)$ satisfying $\Pr_{S \in \binom{[n]}{k}} [G_S = G|_S] = 1 - O_\alpha(\varepsilon)$, where $\alpha = t/k$.

Here too we emphasize that the strength of the statement is in that the conclusion talks about *exact* agreement between the global graph and the local graphs, i.e. $G_S = G|_S$ and not $G_S \approx G|_S$, for a fraction of $1 - O(\varepsilon)$ of the sets S . It is also important that there is no dependence in the $O(\cdot)$ on either n or k . A similar agreement testing statement can be made for hypergraphs of any uniformity $\leq d$.

A technical component in our proof which we wish to highlight is the *reverse union bound*, which may be of independent interest.

Lemma 1.4 (reverse union bound). *For each $c > 0$ and integer $d \geq 1$ there exists a constant $C'(d, c) = 2^{O_c(d^3)}$ such that the following holds for all $d \leq k \leq n$. Let $Y \subseteq \binom{[n]}{k}$ and $B \subseteq \binom{[n]}{d}$ satisfy*

$$\Pr_{S \sim \nu_n(k)} [S \in Y \mid S \supseteq I] \geq c \text{ for all } I \in B.$$

Then

$$\Pr_{S \sim \nu_n(k)} [S \text{ hits } B] \leq C'(d, c) \cdot \Pr_{S \sim \nu_n(k)} [S \in Y],$$

where S hits B if $S \supseteq I$ for some $I \in B$.

We illustrate an application of this lemma later on in the introduction.

Application to Structure Theorems

The agreement theorems we prove allow for a new paradigm for studying the structure of Boolean functions on the *biased* Boolean hypercube, i.e. when the measure is μ_p and p is potentially very small, e.g. $p = o_n(1)$, $p \rightarrow 0$ as $n \rightarrow \infty$. This paradigm is based on the following simple fact: the p -biased hypercube is expressible as a convex combination of many small-dimensional copies of the uniform hypercube. To uncover structure for μ_p , we invoke known structure theorems for $\mu_{1/2}$, obtaining a structured approximation for each copy separately. We then sew these approximations together using the agreement theorem. This strategy allows us to lift structure theorems from $\mu_{1/2}$ to μ_p .

GF(2)-low degree testing in the biased hypercube As an illustration of this paradigm, we lift the low degree test [AKKLR05, BKSSZ10] to the biased setting using a straightforward application of the agreement theorem. This is similar to the way in which the analysis of the “uniform BLR” test was lifted from the middle slice to an arbitrary slice by David *et al.* [DDGKS17].

Alon *et al.* [AKKLR05] studied a 2^{d+1} -query test T_d for being of low degree. Bhattacharyya *et al.* [BKSSZ10] gave an optimal analysis of this test, showing that $\delta_d(f) = O_d(\text{rej}_d(f))$, where $\delta_d(f)$ refers to the distance of f to the closest degree d function under the $\mu_{1/2}$ measure (in other words, $\delta_d(f) = \min_{\text{bdeg}(g) \leq d} \Pr_{\mu_{1/2}}[f \neq g]$), and $\text{rej}_d(f)$ is the rejection probability of the test T_d on input function f . We would like to extend the test T_d to the p -biased setting, wherein we measure closeness of f to low degree functions with respect to the μ_p measure instead of $\mu_{1/2}$ measure. More precisely, $\delta_d^{(p)}(f) := \min_{\text{bdeg}(g) \leq d} \Pr_{\mu_p}[f \neq g]$. To this end, we study the following test $T_{p,d}$.

Algorithm 1: p -BIASED LOW DEGREE TEST $T_{p,d}$

Input: $f: \{0,1\}^n \rightarrow \{0,1\}$ (provided as an oracle)

- 1 Pick $S \subseteq [n]$ according to the distribution μ_{2p} .
- 2 If $|S| \leq d$, then accept.
- 3 Let $f|_S: \{0,1\}^S \rightarrow \{0,1\}$ denote the restriction of f to $\{0,1\}^S$ by zeroing out all the coordinates outside S .
- 4 Pick $x, a_1, \dots, a_{d+1} \in \{0,1\}^S$ independently from the distribution $\mu_{1/2}^{\otimes S}$, subject to the constraint that a_1, \dots, a_{d+1} are linearly independent.
- 5 Accept iff

$$\sum_{I \subseteq [d+1]} f|_S \left(x + \sum_{i \in I} a_i \right) = 0 \pmod{2} .$$

We use the agreement theorem to show that this natural extension is a valid low degree test for the p -biased setting.

Theorem 1.5 (p -biased version of the BKSSZ Theorem). *For every d and $p \in (0, 1/2)$, the 2^{d+1} -query test $T_{p,d}$ satisfies the following properties:*

- *Completeness: if $\text{bdeg}(f) \leq d$ then $\text{rej}_{T_{d,p}}(f) = 0$.*
- *Soundness: $\delta_d^{(p)}(f) = O_d(\text{rej}_{T_{d,p}}(f))$, where the hidden constant is independent of p (but depends on d).*

Kindler–Safran Junta theorem for the biased hypercube A second (and more involved) illustration of this paradigm is lifting the Kindler–Safran Junta theorem to the p -biased hypercube, even when p is very small, proved by the authors [DFH17] in a parallel work to the current paper. Previous proofs of such statements deteriorated as p becomes smaller due to the use of hypercontractivity.

The application to the Kindler–Safran Junta theorem for the p -biased hypercube served as the primary motivation for the agreement theorem proved in this work. We remark that while the original version of [DFH17] used Theorem 1.2, a subsequent revision observed that a *simpler* agreement theorem that worked only for juntas sufficed. The revised version of [DFH17] contains this simpler “junta” agreement theorem with a self-contained proof.

Context and Motivation

Agreement tests were first studied implicitly in the context of PCP theorems. In fact, every PCP construction that has a composition step invariably relies on an agreement theorem. This is because in a typical PCP construction, the proof is broken into small pieces that are further encoded e.g. by composition or by a gadget. The soundness analysis decodes each gadget separately, thereby obtaining a collection of local views. Then, essentially through an agreement theorem, these are stitched together into one global NP witness. Similar to locally testable codes, agreement tests are a combinatorial question that is related to PCPs. Interestingly, this relation has recently been made formal by Dinur *et al.* [DKKMS18], where it is proved that a certain agreement test (whose correctness was proved by Khot *et al.* [KMS18]) *formally implies* a certain rather strong unique games PCP theorem. Such a formal connection is not known to exist between LTCs and PCPs. For example, even if someone manages to construct the “ultimate” locally testable codes with linear length and distance, and testable with a constant number of

queries, this is not known to have any implications for constructing linear size PCPs (although one may hope that such codes will be useful toward that goal).

Beyond their role in PCPs, we believe that agreement tests capture a fundamental local-to-global phenomenon, and merit study on their own. Exploring new structures that support agreement theorems seems to be an important direction. Several steps in this direction have been taken by Dinur and Kaufman [DK17] and by Dikstein and Dinur [DD19].

Relation to Property Testing Agreement testing is similar to property testing in that we study the relation between a global object and its local views. In property testing we have access to a single global object, and we restrict ourselves to look only at random local views of it. In agreement tests, we don't get access to a global object, but rather to an *ensemble of local functions* that are not apriori guaranteed to come from a single global object. Another difference is that unlike in property testing, in an agreement test the local views are pre-specified and are a part of the problem description, rather than being part of the algorithmic solution.

Still, there is an interesting interplay between Corollary 1.3, which talks about combining an ensemble of local graphs into one global graph, and graph property testing. Suppose we focus on some testable graph property, and suppose further that the test proceeds by choosing a random set of vertices and reading all of the edges in the induced subgraph, and checking that the property is satisfied there (many graph properties are testable this way, for example bipartiteness [CGR98]). Suppose we only allow ensembles $\{G_S\}$ where for each subset S , the local graph G_S satisfies the property (e.g. it is bipartite). This fits into our formalism by specifying the space of allowed functions \mathcal{F}_S to consist only of accepting local views. This is analogous to requiring, in the low degree test, that the local function on each line has low degree as a univariate polynomial. By Corollary 1.3, we know that if these local graphs agree with each other with probability $1 - \epsilon$, there is a global graph G that agrees with $1 - O(\epsilon)$ of them. In particular, this graph *passes the property test*, so must itself be close to having the property! At this point it is absolutely crucial that the agreement theorem provides the stronger guarantee that $G|_S = G_S$ (and not $G|_S \approx G_S$) for $1 - O(\epsilon)$ of the S 's. We can thus conclude that not only is there a global graph G , but actually that this global G is close to having the property.

This should be compared to the low degree agreement test, where we only allow local functions with low degree, and the conclusion is that there is a global function that itself has low degree.

Technical Contribution

Instead of proving Theorem 1.2 directly, we first prove a version which employs a different test distribution $\nu_n(k, t)$. For comparison, here are the two distributions:

- $(T, S_1, S_2) \sim \mu_n(k, t)$ if T is a random set of size t , and S_1, S_2 are two random extensions of T of size k such that $S_1 \cap S_2 = T$.
- $(T, S_1, S_2) \sim \nu_n(k, t)$ if T is a random set of size T , and S_1, S_2 are two random extensions of T of size k .

In both cases, the test is $f_{S_1}|_T = f_{S_2}|_T$. For $\mu_n(k, t)$, this is the same as $f_{S_1}|_{S_1 \cap S_2} = f_{S_2}|_{S_1 \cap S_2}$, but for $\nu_n(k, t)$ it could be that $S_1 \cap S_2 \supsetneq T$.

We deduce Theorem 1.2 from its $\nu_n(k, t)$ -version via a coupling argument.

To explain the proof, we start with the easier case $d = 1$, which is just the $\nu_n(k, t)$ -analog of Theorem 1.1. Given an ensemble $\{f_S\}$, it is easy to define the global function G , by popular vote ("majority

decoding”). The main difficulty is to prove that for a typical set S , f_S agrees with $G|_S$ on all elements $i \in S$ (and later on all d -sets).

Our proof doesn’t proceed by defining G as majority vote right away. Instead, like in many previous proofs [DG08, IKW12, DS14], we condition on a certain event (focusing, say, on all subsets that contain a certain set T , and such that $f_S|_T = \alpha$ for a certain value of α), and define a “restricted global” function, for each T , by taking majority just among the sets in the conditioned event. This boosts the probability of agreement inside this event. After this boost, we can afford to take a union bound and safely get agreement with the restricted global function G_T . The proof then needs to perform another agreement step which stitches the restricted global functions $\{G_T\}_T$ into a completely global function. The resulting global function does not necessarily equal the majority vote function G , and a separate argument is then carried out to show that the conclusion is correct also for G .

In higher dimensions $d > 1$, these two steps of agreement (first to restricted global and then to global) become a longer sequence of steps, where at each step we are looking at restricted functions that are defined over larger and larger parts of the domain.

The main technical difficulty is that a single event $f_S = F|_S$ consists of $\binom{k}{d}$ little events, namely $f_S(A) = F(A)$ for all $A \in \binom{S}{d}$, that each have some probability of failure. We thus need to boost the failure probability from ε to ε/k^d so that we can afford to take a union bound on the $\binom{k}{d}$ different sub-events. How do we get this large boost? Our strategy is to proceed in steps, where at each stage, we condition on the global function from the previous stage, boosting the probability of success further.

Majority decoding The most natural choice for the global function F in the conclusion of [Theorem 1.2](#) is the majority decoding, where $F(A)$ is the most common value of $f_S(A)$ over all S containing A . This is the content of the “furthermore” clause in the statement of the theorem. Neither the proof strategy of Dinur and Steurer [DS14] nor our generalization promises that the produced global function F is the majority decoding.

Our strategy produces a global function which agrees with most local functions, but we cannot guarantee immediately that this global function corresponds to majority decoding. What we are able to show is that *if* there is a global function agreeing with most of the local functions *then* the function obtained via majority decoding also agrees with most of the local functions. We outline the argument below.

Suppose that $\{f_S\}$ is an ensemble of local functions that mostly agree with each other, and suppose that they also mostly agree with some global function F . Let G be the function obtained by majority decoding: $G(A)$ is the most common value of $f_S(A)$ over all S containing A . Our goal is to show that G also mostly agrees with the local functions, and we do this by showing that F and G mostly agree.

Suppose that $F(A) \neq G(A)$. We consider two cases. If the distribution of $f_S(A)$ is very skewed toward $G(A)$, then $f_S(A) \neq F(A)$ will happen very often. If the distribution of $f_S(A)$ is very spread out, then $f_{S_1}(A) \neq f_{S_2}(A)$ will happen very often. Since both events $f_S(A) \neq F(A)$ and $f_{S_1}(A) \neq f_{S_2}(A)$ are known to be rare, we would like to conclude that $F(A) \neq G(A)$ happens for very few A ’s.

Here we face a problem: the bad events (either $f_S(A) \neq F(A)$ or $f_{S_1}(A) \neq f_{S_2}(A)$) corresponding to different A ’s are not necessarily disjoint. A priori, there might be many different A ’s such that $F(A) \neq G(A)$, but the bad events implied by them could all coincide.

The missing piece is the reverse union bound, [Lemma 1.4](#). We choose $Y = \{S : f_S \neq F|_S\}$ and $B = \{A : F(A) \neq G(A)\}$. If $A \in B$ then the probability that $f_S(A) \neq F(A)$ (over $S \supseteq A$) is at least $1/2$, since otherwise $F(A)$ would have been the majority decoded value. Therefore the premise of [Lemma 1.4](#)

is satisfied (with $c = 1/2$), and we conclude that the probability of $F|_S \neq G|_S$, which is the same as the probability that S hits B , is only $O(\Pr[f_S \neq F|_S]) = O(\varepsilon)$.

Reverse union bound The reverse union bound follows from a *hypergraph pruning lemma*, which allows approximating a given hypergraph H by a subhypergraph $H' \subset H$ which is sparse and has certain “disjointness” properties.

Lemma 1.6 (hypergraph pruning lemma). *For each $\varepsilon > 0$ and integer $d \geq 1$ there exists a constant $C''(d, \varepsilon) = 2^{O_\varepsilon(d^3)}$ such that for every d -uniform hypergraph H there exists a d -uniform hypergraph $H' \subseteq H$ such that:*

- $\Pr_{S \sim \nu_n(k)}[S \text{ hits } H'] \geq \Pr_{S \sim \nu_n(k)}[S \text{ hits } H] / C''(d, \varepsilon)$.
- For every $I \in H'$,

$$\Pr_{S \sim \nu_n(k)}[\exists I' \in H' \setminus \{I\}, S \supseteq I' \mid S \supseteq I] \geq 1 - \varepsilon.$$

The hypergraph pruning lemma states that we can prune a given d -uniform hypergraph H to a subhypergraph H' such that the events “ $S \supseteq I$ ” (where $S \sim \nu_n(k)$ and $I \in H'$) are nearly disjoint, while maintaining the property that the probability that S contains some hyperedge in H' does not drop too much compared to the corresponding probability with respect to H .

The pruned hypergraph H' satisfies the following crucial property, for each $I \in H'$:

$$\Pr_{S \sim \nu_n(k)}[S \text{ hits } H' \text{ only at } I] = \Theta \left(\Pr_{S \sim \nu_n(k)}[S \supseteq I] \right).$$

Crucially, the events “ S hits H' only at I ” are disjoint. This allows us to quickly prove [Lemma 1.4](#): applying the hypergraph pruning lemma, we prune B to a smaller hypergraph B' . The disjointness of the aforementioned events implies that

$$\begin{aligned} \Pr[S \in Y] &\geq \sum_{I \in B'} \Pr[S \text{ hits } B' \text{ only at } I \text{ and } S \in Y] \\ &\stackrel{(*)}{\geq} \Omega(1) \cdot \sum_{I \in B'} \Pr_{S \sim \nu_n(k)}[S \supseteq I] = \Omega(\Pr[S \text{ hits } B']) = \Omega(\Pr[S \text{ hits } B]), \end{aligned}$$

where $(*)$ follows from the crucial property:

$$\Pr[S \text{ hits } B' \text{ only at } I \text{ and } S \in Y] \geq \Pr[S \text{ hits } B' \text{ only at } I] - \Pr[S \supseteq I \text{ and } S \in Y] = \Omega(\Pr[S \supseteq I]).$$

The proof of [Lemma 1.6](#) is inductive and elementary. It is proved by introducing the notion of *branching factor*; a hypergraph H over a vertex set V is said to have *branching factor* ρ if for all subsets $A \subset V$ and integers $r \geq 0$, there are at most ρ^r hyperedges in H of cardinality $|A| + r$ containing A (see [Section 7](#) and [Definition 7.1](#) therein for further details).

Organization

The rest of this paper is organized as follows. We start in [Section 3](#) by proving the version of [Theorem 1.1](#) which uses the test distribution $\nu_n(k, t)$. We then generalize this to higher dimensions in [Section 4](#). We deduce [Theorem 1.2](#) in [Section 5](#), in which we also deduce a p -biased agreement theorem and an

auxiliary agreement theorem required in the subsequent section. The application to $GF(2)$ -low degree testing appears in [Section 6](#). We conclude the paper by proving the reverse union bound in [Section 7](#).

2 Notation and Preliminaries

We begin with some basic notation. Let $[n] = \{1, \dots, n\}$. For $k \leq n$, we will refer to a set $S \subseteq [n]$ of size k as a k -set of $[n]$.

We will be working with several distributions on subsets of $[n]$ (and pairs, triples of subsets on $[n]$), which are defined in detail below.

$\nu_n(k)$: The *slice* $\binom{[n]}{k}$ consists of all k -subsets of $[n]$. We denote the uniform distribution over $\binom{[n]}{k}$ by $\nu_n(k)$.

$\nu_n(k, t)$: For $1 \leq t \leq k \leq n$, let $\nu_n(k, t)$ denote the distribution over triples of sets (T, S_1, S_2) obtained as follows: pick a uniformly random t -set T of $[n]$ and then pick independently two random k -sets S_1, S_2 of $[n]$ such that $T \subset S_1, S_2$.

$\mu_n(k, t)$: The distribution $\mu_n(k, t)$, on the other hand, is the distribution over pair of sets (S_1, S_2) obtained as follows: pick a uniformly random t -set T and then two k -sets S_1 and S_2 uniformly at random conditioned on $S_1 \cap S_2 = T$. We will refer to the latter distribution $\mu_n(k, t)$ as the exact-intersection distribution.

We now define the p -biased version of the above distributions.

$\mu_p(A)$: For a set A , the distribution $\mu_p(A)$ denotes a random subset of A obtained by putting it each element with probability p independently. Stated differently, $\Pr[\mu_p(A) = B] = p^{|B|}(1-p)^{|A \setminus B|}$.

$\mu_{p,\alpha}(A)$: The distribution $\mu_{p,\alpha}(A)$, defined whenever $p(2-\alpha) \leq 1$, is a distribution on pairs of subsets $B_1, B_2 \subseteq A$ which are sampled as follows: for each element $x \in A$, we put x in both B_1 and B_2 with probability $p\alpha$; we put x only in B_1 or only in B_2 with probability $p(1-\alpha)$ each; and we put x in neither with probability $1 - p(2-\alpha)$.

2.1 Plurality vs. collision

Let X be a random variable ranging over a finite set, and let x_0 be the most probably value. Let Y be an i.i.d. copy of X . Then

$$\Pr[X \neq x_0] \leq \Pr[X \neq Y].$$

To see this, let p_x be the probability of x . Then

$$\Pr[X \neq Y] = \sum_x p_x(1 - p_x) \geq \sum_x p_x(1 - p_{x_0}) = 1 - p_{x_0} = \Pr[X \neq x_0].$$

2.2 Reverse union bound

We recall the statement of the reverse union bound from the introduction.

Lemma 1.4 (reverse union bound). *For each $c > 0$ and integer $d \geq 1$ there exists a constant $C'(d, c) = 2^{O_c(d^3)}$ such that the following holds for all $d \leq k \leq n$. Let $Y \subseteq \binom{[n]}{k}$ and $B \subseteq \binom{[n]}{d}$ satisfy*

$$\Pr_{S \sim \nu_n(k)}[S \in Y \mid S \supseteq I] \geq c \text{ for all } I \in B.$$

Then

$$\Pr_{S \sim \nu_n(k)}[S \text{ hits } B] \leq C'(c, d) \cdot \Pr_{S \sim \nu_n(k)}[S \in Y],$$

where S hits B if $S \supseteq I$ for some $I \in B$.

We prove [Lemma 1.4](#) in [Section 7](#). For now, let us explain the name *reverse union bound*. One could try to lower bound $\Pr[S \in Y]$ in terms of $\Pr[S \text{ hits } B]$ using the following simple argument: if S hits B , then it contains some $I \in B$, and so the probability that $S \in Y$ should be at least c . In symbols, this would be

$$\Pr[S \in Y] \stackrel{(*)}{\geq} \sum_{I \in B} \Pr[S \supseteq I] \Pr[S \in Y \mid S \supseteq I] \geq c \sum_{I \in B} \Pr[S \supseteq I] \geq c \Pr[S \text{ hits } B].$$

This argument is, of course, fallacious: $(*)$ fails since the events in question are not disjoint. The reverse union bound shows that if you are willing to pay in the constant c , then you can make sense of this fallacious argument.

3 One-dimensional agreement theorem

In this section, we prove the following direct product agreement testing theorem for one dimension. This theorem is a special case of the more general theorem ([Theorem 4.1](#)) proved in the next section, and a slight strengthening of the result of Dinur and Steurer [[DS14](#)] ([Theorem 1.1](#)). We give the proof for the one-dimensional case as it serves as a warmup to the general case.

Theorem 3.1 (agreement theorem, dimension 1 (strengthened version of [Theorem 1.1](#))). *For all positive integers n, k, t satisfying $1 \leq t < k \leq n$ and alphabet Σ , the following holds. Let $\{f_S : S \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)}[f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon.$$

Then the majority-decoded function $G : [n] \rightarrow \Sigma$ satisfies

$$\Pr_{S \sim \nu_n(k)}[f_S \neq G|_S] = O\left(\frac{\varepsilon}{\min\left(\frac{t}{k}, 1 - \frac{t}{k}\right)}\right),$$

where the majority-decoded function G is the one given by “popular vote”, namely for each $i \in [n]$, $G(i)$ is the most frequently occurring value among $\{f_S(i) \mid S \ni i\}$ (breaking ties arbitrarily).

3.1 Two examples

Consider the following two examples.

Example 1: “Small non-expanding set of k -tuples” Let B be an arbitrary set of size $(\varepsilon/2k)n$. We define f_S as follows: $f_S(i) = 1_{S \cap B \neq \emptyset}$. Majority decoding gives the function $g(i) = 0$. In this example, the probability that a random k -set S_1 intersects B is $\Theta(\varepsilon)$, and when that happens, the intersection with B has constant size. Therefore, given that S_1 intersects B , the probability that another random k -set S_2 also intersects B is roughly t/k . Overall, the probability that f_{S_1}, f_{S_2} disagree is $\Theta((1 - t/k)\varepsilon)$. In contrast, $\Pr[f_S \neq g|_S] = \Theta(\varepsilon)$.

Example 2: “Bernoulli noise” Consider the following function: $f_S(i) \sim \text{Ber}(\varepsilon/2t)$, where majority decoding again gives the function $g(i) = 0$. In this example, the probability that f_{S_1} and f_{S_2} disagree is $\Theta(\varepsilon)$, while $\Pr[f_S \neq g|_S] \approx 1 - \exp(-(k/2t)\varepsilon)$, which is close to $\Theta((k/t)\varepsilon)$ for ε much smaller than k/t .

These two examples show that the statement of [Theorem 3.1](#) is tight.

We begin by considering a natural approach¹ to proving the above theorem using the majority (or plurality) decoding, and show how this approach only obtains an error bound of $O(k\varepsilon)$ instead of $O(\varepsilon)$ as promised by the above result.

For the purpose of this section, we assume that $0.1 \leq t/k \leq 0.5$. Given a local ensemble of functions $\{f_S: S \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$, define the majority decoded function $g: [n] \rightarrow \Sigma$ as follows: $g(i)$ is the most popular value of $f_S(i)$ among sets S containing i . Plurality vs. collision implies that for each $i \in [n]$,

$$\Pr_{S \ni i}[f_S(i) \neq g(i)] \leq \Pr_{S_1, S_2 \ni i}[f_{S_1}(i) \neq f_{S_2}(i)].$$

The distribution of the pair of sets S_1, S_2 in the above expression is not the same as that used in the test. To relate this to the success probability of the test, we construct the following coupled distribution (T_1, T_2, S_1, S_2, S) for each $i \in [n]$: choose two t -sets T_1, T_2 containing i , choose two random k -sets S_1 and S_2 such that $T_1 \subset S_1$ and $T_2 \subset S_2$, and finally choose a random k -set S such that $T_1 \cup T_2 \subset S$. (The last step is possible if $1 + 2(t - 1) = 2t - 1 \leq k$, which is true since $0.1 \leq t/k \leq 0.5$). This coupling has the following property: the pair of sets (S_1, S_2) are two random k -sets containing i , while for each $j \in [2]$ the pair of sets (S_j, S) are two random k -sets containing T_j , which is itself a random t -set containing i .

If $f_{S_1}(i) \neq f_{S_2}(i)$ then $f_{S_j}(i) \neq f_S(i)$ for some $j \in [2]$, and so

$$\Pr_{S \ni i}[f_S(i) \neq g(i)] \leq \Pr_{S_1, S_2 \ni i}[f_{S_1}(i) \neq f_{S_2}(i)] \leq 2 \Pr_{\substack{T \ni i \\ S_1, S_2 \supset T}}[f_{S_1}(i) \neq f_{S_2}(i)].$$

Undoing the conditioning, and using the fact that for any fixed i , $\Pr[S \ni i] = \Theta(\Pr[T \ni i])$ (since $0.1 \leq t/k \leq 0.5$), this shows that

$$\Pr_S[S \ni i \text{ and } f_S(i) \neq g(i)] \leq O\left(\Pr_{T; S_1, S_2 \supset T}[T \ni i \text{ and } f_{S_1}(i) \neq f_{S_2}(i)]\right).$$

Summing over all i , this gives

$$\begin{aligned} \Pr_S[f_S \neq g|_S] &\stackrel{(1)}{\leq} \sum_{i=1}^n \Pr_S[S \ni i \text{ and } f_S(i) \neq g(i)] \\ &= O\left(\sum_{i=1}^n \Pr_{T; S_1, S_2 \supset T}[T \ni i \text{ and } f_{S_1}(i) \neq f_{S_2}(i)]\right) = O\left(\mathbb{E}_T\left[\sum_{T \ni i} \Pr_{S_1, S_2 \supset T}[f_{S_1}(i) \neq f_{S_2}(i)]\right]\right) \\ &= O\left(\mathbb{E}_{T; S_1, S_2 \supset T}[\#\{T \ni i \mid f_{S_1}(i) \neq f_{S_2}(i)\}]\right) \stackrel{(2)}{\leq} O(t\varepsilon). \end{aligned}$$

Observe that this simple argument gives a bound which is off by a multiplicative factor of t . Inequality (1) is loose whenever $f_S \neq g|_S$ on many coordinates, and inequality (2) is loose whenever f_{S_1}, f_{S_2} disagree on a few coordinates.

Can it really be that conditioned on $f_{S_1} \neq f_{S_2}$, the number of bad i , i.e., the number of $i \in T$ such

¹This argument is adapted from the introduction of Yotam Dikstein’s Master’s thesis [\[Dik19\]](#)

that $f_{S_1}(i) \neq f_{S_2}(i)$, be as large as t ? Yes! In Example 1, when $f_{S_1} \neq f_{S_2}$, it is always the case that f_{S_1}, f_{S_2} disagree on all of their intersection. Hence (2) is tight, however (1) is loose: the contribution of $i \in B$ to the sum is $\varepsilon/2$, while each of the $(1 - \varepsilon/(2k))n$ other indices contributes roughly $(\varepsilon/2)(k/n)$, for a total of roughly $(\varepsilon/2)k$. On the other hand, in example 2, (1) is tight but (2) is loose, since the expected number of bad i is constant. Our main result shows that (1) and (2) cannot both be tight at the same time.

3.2 Proof of Theorem 3.1

Let $t + 1 \leq k \leq n$ as in the hypothesis. We will assume the following convention throughout this section.

- S will always denote a set of size k .
- T will always denote a set of size t .
- U will always denote a set of size $t - 1$.
- $n' = n - (t - 1)$ and $k' = k - (t - 1)$.

Define

$$\zeta := t/(k - t + 1) = t/k' \text{ and } \xi := 1 - t/k.$$

Given a global function $g: [n] \rightarrow \Sigma$ and a local function $f_S: S \rightarrow \Sigma$, we say that f_S and g *disagree* if $f_S \neq g|_S$.

We now turn to the actual proof. Let U be a $(t - 1)$ -set. In the first step (which we refer to as the relative decoding step), we define a “global” function $g_U: [n] \rightarrow \Sigma$ relative to the set U and show that for a typical $(t - 1)$ -set U and k -set S such that $U \subset S$, we have $\Pr_{U, S \supset U}[f_S \neq g_U|_S] = O(\varepsilon)$. In the next step (which we refer to as the sewing step), we sew together these different relative functions g_U and show that there exists a U such that for most k -sets S (not necessarily k -sets that contain U), we have that $f_S = g_U|_S$. In the final step, we conclude that if such a “global” g_U exists, then the majority decoded function also works.

3.2.1 Part 1: relative decoding

Let U be a set of size $t - 1$. We define a function $g = g_U: [n] \rightarrow \Sigma$ as follows (we drop the subscript U as U is fixed for most of the argument below).

1. $g|_U$ is the most popular value of $f_S|_U$ among sets S containing U .
2. For $i \notin U$, $g(i)$ is the most popular value of $f_S(i)$ among sets S containing U and satisfying $f_S|_U = g|_U$.

Our goal in this part is to bound $\mathbb{E}_U \Pr_{S \supset U}[f_S, g \text{ disagree}]$. More precisely,

Lemma 3.2. *For all positive integers n, k, t satisfying $1 \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: S \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon.$$

Then there exists an ensemble $\{g_U: [n] \rightarrow \Sigma \mid U \in \binom{[n]}{t-1}\}$ of global functions such that when a random $U \in \binom{[n]}{t-1}$ and $S \in \binom{[n]}{k}$ are chosen satisfying $S \supset U$, we have

$$\Pr_{U, S \supset U} [g_U|_S \neq f_S] = O(\epsilon/\zeta),$$

where $\zeta = t/(k-t+1)$.

Proof. We begin by defining the notion of a *good* element i and a *good* k -set S . We say that $i \notin U$ is *good* if

$$\Pr_{S \supset U+i} [f_S|_U = g|_U] \geq \frac{1}{2}.$$

We say that a k -set $S \supset U$ is *good* if all $i \in S \setminus U$ are good.

We bound the quantity of interest $\Pr_{S \supset U} [f_S \neq g|_S]$ for a fixed U as follows:

$$\begin{aligned} \Pr_{S \supset U} [f_S \neq g|_S] &\leq \Pr_{S \supset U} [S \text{ is bad}] + \Pr_{S \supset U} [S \text{ is good and } f_S \neq g|_S] \\ &\leq \Pr_{S \supset U} [S \text{ is bad}] + \Pr_{S \supset U} [f_S|_U \neq g|_U] + \Pr_S [S \text{ is good and } f_S|_U = g|_U \text{ and } \exists i \in S \setminus U, f_S(i) \neq g(i)] \\ &\leq \Pr_{S \supset U} [S \text{ is bad}] + \Pr_{S \supset U} [f_S|_U \neq g|_U] + \sum_{\substack{i \notin U \\ i \text{ good}}} \Pr_{S \supset U} [f_S|_U = g|_U \text{ and } i \in S \text{ and } f_S(i) \neq g(i)] \\ &= \underbrace{\Pr_{S \supset U} [S \text{ is bad}]}_{\delta} + \underbrace{\Pr_{S \supset U} [f_S|_U \neq g|_U]}_{\gamma} + \frac{k'}{n'} \sum_{\substack{i \notin U \\ i \text{ good}}} \underbrace{\Pr_{S \supset U+i} [f_S|_U = g|_U \text{ and } f_S(i) \neq g(i)]}_{\gamma_i}. \end{aligned}$$

The penultimate step follows from a union bound and the fact that for good S , every $i \in S \setminus U$ is good. The last step follows from the fact that $\Pr_{S \supset U} [i \in S] = k'/n'$. Finally, we can extend the last summation to all i (not necessarily good i) by defining $\gamma_i = 0$ for bad i . We bound each of the error terms γ, δ and γ_i in the following three claims.

Claim 3.3 (bounding γ). *If $1 \leq t < k \leq n$, then $\mathbb{E}_U [\gamma] \leq 2\epsilon$.*

Proof. By plurality vs. collision, $\gamma \leq \Pr_{S_1, S_2 \supset U} [f_{S_1}|_U \neq f_{S_2}|_U]$.

Now consider the following coupling. We choose random independent $i_1, i_2 \notin U$, a random k -set $S_1 \supset U + i_1$, a random k -set $S_2 \supset U + i_2$, and a random k -set $S \supset U \cup \{i_1, i_2\}$. Here, we are using the fact that $k \geq (t-1) + 2 = t+1$. Also, note i_1 might be equal to i_2 in the above coupling. Since i_1, i_2 are independent, the marginal distribution of (S_1, S_2) is the same as two independent sets containing U , and the marginal distribution of (S, S_j) is the same as two independent sets containing U and one additional element. Thus if $f_{S_1}|_U \neq f_{S_2}|_U$, it must be the case that for some $j \in \{1, 2\}$, we have $f_{S_j}|_U \neq f_S|_U$. Therefore

$$\gamma \leq \Pr_{S_1, S_2 \supset U} [f_{S_1}|_U \neq f_{S_2}|_U] \leq 2 \Pr_{\substack{T \supset U \\ S_j, S \supset T}} [f_{S_j}|_U \neq f_S|_U] \leq 2 \Pr_{\substack{T \supset U \\ S_j, S \supset T}} [f_{S_j}|_T \neq f_S|_T].$$

This shows that $\mathbb{E}_U [\gamma] \leq 2\epsilon$. □

Claim 3.4 (bounding δ). $\mathbb{E}_U [\delta] = O(\epsilon)$.

Proof. Let $Y = \{S \supset U \mid f_S|_U \neq g|_U\}$, so that $\Pr[S \in Y] = \gamma$. Let B be the set of bad i . If i is bad then

$$\Pr_{S \supset U} [S \in Y \mid i \in S] \geq \frac{1}{2}.$$

Hence the reverse union bound shows that

$$\delta = \Pr_{S \supset U} [S \text{ is bad}] = \Pr_{S \supset U} [S \text{ hits } B] = O\left(\Pr_{S \supset U} [S \in Y]\right) = O(\gamma).$$

Hence $\mathbb{E}_U [\delta] = O(\mathbb{E}_U [\gamma]) = O(\varepsilon)$. □

Claim 3.5 (bounding γ_i). *For all $i \notin U$, $\mathbb{E}_{U, i \notin U} [\gamma_i] = O(\varepsilon/t)$.*

Proof. We first consider the case when $i \notin U$ is good.

$$\begin{aligned} \gamma_i &= \Pr_{S \supset U+i} [f_S|_U = g|_U] \cdot \Pr_{\substack{S \supset U+i \\ f_S|_U = g|_U}} [f_S(i) \neq g(i)] \\ &\leq \Pr_{S \supset U+i} [f_S|_U = g|_U] \cdot \Pr_{\substack{S_1, S_2 \supset U+i \\ f_{S_1}|_U = f_{S_2}|_U = g|_U}} [f_{S_1}(i) \neq f_{S_2}(i)] \quad [\text{plurality vs. collision}] \\ &= \left(\Pr_{S \supset U+i} [f_S|_U = g|_U] \right)^{-1} \cdot \Pr_{S_1, S_2 \supset U+i} [f_{S_1}|_U = f_{S_2}|_U = g|_U \text{ and } f_{S_1}(i) \neq f_{S_2}(i)] \\ &\leq 2 \Pr_{S_1, S_2 \supset U+i} [f_{S_1}|_U = f_{S_2}|_U \text{ and } f_{S_1}(i) \neq f_{S_2}(i)], \end{aligned}$$

where we have used the fact that i is good at the very last step. Clearly, the same bound also holds for bad i .

This shows that

$$\begin{aligned} \mathbb{E}_{U, i \notin U} [\gamma_i] &\leq 2 \mathbb{E}_T \Pr_{\substack{S_1, S_2 \supset T \\ i \in T}} [f_{S_1}|_T \text{ and } f_{S_2}|_T \text{ disagree only on } i] \\ &= \frac{2}{t} \mathbb{E}_T \Pr_{S_1, S_2 \supset T} [f_{S_1}|_T \text{ and } f_{S_2}|_T \text{ disagree on a single element}] \\ &\leq \frac{2}{t} \mathbb{E}_T \Pr_{S_1, S_2 \supset T} [f_{S_1}|_T \text{ and } f_{S_2}|_T \text{ disagree}] \\ &\leq \frac{2}{t} \varepsilon. \quad \square \end{aligned}$$

Combining the three bounds, we conclude that

$$\mathbb{E}_{U, S \supset U} [f_S \neq g|_U] \leq O(\varepsilon) + k' \frac{2}{t} \varepsilon = O\left(\frac{\varepsilon}{\zeta}\right).$$

This completes the proof of [Lemma 3.2](#). □

3.2.2 Part 2: sewing

In this step, we show that the global function g_U corresponding to some $(t-1)$ -set U explains most local functions f_S corresponding to S 's not necessarily containing U . We will first prove this under the assumption that $k \geq 4t$ (or equivalently $\zeta \geq 3/4$), and then extend it to all $\zeta \in (0, 1)$.

Lemma 3.6. For all positive integers n, k, t satisfying $1 \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: S \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon, \quad (1)$$

then there exists a global function $g: [n] \rightarrow \Sigma$ such that $\Pr_S [f_S \neq g|_S] = O(\varepsilon/\xi\zeta)$, where $\zeta = t/(k - t + 1)$ and $\xi = 1 - t/k$.

Proof. Let us first assume that $k \geq 4t$. Lemma 3.2 implies that

$$\Pr_{\substack{U_1, U_2 \\ S \supset U_1 \cup U_2}} [g_{U_1}|_S \neq g_{U_2}|_S] = O(\varepsilon/\zeta).$$

We now define a coupling (U_1, U_2, S, S_1, S_2) as follows: Choose two random $(t-1)$ -sets U_1, U_2 . Then, choose a random k -set S containing U_1 . Choose a random ordering of $S \setminus U_1$, let X_1 be the first $\lfloor (k - t + 1)/2 \rfloor$ elements, and let X_2 be the subsequent $\lceil (k - t + 1)/2 \rceil$ elements. Finally, choose a random k -set S_1 containing $U_1 \cup U_2 \cup X_1$, and a random k -set S_2 containing $U_1 \cup U_2 \cup X_2$. The last step is possible since $2(t-1) + \lceil (k - t + 1)/2 \rceil \leq 2t + k/2 \leq k$. The marginal distribution of (S_j, U_1, U_2) is random sets U_1, U_2 and a random set $S_j \supset U_1 \cup U_2$. The coupling ensures that $S \subset S_1 \cup S_2$, and so $g_{U_1}|_S \neq g_{U_2}|_S$ implies that $g_{U_1}|_{S_j} \neq g_{U_2}|_{S_j}$ for some $j \in \{1, 2\}$. Therefore

$$\Pr_{\substack{U_1, U_2 \\ S \supset U_1}} [g_{U_1}|_S \neq g_{U_2}|_S] = O(\varepsilon/\zeta).$$

Fix $U := U_2$ such that

$$\Pr_{U_1, S \supset U_1} [g_{U_1}|_S \neq g_U|_S] = O(\varepsilon/\zeta).$$

Then $g = g_U$ satisfies

$$\Pr_{U_1, S \supset U_1} [f_S \neq g|_S] \leq \Pr_{U_1, S \supset U_1} [f_S \neq g_{U_1}|_S] + \Pr_{U_1, S \supset U_1} [g_{U_1}|_S \neq g|_S] = O(\varepsilon/\zeta).$$

This completes the case of small t .

Suppose now that $t \geq k/4$, and let $t' = \lfloor k/4 \rfloor$. For any $0 < \ell \leq t$, define

$$\varepsilon_\ell := \Pr_{(T, S_1, S_2) \sim \nu_n(k, \ell)} [f_{S_1}|_T \neq f_{S_2}|_T].$$

Given the hypothesis (1) that $\varepsilon_t \leq \varepsilon$, we will show that $\varepsilon_{t'} = O(\varepsilon/\xi)$. We can then apply the preceding case and complete the proof.

Let $r < t$ such that $r \geq 2t - k$. Consider the following coupling argument. Consider the joint distribution $(R, T_1, T_2, S_1, S_2, S)$ obtained as follows: Choose a random r -set R . Choose two random t -sets $T_1, T_2 \supset R$. Choose a random k -set S_1 containing T_1 , a random k -set S_2 containing T_2 , and a random k -set S containing $T_1 \cup T_2$. The last step is possible if $k \geq 2t - r$, i.e., if $r \geq 2t - k$.

The sets S_1, S_2 are two random sets containing R , and the sets S, S_j are two random sets containing T_j . Furthermore, if $f_{S_1}|_R \neq f_{S_2}|_R$ then $f_{S_j}|_{T_j} \neq f_S|_{T_j}$ for some $j \in \{1, 2\}$. Therefore

$$\Pr_{R, S_1, S_2 \supset R} [f_{S_1}|_R \neq f_{S_2}|_R] \leq \Pr_{\substack{T_1 \\ S_1, S \supset T_1}} [f_{S_1}|_{T_1} \neq f_S|_{T_1}] + \Pr_{\substack{T_2 \\ S_2, S \supset T_2}} [f_{S_2}|_{T_2} \neq f_S|_{T_2}] \leq 2\varepsilon.$$

This demonstrates that if the hypothesis for the agreement theorem is true for a particular choice of n, k, t (i.e., $\varepsilon_t \leq \varepsilon$), then the hypothesis is also true for n, k, r by increasing ε to 2ε (i.e., $\varepsilon_r \leq 2\varepsilon$) provided $r \geq 2t - k$ or equivalently $2(1 - t/k) \geq (1 - r/k)$. Thus, given the hypothesis is true for some t, k satisfying $\zeta = 1 - t/k$, we can perform the above coupling argument $\lceil \log_2 1/\zeta \rceil$ times to reduce t to less than $k/4$. Each step doubles the error, and so the overall error probability becomes $O(\varepsilon/\zeta)$. Now applying the argument for small t , we obtain that there exists a global g such that $\Pr_S [f_S \neq g|_U] = O(\varepsilon/\zeta\tilde{\zeta})$. \square

3.2.3 Part 3: majority decoding

In this final step of the argument, we show that [Lemma 3.6](#) can be further strengthened by showing that the global function g can in fact be the majority-decoded function.

Proof of Theorem 3.1. Let g be the global function given by [Lemma 3.6](#). Define $G: [n] \rightarrow \Sigma$ as follows: let $G(i)$ be a plurality value of $f_S(i)$ over all $S \ni i$ (breaking ties arbitrarily). Let $B = \{i \mid g(i) \neq G(i)\}$, and let $Y = \{S \mid f_S \neq g|_S\}$.

For each $i \in B$,

$$\Pr_S [S \in Y \mid i \in S] \geq \Pr_S [f_S(i) \neq g(i) \mid i \in S] \geq \frac{1}{2},$$

since otherwise $g(i)$ would be the unique majority value. Applying the reverse union bound, we conclude that

$$\Pr_S [S \text{ hits } B] = O\left(\Pr_S [f_S \neq g|_S]\right) = O(\varepsilon/\zeta\tilde{\zeta}).$$

Therefore

$$\Pr_S [f_S \neq G|_S] \leq \Pr_S [f_S \neq g|_S] + \Pr_S [g|_S \neq G|_S] = \Pr_S [f_S \neq g|_S] + \Pr_S [S \text{ hits } B] = O(\varepsilon/\zeta\tilde{\zeta}).$$

[Theorem 3.1](#) follows by observing that $1/\zeta\tilde{\zeta} = O\left(\frac{1}{\min(\frac{t}{k}, 1 - \frac{t}{k})}\right)$. \square

4 Agreement theorem for high dimensions

Theorem 4.1 (agreement theorem for high dimensions). *For every positive integer d there exists a constant $M_d = 2^{O(d^4)}$ such that for all positive integers n, k, t satisfying $d \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: \binom{[n]}{d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon,$$

then the majority decoded function $G: \binom{[n]}{d} \rightarrow \Sigma$ satisfies

$$\Pr_{S \sim \nu_n(k)} [f_S \neq G|_S] \leq \frac{M_d}{\zeta\tilde{\zeta}^d} \cdot \varepsilon,$$

where $\zeta = \min\{1, t/(k - t + 1)\}$ and $\tilde{\zeta} = 1 - t/k$, and the majority-decoded function G is the one given by “popular vote”, namely for each $A \in \binom{[n]}{d}$ set $G(A)$ to be the most frequently occurring value among $\{f_S(A) \mid S \supset A\}$ (breaking ties arbitrarily).

The same two examples showing the tightness of [Theorem 3.1](#) also show that an inverse polynomial dependence on ξ and ζ are needed in the above theorem. However, it is unclear if the exponential dependence on d (in both the ζ^d term and the constant M_d) is required.

Let $t + 1 \leq k \leq n$ be as in the hypothesis. We will assume the following convention throughout this section.

- S will always denote a set of size k .
- T will always denote a set of size t .
- U will always denote a set of size $t - d$.
- $n' = n - (t - d)$, $k' = k - (t - d)$, $t' = t - (t - d) = d$.

As in the 1-dimensional case, we will first (in part 1) define for each U a global function $g_U: \binom{[n]}{d} \rightarrow \Sigma$ such that $\Pr_{U, S \supseteq U} [f_S \neq g_U|_S] = O(\epsilon)$, then show (in part 2) that g_U corresponding to a typical U explains most S (not necessarily ones that contain U) and finally (in part 3) show that the majority-decoded function works.

4.1 Part 1: relative decoding

In this part, for each $(t - d)$ -set U , we define a function $g_U: \binom{[n]}{d} \rightarrow \Sigma$ such that

$$\Pr_{U, S \supseteq U} [f_S \neq g_U|_S] = O(\epsilon).$$

For a large part of the proof, we will focus on a single U , and use g for g_U . The definition of such a $g = g_U$ will be inspired by the corresponding definition in the 1-dimensional case. We will define g incrementally. We will first define g for all d -sets fully contained in U , then for d -sets which have at most 1 element outside U , then for sets which have at most 2 elements outside U , and so on.

First for some notation. Let $h: \binom{[n]}{d} \rightarrow \Sigma$. For any $0 \leq \ell \leq d$ and any ℓ -set $L \subset [n] \setminus U$, we define the function $h^U[L]: \binom{U}{d-\ell} \rightarrow \Sigma$ as follows:

$$h^U[L](A) := h(A \cup L).$$

For any k -set S such that $S \supset U \cup L$, we can similarly define $f_S^U[L]: \binom{U}{d-\ell} \rightarrow \Sigma$. We will say that $f_S^U \stackrel{\ell}{\approx} h^U$ if $f_S^U[L] = h^U[L]$ whenever $|L| \leq \ell$. If the set U is clear from context, we will drop the superscript U from the above notation (as in $f_S[L]$ and $f_S \stackrel{\ell}{\approx} h$).

Lemma 4.2. *For every positive integer d , there exists a positive constant $\kappa_d = 2^{O(d^4)}$ such that the following is true. For all positive integers n, k, t satisfying $d \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: \binom{S}{d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \epsilon,$$

then there exists an ensemble $\{g_U: \binom{[n]}{d} \rightarrow \Sigma \mid U \in \binom{[n]}{t-d}\}$ of global functions such that when a random $U \in \binom{[n]}{t-d}$ and $S \in \binom{[n]}{k}$ are chosen satisfying $S \supset U$, then

$$\Pr_{U, S \supset U} [g_U|_S \neq f_S] \leq \frac{\kappa_d}{\zeta^d} \cdot \epsilon,$$

where $\zeta = \min\{1, t/(k-t+1)\}$.

Proof. We begin with the definition of g_U . Fix some $(t-d)$ -set U . We will define $g = g_U$ by defining $g[L]$ incrementally for $|L| = 0 \dots d$. To this end, we need the notions of ℓ -good and ℓ -excellent sets for $-1 \leq \ell \leq d$. To begin with, all k -sets S are (-1) -excellent. For larger ℓ , we define inductively as follows:

- For $\ell = 0$ to d :

- For all ℓ -sets L ,

$$g[L] := \text{popular}\{f_S[L] \mid S \supset U \cup L, S \text{ is } (\ell-1)\text{-excellent}\}.$$

If there are no $(\ell-1)$ -excellent k -sets S , we define $g[L]$ arbitrarily.

- An ℓ -set L is said to be *good* if it satisfies the following condition (and said to be *bad* otherwise).

$$\Pr_{S \supset U} [S \text{ is } (\ell-1)\text{-excellent} \mid S \supseteq L] \geq \frac{1}{2}.$$

- A k -set S is said to be ℓ -good if it is $(\ell-1)$ -excellent and does not contain any bad ℓ -set L .

- A k -set S is said to be ℓ -excellent if it is ℓ -good and $f_S \stackrel{\ell}{\approx} g$.

If S is ℓ -excellent then $f_S \stackrel{\ell}{\approx} g$. In particular, if S is d -excellent then $f_S \stackrel{d}{\approx} g$ which implies that $f_S = g|_S$. We thus have,

$$\Pr_{S \supset U} [f_S \neq g|_S] \leq \Pr_{S \supset U} [S \text{ is not } d\text{-excellent}] \quad (2)$$

Towards bounding this probability, we prove the following claims. For $\ell = 0, \dots, d$, let

$$\gamma_\ell := \Pr_{S \supset U} [S \text{ is } \ell\text{-good but not } \ell\text{-excellent}]$$

Claim 4.3. *If n, k, t, d, ℓ are 5 positive integers satisfying $0 \leq \ell \leq d \leq t < k \leq n$, then*

$$\Pr_{\substack{R: |R|=t-\ell \\ S_1, S_2 \supset R}} [f_{S_1}|_R \neq f_{S_2}|_R] \leq 2^\ell \cdot \Pr_{\substack{T: |T|=t \\ S_1, S_2 \supset T}} [f_{S_1}|_T \neq f_{S_2}|_T].$$

Proof. The proof of this claim employs a coupling argument identical to the one used in the proof of [Claim 3.3](#). Sample $(d+1)$ k -sets S_{-1}, S_0, \dots, S_d using the following process.

1. Choose $t+1$ distinct random elements a_0, a_1, \dots, a_t .
2. For $\ell \in \{0, 1, \dots, d\}$
 - Choose a random element a'_ℓ distinct from the t -set $\{a'_i \mid 0 \leq i < \ell\} \cup \{a_i \mid \ell < i \leq t\}$.
3. Choose a random $(k-t-1)$ -set W_{-1} disjoint from the $(t+1)$ -set $\{a_0, \dots, a_t\}$.
4. Set $S_{-1} \leftarrow \{a_0, \dots, a_t\} \cup W_{-1}$.
5. For $\ell \in \{0, 1, \dots, d\}$
 - Choose a random $(k-t-1)$ -set W_ℓ disjoint from the $(t+1)$ -set $\{a'_i \mid i \leq \ell\} \cup \{a_i \mid i > \ell\}$.
 - Set $S_\ell \leftarrow \{a'_i \mid i \leq \ell\} \cup \{a_i \mid i > \ell\} \cup W_\ell$.

- Set $T_\ell \leftarrow \{a'_i \mid i < \ell\} \cup \{a_i \mid i > \ell\}$.
- Set $R_\ell \leftarrow \{a_i \mid i > \ell\}$.

Such a coupling is feasible as long as $k \geq t + 1$. It follows from the definition that for any $\ell \in \{0, 1, \dots, d\}$, the triple $(T_\ell, S_{\ell-1}, S_\ell)$ is distributed according to $\nu_n(k, t)$ while the triple (R_ℓ, S_{-1}, S_ℓ) is distributed according to $\nu_n(k, t - \ell)$. Furthermore, observe that $R_\ell \subseteq T_i$ for all $i \leq \ell$. Hence, if $f_{S_{-1}}|_{R_\ell} \neq f_{S_\ell}|_{R_\ell}$, it must be the case that for some $i \in \{1, 2, \dots, \ell\}$, we have $f_{S_{i-1}}|_{T_i} \neq f_{S_i}|_{T_i}$. Hence,

$$\Pr[f_{S_{-1}}|_{R_\ell} \neq f_{S_\ell}|_{R_\ell}] \leq \sum_{i=0}^{\ell} \Pr[f_{S_{i-1}}|_{T_i} \neq f_{S_i}|_{T_i}] = (\ell + 1) \Pr_{(T, S, S') \sim \nu_n(k, t)}[f_S|_T \neq f_{S'}|_T]. \quad \square$$

Claim 4.4. *For every positive integer d , there exists a positive constant $C'_d = 2^{O(d)}$ such that for integers ℓ, t, k, n satisfying $0 \leq \ell \leq d \leq t < k \leq n$ the following holds.*

$$\mathbb{E}_U[\gamma_\ell] = \mathbb{E}_U \left[\Pr_{S \supset U} [S \text{ is } \ell\text{-good but not } \ell\text{-excellent}] \right] \leq \frac{C'_d \cdot \varepsilon}{\zeta^\ell}.$$

Proof. The proof of this claim is, in some sense, a generalization of the proof of [Claim 3.5](#). To begin with, let us fix a $(t - d)$ -set U .

$$\begin{aligned} \gamma_\ell &= \Pr_{S \supset U} [S \text{ is } \ell\text{-good but not } \ell\text{-excellent}] \\ &= \Pr_{S \supset U} [S \text{ is } \ell\text{-good and } \exists \ell\text{-set } L \subset \bar{U} \text{ such that } S \supset L \text{ and } f_S[L] \neq g[L]] && \text{[Definition of } \ell\text{-excellent set]} \\ &\leq \sum_{L: |L|=\ell, L \subset \bar{U}} \Pr_{S \supset U} [S \supset L \text{ and } S \text{ is } \ell\text{-good and } f_S[L] \neq g[L]] && \text{[Union Bound]} \\ &\leq \sum_{\substack{L: |L|=\ell \\ L \text{ good}, L \subset \bar{U}}} \Pr_{S \supset U} [S \supset L \text{ and } S \text{ is } \ell\text{-good and } f_S[L] \neq g[L]] && [S \text{ is } \ell\text{-good} \ \& \ L \subset S \Rightarrow L \text{ is good}] \\ &\leq \sum_{\substack{L: |L|=\ell \\ L \text{ good}, L \subset \bar{U}}} \underbrace{\Pr_{S \supset U} [S \supset L \text{ and } S \text{ is } (\ell - 1)\text{-excellent and } f_S[L] \neq g[L]]}_{\gamma_\ell(L)} && [S \text{ is } \ell\text{-good} \Rightarrow S \text{ is } (\ell - 1)\text{-excellent}] \end{aligned}$$

We now turn to bounding each term $\gamma_\ell(L)$ in the above sum for good L .

$$\begin{aligned} \gamma_\ell(L) &:= \Pr_{S \supset U} [S \supset L \text{ and } S \text{ is } (\ell - 1)\text{-excellent and } f_S[L] \neq g[L]] \\ &= \Pr_{S \supset U} [S \supset L] \cdot \Pr_{S \supset U \cup L} [S \text{ is } (\ell - 1)\text{-excellent}] \cdot \Pr_{S \supset U \cup L} [f_S[L] \neq g[L] \mid S \text{ is } (\ell - 1)\text{-excellent}] \\ &\leq \Pr_{S \supset U} [S \supset L] \cdot \Pr_{S \supset U \cup L} [S \text{ is } (\ell - 1)\text{-excellent}] \cdot \Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \mid S_1, S_2 \text{ are } (\ell - 1)\text{-excellent}] \\ &= \Pr_{S \supset U} [S \supset L] \cdot \frac{\Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \text{ and } S_1, S_2 \text{ are } (\ell - 1)\text{-excellent}]}{\Pr_{S \supset U \cup L} [S \text{ is } (\ell - 1)\text{-excellent}]} \\ &\leq 2 \Pr_{S \supset U} [S \supset L] \cdot \Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \text{ and } S_1, S_2 \text{ are } (\ell - 1)\text{-excellent}], \end{aligned}$$

where in the last step we have used the fact that L is good. Plugging this back into the earlier expression,

we have

$$\begin{aligned}
\gamma_\ell &\leq \sum_{\substack{L:|L|=\ell \\ L \text{ good}, L \subset \bar{U}}} \gamma_\ell(L) \leq 2 \sum_{\substack{L:|L|=\ell \\ L \text{ good}, L \subset \bar{U}}} \Pr[S \supset L] \cdot \Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \text{ and } S_1, S_2 \text{ are } (\ell-1)\text{-excellent}] \\
&\leq 2 \sum_{L:|L|=\ell, L \subset \bar{U}} \Pr[S \supset L] \cdot \Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \text{ and } S_1, S_2 \text{ are } (\ell-1)\text{-excellent}] \\
&= 2 \binom{k'}{\ell} \mathbb{E}_{L:|L|=\ell, L \subset \bar{U}} \left[\Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \text{ and } S_1, S_2 \text{ are } (\ell-1)\text{-excellent}] \right] \\
&\leq 2 \binom{k'}{\ell} \mathbb{E}_{L:|L|=\ell, L \subset \bar{U}} \left[\Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}[L] \neq f_{S_2}[L] \text{ and } f_{S_1}^{\ell-1} \approx f_{S_2}^{\ell-1}] \right].
\end{aligned}$$

We now average over U :

$$\begin{aligned}
\mathbb{E}_U [\gamma_\ell] &\leq 2 \binom{k'}{\ell} \mathbb{E}_{U, L \in \binom{\bar{U}}{\ell}} \left[\Pr_{S_1, S_2 \supset U \cup L} [f_{S_1}^U[L] \neq f_{S_2}^U[L] \text{ and } f_{S_1}^U \approx^{\ell-1} f_{S_2}^U] \right] \\
&\leq 2 \binom{k'}{\ell} \mathbb{E}_{\substack{R:|R|=t-d+\ell \\ L \in \binom{R}{\ell}}} \left[\Pr_{S_1, S_2 \supset R} [f_{S_1}|_R \neq f_{S_2}|_R \text{ and } f_{S_1}^{R \setminus L} \approx^{\ell-1} f_{S_2}^{R \setminus L}] \right] \\
&= 2 \binom{k'}{\ell} \Pr_{\substack{R:|R|=t-d+\ell \\ S_1, S_2 \supset R}} [f_{S_1}|_R \neq f_{S_2}|_R] \cdot \mathbb{E}_{\substack{R:|R|=t-d+\ell \\ S_1, S_2 \supset R}} \left[\Pr_{L \in \binom{R}{\ell}} [f_{S_1}^{R \setminus L} \approx^{\ell-1} f_{S_2}^{R \setminus L} \mid f_{S_1}|_R \neq f_{S_2}|_R] \right]
\end{aligned}$$

The first probability expression in the final expression above is upper bounded by $(\ell+1)\varepsilon$. The second probability expression is bounded as follows: Consider any $R, S_1, S_2 \supset R$ such that $f_{S_1}|_R \neq f_{S_2}|_R$. In particular, let $A \subset R$ be any d -set such that $f_{S_1}(A) \neq f_{S_2}(A)$. Now, if $f_{S_1}^{R \setminus L} \approx^{\ell-1} f_{S_2}^{R \setminus L}$, we must have that $L \subset A$. In other words, the second probability expression is bounded above $\Pr_{L: L \supset R, |L|=\ell} [L \subset A] = \binom{d}{\ell} / \binom{t-d+\ell}{\ell}$. We thus, have

$$\mathbb{E}_U [\gamma_\ell] \leq \frac{2^\ell \binom{k'}{\ell} \cdot \binom{d}{\ell}}{\binom{t-d+\ell}{\ell}} \cdot O(\varepsilon) = \frac{2^{O(d)}}{\zeta^\ell} \cdot \varepsilon. \quad \square$$

Claim 4.5. For every positive integer d , there is a positive constant $C_d = 2^{O(d^3)}$ such that for all $\ell = 0, \dots, d$, we have

$$\Pr_{S \supset U} [S \text{ is not } \ell\text{-good}] \leq C_d \cdot \Pr_{S \supset U} [S \text{ is not } (\ell-1)\text{-excellent}].$$

Proof. This claim is proved using the Reverse Union Bound as in the proof of [Claim 3.4](#). Let $C_d := \max_\ell \{c'(\ell, 1/2)\} + 1 = 2^{O(d^3)}$, where $c'(d, c)$ is the constant guaranteed by the Reverse Union Bound, [Lemma 1.4](#). Define $Y \subset \binom{[n]}{k}$ and $B \subset \binom{[n]}{\ell}$ as follows:

$$\begin{aligned}
Y &:= \left\{ S \in \binom{[n]}{k} \mid S \text{ is not } (\ell-1)\text{-excellent} \right\}, \\
B &:= \left\{ L \in \binom{[n]}{\ell} \mid L \text{ is bad} \right\}.
\end{aligned}$$

By definition of bad set, we have that

$$\Pr_{S \supset U} [S \in Y \mid S \supseteq L] \geq \frac{1}{2}.$$

Hence, by reverse union bound,

$$\begin{aligned} \Pr_{S \supset U} [S \text{ is not } \ell\text{-good}] &\leq \Pr[S \text{ hits } B] + \Pr_{S \supset U} [S \text{ is not } (\ell - 1)\text{-excellent}] \\ &\leq (C_d - 1) \cdot \Pr_{S \supset U} [S \text{ is not } (\ell - 1)\text{-excellent}] + \Pr_{S \supset U} [S \text{ is not } (\ell - 1)\text{-excellent}]. \quad \square \end{aligned}$$

We now return to our goal of bounding $\Pr_{U, S \supset U} [f_S \neq g|_S]$ via the inequality (2):

$$\begin{aligned} \mathbb{E}_U \Pr_{S \supset U} [f_S \neq g|_S] &\leq \mathbb{E}_U \Pr_{S \supset U} [S \text{ is not } d\text{-excellent}] \\ &\leq \mathbb{E}_U \Pr_{S \supset U} [S \text{ is not } d\text{-good}] + \mathbb{E}_U \Pr_{S \supset U} [S \text{ is } d\text{-good but not } d\text{-excellent}] \\ &\leq C_d \cdot \mathbb{E}_U \Pr_{S \supset U} [S \text{ is not } (d - 1)\text{-excellent}] + \frac{C'_d \cdot \varepsilon}{\zeta^d} \\ &\leq C_d^2 \cdot \mathbb{E}_U \Pr_{S \supset U} [S \text{ is not } (d - 2)\text{-excellent}] + C_d \cdot \frac{C'_d \cdot \varepsilon}{\zeta^{d-1}} + \frac{C'_d \cdot \varepsilon}{\zeta^d} \\ &\leq C_d^\ell \cdot \mathbb{E}_U \Pr_{S \supset U} [S \text{ is not } (d - \ell)\text{-excellent}] + \sum_{i=1}^{\ell} C_d^{\ell-i} \cdot \frac{C'_d \cdot \varepsilon}{\zeta^{d-(\ell-i)}} \\ &\leq C_d^{d+1} \cdot \mathbb{E}_U \Pr_{S \supset U} [S \text{ is not } (-1)\text{-excellent}] + \sum_{i=1}^{d+1} C_d^{d+1-i} \cdot \frac{C'_d \cdot \varepsilon}{\zeta^{i-1}} \\ &= \frac{2^{O(d^4)} \cdot \varepsilon}{\zeta^d}. \quad \square \end{aligned}$$

4.2 Part 2: sewing

In this step, we show that the global function g_U corresponding to some $(t - d)$ -set U explains most local functions f_S corresponding to S 's not necessarily containing U . We will first prove this under the assumption that $k \geq 2(d + 1)t$ (or equivalently $\zeta \geq 1 - 1/2(d+1)$) and then extend it to all $\zeta \in (0, 1)$.

Lemma 4.6. *For every positive integer d , let $\kappa_d = 2^{O(d^4)}$ be the positive constant from Lemma 4.2. For all positive integers n, k, t satisfying $d \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: \binom{S}{d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon, \quad (3)$$

then there exist a global function $g: \binom{[n]}{d} \rightarrow \Sigma$ such that

$$\Pr_S [f_S \neq g|_S] \leq \frac{2(2d + 4)\kappa_d}{\xi \zeta^d} \cdot \varepsilon,$$

where $\zeta = \min\{1, t/(k - t + 1)\}$ and $\xi = 1 - t/k$.

Proof. Let us first assume that $k \geq 2(d + 1)t$. Lemma 4.2 implies that

$$\Pr_{\substack{U_1, U_2 \\ S \supset U_1 \cup U_2}} [g_{U_1}|_S \neq g_{U_2}|_S] \leq \frac{2\kappa_d}{\zeta^d} \cdot \varepsilon.$$

We now define a coupling (U_1, U_2, S, S_1, S_2) as follows: Choose two random $(t - d)$ -sets U_1, U_2 .

Then, choose a random k -set S containing U_1 . Randomly partition $S \setminus U_1$ into $d+1$ sets R_1, \dots, R_{d+1} of (almost) equal size $\frac{k-t+d}{d+1}$. Finally, for each $j \in [d+1]$, choose a random k -set S_j containing $U_1 \cup U_2 \cup [(S \setminus U_1) \setminus R_j]$. The last step is possible since $2(t-d) + \frac{d}{d+1}(k-t+d) \leq k$. The marginal distribution of (S_j, U_1, U_2) is that of two random sets U_1, U_2 and a random set $S_j \supset U_1 \cup U_2$. The coupling ensures that $\binom{S}{d}$ is covered by $\bigcup_j \binom{S_j}{d}$ (since for any d elements in S , there must be some R_j missing). Hence, $g_{U_1}|_S \neq g_{U_2}|_S$ implies that $g_{U_1}|_{S_j} \neq g_{U_2}|_{S_j}$ for some $j \in \{1, \dots, d+1\}$. Therefore

$$\Pr_{\substack{U_1, U_2 \\ S \supset U_1}} [g_{U_1}|_S \neq g_{U_2}|_S] = \sum_{j=1}^{d+1} \Pr_{\substack{U_1, U_2 \\ S_j \supset U_1 \cup U_2}} [g_{U_1}|_{S_j} \neq g_{U_2}|_{S_j}] \leq \frac{2(d+1)\kappa_d}{\zeta^d} \cdot \varepsilon.$$

Fix $U := U_2$ such that

$$\Pr_{U_1, S \supset U_1} [g_{U_1}|_S \neq g_U|_S] \leq \frac{2(d+1)\kappa_d}{\zeta^d} \cdot \varepsilon.$$

Then $g = g_U$ satisfies

$$\begin{aligned} \Pr_{U_1, S \supset U_1} [f_S \neq g|_S] &\leq \Pr_{U_1, S \supset U_1} [f_S \neq g_{U_1}|_S] + \Pr_{U_1, S \supset U_1} [g_{U_1}|_S \neq g|_S] \\ &\leq \frac{2\kappa_d}{\zeta^d} \cdot \varepsilon + \frac{2(d+1)\kappa_d}{\zeta^d} \cdot \varepsilon = \frac{(2d+4)\kappa_d}{\zeta^d} \cdot \varepsilon. \end{aligned}$$

This completes the case of small t .

Suppose now that $t \geq k/2(d+1)$, and let $t' = \lfloor k/2(d+1) \rfloor$. This case is handled identically to the corresponding case in dimension one. For any $0 < \ell \leq t$, define

$$\varepsilon_\ell := \Pr_{(T, S_1, S_2) \sim \nu_n(k, \ell)} [f_{S_1}|_T \neq f_{S_2}|_T].$$

Mirroring the argument for dimension one, we obtain that if the hypothesis for the agreement theorem is true for a particular choice of n, k, t (i.e., $\varepsilon_t \leq \varepsilon$), then the hypothesis is also true for n, k, r by increasing ε to 2ε (i.e., $\varepsilon_r \leq 2\varepsilon$) provided $r \geq 2t - k$, or equivalently $2(1 - t/k) \geq (1 - r/k)$. Thus, given the hypothesis is true for some t, k satisfying $\zeta = 1 - t/k$, we can perform this argument $m := \lceil \log_2 1/\zeta \cdot (1 - 1/2(d+1)) \rceil$ times to reduce t to less than $k/2(d+1)$. Each step doubles the error, and so the overall error probability becomes $2^m \cdot \varepsilon \leq 2/\zeta(1 - 1/2(d+1)) \cdot \varepsilon$. Now applying the argument for small t , we obtain that there exists a global g such that

$$\Pr_S [f_S \neq g|_S] \leq \frac{(2d+4)\kappa_d}{\zeta^d} \cdot \frac{2}{\zeta} \left(1 - \frac{1}{2(d+1)}\right) \cdot \varepsilon \leq \frac{2(2d+4)\kappa_d}{\zeta \zeta^d} \cdot \varepsilon. \quad \square$$

4.3 Part 3: majority-decoding

The proof is identical to the dimension one case.

Proof of Theorem 4.1. Let g be the global function given by Lemma 4.6 that satisfies

$$\Pr_S [f_S \neq g|_S] \leq \frac{2(2d+4)\kappa_d}{\zeta \zeta^d} \cdot \varepsilon =: \eta.$$

Define $G: \binom{[n]}{d} \rightarrow \Sigma$ as follows: for each $A \in \binom{[n]}{d}$ set $G(A)$ to be the most frequently occurring value among $\{f_S(A) \mid S \supset A\}$ (breaking ties arbitrarily).

Let $B = \left\{ A \in \binom{[n]}{d} \mid g(A) \neq G(A) \right\}$, and let $Y = \{S \mid f_S \neq g|_S\}$.
 For each $A \in B$,

$$\Pr_S [S \in Y \mid A \subset S] \geq \Pr_S [f_S(A) \neq g(A) \mid A \subset S] \geq \frac{1}{2},$$

since otherwise $g(A)$ would be the unique majority value. Applying the Reverse Union Bound [Lemma 1.4](#), we conclude that

$$\Pr_S [S \text{ hits } B] \leq C'(d, 1/2) \cdot \Pr_S [f_S \neq g|_S] \leq C'(d, 1/2) \cdot \eta.$$

Therefore

$$\Pr_S [f_S \neq G|_S] \leq \Pr_S [f_S \neq g|_S] + \Pr_S [g|_S \neq G|_S] = \Pr_S [f_S \neq g|_S] + \Pr_S [S \text{ hits } B] = (C'(d, 1/2) + 1) \cdot \eta.$$

[Theorem 4.1](#) follows. \square

5 More agreement theorems

In this section, we present several variants of the d -dimensional agreement theorem [Theorem 4.1](#).

The first variant will be the simplest variant in which the local function is of the form

$$f_S: \binom{S}{\leq d} \rightarrow \Sigma$$

I.e., it specifies a value for all ℓ -subsets A of S such that $\ell \leq d$ instead of just d -subsets of S . This variation is obtained by a simple union bound over [Theorem 4.1](#), one for each integer $\ell \in [0, d]$.

Theorem 5.1. *For every positive integer d , let $M_d = 2^{O(d^4)}$ be the positive constant from [Theorem 4.1](#). For all positive integers n, k, t satisfying $d \leq t < k \leq n$ and alphabet Σ the following holds. Let $\{f_S: \binom{S}{\leq d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \nu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon,$$

then the majority decoded function $G: \binom{[n]}{\leq d} \rightarrow \Sigma$ satisfies

$$\Pr_{S \sim \nu_n(k)} [f_S \neq G|_S] \leq \frac{(d+1)M_d}{\zeta \zeta^d} \cdot \varepsilon,$$

where $\zeta = \min\{1, t/(k-t+1)\}$ and $\xi = 1 - t/k$, and the majority-decoded function G is the one given by “popular vote”, namely for each $A \in \binom{[n]}{\leq d}$ set $G(A)$ to be the most frequently occurring value among $\{f_S(A) \mid S \supset A\}$ (breaking ties arbitrarily).

5.1 The exact intersection setting

The d -dimensional agreement theorems ([Theorem 4.1](#) and [Theorem 5.1](#)) are both proved when the triple of sets (S_1, S_2, T) are picked according to the distribution $\nu_n(k, t)$. We now extend the agreement to the exact intersection setting when the triple of sets (S_1, S_2, T) are instead picked according to the distribution $\mu_n(k, t)$ as follows:

- Pick a random t -set $T \subseteq [n]$.

- Pick two k -sets S_1 and S_2 uniformly from the set of pairs of k -sets that have exact intersection T .

In other words,

$$(S_1, S_2) \leftarrow_u \left\{ (S'_1, S'_2) \in \binom{[n]}{k} \times \binom{[n]}{k} \mid S'_1 \cap S'_2 = T \right\}.$$

Note that for the distribution $\mu_n(k, t)$ to have non-zero support, we must have $t + 2(k - t) = 2k - t \leq n$.

Theorem 5.2 (exact-intersection setting (strengthened form of [Theorem 1.2](#))). *For every positive integer d , let $M_d = 2^{O(d^4)}$ be the positive constant from [Theorem 4.1](#). Let $\rho \in (0, 1)$. For all positive integers n, k, t satisfying $d \leq t < k \leq n$ and $2k - t \leq (1 - \rho)n$ and any alphabet Σ the following holds. Let $\{f_S : \binom{[n]}{\leq d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ be an ensemble of local functions satisfying*

$$\Pr_{(T, S_1, S_2) \sim \mu_n(k, t)} [f_{S_1}|_T \neq f_{S_2}|_T] \leq \varepsilon,$$

then the majority decoded function $G : \binom{[n]}{\leq d} \rightarrow \Sigma$ satisfies

$$\Pr_{S \sim \nu_n(k)} [f_S \neq G|_S] \leq (1/\rho)^{O(1)} \cdot \frac{(d+1)M_d}{\zeta \zeta^d} \cdot \varepsilon,$$

where $\zeta = \min\{1, t/(k - t + 1)\}$ and $\zeta = 1 - t/k$, and the majority-decoded function G is the one given by “popular vote”, namely for each $A \in \binom{[n]}{\leq d}$ set $G(A)$ to be the most frequently occurring value among $\{f_S(A) \mid S \supset A\}$ (breaking ties arbitrarily).

Proof. For every $0 \leq r \leq k - t$, we define the following intermediate distributions $\nu_n(k, t, r)$ on triples of sets as follows.

- Choose a random t -set T .
- Choose a disjoint random r -set R .
- Choose two random k -sets $S_1, S_2 \subseteq T \cup R$ whose intersection is exactly $T \cup R$.
- Output (S_1, S_2, T) .

Observe that $\mu_n(k, t) = \nu_n(k, t, 0)$.

Given an ensemble $\mathcal{F} = \{f_S : \binom{[n]}{\leq d} \rightarrow \Sigma \mid S \in \binom{[n]}{k}\}$ of local functions, define for each $r \in [0, k - t]$,

$$\varepsilon_{n, k, t}^r(\mathcal{F}) := \Pr_{(T, S_1, S_2) \sim \nu_n(k, t, r)} [f_{S_1}|_T \neq f_{S_2}|_T]. \quad (4)$$

We are given that $\varepsilon_{n, k, t}^0(\mathcal{F}) \leq \varepsilon$. Applying [Claim 5.3](#) t many times, we get that $\varepsilon_{n, k, t}^r(\mathcal{F}) \leq 3^t \varepsilon$ for all $r \leq (2^t - 1)\rho n$, as a simple induction shows. We are interested in a bound that works for all $r \leq k - t$, and for that it suffices to take $t_0 = O(\log \frac{k-t}{\rho n}) = O(\log(1/\rho))$. Therefore $\varepsilon_{n, k, t}^r(\mathcal{F}) \leq 3^{t_0} \varepsilon = (1/\rho)^{O(1)} \varepsilon$. The theorem follows since $\nu_n(k, t)$ is a convex combination of $\nu_n(k, t, r)$ for $r \leq k - t$. \square

Claim 5.3. *Suppose $2k - t \leq (1 - \rho)n$. If $t + r + \Delta \leq k$ and $\Delta - r \leq \rho n$, then $\varepsilon_{n, k, t}^{r+\Delta}(\mathcal{F}) \leq 3 \cdot \varepsilon_{n, k, t}^r(\mathcal{F})$.*

Proof. We prove this using the following coupling argument:

- Choose disjoint sets T of size t , R of size r , D_1, D_2, D_3 of size Δ , and A, B of size $k - t - r - \Delta$.
- Set $S_1 \leftarrow T \cup R \cup D_1 \cup A$.

- Let $S_2 \leftarrow T \cup R \cup D_2 \cup B$.
- Let $S_3 \leftarrow T \cup R \cup D_3 \cup A$.
- Let $S_4 \leftarrow T \cup R \cup D_1 \cup B$.

We can choose such sets if $t + r + 3\Delta + 2(k - t - r - \Delta) \leq n$, which translates to $2k - t + (\Delta - r) \leq n$ which is true since $2k - t \leq (1 - \rho)n$ and $\Delta - r \leq \rho n$.

All the sets contain T . The sets S_1, S_4 are random supersets of T whose intersection is exactly $t + r + \Delta$. Thus, the marginal of the triple (S_1, S_4, T) is exactly $\nu_n(k, t, r + \Delta)$. Every two adjacent sets S_i, S_{i+1} ($i \in \{1, 2, 3\}$) are random supersets of T whose intersection is exactly $t + r$. Thus, the marginal of the triple (S_i, S_{i+1}, T) is exactly $\nu_n(k, t, r)$. Clearly, if $f_{S_1}|_T \neq f_{S_4}|_T$, then it must be the case that for some $i \in \{1, 2, 3\}$ we have $f_{S_i}|_T \neq f_{S_{i+1}}|_T$. Hence, $\epsilon_{n,k,t}^{r+\Delta}(\mathcal{F}) \leq 3 \cdot \epsilon_{n,k,t}^r(\mathcal{F})$. \square

5.2 The p -biased distribution

The d -dimensional agreement theorems (Theorems 4.1, 5.1 and 5.2) are all proved with respect to the distribution $\nu_n(k)$, the uniform distribution over k -sized subsets of $[n]$. In this section, we extend these results to the p -biased setting μ_p . In this setting, the distribution $\nu_n(k, t)$ is replaced by the distribution $\mu_{p,\alpha}$, which is a distribution over pairs S_1, S_2 of subsets of $[n]$ defined as follows. For each element x independently, we put x only in S_1 or only in S_2 with probability $p(1 - \alpha)$ (each), and we put x in both with probability $p\alpha$. This is possible if $p(2 - \alpha) \leq 1$. Note that if sets S_1, S_2 are picked according to the distribution $\mu_{p,\alpha}$ then the marginal distribution of each of S_1 and S_2 is μ_p , while the distribution of the intersection $S_1 \cap S_2$ is $\mu_{p\alpha}$. We present below the p -biased version of Theorem 5.2.

Theorem 5.4 (p -biased setting). *For every positive integer d , let $M_d = 2^{O(d^4)}$ be the positive constant from Theorem 4.1 and $\rho \in (0, 1)$. For all positive numbers $p, \alpha \in (0, 1)$ satisfying $p(2 - \alpha) \leq 1 - \rho$, positive integer n and alphabet Σ the following holds. Let $\{f_S : \binom{[n]}{\leq d} \rightarrow \Sigma \mid S \subseteq [n]\}$ be an ensemble of local functions satisfying*

$$\Pr_{(S_1, S_2) \sim \mu_{p,\alpha}} [f_{S_1}|_{S_1 \cap S_2} \neq f_{S_2}|_{S_1 \cap S_2}] \leq \epsilon,$$

then the majority decoded function $G : \binom{[n]}{\leq d} \rightarrow \Sigma$ satisfies

$$\Pr_{S \sim \mu_p} [f_S \neq G|_S] \leq (1/\rho)^{O(1)} \cdot \frac{3(d+1)M_d}{\min\{(1-\alpha), \alpha^d\}} \cdot \epsilon,$$

where the majority-decoded function G is the one given by “popular vote”, namely for each $A \in \binom{[n]}{\leq d}$ set $G(A)$ to be the most frequently occurring value among $\{f_S(A) \mid S \supset A\}$ (breaking ties arbitrarily).

Proof of Theorem 5.4. Let N be a large integer and define $K = \lfloor pN \rfloor$, $T = \lfloor p\alpha N \rfloor$. For every $S \in \binom{[N]}{K}$, define $\tilde{f}_S := f_{S \cap [n]}$. In other words, for all $A \subset S \in \binom{[N]}{K}$, $|A| \leq d$, let $\tilde{f}_S(A) := f_{S \cap [n]}(A \cap [n])$. If $S_1, S_2 \sim \mu_N(K, T)$ then the distribution of $S_1 \cap [n], S_2 \cap [n]$ is close to $\mu_{p,\alpha}$, and so for large enough N we have

$$\Pr_{S_1, S_2 \sim \mu_N(K, T)} [\tilde{f}_{S_1}|_{S_1 \cap S_2} \neq \tilde{f}_{S_2}|_{S_1 \cap S_2}] \leq 2\epsilon.$$

Hence, the ensemble of functions $\{\tilde{f}_S\}_{S \in \binom{[N]}{K}}$ satisfies the hypothesis of the agreement theorem (Theorem 4.1) with ϵ replaced by 2ϵ . Hence, by Theorem 4.1, if we define $\tilde{G} : \binom{[N]}{\leq d} \rightarrow \Sigma$ by plurality decoding then $\Pr_{S \sim \mu_N(K)} [\tilde{f}_S \neq \tilde{G}|_S] = O_d(\epsilon)$. Since \tilde{f}_S depends only on $S \cap [n]$, there exists a function $\hat{G} : \binom{[n]}{\leq d} \rightarrow \Sigma$

such that $\tilde{G}(T) = \hat{G}(T \cap [n])$. Moreover, for large enough N the distribution of $S \cap [n]$ approaches μ_p , and so $\hat{G} = G$.² This completes the proof. \square

yf: Expanded on the issue of several most common values.

5.3 Product distributions (for large p)

The multiplicative decay in the d -dimensional agreement theorem, [Theorem 5.4](#), deteriorates as α goes very close to 1. For the application to testing Reed–Muller codes, we need an agreement theorem that holds for all $\alpha \in (0, 1)$ but for $p \geq p_0$ some fixed constant. The following agreement theorem over the product distribution $\mu_{p,p}$ shows that it is easy to prove a counterpart of [Theorem 5.4](#) if one is allowed a multiplicative decay of p^{-d} . Note that for $p \geq p_0$, this loss is not an issue for our applications (as we think of d as a constant and $p \geq p_0$ is a constant). This is no longer true when $p = o(1)$, a regime in which we need the stronger result proved in [Theorem 5.4](#), which is independent of p (but still depends on d).

Theorem 5.5. *For every positive integer d and alphabet Σ , the following holds for all p . If $\{f_S: \binom{S}{\leq d} \rightarrow \Sigma \mid S \in \{0, 1\}^n\}$ is an ensemble of functions satisfying*

$$\Pr_{S_1, S_2 \sim \mu_{p,p}} [f_{S_1}|_{S_1 \cap S_2} \neq f_{S_2}|_{S_1 \cap S_2}] = \varepsilon$$

then the global function $G: \binom{[n]}{\leq d} \rightarrow \Sigma$ defined by plurality decoding satisfies

$$\Pr_{S \sim \mu_p} [f_S \neq G|_S] \leq 2p^{-d}\varepsilon.$$

Proof. The main observation behind the proof is this: if we choose $S_1, S_2 \sim \mu_p$ independently, then $(S_1, S_2) \sim \mu_{p,p}$.

Consider now an arbitrary f_{S_1} . Suppose that $f_{S_1} \neq G|_{S_1}$, and choose an entry T such that $f_{S_1}(T) \neq G(T)$. If we choose $S_2 \sim \mu_p$, then $T \subseteq S_2$ with probability $p^{|T|} \geq p^d$. Moreover, since $f_{S_1}(T) \neq G(T)$, the probability that $f_{S_2}(T) = f_{S_1}(T)$ is at most $1/2$. Therefore the probability that f_{S_1} and f_{S_2} disagree on their intersection is at least $p^d/2$. This shows that

$$\varepsilon = \Pr_{S_1, S_2 \sim \mu_{p,p}} [f_{S_1}|_{S_1 \cap S_2} \neq f_{S_2}|_{S_1 \cap S_2}] \geq \frac{p^d}{2} \Pr_{S_1 \sim \mu_p} [f_{S_1} \neq G|_{S_1}]. \quad \square$$

6 Testing Reed–Muller codes

Every Boolean function $f: \{0, 1\}^n \rightarrow \{0, 1\}$ can be written in a unique way as $P \bmod 2$, where P is a *Boolean polynomial*, that is, a sum of distinct multilinear monomials. The *Boolean degree* of f , denoted $\text{bdeg}(f)$, is the degree of this polynomial.

The well-known BLR test [[BLR93](#), [BCHKS96](#)] checks whether a given Boolean function has Boolean degree 1. Alon *et al.* [[AKKLR05](#)] developed the following 2^{d+1} -query test T_d , which is a generalization of the BLR test to large Boolean degrees.

²There's a fine point here: there could be several most common values, only some of which could be strictly most common for \tilde{f} . However, any most common value of f is almost most common \tilde{f} , and the proof for the uniform setting can be straightforwardly adapted to handle such values.

Algorithm 2: AKKLR LOW DEGREE TEST T_d

Input: $f: \{0,1\}^n \rightarrow \{0,1\}$ (provided as an oracle)

- 1 Pick $x, a_1, \dots, a_{d+1} \in \{0,1\}^n$ independently from the distribution $\mu_{1/2}^{\otimes n}$, subject to the constraint that a_1, \dots, a_{d+1} are linearly independent.
- 2 Accept iff

$$\sum_{I \subseteq [d+1]} f\left(x + \sum_{i \in I} a_i\right) = 0 \pmod{2}.$$

This test is closely related to the Gowers norms. An optimal analysis of the test was provided by Bhattacharyya *et al.* [BKSSZ10]. We need a few definitions to state their result.

Definition 6.1. Let $f: \{0,1\}^n \rightarrow \{0,1\}$ and let $d \geq 0$. The distance to degree d of f is defined as follows:

$$\delta_d(f) := \min_{\text{bdeg}(g) \leq d} \left\{ \Pr_{x \sim \mu_{1/2}^{\otimes n}} [f(x) \neq g(x)] \right\}.$$

Definition 6.2. Let $f: \{0,1\}^n \rightarrow \{0,1\}$ and let $d \geq 0$. The failure probability of test T_d is

$$\text{rej}_d(f) := \Pr_{\substack{x, a_1, \dots, a_{d+1} \sim \mu_{1/2}^{\otimes n} \\ a_1, \dots, a_{d+1} \text{ linearly independent}}} \left[\sum_{I \subseteq [d+1]} f\left(x + \sum_{i \in I} a_i\right) \neq 0 \pmod{2} \right].$$

Remark 6.3. Another variant of T_d , which is closer to the definition of the Gowers norms, samples a_1, \dots, a_{d+1} without requiring them to be linearly independent. When a_1, \dots, a_{d+1} are linearly dependent, the test always succeeds. On the other hand, when $n \geq d+1$, the probability that a_1, \dots, a_{d+1} are linearly dependent is upper-bounded by a positive constant. Therefore when $n \geq d+1$, removing the constraint of linear independence only affects the rejection probability by a constant factor. Finally, when $n \leq d$ the test is pointless, since every function has Boolean degree at most d .

Theorem 6.4 ([BKSSZ10]). For every integer $d \geq 1$ there exists a constant $\varepsilon_d > 0$ such that for all Boolean functions $f: \{0,1\}^n \rightarrow \{0,1\}$,

$$\text{rej}_d(f) \geq \min(2^d \delta_d(f), \varepsilon_d).$$

Corollary 6.5. For every integer $d \geq 1$ and all Boolean functions $f: \{0,1\}^n \rightarrow \{0,1\}$,

$$\delta_d(f) = O_d(\text{rej}_d(f)).$$

Proof. If $2^d \delta_d(f) \leq \varepsilon_d$ then $\delta_d(f) \leq 2^{-d} \text{rej}_d(f)$. Else, $\text{rej}_d(f) \geq \varepsilon_d$, and so $\delta_d(f) \leq 1 \leq \varepsilon_d^{-1} \text{rej}_d(f)$. \square

Our goal in this section is to extend the analysis of Bhattacharyya *et al.* to the μ_p setting (wherein we measure closeness of f to degree d with respect to the μ_p measure instead of the $\mu_{1/2}$ measure). More precisely:

Definition 6.6. Let $f: \{0,1\}^n \rightarrow \{0,1\}$ and let $d \geq 0$. The μ_p -distance to degree d of f is defined as follows:

$$\delta_d^{(p)}(f) := \min_{\text{bdeg}(g) \leq d} \left\{ \Pr_{x \sim \mu_p^{\otimes n}} [f(x) \neq g(x)] \right\}.$$

To this end, we consider the natural extension $T_{p,d}$ (Algorithm 1) of the AKLR test T_d to the μ_p measure (restated below from the introduction).

Algorithm 3: p -BIASED LOW DEGREE TEST $T_{p,d}$

Input: $f: \{0,1\}^n \rightarrow \{0,1\}$ (provided as an oracle)

- 1 Pick $S \subseteq [n]$ according to the distribution μ_{2p} .
- 2 If $|S| \leq d$, then accept.
- 3 Let $f|_S: \{0,1\}^S \rightarrow \{0,1\}$ denote the restriction of f to $\{0,1\}^S$ by zeroing out all the coordinates outside S .
- 4 Pick $x, a_1, \dots, a_{d+1} \in \{0,1\}^S$ independently from the distribution $\mu_{1/2}^{\otimes S}$, subject to the constraint that a_1, \dots, a_{d+1} are linearly independent.
- 5 Accept iff

$$\sum_{I \subseteq [d+1]} f|_S \left(x + \sum_{i \in I} a_i \right) = 0 \pmod{2}.$$

Observe that the points $(x + \sum_{i \in I} a_i)$, when viewed as points in $\{0,1\}^n$ (by filling the coordinates outside S with 0's), are distributed individually according to $\mu_p^{\otimes n}$.

Let $\text{rej}_T(f)$ denote the rejection probability of a test T on input function f . In the rest of this section, we prove Theorem 1.5 (restated below from the introduction).

Theorem 1.5 (p -biased version of the BKSSZ Theorem). *For every d and $p \in (0, 1/2)$, the 2^{d+1} -query test $T_{p,d}$ satisfies the following properties:*

- *Completeness: if $\text{bdeg}(f) \leq d$ then $\text{rej}_{T_{d,p}}(f) = 0$.*
- *Soundness: $\delta_d^{(p)}(f) = O_d(\text{rej}_{T_{d,p}}(f))$, where the hidden constant is independent of p (but depends on d).*

Remark 6.7. The case for $p \in (1/2, 1)$ is proved by reversing the roles of 0 and 1. In particular, by running the test $T_{p,d}$ on the function $\bar{f} = (x_1, \dots, x_n) \mapsto f(1 - x_1, \dots, 1 - x_n)$.

The main ingredients that go into proving Theorem 1.5 are the following.

d -dimensional agreement theorem: We will use the d -dimensional agreement theorems (Theorems 5.4 and 5.5) to reduce the analysis of the test $T_{p,d}$ in the p -biased setting to the analysis of the standard test T_d . Given $p, \alpha \in (0, 1)$ such that $p(2 - \alpha) \leq 1$ and $\kappa \geq 1$, we will say that *agreement κ -holds for (p, α)* if the hypothesis

$$\Pr_{(S_1, S_2) \sim \mu_{p,\alpha}} [P_{S_1}|_{S_1 \cap S_2} \neq P_{S_2}|_{S_1 \cap S_2}] \leq \varepsilon$$

implies the conclusion that there exists a global function $P: \binom{[n]}{\leq d} \rightarrow \{0,1\}$ such that

$$\Pr_{S \sim \mu_p} [P_S \equiv P|_S] \leq \kappa \cdot \varepsilon.$$

Stated in this language, Theorem 5.4 implies that “agreement $\kappa_{d,\alpha,\rho}$ -holds for (p, α) whenever $p(2 - \alpha) \leq 1 - \rho$ ”, where $\kappa_{d,\alpha}$ is some constant depending on d, α, ρ (but independent of p and n). Similarly, Theorem 5.5 implies “agreement $2p^{-d}$ -holds for (p, p) ”.

Biased Schwartz–Zippel Lemma: The standard Schwartz–Zippel Lemma states that if P is a non-zero Boolean polynomial with $\text{bdeg}(P) \leq d$, then $\Pr_{x \sim \mu_{1/2}^{\otimes n}}[P(x) \bmod 2 = 1] \geq (\frac{1}{2})^d$. The following is an extension of this lemma to the p -biased setting.

Claim 6.8. Suppose that P is a non-zero polynomial with $\text{bdeg}(f) \leq d$. Then for all $\theta \in (0, 1)$,

$$\Pr_{\mu_\theta}[P \bmod 2 = 1] \geq \min(\theta, 1 - \theta)^d =: \delta_\theta^{(d)}.$$

Proof. Let I be an inclusion-maximal set such that P contains the monomial $\prod_{i \in I} x_i$. For every setting σ of the variables outside I , we obtain a nonzero polynomial $P|_\sigma$ (since the monomial $\prod_{i \in I} x_i$ necessarily survives). Therefore there is at least one input on which $P|_\sigma \bmod 2 = 1$. Since $P|_\sigma$ is a polynomial on at most d variables, this input occurs with probability at least $\min(\theta, 1 - \theta)^d$ in μ_θ . \square

BKSSZ Theorem for $p = 1/2$: We will say that a test T is ρ -valid for p if the following holds.

- Completeness: $\text{rej}_T(f) = 0$ whenever $\text{bdeg}(f) \leq d$ and
- Soundness: $\delta_d^{(p)}(f) \leq \rho \cdot \text{rej}_T(f)$.

[Corollary 6.5](#) states that T_d (modified so that it always accepts when the dimension is at most d) is ρ_d -valid for $1/2$ for some constant ρ_d dependent on d (but independent of n).

The following lemma shows how we can use the agreement theorem to reduce the analysis of the test $T_{p,d}$ to that of T_d .

Lemma 6.9. Suppose that T_d is ρ_d -valid for $1/2$ and agreement κ -holds for $(2p, \alpha)$ for some $p \in (0, 1/2)$ and $\alpha \in (1/2, 1)$ satisfying $2p(2 - \alpha) \leq 1$, then $T_{p,d}$ is $\left(\frac{4\kappa}{\delta_{1/2\alpha}^{(d)}} + 1\right) \cdot \rho_d$ -valid for p , where $\delta_\theta^{(d)}$ is as defined in [Claim 6.8](#).

Proof. Let us start by noticing that completeness is clear, since $\text{bdeg}(f|_S) \leq \text{bdeg}(f)$. It remains to prove soundness.

By construction, $\text{rej}_{T_{p,d}}(f) = \mathbb{E}_{S \sim \mu_{2p}}[\text{rej}_{T_d}(f|_S)]$. The assumption that T_d is ρ_d -valid for $1/2$ guarantees for each set $S \subseteq [n]$ the existence of a Boolean polynomial P_S over $\{x_i : i \in S\}$ of degree at most d satisfying

$$\delta_S := \Pr_{x \sim \mu_{1/2}^{\otimes S}}[f|_S(x) \neq P_S(x) \bmod 2] \leq \rho_d \cdot \text{rej}_{T_d}(f|_S).$$

Hence, $\delta := \mathbb{E}_{S \sim \mu_{2p}}[\delta_S] \leq \rho_d \cdot \text{rej}_{T_{p,d}}(f)$.

Our goal now is to use agreement in order to sew the various polynomials P_S into a global polynomial P . To this end, we will view each polynomial P_S as a list of coefficients. In other words, $P_S : \binom{S}{\leq d} \rightarrow \{0, 1\}$. To apply the agreement theorem, we need to show that the polynomials P_S agree with each other. We show this using the p -biased variant of the Schwartz–Zippel Lemma, [Claim 6.8](#).

Let $\theta = 1/2\alpha$. Note that $\theta \in (1/2, 1)$ since $\alpha \in (1/2, 1)$. For two sets $S \supseteq T$, the polynomial $P_S|_T$ is the polynomial obtained by zeroing out all the coefficients of monomials involving variables outside T . Note that this matches with the definition $P_S|_T : \binom{T}{\leq d} \rightarrow \{0, 1\}$ when P_S is viewed as a list of coefficients, $P_S : \binom{S}{\leq d} \rightarrow \{0, 1\}$. We define

$$\delta_{S,T} := \Pr_{x \sim \mu_\theta^{\otimes T}}[f|_T(x) \neq P_S|_T(x) \bmod 2].$$

Since $\alpha\theta = 1/2$, we have

$$\begin{aligned}\mathbb{E}_{S \sim \mu_{2p}} \mathbb{E}_{T \sim \mu_\alpha(S)} [\delta_{S,T}] &= \mathbb{E}_{S \sim \mu_{2p}} \mathbb{E}_{T \sim \mu_\alpha(S)} \Pr_{x \sim \mu_\theta^{\otimes T}} [f|_T(x) \neq P_S|_T(x) \bmod 2] \\ &= \mathbb{E}_{S \sim \mu_{2p}} \Pr_{x \sim \mu_{1/2}^{\otimes S}} [f|_S(x) \neq P_S(x) \bmod 2] = \mathbb{E}_{S \sim \mu_{2p}} [\delta_S] = \delta.\end{aligned}$$

Hence, we have by Markov's inequality that

$$\Pr_{S \sim \mu_{2p}, T \sim \mu_\alpha(S)} \left[\delta_{S,T} \geq \frac{\delta_\theta^{(d)}}{2} \right] \leq \frac{2}{\delta_\theta^{(d)}} \cdot \delta.$$

By the union bound

$$\begin{aligned}\Pr_{(S_1, S_2) \sim \mu_{2p, \alpha}} \left[\delta_{S_1, S_1 \cap S_2} \geq \frac{\delta_\theta^{(d)}}{2} \text{ or } \delta_{S_2, S_1 \cap S_2} \geq \frac{\delta_\theta^{(d)}}{2} \right] &\leq 2 \cdot \Pr_{(S_1, S_2) \sim \mu_{2p, \alpha}} \left[\delta_{S_1, S_1 \cap S_2} \geq \frac{\delta_\theta^{(d)}}{2} \right] \\ &= 2 \cdot \Pr_{S \sim \mu_{2p}, T \sim \mu_\alpha(S)} \left[\delta_{S,T} \geq \frac{\delta_\theta^{(d)}}{2} \right] \leq \frac{4}{\delta_\theta^{(d)}} \cdot \delta.\end{aligned}$$

If $\delta_{S_1, S_1 \cap S_2} < \delta_\theta^{(d)}/2$ and $\delta_{S_2, S_1 \cap S_2} < \delta_\theta^{(d)}/2$, then we have

$$\Pr_{x \sim \mu_\theta^{\otimes S_1 \cap S_2}} [P_{S_1}|_{S_1 \cap S_2}(x) \neq P_{S_2}|_{S_1 \cap S_2}(x)] < \delta_\theta^{(d)}.$$

It then follows from the p -biased variant of the Schwartz–Zippel Lemma, [Claim 6.8](#), that the two Boolean degree d polynomials $P_{S_1}|_{S_1 \cap S_2}$ and $P_{S_2}|_{S_1 \cap S_2}$ must be identical. We thus have

$$\Pr_{(S_1, S_2) \sim \mu_{2p, \alpha}} [P_{S_1}|_{S_1 \cap S_2} \neq P_{S_2}|_{S_1 \cap S_2}] \leq \frac{4}{\delta_\theta^{(d)}} \cdot \delta.$$

We have thus satisfied the hypothesis of the agreement theorem. Using the fact that “agreement κ -holds for $(2p, \alpha)$ ”, we can conclude that there exists a global polynomial $P: \binom{[n]}{\leq d} \rightarrow \{0, 1\}$ (and hence of Boolean degree at most d) such that

$$\Pr_{S \sim \mu_{2p}} [P_S \neq P|_S] \leq \frac{4\kappa}{\delta_\theta^{(d)}} \cdot \delta.$$

The following argument now shows that $P \bmod 2$ is close to f under the μ_p -measure, completing the

proof of the lemma.

$$\begin{aligned}
\Pr_{x \sim \mu_p^{\otimes n}} [f(x) \neq P(x) \bmod 2] &= \mathbb{E}_{S \sim \mu_{2p}} \left[\Pr_{x \sim \mu_{1/2}^{\otimes S}} [f|_S(x) \neq P|_S(x) \bmod 2] \right] \\
&\leq \mathbb{E}_{S \sim \mu_{2p}} \left[\Pr_{x \sim \mu_{1/2}^{\otimes S}} [f|_S(x) \neq P_S(x) \bmod 2] \right] + \mathbb{E}_{S \sim \mu_{2p}} \left[\Pr_{x \sim \mu_{1/2}^{\otimes S}} [P_S(x) \neq P|_S(x) \bmod 2] \right] \\
&\leq \mathbb{E}_{S \sim \mu_{2p}} [\delta_S] + \Pr_{S \sim \mu_{2p}} [P_S \neq P|_S] \\
&\leq \delta + \frac{4\kappa}{\delta_\theta^{(d)}} \cdot \delta \leq \left(1 + \frac{4\kappa}{\delta_\theta^{(d)}}\right) \cdot \rho_d \cdot \text{rej}_{T_{p,d}}(f). \quad \square
\end{aligned}$$

An immediate application of [Lemma 6.9](#) using the agreement theorem [Theorem 5.4](#) is that $T_{p,d}$ is $C_{p,d}$ -valid for p for some constant $C_{p,d}$ (depending on p and d). We will do a slightly more careful analysis to show that $T_{p,d}$ is C_d -valid for p for some constant C_d (depending on d but not on p).

Proof of [Theorem 1.5](#). Our starting point is [Corollary 6.5](#), which states that T_d is ρ_d -valid for $1/2$. The proof then follows two cases: p small and p large.

$p \in (0, 0.3)$: We set $\alpha = 3/4$. Observe that $2p(2 - \alpha) = 5p/2 < 3/4 < 1$. [Theorem 5.4](#) implies that there exists a constant κ_d such that for all $p \in (0, 0.3)$, “agreement κ_d -holds for $(2p, \alpha)$ ”. Applying [Lemma 6.9](#), we get that $T_{p,d}$ is C_d^{small} -valid for all $p \in (0, 0.3)$, where $C_d^{\text{small}} := \left(\frac{4\kappa_d}{\delta_{2/3}^{(d)}} + 1\right) \cdot \rho_d$.

$p \in [0.3, 0.5)$: We set $\alpha = 2p$. Observe that $2p(2 - \alpha) = 4p(1 - p) < 1$ and $\alpha \in [0.6, 1)$. Hence, $1/2\alpha \in (1/2, 5/6]$ and $\delta_{1/2\alpha}^{(d)} = \min(1/2\alpha, 1 - 1/2\alpha)^d \geq (1/6)^d$. [Theorem 5.5](#) implies that there exists a constant $\kappa'_d := 2(0.3)^{-d}$ such that for all $p \in [0.3, 0.5)$, “agreement κ'_d -holds for $(2p, \alpha)$ ”. Applying [Lemma 6.9](#), we get that $T_{p,d}$ is C_d^{large} -valid for all $p \in [0.3, 0.5)$ where $C_d^{\text{large}} := \left(\frac{4\kappa'_d}{\delta_{1/2\alpha}^{(d)}} + 1\right) \cdot \rho_d \leq (4 \cdot 6^d \cdot \kappa'_d + 1) \cdot \rho_d$.

Thus, $T_{p,d}$ is C_d -valid for all $p \in (0, 0.5)$ where $C_d = \max(C_d^{\text{small}}, C_d^{\text{large}})$. \square

7 Reverse union bound and hyper graph pruning lemma

In this section we prove [Lemma 1.4](#), which we now restate.

Lemma 1.4 (reverse union bound). *For each $c > 0$ and integer $d \geq 1$ there exists a constant $C'(d, c) = 2^{O_c(d^3)}$ such that the following holds for all $d \leq k \leq n$. Let $Y \subseteq \binom{[n]}{k}$ and $B \subseteq \binom{[n]}{d}$ satisfy*

$$\Pr_{S \sim \nu_n(k)} [S \in Y \mid S \supseteq I] \geq c \text{ for all } I \in B.$$

Then

$$\Pr_{S \sim \nu_n(k)} [S \text{ hits } B] \leq C'(c, d) \cdot \Pr_{S \sim \nu_n(k)} [S \in Y],$$

where S hits B if $S \supseteq I$ for some $I \in B$.

We will deduce [Lemma 1.4](#) from another result, which we call the *hypergraph pruning lemma* (restated below from the introduction):

Lemma 1.6 (hypergraph pruning lemma). *For each $\varepsilon > 0$ and integer $d \geq 1$ there exists a constant $C''(d, \varepsilon) = 2^{O_\varepsilon(d^3)}$ such that for every d -uniform hypergraph H there exists a d -uniform hypergraph $H' \subseteq H$ such that:*

- $\Pr_{S \sim \nu_n(k)}[S \text{ hits } H'] \geq \Pr_{S \sim \nu_n(k)}[S \text{ hits } H] / C''(d, \varepsilon).$
- For every $I \in H'$,

$$\Pr_{S \sim \nu_n(k)}[\Delta I' \in H' \setminus \{I\}, S \supseteq I' \mid S \supseteq I] \geq 1 - \varepsilon.$$

We use the following notation to prove [Lemmas 1.4](#) and [1.6](#). If $H \subseteq \binom{[n]}{d}$ and $S \subseteq [n]$ then

$$\text{Hit}_H(S) = \{I \in H : I \subseteq S\}.$$

We think of $\text{Hit}_H(S)$ as a random variable that depends on the distribution of S . With this notation, the conclusion of [Lemma 1.4](#) can be rephrased as follows.

$$\Pr_{S \in \nu_n(k)}[\text{Hit}_B(S) \neq \emptyset] \leq C'(c, d) \cdot \Pr_{S \in \nu_n(k)}[S \in Y],$$

while the conclusion of [Lemma 1.6](#) can be rewritten as

- $\Pr_{S \sim \nu_n(k)}[\text{Hit}_{H'}(S) \neq \emptyset] \geq \Pr_{S \sim \nu_n(k)}[\text{Hit}_H(S) \neq \emptyset] / C''(d, \varepsilon).$
- For every $I \in H'$,

$$\Pr_{S \sim \nu_n(k)}[\text{Hit}_{H'}(S) = \{I\} \mid \text{Hit}_{H'}(S) \ni I] \geq 1 - \varepsilon.$$

The hypergraph pruning lemma states that you can prune a given d -uniform hypergraph H to a subhypergraph H' for which the events “ $S \supseteq I$ ” (where $S \sim \nu_n(k)$ and $I \in H'$) are nearly disjoint. For this to be nontrivial, we also require that the probability that S contains *some* hyperedge in H' does not drop too much compared to H .

Let us now deduce [Lemma 1.4](#) from [Lemma 1.6](#).

Proof of Lemma 1.4. We apply [Lemma 1.6](#) to $H := B$ with $\varepsilon := c/2$, obtaining a subset $B' \subseteq B$. For every $I \in B'$, we have

$$\begin{aligned} \Pr_{S \sim \nu_n(k)}[S \in Y \text{ and } \text{Hit}_{B'}(S) = \{I\}] &\geq \Pr_{S \sim \nu_n(k)}[S \supseteq I] \left(\Pr_{S \sim \nu_n(k)}[S \in Y \mid S \supseteq I] - \Pr_{S \sim \nu_n(k)}[\text{Hit}_{B'}(S) \supsetneq \{I\} \mid S \supseteq I] \right) \\ &\geq \Pr_{S \sim \nu_n(k)}[S \supseteq I] \cdot (c - \varepsilon) = \Pr_{S \sim \nu_n(k)}[S \supseteq I] \cdot \frac{c}{2}. \end{aligned}$$

The events “ $S \in Y$ and $\text{Hit}_{B'}(S) = \{I\}$ ” are disjoint for different I , and so

$$\Pr_{S \sim \nu_n(k)}[S \in Y] \geq \sum_{I \in B'} \Pr_{S \sim \nu_n(k)}[S \in Y \text{ and } \text{Hit}_{B'}(S) = \{I\}] \geq \frac{c}{2} \sum_{I \in B'} \Pr_{S \sim \nu_n(k)}[S \supseteq I] \geq \frac{c}{2} \Pr_{S \sim \nu_n(k)}[\text{Hit}_{B'}(S) \neq \emptyset].$$

Together with $\Pr_{S \sim \nu_n(k)}[\text{Hit}_{B'}(S) \neq \emptyset] \geq \Pr_{S \sim \nu_n(k)}[\text{Hit}_B(S) \neq \emptyset] / C''(c, d)$, this completes the proof. \square

It will be slightly simpler to present the proof of [Lemma 1.6](#) first in the product setting, that is, when $\nu_n(k)$ is replaced with the distribution μ_p , which we do in [Section 7.1](#). We then explain the changes needed to adapt it to the uniform setting in [Section 7.2](#).

7.1 Proof in product setting

The analog of [Lemma 1.6](#) in the product setting has exactly the same statement, but instead of choosing $S \sim \nu_n(k)$, we now choose $S \sim \mu_p([n])$.

Instead of proving the product version of [Lemma 1.6](#) directly, we will deduce it from a more detailed lemma which involves the concept of *branching factor*.

Definition 7.1 (branching factor). *A d -uniform hypergraph H has branching factor ρ if for all sets A of size at most d , the number of hyperedges in H extending A is at most $\rho^{d-|A|}$.*

If this holds only for sets A of size at least $d - k$, we say that H has k -bounded branching factor ρ .

Lemma 7.2. *For each $d \geq 1$ there exist constants $c_{\max}(d) = 2^{-O(d^2)}$ and $\kappa(d) = 2^{-O(d^2)}$ such that for every $c \leq c_{\max}(d)$ there exists a constant $K(d, c) = 2^{O_c(d^3)}$ such that for all positive $p \leq \kappa(d)c$, every d -uniform hypergraph H has a subhypergraph H' such that:*

- *Hitting property:* $\Pr_{S \sim \mu_p}[\text{Hit}_{H'}(S) \neq \emptyset] \geq \Pr_{S \sim \mu_p}[\text{Hit}_H(S) \neq \emptyset] / K(d, c)$.
- *Branching factor property:* H' has branching factor c/p .

[Lemma 1.6](#) (in its product version) follows almost immediately from [Lemma 7.2](#) given the following lemma.

Lemma 7.3. *If $c \leq 1$ and H is a d -uniform hypergraph with branching factor c/p then for all $I \in H$,*

$$\Pr_{S \sim \mu_p}[\text{Hit}_H(S) = \{I\} \mid \text{Hit}_H(S) \ni I] \geq 1 - 2^d c.$$

Proof. Recall that $\text{Hit}_H(S)$ consists of all sets in H that are contained in S . Conditioning on $\text{Hit}_H(S) \ni I$ is the same as conditioning on $S \supseteq I$. If $\text{Hit}_H(S) \supsetneq \{I\}$ then $S \supset J$ for some $J \in H$. For a fixed J , conditioned on $S \supseteq I$, this happens iff $S \setminus I \supseteq J \setminus I$ for some $J \in H$, which happens with probability $p^{|J \setminus I|} = p^{d-|I \cap J|}$. Therefore

$$\Pr_{S \sim \mu_p}[\text{Hit}_H(S) \supsetneq \{I\} \mid \text{Hit}_H(S) \ni I] \leq \sum_{\substack{J \in H \\ J \neq I}} p^{|J \setminus I|} \leq \sum_{K \subsetneq I} \sum_{\substack{J \in H \\ J \cap I = K}} p^{d-|K|}.$$

For each $K \subsetneq I$, the number of $J \in H$ extending K by $d - |K|$ elements is at most $(c/p)^{d-|K|}$, and so

$$\Pr_{S \sim \mu_p}[\text{Hit}_H(S) \supsetneq \{I\} \mid \text{Hit}_H(S) \ni I] \leq \sum_{K \subsetneq I} \sum_{\substack{J \in H \\ J \cap I = K}} p^{d-|K|} \leq \sum_{K \subsetneq I} c^{d-|K|} < 2^d c,$$

since there are $2^d - 1$ summands, and each one is at most c (since $c \leq 1$). □

We can now deduce the product version of [Lemma 1.6](#) from [Lemma 7.2](#).

Proof of product version of [Lemma 1.6](#). Let $c = \min(\varepsilon/2^d, c_{\max}(d), 1)$. If $p \leq \kappa(d)c$ then the lemma follows directly from [Lemma 7.2](#) via [Lemma 7.3](#), so suppose that $p > \kappa(d)c$. If H is empty then we can take H'

to be empty as well, so suppose that H contains some hyperedge I . We take $H' = \{I\}$. This works since

$$\Pr_{S \sim \mu_p} [\text{Hit}_{H'}(S) \neq \emptyset] = p^d \geq (\kappa(d)c)^d \geq \Pr_{S \sim \mu_p} [\text{Hit}_H(S) \neq \emptyset] / (1/\kappa(d)c)^d,$$

and $1/\kappa(d)c = 2^{O(d^2)} \max(2^d/\varepsilon, 1/c_{\max}(d), 1)$. \square

In the sequel we will need the following simple corollary of [Lemma 7.3](#).

Corollary 7.4. *If $c \leq 1$ and H is a d -uniform hypergraph with branching factor c/p then*

$$(1 - 2^d c) p^d |H| \leq \Pr_{S \sim \mu_p} [\text{Hit}_H(S) \neq \emptyset] \leq p^d |H|.$$

Proof. The upper bound follows from the union bound, since the probability of hitting any particular set is p^d . The lower bound follows from

$$\Pr_{S \sim \mu_p} [\text{Hit}_H(S) \neq \emptyset] \geq \sum_{I \in H} \Pr_{S \sim \mu_p} [\text{Hit}_H(S) = \{I\}] = p^d \sum_{I \in H} \Pr_{S \sim \mu_p} [\text{Hit}_H(S) = \{I\} \mid \text{Hit}_H(S) \ni I]$$

via an application of [Lemma 7.3](#). \square

The proof of [Lemma 7.2](#) is by induction. The case $d = 1$ is almost trivial, and the case $d = 2$ illustrates most of the ideas appearing in the general case. We therefore start by proving the lemma in these two cases, and only then present the general inductive argument.

Below, we will write $\Pr[\text{Hit}_H \neq \emptyset]$ as a shortcut for $\Pr_{S \sim \mu_p} [\text{Hit}_H(S) \neq \emptyset]$.

7.1.1 Proof for $d = 1$

The idea is that there are two cases: either H is *sparse* (contains at most c/p elements) or it is *dense*. In the first case, take $H' = H$ and there is nothing to prove. In the second case, if we retain an arbitrary subset of $\lfloor c/p \rfloor$ elements (“pruning”), then a simple calculation shows that the hitting property is satisfied, essentially since the probability of hitting a set of $\Theta(1/p)$ elements is constant.

Proof of [Lemma 7.2](#) for $d = 1$. We choose $c_{\max}(1) = 1/3$ and $\kappa(1) = 1$. If H has branching factor c/p then $H' := H$ satisfies the conditions of the lemma, so suppose that $|H| > c/p$. Let H' be an arbitrary choice of $\lfloor c/p \rfloor$ elements of H . Note that $c/p \geq 1$ (since we assume that $p \leq c$), and so $\lfloor c/p \rfloor \geq c/(2p)$ (the worst case is when $c/p = 2 - \delta$). Clearly H' has branching factor c/p , and by [Corollary 7.4](#),

$$\Pr[\text{Hit}_{H'} \neq \emptyset] \geq (1 - 2c)p|H| \geq \frac{1}{3} \cdot p \cdot \frac{c}{2p} = \frac{c}{6}. \quad \square$$

This argument requires $c_{\max}(1) < 1/2$, in order for the lower bound in [Corollary 7.4](#) to be nontrivial. However, it is clear that we get a nontrivial bound even if $c \geq 1/2$. Nevertheless, we use [Corollary 7.4](#) to simplify calculations in more complicated situations.

7.1.2 Proof for $d = 2$

When $d = 2$, the hypergraph H is a graph. There are two ways for H to fail to satisfy the branching factor condition. First, H could have vertices of high degree (more than c/p). Second, it could simply contain too many edges (more than $(c/p)^2$).

Let us first give a sketch before the formal proof. For the first case, assuming the second case doesn't hold, there are at most $2c/p$ high degree vertices. For each one choose an arbitrary subset of c/p edges touching it and remove the rest. The resulting graph will actually have degree up to $4c/p$, so we will have to dial down c . We handle the second case by arbitrarily retaining $(c/p)^2$ edges. Since we can assume that there are no high-degree vertices, we can appeal to [Corollary 7.4](#) in order to lower-bound the hitting probability of the resulting graph.

Proof of [Lemma 7.2](#) for $d = 2$. We choose $c_{\max}(2) = 1/5$ and $\kappa(2) = 1/2$.

Let $\gamma = c/2$, and note that $\gamma/p \geq 1$. Let B be the set of vertices of degree at least γ/p , and let K be the graph resulting from H by removing all edges incident to a vertex in B . Clearly

$$\Pr[\text{Hit}_H \neq \emptyset] \leq \Pr[\text{Hit}_B \neq \emptyset] + \Pr[\text{Hit}_K \neq \emptyset],$$

and so one of the two latter probabilities is at least $\Pr[\text{Hit}_H \neq \emptyset]/2$. We consider these two cases separately.

Case 1: $\Pr[\text{Hit}_K \neq \emptyset] \geq \Pr[\text{Hit}_H \neq \emptyset]/2$. If $|K| \leq (\gamma/p)^2$ then we can take $H' := K$. Otherwise, let H' be an arbitrary subset of K consisting of $\lfloor (\gamma/p)^2 \rfloor \geq \gamma^2/(2p^2)$ edges. By construction, H' has branching factor $\gamma/p \leq c/p$, and so [Corollary 7.4](#) shows that

$$\Pr[\text{Hit}_K \neq \emptyset] \geq (1 - 4\gamma)p^2 \frac{\gamma^2}{2p^2} = \Omega(\gamma^2) \geq \Omega(\gamma^2) \Pr[\text{Hit}_H \neq \emptyset],$$

since $4\gamma \leq 2/5$.

Case 2: $\Pr[\text{Hit}_B \neq \emptyset] \geq \Pr[\text{Hit}_H \neq \emptyset]/2$. We start by applying [Lemma 7.2](#) inductively on B with the parameter γ ; note $\gamma \leq c_{\max}(2)/2 \leq c_{\max}(1)$ and $p \leq c/2 \leq \gamma = \kappa(1)\gamma$. This results in a set $B' \subseteq B$ of size at most γ/p such that

$$\Pr[\text{Hit}_{B'} \neq \emptyset] \geq \Pr[\text{Hit}_B \neq \emptyset]/K(1, \gamma) \geq \Pr[\text{Hit}_H \neq \emptyset]/(2K(1, \gamma)).$$

Each vertex $i \in B'$ had degree at least γ/p in the original graph H . We choose a set N_i of $\lfloor \gamma/p \rfloor \geq \gamma/(2p)$ edges adjacent to i for each $i \in B'$, and form H' by taking the union of the sets N_i for all $i \in B'$. Since every edge is adjacent to at most two vertices, each edge appears in at most two sets N_i , and so $|H'| \geq \gamma/(4p) \cdot |B'|$.

The graph H' has branching factor c/p . Indeed, the degree of each vertex $i \in B'$ is at most $N_i + |B'| \leq 2(\gamma/p) = c/p$, since each other N_j contributes at most one neighbor of i . Similarly, the degree of any other vertex is at most $|B'| \leq c/p$. The total number of edges is at most $\gamma/p \cdot |B'| \leq (\gamma/p)^2 \leq (c/p)^2$. Therefore H' has branching factor c/p .

Applying [Corollary 7.4](#) twice, we see that

$$\Pr[\text{Hit}_{H'} \neq \emptyset] \stackrel{(1)}{\geq} (1 - 4c)p^2 |H'| \geq \frac{p^2}{5} \frac{c}{8p} |B'| \stackrel{(2)}{\geq} \frac{c}{40} \Pr[\text{Hit}_{B'} \neq \emptyset] \geq \frac{c}{80K(1, \gamma)} \Pr[\text{Hit}_H \neq \emptyset],$$

completing the proof. □

7.1.3 General case

The general case is quite similar to the case $d = 2$. There are now d ways for the hypergraph to fail to satisfy the branching factor property, and accordingly we decompose H into d parts (in fact, we use $d + 1$ parts).

Let B_1 be all $(d - 1)$ -sets with high degree in H . Let H_1 be the result of removing from H all extensions of sets in B_1 . Let B_2 be all $(d - 2)$ -sets with high degree in H_1 , and let H_2 be the result of removing from H_1 all extensions of sets from B_2 . We continue in this manner and consider $B_1, B_2, \dots, B_d, H_d$. Since $\Pr[\text{Hit}_H \neq \emptyset] \geq \Pr[\text{Hit}_{B_1} \neq \emptyset] + \dots + \Pr[\text{Hit}_{B_d} \neq \emptyset] + \Pr[\text{Hit}_{H_d} \neq \emptyset]$ at least one of these is non-negligible and we use it to inductively construct the hypergraph H' with bounded branching factor.

Proof of Lemma 7.2. The proof is by induction on d . In the base case $d = 0$ we can take $c_{\max}(0) = \kappa(0) = 1$ and $K(0, c) = 1$, since $H' = H$ always works.

Suppose now that $d \geq 1$. We choose

$$\begin{aligned} c_{\max}(d) &= \min(c_{\max}(0)/2^{d-1}, \dots, c_{\max}(d-1)/2^{d-1}, 1/2^{d+1}), \\ \kappa(d) &= \min(\kappa(0), \dots, \kappa(d-1))/2^{d-1}, \\ K(d, c) &= 2^{3d} \max(K(0, c/2^{d-1}), \dots, K(d-1, c/2^{d-1}))/c. \end{aligned}$$

It is not hard to verify that $c_{\max}(d), \kappa(d) \geq 2^{-O(d^2)}$. As for $K(d, c)$, it is not hard to check that $K(d, c)$ is monotone in d , and so

$$K(d, c) = \frac{2^{3d}}{c} K\left(d-1, \frac{c}{2^{d-1}}\right) = \frac{2^{3d}}{c} \cdot \frac{2^{3(d-1)} \cdot 2^{d-1}}{c} K\left(d-2, \frac{c}{2^{(d-1)+(d-2)}}\right),$$

and so on. There are d factors, each of which is $2^{O(d^2)}/c$, and so $K(d, c) = 2^{O(d^3)}/c^d$.

Let $\gamma = c/2^{d-1}$, and note that $\gamma/p \geq 1$. We decompose H into $d + 1$ parts, as follows. Let $H_0 := H$. For $e = 1, \dots, d$:

1. Let B_e be the $(d - e)$ -uniform hypergraph consisting of all sets A with at least $(\gamma/p)^e$ extensions in H_{e-1} .
2. Let H_e be the subhypergraph of H_{e-1} obtained by removing all extensions of hyperedges in B_e .

Clearly $\Pr[\text{Hit}_{H_{e-1}} \neq \emptyset] \leq \Pr[\text{Hit}_{B_e} \neq \emptyset] + \Pr[\text{Hit}_{H_e} \neq \emptyset]$. Expanding this, we get

$$\Pr[\text{Hit}_H \neq \emptyset] \leq \sum_{e=1}^d \Pr[\text{Hit}_{B_e} \neq \emptyset] + \Pr[\text{Hit}_{H_d} \neq \emptyset].$$

There are $d + 1$ summands on the right-hand side, and one of them needs to be at least $\Pr[\text{Hit}_H \neq \emptyset]/(d + 1)$. If $\Pr[\text{Hit}_{H_d} \neq \emptyset] \geq \Pr[\text{Hit}_H \neq \emptyset]/(d + 1)$ then we choose $H' := H_d$. By construction, H' has branching factor $\gamma/p \leq c/p$, completing the proof in this case.

Suppose next that $\Pr[\text{Hit}_{B_e} \neq \emptyset] \geq \Pr[\text{Hit}_H \neq \emptyset]/(d + 1)$ for some $e \in \{1, \dots, d\}$. We apply the lemma inductively on B_e and $c := \gamma$ to get a $(d - e)$ -uniform hypergraph $B'_e \subseteq B_e$ satisfying

$$\Pr[\text{Hit}_{B'_e} \neq \emptyset] \geq \Pr[\text{Hit}_H \neq \emptyset]/((d + 1)K(d - e, \gamma)).$$

We can apply the lemma since $\gamma = c/2^{d-1} \leq c_{\max}(d - e)$ and $p \leq 2^{-(d-1)}\kappa(d - e)c = \kappa(d - e)\gamma$.

For each $I \in B'_e$, we choose a set N_I of $\lfloor (\gamma/p)^e \rfloor \geq (\gamma/p)^e/2$ extensions in H_{e-1} . We let H' be the union of the sets N_I for all $I \in B'_e$. Each hyperedge can appear in at most $\binom{d}{d-e} \leq 2^{d-1}$ many N_I (since a hyperedge only has so many subsets of size $d-e$), and so $|H'| \geq (\gamma/p)^e |B'_e|/2^d$.

We start by showing that H' has branching factor c/p . We have to show that each set A has at most $(c/p)^{d-|A|}$ extensions in H' . We distinguish between two cases, depending on the size of A , which we can assume is at most $d-1$ (the condition always holds for larger A for trivial reasons).

Note first that by construction, H_{e-1} has $(e-1)$ -bounded branching factor γ/p , which means that every set A of size at least $d-(e-1)$ has at most $(\gamma/p)^{d-|A|}$ extensions in H_{e-1} . Since $H' \subseteq H_{e-1}$, this shows that each set A of size more than $d-e$ has at most $(\gamma/p)^{d-|A|} \leq (c/p)^{d-|A|}$ extensions in H' as well.

Suppose next that $|A| \leq d-e$. For a set of hyperedges X , let $A(X)$ be the number of hyperedges in X containing A . Each hyperedge in H' containing A belongs to some N_I , and so

$$A(H') \leq \sum_{I \in B'_e} A(N_I) = \sum_{J \subseteq A} \sum_{\substack{I \in B'_e \\ I \cap A = J}} A(N_I).$$

Fix some particular $J \subseteq A$ and consider all I such that $A \cap I = J$. Since B'_e has branching factor γ/p , it contains at most $(\gamma/p)^{d-e-|J|}$ many hyperedges I extending J .

If $J \subsetneq A$ then each hyperedge counted in $A(N_I)$ contains $I \cup A$, a set of size $|I \cup A| = d-e+|A \setminus J| \geq d-e+1$. Since $N_I \subseteq H_{e-1}$ and H_{e-1} has $(e-1)$ -bounded branching factor γ/p , we see that $A(N_I) \leq (\gamma/p)^{d-(d-e+|A \setminus J|)} = (\gamma/p)^{e-|A \setminus J|}$. In total, the contribution of this J to the sum is at most

$$(\gamma/p)^{d-e-|J|} \cdot (\gamma/p)^{e-|A \setminus J|} = (\gamma/p)^{d-|A|}.$$

If $J = A$ then $A(N_I) \leq (\gamma/p)^e = (\gamma/p)^{e-|A \setminus J|}$ by construction, and so we reach the same conclusion.

Since there are $2^{|A|} \leq 2^{d-e} \leq 2^{d-1}$ summands, we conclude that $A(H') \leq 2^{d-1} (\gamma/p)^{d-|A|} \leq (c/p)^{d-|A|}$, completing the proof that H' has branching factor c/p .

Having shown that H' has branching factor c/p , we can now appeal to [Corollary 7.4](#), which shows that

$$\begin{aligned} \Pr[\text{Hit}_{H'} \neq \emptyset] &\geq (1-2^d c) p^d |H'| \geq (1/2) p^d (\gamma/p)^e |B'_e|/2^d \geq (1/2) (\gamma^e/2^d) \Pr[\text{Hit}_{B'_e} \neq \emptyset] \\ &\geq \frac{\gamma^e/2^{d+1}}{(d+1)K(d-e, \gamma)} \Pr[\text{Hit}_H \neq \emptyset] \geq \frac{\gamma^e}{2^{3d}K(d-e, \gamma)} \Pr[\text{Hit}_H \neq \emptyset], \end{aligned}$$

using $d+1 \leq 2^d$ and $d \geq 1$. This completes the proof. \square

7.2 Proof in uniform setting

The only two places at which the proof in [Section 7.1](#) accesses the measure directly are [Lemma 7.3](#) and [Corollary 7.4](#), so we start by generalizing these two results to the uniform setting: $S \sim \nu_n(k)$, and $p = k/n$. To avoid trivialities, we also assume that $k \geq d$.

Lemma 7.5. *If $c \leq 1$ and H is a d -uniform hypergraph with branching factor c/p then for all $I \in H$,*

$$\Pr_{S \in \nu_n(k)} [\text{Hit}_H(S) = \{I\} \mid \text{Hit}_H(S) \ni I] \geq 1 - 2^d c.$$

Proof. The original proof uses the bound

$$\Pr_{S \sim \mu_p} [S \supseteq J \mid S \supseteq I] \leq p^{|J \setminus I|}.$$

We claim that this bound holds also for the uniform setting. Indeed, there are $\binom{n-|I|}{k-|I|}$ sets containing I , and out of those, $\binom{n-|I \cup J|}{k-|I \cup J|}$ also contain J . Therefore the probability is exactly

$$\frac{\binom{n-|I \cup J|}{k-|I \cup J|}}{\binom{n-|I|}{k-|I|}} = \frac{(k-|I \cup J|+1) \cdots (k-|I|)}{(n-|I \cup J|+1) \cdots (n-|I|)} = \frac{k-|I \cup J|+1}{n-|I \cup J|+1} \cdots \frac{k-|I|}{n-|I|}.$$

Since $(k-\ell)/(n-\ell) \leq k/n$, we see that the probability is at most $p^{(n-|I|)-(n-|I \cup J|)} = p^{|J \setminus I|}$, as needed.

Using this bound, the same statement follows with an identical proof. \square

Generalizing [Corollary 7.4](#) is messier, since we need a *lower bound* on the event “ $S \supseteq I$ ”, where $|I| = d$.

Corollary 7.6. *If $c \leq 1$ and H is a d -uniform hypergraph with branching factor c/p then for all $I \in H$,*

$$\Omega(e^{-d}) \cdot (1 - 2^d c) p^d |H| \leq \Pr_{S \in \mathcal{V}_n(k)} [\text{Hit}_H \neq \emptyset] \leq p^d |H|.$$

Proof. The upper bound follows from the union bound, since the probability of hitting any particular set is at most p^d , as shown in the proof of [Lemma 7.5](#).

The original proof of the lower bound uses $\Pr_{S \sim \mu_p} [S \supseteq I] \geq p^d$ for a set I of size d . This inequality no longer holds in the uniform case, instead, we have (for the formula, see the proof of [Lemma 7.5](#))

$$\Pr_{S \sim \nu_n(k)} [S \supseteq I] = \frac{k-d+1}{n-d+1} \cdots \frac{k}{n} \geq \frac{k(k-1) \cdots (k-d+1)}{k^d} p^d.$$

We can lower bound the factor in front of p^d by

$$\left(1 - \frac{1}{k}\right) \cdots \left(1 - \frac{d-1}{k}\right) \geq \left(1 - \frac{1}{d}\right) \cdots \left(1 - \frac{d-1}{d}\right) = \frac{d!}{d^d},$$

which according to Stirling’s approximation is $\Omega(e^{-d})$. \square

This adds an additional $\Omega(e^d)$ factor to the very last display of the proof of [Lemma 7.2](#), but only affects some hidden constants in the statement of the Lemma.

References

- [AKKLR05] NOGA ALON, TALI KAUFMAN, MICHAEL KRIVELEVICH, SIMON LITSYN, and DANA RON. [Testing Reed-Muller codes](#). IEEE Trans. Inform. Theory, 51(11):4032–4039, 2005. (Preliminary version in 7th RANDOM, 2003). [4, 26](#)
- [ALMSS98] SANJEEV ARORA, CARSTEN LUND, RAJEEV MOTWANI, MADHU SUDAN, and MARIO SZEGEDY. [Proof verification and the hardness of approximation problems](#). J. ACM, 45(3):501–555, May 1998. (Preliminary version in 33rd FOCS, 1992). [eccc:1998/TR98-008. 1](#)

- [AS98] SANJEEV ARORA and SHMUEL SAFRA. *Probabilistic checking of proofs: A new characterization of NP*. J. ACM, 45(1):70–122, January 1998. (Preliminary version in 33rd FOCS, 1992). 1
- [BCHKS96] MIHIR BELLARE, DON COPPERSMITH, JOHAN HÅSTAD, MARCOS A. KIWI, and MADHU SUDAN. *Linearity testing in characteristic two*. IEEE Trans. Inform. Theory, 42(6):1781–1795, November 1996. (Preliminary version in 36th FOCS, 1995). 26
- [BKSSZ10] ARNAB BHATTACHARYYA, SWASTIK KOPPARTY, GRANT SCHOENEBECK, MADHU SUDAN, and DAVID ZUCKERMAN. *Optimal testing of Reed-Muller codes*. In Proc. 51st IEEE Symp. on Foundations of Comp. Science (FOCS), pages 488–497. 2010. [arXiv:0910.0641](#), [eccc:2009/TR09-086](#). 4, 27
- [BLR93] MANUEL BLUM, MICHAEL LUBY, and RONITT RUBINFELD. *Self-testing/correcting with applications to numerical problems*. J. Comput. Syst. Sci., 47(3):549–595, December 1993. (Preliminary version in 22nd STOC, 1990). 26
- [DD19] YOTAM DIKSTEIN and IRIT DINUR. *Agreement testing theorems on layered set systems*. In Proc. 60th IEEE Symp. on Foundations of Comp. Science (FOCS), pages 1495–1524. 2019. [arXiv:1909.00638](#), [eccc:2019/TR19-112](#). 6
- [DDGKS17] ROEE DAVID, IRIT DINUR, ELAZAR GOLDENBERG, GUY KINDLER, and IGOR SHINKAR. *Direct sum testing*. SIAM J. Comput., 46(4):1336–1369, 2017. (Preliminary version in 6th ITCS, 2015). [eccc:2014/TR14-002](#). 4
- [DFH17] IRIT DINUR, YUVAL FILMUS, and PRAHLADH HARSHA. *Low degree almost Boolean functions are sparse juntas*, 2017. (manuscript). [arXiv:1711.09428](#), [eccc:2017/TR17-180](#). , 5
- [DFH19] ———. *Analyzing Boolean functions on the biased hypercube via higher-dimensional agreement tests*. In Proc. 30th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA), pages 2124–2133. 2019. [arXiv:1711.09428](#), [1711.09426](#), [eccc:2017/TR17-180](#), [2017/TR17-181](#).
- [DG08] IRIT DINUR and ELAZAR GOLDENBERG. *Locally testing direct product in the low error range*. In Proc. 49th IEEE Symp. on Foundations of Comp. Science (FOCS), pages 613–622. 2008. 7
- [Dik19] YOTAM DIKSTEIN. *Agreement Theorems and Fourier Decomposition on High Dimensional Expanders*. Master’s thesis, Weizmann Institute of Science, Israel, 2019. 11
- [DK17] IRIT DINUR and TAL KAUFMAN. *High dimensional expanders imply agreement expanders*. In Proc. 58th IEEE Symp. on Foundations of Comp. Science (FOCS), pages 974–985. 2017. [eccc:2017/TR17-089](#). 2, 6
- [DKKMS18] IRIT DINUR, SUBHASH KHOT, GUY KINDLER, DOR MINZER, and MULI SAFRA. *Towards a proof of the 2-to-1 games conjecture?* In Proc. 50th ACM Symp. on Theory of Computing (STOC), pages 376–389. 2018. [eccc:2016/TR16-198](#). 5
- [DS14] IRIT DINUR and DAVID STEURER. *Direct product testing*. In Proc. 29th IEEE Conf. on Comput. Complexity, pages 188–196. 2014. [eccc:2013/TR13-179](#). 2, 7, 10
- [GGR98] ODED GOLDBREICH, SHAFI GOLDWASSER, and DANA RON. *Property testing and its connection to learning and approximation*. J. ACM, 45(4):653–750, July 1998. (Preliminary version in 37th FOCS, 1996). [eccc:1996/TR96-057](#). 6
- [GLRSW91] PETER GEMMELL, RICHARD J. LIPTON, RONITT RUBINFELD, MADHU SUDAN, and AVI WIGDERSON. *Self-testing/correcting for polynomials and for approximate functions*. In Proc. 23rd ACM Symp. on Theory of Computing (STOC), pages 32–42. 1991. 1
- [GS00] ODED GOLDBREICH and SHMUEL SAFRA. *A combinatorial consistency lemma with application to proving the PCP theorem*. SIAM J. Comput., 29(4):1132–1154, 2000. (Preliminary version in 1st RANDOM, 1997). [eccc:1996/TR96-047](#). 1
- [IKW12] RUSSELL IMPAGLIAZZO, VALENTINE KABANETS, and AVI WIGDERSON. *New direct-product testers and 2-query PCPs*. SIAM J. Comput., 41(6):1722–1768, 2012. (Preliminary version in 41st STOC, 2009). [eccc:2009/TR09-090](#). 7

- [KMS18] SUBHASH KHOT, DOR MINZER, and MULI SAFRA. *Pseudorandom sets in Grassmann graph have near-perfect expansion*. In *Proc. 59th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 592–601. 2018. [eccc:2018/TR18-006](#). 5
- [Raz98] RAN RAZ. *A parallel repetition theorem*. *SIAM J. Comput.*, 27(3):763–803, June 1998. (Preliminary version in *27th STOC*, 1995). 2
- [RS96] RONITT RUBINFELD and MADHU SUDAN. *Robust characterizations of polynomials with applications to program testing*. *SIAM J. Comput.*, 25(2):252–271, April 1996. (Preliminary version in *23rd STOC*, 1991 and *3rd SODA*, 1992). 1