

# Relative-error testing of conjunctions and decision lists

Xi Chen<sup>\*</sup>      William Pires<sup>†</sup>      Toniann Pitassi<sup>‡</sup>      Rocco A. Servedio<sup>§</sup>

September 14, 2025

## Abstract

We study the *relative-error* property testing model for Boolean functions that was recently introduced in the work of [CDH<sup>+</sup>25]. In relative-error testing, the testing algorithm gets uniform random satisfying assignments as well as black-box queries to  $f$ , and it must accept  $f$  with high probability whenever  $f$  has the property that is being tested and reject any  $f$  that is *relative-error* far from having the property. Here the relative-error distance from  $f$  to a function  $g$  is measured with respect to  $|f^{-1}(1)|$  rather than with respect to the entire domain size  $2^n$  as in the Hamming distance measure that is used in the standard model; thus, unlike the standard model, relative-error testing allows us to study the testability of *sparse* Boolean functions that have few satisfying assignments. It was shown in [CDH<sup>+</sup>25] that relative-error testing is at least as difficult as standard-model property testing, but for many natural and important Boolean function classes the precise relationship between the two notions is unknown.

In this paper we consider the well-studied and fundamental properties of being a *conjunction* and being a *decision list*. In the relative-error setting, we give an efficient one-sided error tester for conjunctions with running time and query complexity  $O(1/\varepsilon)$ .

Secondly, we give a two-sided relative-error  $\tilde{O}(1/\varepsilon)$  tester for decision lists, matching the query complexity of the state-of-the-art algorithm in the standard model [Bsh20, DLM<sup>+</sup>07].

---

<sup>\*</sup>Columbia University. Email: [xc2198@columbia.edu](mailto:xc2198@columbia.edu).

<sup>†</sup>Columbia University. Email: [wp2294@columbia.edu](mailto:wp2294@columbia.edu).

<sup>‡</sup>Columbia University. Email: [tonipitassi@gmail.com](mailto:tonipitassi@gmail.com).

<sup>§</sup>Columbia University. Email: [rocco@cs.columbia.edu](mailto:rocco@cs.columbia.edu).

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Our results and techniques . . . . .	3
<b>2</b>	<b>Notation, Definitions, and Preliminaries</b>	<b>5</b>
2.1	Notation . . . . .	5
2.2	Background on property testing and relative-error testing . . . . .	5
2.3	From relative-error testing to standard testing . . . . .	6
<b>3</b>	<b>A relative-error testing algorithm for conjunctions</b>	<b>8</b>
3.1	Preliminaries . . . . .	9
3.2	Intuition: the tester of [PRS02] and our tester . . . . .	9
3.3	The relative-error anti-monotone conjunction tester. . . . .	10
3.4	Completeness . . . . .	11
3.5	Soundness . . . . .	11
3.5.1	Analysis of phase 1 (steps 1-3) . . . . .	11
3.5.2	Analysis of phase 2 (lines 4-9) . . . . .	15
3.5.3	Putting the pieces together . . . . .	16
3.6	Robustness of CONJ-TEST to faulty oracles . . . . .	17
<b>4</b>	<b>A relative-error testing algorithm for decision lists</b>	<b>17</b>
4.1	Basics about decision lists . . . . .	18
4.2	The Algorithm . . . . .	19
4.3	Query Complexity . . . . .	21
4.4	Completeness: Correctness of Algorithm 3 when $f$ is a decision list . . . . .	23
4.4.1	Case 1 . . . . .	23
4.4.2	Case 2 . . . . .	24
4.5	Soundness: Correctness of Algorithm 3 when $f$ is far from decision lists . . . . .	27
<b>A</b>	<b>Lower Bounds</b>	<b>33</b>

# 1 Introduction

**Background.** Over the past few decades, property testing of Boolean functions has grown into a rich and active research area (see e.g. the extensive treatment in [Gol17, BY22]). Emerging from the study of program testing and self-correction [BLR93, GLR<sup>+</sup>91], the field has developed strong connections to many other topics in theoretical computer science including complexity theory, Boolean function analysis, and computational learning theory, see e.g. [Gol20, O’D14, Ron08].

In the “standard” model of Boolean function property testing, the algorithm for testing an unknown Boolean function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  has black-box oracle (also known as “membership query” and will be referred to as  $\text{MQ}(f)$  in this paper) access as its only mode of access to  $f$ . The goal of a tester in this standard model is to distinguish between the two alternatives that (a)  $f$  has some property of interest, versus (b)  $f$  is “far” (in the sense of disagreeing on at least an  $\varepsilon$ -fraction of all the  $2^n$  inputs) from every function with the property. A wide range of properties (or equivalently, classes of Boolean functions) have been studied from this perspective, including but not limited to monotone Boolean functions [GGL<sup>+</sup>00, FLN<sup>+</sup>02, CS13a, CS13b, CST14, CDST15, BB16, CWX17a, KMS18], unate Boolean functions [KS16, CS16, LW19, CW19, CWX17a, CWX17b, CW19], literals [PRS02], conjunctions [PRS02, GR20], decision lists [DLM<sup>+</sup>07, Bsh20], linear threshold functions [MORS10],  $s$ -term monotone DNF [PRS02, DLM<sup>+</sup>07, CGM11, Bsh20],  $s$ -term DNF, size- $s$  decision trees, size- $s$  branching programs, size- $s$  Boolean formulas, size- $s$  Boolean circuits, functions with Fourier degree at most  $d$ ,  $s$ -sparse polynomials over  $\mathbb{F}_2$  [DLM<sup>+</sup>07, CGM11, Bsh20], and many others (see e.g. Figure 1 of [Bsh20]).

As the field of Boolean function property testing has grown more mature, an interest has developed in considering alternative variants of the “standard” property testing model discussed above. One early variant was the *distribution-free* property testing model [GGR98, HK03], inspired by Valiant’s distribution-free PAC learning model [Val84]. In this model the distance between two functions  $f$  and  $g$  is measured according to  $\Pr_{\mathbf{x} \sim \mathcal{D}}[f(\mathbf{x}) \neq g(\mathbf{x})]$ , where  $\mathcal{D}$  is an unknown and arbitrary distribution over the domain  $\{0, 1\}^n$ . In addition to making black-box oracle queries, distribution-free testing algorithms can also draw i.i.d. samples from the unknown distribution  $\mathcal{D}$ . The motivation behind this model was to align property testing more closely with the influential distribution-free PAC learning model and to capture scenarios of interest in which there is an underlying non-uniform distribution over the domain. Unfortunately, many of the results that have been established for the distribution-free PAC model are negative in nature [HK05, GS09, CX16, CP22, CFP24], often showing that almost as many oracle calls are needed for distribution-free testing a class  $\mathcal{C}$  as would be required for PAC learning  $\mathcal{C}$ . (As shown in [GGR98], the PAC learning complexity of any class  $\mathcal{C}$  gives an easy upper bound on its distribution-free testing complexity.)

**Relative-error testing.** Very recent work [CDH<sup>+</sup>25] introduced a new model of Boolean function property testing, called *relative-error testing*. The motivation for the model comes from the observation that the standard testing framework is not well suited for testing *sparse* Boolean functions, i.e., functions  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  that have  $|f^{-1}(1)| = p2^n$  where  $p$  is very small<sup>1</sup> — perhaps as small as  $1/n^{\omega(1)}$  or even  $2^{-\Theta(n)}$  — simply because any such function will be  $p$ -close to the constant-0 function. (This is closely analogous to the well-known fact that the standard dense graph property testing model, where error  $\varepsilon$  corresponds to adding or deleting  $\varepsilon n^2$  potential edges in an  $n$ -vertex graph, is not well suited to testing sparse graphs, because any sparse graph trivially has small distance from the empty graph on  $n$  vertices.)

<sup>1</sup>Such functions can be of significant interest for practical as well as theoretical reasons; since Boolean functions correspond to classification rules, a sparse Boolean function corresponds to a classification rule which outputs 1 only on positive examples of some rare but potentially desirable phenomenon.

Motivated by the above considerations, the relative-error property testing model of [CDH<sup>+</sup>25] defines the distance between the unknown  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  and another function  $g$  to be

$$\text{rel-dist}(f, g) := \frac{|f^{-1}(1) \triangle g^{-1}(1)|}{|f^{-1}(1)|},$$

so relative distance is measured *at the scale of*  $|f^{-1}(1)|$  rather than at the “absolute” scale of  $2^n = |\{0, 1\}^n|$ . In addition to the black-box oracle  $\text{MQ}(f)$ , the model also allows the testing algorithm to sample i.i.d. uniform random satisfying assignments of  $f$  via a sampling oracle  $\text{Samp}(f)$ : each call to  $\text{Samp}(f)$  returns a uniformly random sample from  $f^{-1}(1)$  (see Section 2.2 and Remark 4 for a detailed description of the model and comparison with the distribution-free testing model).

The initial work [CDH<sup>+</sup>25] established some relations between the query complexity of standard-model testing and relative-error testing. In more detail, letting  $q_{\text{std}}(\varepsilon, n)$  and  $q_{\text{rel}}(\varepsilon, n)$  denote the number of oracle calls needed to test a property  $\mathcal{C}$  in the standard model and the relative-error model, respectively, they showed (as Facts 8 and 9 in [CDH<sup>+</sup>25]) that

$$q_{\text{std}}(\varepsilon, n) \lesssim q_{\text{rel}}(\varepsilon, n)/\varepsilon \quad \text{and} \quad q_{\text{rel}}(\varepsilon, n) \lesssim q_{\text{std}}(p\varepsilon, n). \quad (1)$$

The first inequality shows that standard-model testing is never significantly harder than relative-error testing, but since the value of  $p = |f^{-1}(1)|/2^n$  can be extremely small, the second inequality leaves open the possibility that relative-error testing can be much harder than standard-model testing. Indeed, [CDH<sup>+</sup>25] gave an example of an property which is constant-query testable in the standard model but requires *exponentially* many queries for relative-error testing. However, the class of functions exhibiting this separation<sup>2</sup> is arguably contrived. This leads to the following question, which is the motivation for our work:

What is the complexity of relative-error testing  
for *natural and well-studied* properties of Boolean functions?

Progress in this direction was made by [CDH<sup>+</sup>25], which give a relative-error tester for monotonicity which makes at most quadratically more queries than the best standard-model tester.

**The problems we consider.** We attack the above broad question by studying the relative-error testability of two basic and natural classes of Boolean functions, namely *conjunctions* and *decision lists*.<sup>3</sup> Conjunctions and decision lists have been thoroughly studied in both the standard model of Boolean function property testing and the distribution-free model [PRS02, DLM<sup>+</sup>07, GS09, CX16, GR20, Bsh20, CFP24]. For the standard model, unlike monotonicity, both problems belong to the regime of “efficiently testable” properties, where the number of queries needed depends on  $\varepsilon$  only but not the dimension  $n$ : [PRS02, GR20] gave  $O(1/\varepsilon)$ -query algorithms for testing conjunctions and [Bsh20] gave an  $\tilde{O}(1/\varepsilon)$ -query algorithm for testing decision lists (improving on an earlier  $\tilde{O}(1/\varepsilon^2)$ -query algorithm of [DLM<sup>+</sup>07]). In contrast, distribution-free testing is provably much more difficult: improving on [GS09], [CX16] gave an  $\tilde{\Omega}(n^{1/3})$  lower bound for conjunctions (and a matching upper bound); [CFP24] gave an  $\tilde{\Omega}(n^{1/2})$  lower bound for decision lists (and an  $\tilde{O}(n^{11/12})$  upper bound).

<sup>2</sup>The class consists of all functions  $g$  for which  $|g^{-1}(1)|$  is a multiple of  $2^{2n/3}$  (see Appendix A of [CDH<sup>+</sup>25]).

<sup>3</sup>Recall that a *conjunction* is an AND of literals  $\ell_1 \wedge \dots \wedge \ell_k$ , where each  $\ell_i$  is either some Boolean variable  $x_j$  or its negation  $\bar{x}_j$ . A *decision list* is a sequence of nested “if (condition) holds, then output (bit), else ...” rules, where each (condition) is a Boolean literal (see the beginning of Section 4 for the formal definition of decision lists).

One intuitive reason for this disparity between the standard model and the distribution-free model is that, *under the uniform distribution*, (i) *every conjunction is either  $\varepsilon$ -close to the constant-0 function or else is a  $\log(1/\varepsilon)$ -junta*, and (ii) *every decision list is  $\varepsilon$ -close to a  $\log(1/\varepsilon)$ -junta*.<sup>4</sup> In the distribution-free setting, though, these facts fail, and notably, they fail in the relative-error setting as well. Given this, it is *a priori* unclear whether one should expect the relative-error testability of conjunctions and decision lists to be more like the standard setting, or more like the distribution-free setting. Resolving this is one of the motivations of the present work.

## 1.1 Our results and techniques

**Results.** As our main results, we show that both classes of conjunctions and decision lists are efficiently testable in the relative-error model; indeed we give relative-error testing algorithms that have the same query complexity as the best standard-model testing algorithms.

For conjunctions we prove the following theorem:

**Theorem 1** (Relative-error testing of conjunctions). There is a one-sided non-adaptive algorithm which is an  $\varepsilon$ -relative-error tester for conjunctions on  $\{0, 1\}^n$ . The algorithm makes  $O(1/\varepsilon)$  calls to the sampling oracle  $\text{Samp}(f)$  and the membership query oracle  $\text{MQ}(f)$ .

Readers who are familiar with the literature on testing conjunctions in the standard setting may have noticed that the algorithm above is one-sided and uses  $O(1/\varepsilon)$  queries, while only two-sided algorithms are known for testing conjunctions in the standard model with  $O(1/\varepsilon)$  queries [PRS02] and the current best one-sided algorithm needs  $\tilde{O}(1/\varepsilon)$  queries [GR20]. It was posed as an open problem in [Bsh20] whether there exists a one-sided,  $O(1/\varepsilon)$ -query tester for conjunctions in the standard model. We resolve this question by giving such a tester as a corollary of **Theorem 1**:

**Theorem 2.** There is a one-sided, adaptive algorithm which is an  $\varepsilon$ -error tester for conjunctions over  $\{0, 1\}^n$  in the standard setting. The algorithm makes  $O(1/\varepsilon)$  calls to  $\text{MQ}(f)$ .

Note that combining the first part of **Equation (1)** and **Theorem 1** will only lead to a one-sided upper bound of  $O(1/\varepsilon^2)$  for the standard model, where the extra  $1/\varepsilon$  factor comes from the natural fact that whenever  $f$  has density at least  $\varepsilon$ , it takes  $1/\varepsilon$  uniformly random samples to simulate one sample from  $f^{-1}(1)$ . To prove **Theorem 2**, we obtain a more efficient reduction from relative-error testing to standard testing which, informally, states

$$q_{\text{std}}(\varepsilon, n) \lesssim 1/\varepsilon + q_{\text{rel}}(\varepsilon/2, n), \quad (2)$$

from which **Theorem 2** follows. The reduction is presented in **Section 2.3**.

Our second, and more involved, main result obtains an efficient relative-error testing algorithm for decision lists:

**Theorem 3** (Relative-error testing of decision lists). There is an algorithm which is an  $\varepsilon$ -relative-error tester for decision lists on  $\{0, 1\}^n$ . The algorithm makes  $\tilde{O}(1/\varepsilon)$  calls to  $\text{Samp}(f)$  and  $\text{MQ}(f)$ .

We remark that the algorithm of **Theorem 3** crucially uses the relative-error testing algorithm for conjunctions, i.e. **Theorem 1**, as a subroutine; this is explained more below and a more detailed overview of the algorithm can be found in **Section 4.2**.

We also note that the algorithm of **Theorem 3** is adaptive. However, an easy variant of our algorithm and analysis yields a non-adaptive  $\varepsilon$ -relative-error tester for decision lists that makes  $O(1/\varepsilon)$  calls to  $\text{Samp}(f)$  and  $\tilde{O}(1/\varepsilon^2)$  calls to  $\text{MQ}(f)$ .

---

<sup>4</sup>A  $k$ -junta is a function that only depends on at most  $k$  of its variables.

Finally, in [Appendix A](#), we give a lower bound of  $\Omega(1/\varepsilon)$  for any two-sided, adaptive algorithm for the relative-error testing of both conjunctions and decision lists, thus showing that [Theorem 1](#) is optimal and [Theorem 3](#) is nearly optimal.

**Techniques.** We give a brief and high-level overview of our techniques, referring the reader to [Section 3.2](#) and [Section 4.2](#) for more details. Our relative-error testing algorithm for conjunctions begins with a simple reduction to the problem of relative-error testing whether an unknown  $f$  is an anti-monotone conjunction<sup>5</sup>. We solve the relative-error anti-monotone conjunction testing problem with an algorithm which, roughly speaking, first checks whether  $f^{-1}(1)$  is close to a linear subspace (as would be the case if  $f$  were an anti-monotone conjunction), and then checks whether  $f$  is close to anti-monotone (in a suitable sense to meet our requirements). This is similar, at a high level, to the approach from [\[PRS02\]](#), but as discussed in [Section 3.2](#) there are various technical differences; in particular we are able to give a simpler analysis, even for the more general relative-error setting, which achieves one-sided rather than two-sided error.

For decision lists, our starting point is the simple observation that when  $f$  is a sparse decision list then  $f$  can be written as  $C \wedge L$  where  $C$  is a conjunction and  $L$  is a decision list (intuitively, since  $f$  is sparse, a large prefix of the list of output bits of the decision list must all be 0, which corresponds to a conjunction  $C$ ); this decomposition plays an important role for us. Another simple but useful observation is that all of the variables in the conjunction  $C$  must be “unanimous,” in the sense that each one of them must always take the same value across all of the positive examples. Our algorithm draws an initial sample of positive examples from  $\text{Samp}(f)$  and then works with a restriction of  $f$  which fixes the variables which were seen to be “unanimous” in that initial sample. Intuitively, if  $f$  is a decision list, then it should be the case that the resulting restricted function is close to a *non-sparse* decision list, so we can check that the restricted function is non-sparse, and, if the check passes, apply a standard-model decision list tester to the restricted function.

This gives some intuition for how to design an algorithm that accepts in the yes-case, but the fact that we must reject all functions that are relative-error far from decision lists leads to significant additional complications. To deal with this, we need to incorporate additional checks into the test, and we give an analysis showing that if a function  $f$  passes all of these checks then it must be close (in a relative-error sense) to a function of the form  $C' \wedge L'$  where  $C'$  is relative-error close to a conjunction and  $L'$  is relative-error close to a decision list; this lets us conclude that  $f$  is relative-error close to a decision list, as desired. These checks involve applying our relative-error conjunction testing algorithm (which we show to have useful robustness properties against being run with a slightly imperfect sampling oracle — this robustness is crucial for our analysis); we refer the reader to [Section 4.2](#) for more details.

Looking ahead, while our results resolve the relative-error testability of conjunctions and decision lists, there are many open questions remaining about how relative-error testing compares against standard-error testing for other natural and well-studied Boolean function classes, such as  $s$ -term DNF formulas,  $k$ -juntas, subclasses of  $k$ -juntas, unate functions, and beyond. All of these classes have been intensively studied in the standard testing model, see e.g. [\[DLM<sup>+</sup>07, CGM11, KS16, CS16, LW19, CW19, CWX17a, CWX17b, CW19, Bsh20\]](#); it is an interesting direction for future work to establish either efficient relative-error testing algorithms, or relative-error testing lower bounds, for these classes.

**Organization.** [Section 2](#) contains preliminary notation and definitions, and the general reduction from relative-error testing to standard testing is presented in [Section 2.3](#). We prove [Theorem 1](#) in [Section 3](#), and we prove [Theorem 3](#) in [Section 4](#).

---

<sup>5</sup>A conjunction is an anti-monotone conjunction if it only contains negated literals.

## 2 Notation, Definitions, and Preliminaries

### 2.1 Notation

For  $x, y \in \{0, 1\}^n$ , we write  $y \preceq x$  to indicate that  $y_i \leq x_i$  for all  $i \in [n]$ . We write  $x \oplus y$  to denote the string in  $\{0, 1\}^n$  whose  $i$ -th coordinate is the XOR of  $x_i$  and  $y_i$ . Given  $S \subseteq [n]$  and  $z \in \{0, 1\}^n$ , we denote by  $z_S$  the string in  $\{0, 1\}^S$  that agrees with  $z$  on every coordinate in  $S$ .

We use standard notation for restrictions of functions. For  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  and  $u \in \{0, 1\}^S$  for some  $S \subseteq [n]$ , the function  $f|_u : \{0, 1\}^{[n] \setminus S} \rightarrow \{0, 1\}$  is defined as the function

$$f|_u(x) = f(z), \quad \text{where} \quad z_i = \begin{cases} u_i & \text{if } i \in S, \\ x_i & \text{if } i \in [n] \setminus S. \end{cases}$$

We write  $\mathcal{U}_S$  to denote the uniform distribution over a finite set  $S$ , and we write  $\mathcal{U}_n$  to denote the uniform distribution over  $\{0, 1\}^n$ ; we usually skip the subscript  $n$  when it is clear from context.

When we refer to vector spaces/subspaces, we will always be talking about the vector space  $\mathbb{F}_2^n$ , and we will frequently identify  $\{0, 1\}^n$  with the vector space  $\mathbb{F}_2^n$ .

Finally, we write DL to denote the class of all decision lists over  $\{0, 1\}^n$ , and Conj to denote the class of all conjunctions over  $\{0, 1\}^n$ .

### 2.2 Background on property testing and relative-error testing

We recall the basics of the standard property testing model. In the *standard* testing model for a class  $\mathcal{C}$  of  $n$ -variable Boolean functions, the testing algorithm is given oracle access  $\text{MQ}(f)$  to an unknown and arbitrary function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ . The algorithm must output “yes” with high probability (say at least  $2/3$ ; this can be amplified using standard techniques) if  $f \in \mathcal{C}$ , and must output “no” with high probability (again, say at least  $2/3$ ) if  $\text{dist}(f, \mathcal{C}) \geq \varepsilon$ , where

$$\text{dist}(f, \mathcal{C}) := \min_{g \in \mathcal{C}} \text{dist}(f, g) \quad \text{and} \quad \text{dist}(f, g) := \frac{|f^{-1}(1) \triangle g^{-1}(1)|}{2^n}.$$

As defined in [CDH<sup>+</sup>25], a *relative-error* testing algorithm for  $\mathcal{C}$  has oracle access to  $\text{MQ}(f)$  and also has access to a  $\text{Samp}(f)$  oracle which, when called, returns a uniform random element  $\mathbf{x} \sim f^{-1}(1)$ . A relative-error testing algorithm for  $\mathcal{C}$  must output “yes” with high probability (say at least  $2/3$ ; again this can be easily amplified) if  $f \in \mathcal{C}$  and must output “no” with high probability (again, say at least  $2/3$ ) if  $\text{rel-dist}(f, \mathcal{C}) \geq \varepsilon$ , where

$$\text{rel-dist}(f, \mathcal{C}) := \min_{g \in \mathcal{C}} \text{rel-dist}(f, g) \quad \text{and} \quad \text{rel-dist}(f, g) := \frac{|f^{-1}(1) \triangle g^{-1}(1)|}{|f^{-1}(1)|}.$$

Finally, recall that a property testing algorithm for a class of functions  $\mathcal{C}$  (in either the standard model or the relative-error model) is said to be *one-sided* if it returns “yes” with probability 1 when the unknown function belongs to  $\mathcal{C}$ .

**Remark 4** (Relative-error testing versus distribution-free testing). Like the distribution-free model, the relative-error model gives the testing algorithm access to both black-box oracle calls to  $f$  and i.i.d. samples from a distribution  $\mathcal{D}$ . In distribution-free testing  $\mathcal{D}$  is unknown and arbitrary, whereas in relative-error testing  $\mathcal{D}$  is unknown but is guaranteed to be the uniform distribution over  $f^{-1}(1)$ . However, we remark that relative-error testing is *not* a special case of distribution-free testing, because the distance between functions is measured differently between the two settings.



In the distribution-free model, the distance  $\Pr_{\mathbf{x} \sim \mathcal{D}}[f(\mathbf{x}) \neq g(\mathbf{x})]$  between two functions  $f, g$  is measured vis-a-vis the distribution  $\mathcal{D}$  that the testing algorithm can sample from. In contrast, in the relative-error setting the distribution  $\mathcal{D}$  that the tester can sample from is  $\mathcal{D} = \mathcal{U}_{f^{-1}(1)}$ , but  $\text{rel-dist}(f, g)$  is *not* equal to  $\Pr_{\mathbf{x} \sim \mathcal{U}_{f^{-1}(1)}}[f(\mathbf{x}) \neq g(\mathbf{x})]$ ; indeed, a function  $g$  can have large relative distance from  $f$  because  $g$  disagrees with  $f$  on many points outside of  $f^{-1}(1)$  (which are points that have zero weight under the distribution  $\mathcal{D}$ ).

We will need the following simple “approximate triangle inequality” for relative distance:

**Lemma 5** (Approximate triangle inequality for relative distance). Let  $f, g, h : \{0, 1\}^n \rightarrow \{0, 1\}$  be such that  $\text{rel-dist}(f, g) \leq \varepsilon$  and  $\text{rel-dist}(g, h) \leq \varepsilon'$ . Then  $\text{rel-dist}(f, h) \leq \varepsilon + (1 + \varepsilon)\varepsilon'$ .

*Proof.* Since  $\text{rel-dist}(f, g) \leq \varepsilon$  we have that

$$\frac{|f^{-1}(1) \triangle g^{-1}(1)|}{|f^{-1}(1)|} \leq \varepsilon,$$

which means we must have  $|g^{-1}(1)| \leq (1 + \varepsilon)|f^{-1}(1)|$ . Hence

$$\begin{aligned} \text{rel-dist}(f, h) &= \frac{|f^{-1}(1) \triangle h^{-1}(1)|}{|f^{-1}(1)|} \\ &= \frac{1}{|f^{-1}(1)|} |\{z : f(z) \neq h(z)\}| \\ &\leq \frac{1}{|f^{-1}(1)|} |\{z : f(z) \neq g(z) \vee g(z) \neq h(z)\}| \\ &\leq \frac{1}{|f^{-1}(1)|} |\{z : f(z) \neq g(z)\}| + \frac{1}{|f^{-1}(1)|} |\{g(z) \neq h(z)\}| \\ &\leq \text{rel-dist}(f, g) + (1 + \varepsilon) \frac{1}{|g^{-1}(1)|} |\{g(z) \neq h(z)\}| \\ &= \text{rel-dist}(f, g) + (1 + \varepsilon) \text{rel-dist}(g, h) \\ &\leq \varepsilon + (1 + \varepsilon)\varepsilon'. \end{aligned} \quad \square$$

Finally, we recall two results for testing decision lists in the standard model that we will use:

**Theorem 6** ([Bsh20]). There is an adaptive algorithm  $\mathcal{A}$  for testing decision lists in the standard model with query complexity  $\tilde{O}(1/\varepsilon)$ . In particular, if  $f$  is a decision list then  $\mathcal{A}$  accepts with probability at least  $19/20$ ; if  $\text{dist}(f, \text{DL}) \geq \varepsilon$  then  $\mathcal{A}$  rejects with probability at least  $19/20$ .

**Theorem 7** ([DLM<sup>+</sup>07]). There is a non-adaptive algorithm  $\mathcal{B}$  for testing decision lists in the standard model with query complexity  $\tilde{O}(1/\varepsilon^2)$ . In particular, if  $f$  is a decision list then  $\mathcal{B}$  accepts with probability at least  $19/20$ ; if  $\text{dist}(f, \text{DL}) \geq \varepsilon$  then  $\mathcal{B}$  rejects with probability at least  $19/20$ .

### 2.3 From relative-error testing to standard testing

In this section, we use our relative-error tester for conjunctions (see [Theorem 1](#)) to give an optimal one-sided tester for this class in the standard model. Fact 8 of [CDH<sup>+</sup>25] gives a way of transforming any relative-error tester for a property  $\mathcal{C}$  into a tester in the standard model for  $\mathcal{C}$ . However, this transformation incurs a multiplicative  $1/\varepsilon$  blowup in the number of oracle calls made by the tester. We show that if the all-0 function is in  $\mathcal{C}$  and the number of oracle calls made by the relative-error tester grows at least linearly in  $1/\varepsilon$ , we can remove this blowup.



**Input:** Oracle access to  $\text{MQ}(f)$  and an accuracy parameter  $\varepsilon > 0$ .

**Output:** “Reject” or “Accept.”

**Phase 1: Estimating  $|f^{-1}(1)|$ .**

- 1: Draw  $c_1/\varepsilon$  many samples  $\mathbf{z} \sim \mathcal{U}_n$  and let  $\mathbf{S}$  be the set of samples.
- 2: For each  $\mathbf{z} \in \mathbf{S}$  query  $f(\mathbf{z})$  and let  $\hat{p}$  be the fraction of  $\mathbf{z} \in \mathbf{S}$  with  $f(\mathbf{z}) = 1$ .
- 3: If  $\hat{p} \leq \varepsilon/2$ , then halt and accept  $f$ .

**Phase 2: Getting samples**

- 4: Draw  $c_2 \cdot q(\varepsilon/2, n)$  many samples  $\mathbf{z} \sim \mathcal{U}_n$  and let  $\mathbf{G}$  be set of samples with  $f(\mathbf{z}) = 1$ .
- 5: If  $|\mathbf{G}| \leq q(\varepsilon/2\hat{p}, n)$ , then halt and accept  $f$ .

**Phase 3: Running the relative error tester**

- 6: Run  $\mathcal{A}$  using the samples in  $\mathbf{G}$  to simulate  $\text{Samp}(f)$  and distance parameter  $\varepsilon/2\hat{p}$ .
- 7: Output what  $\mathcal{A}$  outputs.

**Algorithm 1:** Standard-model testing algorithm for  $\mathcal{C}$  using a relative-error tester  $\mathcal{A}$ . Here both  $c_1$  and  $c_2$  are sufficiently large absolute constants.

**Lemma 8.** Let  $\mathcal{C}$  be a class of Boolean functions over  $\{0, 1\}^n$  that contains the all-0 function. Suppose that there is a relative-error  $\varepsilon$ -testing algorithm  $\mathcal{A}$  for  $\mathcal{C}$  that makes  $q(\varepsilon, n)$  many calls to  $\text{MQ}(f)$  and  $\text{Samp}(f)$  such that  $q(\varepsilon/\delta, n) \leq \delta q(\varepsilon, n)$  for every  $0 < \delta < 1$ . Then there is an adaptive standard-model  $\varepsilon$ -testing algorithm  $\mathcal{B}$  for  $\mathcal{C}$  which makes at most  $O(1/\varepsilon + q(\varepsilon/2, n))$  calls to  $\text{MQ}(f)$ . Furthermore, if  $\mathcal{A}$  is one-sided, so is  $\mathcal{B}$ .

Before entering into the formal proof, we give some intuition. To prove [Lemma 8](#), we want to run  $\mathcal{A}$  on  $f$  to see if  $f$  is  $\varepsilon$ -far from the class. For this, we will need uniform samples from  $f^{-1}(1)$ ; these can be obtained by randomly querying points of the hypercube to find 1s of  $f$ . However if  $\Pr_{\mathbf{z} \sim \mathcal{U}_n}[f(\mathbf{z}) = 1] =: p_f$  is small, then we would need roughly  $1/p_f$  queries per sample. The following key lemma shows that if  $f$  is sparse and  $\varepsilon$ -far from  $\mathcal{C}$ , then its relative distance to  $\mathcal{C}$  is at least  $\varepsilon/p_f$ . This means we can simulate  $\mathcal{A}$  on  $f$  with the distance parameter for  $\mathcal{A}$  set to a higher value  $\varepsilon$ . Hence while getting samples is hard, this is canceled out by the fact we need less of them.

**Lemma 9.** If  $\text{dist}(f, \mathcal{C}) > \varepsilon$ , then  $\text{rel-dist}(f, \mathcal{C}) > \varepsilon/p_f$  where  $p_f := |f^{-1}(1)|/2^n$ .

*Proof.* Assume for a contradiction that  $\text{rel-dist}(f, \mathcal{C}) \leq \varepsilon/p_f$  for some  $g \in \mathcal{C}$ . Then we have that:

$$\frac{|f^{-1}(1) \triangle g^{-1}(1)|}{|f^{-1}(1)|} = \frac{|f^{-1}(1) \triangle g^{-1}(1)|}{p_f 2^n} \leq \frac{\varepsilon}{p_f},$$

which implies that  $|f^{-1}(1) \triangle g^{-1}(1)| \leq \varepsilon 2^n$  and thus,  $\text{dist}(f, g) \leq \varepsilon$ , a contradiction.  $\square$

We now prove [Lemma 8](#).

*Proof of Lemma 8.* Our goal will be to show that [Algorithm 1](#) is a standard testing algorithm for  $\mathcal{C}$ . Without loss of generality, we assume that  $\mathcal{A}$  is correct with high enough probability, that is, if

$\text{rel-dist}(f, \mathcal{C}) \geq \varepsilon$  the algorithm rejects with probability at least 0.9. Similarly, if  $f \in \mathcal{C}$ , we assume that  $\mathcal{A}$  accepts with probability at least 0.9.

First, consider the case where  $f \in \mathcal{C}$ . Note that  $f$  can only be rejected if  $\mathcal{A}$  rejects  $f$ . Since  $\mathbf{G}$  is a set of uniformly random samples from  $f^{-1}(1)$ ,  $\mathcal{A}$  must accept with probability at least 0.9. And if  $\mathcal{A}$  is one-sided, then  $f$  is always accepted, meaning that Algorithm 1 is also one-sided.

Now, assume that  $f$  is  $\varepsilon$ -far from any function in  $\mathcal{C}$ . Because the all-0 function is in  $\mathcal{C}$  we must have  $p_f := |f^{-1}(1)|/2^n \geq \varepsilon$ . So on line 2, we have by a Chernoff bound that

$$\Pr[|\hat{p} - p_f| > 0.05p_f] \leq 2e^{-\frac{(0.05)^2 \varepsilon}{2+0.05} c_1 / \varepsilon} \leq 0.05,$$

by picking a suitably large constant  $c_1$ . Below we assume that  $\hat{p} \in (1 \pm 0.05)p_f$ .

Since  $p_f \geq \varepsilon$ , this implies  $\hat{p} \geq \varepsilon/2$ . So, the algorithm doesn't accept on line 3. We also have that  $\Pr_{z \sim \mathcal{U}_n}[f(z) = 1] \geq 0.9\hat{p}$ . Hence, the expected size of  $|\mathbf{G}|$  is at least  $0.9c_2\hat{p} \cdot q(\varepsilon/2, n)$ . By our assumption on  $q(\cdot, \cdot)$  this is at least  $0.9c_2 \cdot q(\varepsilon/2\hat{p}, n)$ . Again, by a Chernoff bound and by picking  $c_2$  to be a sufficiently large constant, we have that  $|\mathbf{G}| \leq q(\varepsilon/2\hat{p}, n)$  with probability at most 0.05.

It remains to argue that  $\mathcal{A}$  rejects with high probability on line 6. By Lemma 9, we have that  $\text{rel-dist}(f, \mathcal{C}) \geq \varepsilon/p_f$ . Using  $\hat{p} \geq 0.95p_f$ , we have that  $\text{rel-dist}(f, \mathcal{C}) \geq \varepsilon/2\hat{p}$ . Since we simulate  $\mathcal{A}$  on line 6 using uniform samples from  $f^{-1}(1)$ , it must accept  $f$  with probability at most 0.1. As a result, the probability the algorithm accepts  $f$  is at most  $0.05 + 0.05 + 0.1 \leq 0.2$ .

Line 2 uses  $O(1/\varepsilon)$  many queries, line 4 uses  $O(q(\varepsilon/2, n))$  many queries, and simulating  $\mathcal{A}$  on line 6 uses  $q(\varepsilon/2\hat{p}, n) = O(q(\varepsilon/2, n))$  queries. So, the overall query complexity is  $O(1/\varepsilon + q(\varepsilon/2, n))$ .  $\square$

Using the above, we can obtain a one-sided tester for conjunctions in the standard setting from our one-sided relative-error conjunction tester. The result of [BG22] implies an  $\Omega(1/\varepsilon)$  lower bound for this problem, hence this result is tight.

*Proof of Theorem 2.* Our relative-error tester for conjunctions in Theorem 1 makes  $O(1/\varepsilon)$  calls to  $\text{Samp}(f)$  and  $\text{MQ}(f)$  and is one-sided. As the all-0 function is a conjunction, the theorem follows by Lemma 8.  $\square$

### 3 A relative-error testing algorithm for conjunctions

In this section we prove Theorem 1. At the end of the section, we show that the main algorithm for conjunctions works even when its access to the sampling oracle is not perfect; this will be important when the algorithm is used as a subroutine to test decision lists in the next section.

**Remark 10.** In the rest of this section, we give a relative-error tester for the class of *anti-monotone* conjunctions. That is, for conjunctions of *negated* variables. Using an observation from [PRS02] (see also [GR20]), this implies a relative-error tester for general conjunctions. Specifically, a slightly modified version of Observation 3 from Parnas et al. [PRS02] states that for any conjunction  $f$  (not necessarily anti-monotone) and for any  $y \in f^{-1}(1)$ , the function

$$f_y(x) := f(x \oplus y)$$

is an anti-monotone conjunction. Furthermore, it is easy to see that if  $f$  is  $\varepsilon$ -relative-error far from every conjunction, then  $f_y$  must be  $\varepsilon$ -relative-error far from every anti-monotone conjunction. So to get a relative-error tester for general conjunctions, we simply draw a single  $y \sim f^{-1}(1)$  and then run the tester for anti-monotone conjunctions (Algorithm 2) on  $f_y$ , where  $\text{MQ}(f_y)$  and  $\text{Samp}(f_y)$  can be simulated using  $\text{MQ}(f)$  and  $\text{Samp}(f)$ , respectively; we refer to this algorithm as CONJ-TEST.

### 3.1 Preliminaries

**Definition 11** (Subspaces). A set  $W \subseteq \{0,1\}^n$  is an affine subspace over  $F_2^n$  if and only if  $W$  is the set of solutions to a system of linear equations over  $F_2^n$ . That is, iff  $W$  is the set of solutions to  $Ax = b$ , for some  $b \in \{0,1\}^n$ .  $W$  is a linear subspace over  $F_2^n$  iff it is the set of solutions to  $Ax = b$  where  $b$  is the all-0 vector. A well-known equivalent characterization states that  $W$  is an affine subspace if and only if for all  $x, y, z \in W$ ,  $x \oplus y \oplus z \in W$ . And for linear subspaces, we have that  $W$  is linear if and only if for all  $x, y \in W$ ,  $x \oplus y \in W$ .

We will need the following fact about the intersection of linear subspaces.

**Fact 12.** Let  $H, H'$  be two linear subspaces of  $\{0,1\}^n$  with  $H \not\subseteq H'$ . Then:

$$\frac{|H \cap H'|}{|H|} \leq \frac{1}{2}.$$

**Definition 13** (Anti-monotone set). A set  $H \subseteq \{0,1\}^n$  is said to be *anti-monotone* if and only if whenever  $x \in H$ , we have  $y \in H$  for all  $y \preceq x$ .

We record the following simple result about anti-monotone conjunctions, which will be useful for understanding and analyzing [Algorithm 2](#):

**Lemma 14.** A function  $f : \{0,1\}^n \rightarrow \{0,1\}$  is an anti-monotone conjunction if and only if  $f^{-1}(1)$  is anti-monotone and is a linear subspace of  $\{0,1\}^n$ .

*Proof.* The “if” direction follows immediately from the following easy result of [\[PRS02\]](#):

**Claim 15** (Claim 18 of [\[PRS02\]](#)). If  $H$  is a monotone and affine subspace of  $\{0,1\}^n$ , then  $H = \{x \in \{0,1\}^n : (x_{i_1} = 1) \wedge \dots \wedge (x_{i_k} = 1)\}$  for some  $k$  and some  $i_1, \dots, i_k \in [n]$ . Similarly if  $H$  is an anti-monotone and linear subspace, then  $H = \{x \in \{0,1\}^n : (x_{i_1} = 0) \wedge \dots \wedge (x_{i_k} = 0)\}$  for some  $k$  and some  $i_1, \dots, i_k \in [n]$ .

For the other direction, suppose that  $f(x) = x_{i_1} \wedge \dots \wedge x_{i_k}$  is an anti-monotone conjunction. Then clearly  $f^{-1}(1)$  is an anti-monotone set. Moreover, for  $x, y \in f^{-1}(1)$ ,  $f(x \oplus y) = \neg(x_{i_1} \oplus y_{i_1}) \wedge \dots \wedge \neg(x_{i_k} \oplus y_{i_k})$ . Since  $x, y$  are both in  $f^{-1}(1)$ ,  $x_{i_1}, \dots, x_{i_k}, y_{i_1}, \dots, y_{i_k}$  are all zero, and thus  $f(x \oplus y) = 1$ . Recalling [Definition 11](#), we have that  $f^{-1}(1)$  is a linear subspace, so the only-if direction also holds.  $\square$

### 3.2 Intuition: the tester of [\[PRS02\]](#) and our tester

We begin with an overview of our algorithm and its proof. At a bird’s eye view, our algorithm has two phases: Phase 1 tests whether  $f^{-1}(1)$  is close to a linear subspace, and Phase 2 tests whether  $f$  is anti-monotone. (Note that, while [\[CDH<sup>+</sup>25\]](#) already gave a general tester for monotonicity in the relative error setting, its complexity has a dependence on  $2^n/|f^{-1}(1)|$  which is necessary in the general case. Here we want a much more efficient tester that is linear in  $1/\varepsilon$ , so we cannot use the monotonicity tester of [\[CDH<sup>+</sup>25\]](#).)

If  $f$  is an anti-monotone conjunction then both Phase 1 and Phase 2 tests will clearly pass; the hard part is to prove soundness, showing that when  $f$  is far from every anti-monotone conjunction, then the algorithm will reject with high probability. The crux of the analysis is to show that if  $f$  is close to a linear subspace but far from all anti-monotone conjunctions, then the Phase 2 anti-monotonicity test will fail with high probability. As in [\[PRS02\]](#), we accomplish this with the aid of an intermediate function,  $g_f$  that we will use in the analysis. In Phase 1 of the algorithm, we

**Input:** Oracle access to  $\text{MQ}(f)$ ,  $\text{Samp}(f)$  and an accuracy parameter  $\varepsilon > 0$ .  
**Output:** “Reject” or “Accept.”

---

**Phase 1:**

- 1: Repeat the following  $c_1/\varepsilon$  times :
- 2:        Draw  $\mathbf{x}, \mathbf{y} \sim \text{Samp}(f)$ .
- 3:        If  $\text{MQ}(f)(\mathbf{x} \oplus \mathbf{y}) = 0$ , halt and reject  $f$ .

**Phase 2:**

- 4: Repeat the following  $c_2$  times :
- 5:        Draw  $\mathbf{x} \sim \text{Samp}(f)$ .
- 6:        Draw a uniform random  $\mathbf{y} \preceq \mathbf{x}$ .
- 7:        Draw  $\mathbf{u} \sim \text{Samp}(f)$ .
- 8:        If  $\text{MQ}(f)(\mathbf{y} \oplus \mathbf{u}) = 0$ , halt and reject  $f$ .
- 9: Accept  $f$ .

**Algorithm 2:** Relative-error Anti-monotone Conjunction Tester. Here  $c_1$  and  $c_2$  are two sufficiently large absolute constants.

will actually prove something stronger: if  $f$  passes Phase 1 with good probability, we show that  $f$  is close to a particular function,  $g_f$  (that is defined based on  $f$ ), where  $g_f^{-1}(1)$  is a linear subspace. The function  $g_f$  is chosen to help us downstream in the analysis, to argue that if Phase 2 passes, then  $g_f$  is anti-monotone with good probability; since  $g_f^{-1}(1)$  is a linear subspace, by [Lemma 14](#) this implies  $g_f$  is an anti-monotone conjunction, thereby contradicting the assumption that  $f$  is far from every anti-monotone conjunction.

Finally, we remark that as described in [Section 2.3](#), our relative-error tester yields an  $O(1/\varepsilon)$ -query one-sided tester in the standard model. This strengthened the result of [\[PRS02\]](#), which gave an  $O(1/\varepsilon)$ -query algorithm with two-sided error. Comparing our analysis with [\[PRS02\]](#), while we borrow many of their key ideas, our analysis is arguably simpler in a couple of ways. First since we are forced to create a tester that doesn’t have the ability to estimate the size of the conjunction, this eliminates the initial phase in the [\[PRS02\]](#) which does this estimation, giving a more streamlined argument and enabling us to give an  $O(1/\varepsilon)$ -query algorithm with one-sided error. Second, by a simple transformation (see [Remark 10](#)), we test whether  $f$  is close to a monotone *linear* subspace rather than close to a monotone affine subspace, which makes both the test and analysis easier.

### 3.3 The relative-error anti-monotone conjunction tester.

In the rest of this section we prove the following theorem, which implies [Theorem 1](#) via [Remark 10](#):

**Theorem 16** (Relative-error anti-monotone conjunction tester). [Algorithm 2](#) is a one-sided non-adaptive algorithm which, given  $\text{MQ}(f)$  and  $\text{Samp}(f)$ , is an  $\varepsilon$ -relative-error tester for anti-monotone conjunctions over  $\{0, 1\}^n$ , making  $O(1/\varepsilon)$  oracle calls.

The  $O(1/\varepsilon)$  bound on the number of oracle calls is immediate from inspection of [Algorithm 2](#), so it remains to establish correctness (including soundness and completeness). These are established in [Theorem 17](#) and [Theorem 26](#) respectively.

### 3.4 Completeness

Completeness (with one-sided error) is straightforward to establish:

**Theorem 17.** If  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  is any anti-monotone conjunction, then [Algorithm 2](#) accepts  $f$  with probability 1.

*Proof.* The algorithm can only reject  $f$  on lines 3 or 8. We argue that if  $f$  is an anti-monotone conjunction then the algorithm never rejects at either of these lines. Since  $f$  is an anti-monotone conjunction, by [Lemma 14](#)  $f^{-1}(1)$  is a linear subspace. Hence for any  $x, y \in f^{-1}(1)$  we also have  $f(x \oplus y) = 1$ , so the algorithm never rejects at line 3. Continuing to Phase 2, since  $f$  is anti-monotone,  $x \in f^{-1}(1)$ , and  $y \preceq x$ , we also have  $f(y) = 1$ . Thus, since  $y, u \in f^{-1}(1)$  and  $f^{-1}(1)$  is a linear subspace, recalling [Definition 11](#), the algorithm never rejects at line 8.  $\square$

### 3.5 Soundness

Let  $g_f : \{0, 1\}^n \rightarrow \{0, 1\}$  be the function defined as follows:

$$g_f(a) = \begin{cases} 1 & \text{if } \Pr_{x \sim f^{-1}(1)}[f(a \oplus x) = 1] \geq 1/2 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

For convenience, we will write  $F := f^{-1}(1)$ ,  $G := g_f^{-1}(1)$  and  $N := |F|$ .

#### 3.5.1 Analysis of phase 1 (steps 1-3)

We want to prove the following two lemmas which together say that whenever [Algorithm 2](#) passes Phase 1 with probability at least  $1/10$ , then (1)  $f$  is relative-error-close to  $g_f$  ([Lemma 18](#)) which means that  $|F \Delta G|$  is small relative to  $N$ , and (2)  $G$  is a linear subspace ([Lemma 19](#)).

**Lemma 18.** Suppose that when [Algorithm 2](#) is run on  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ , it passes Phase 1 (that is, it reaches line 4) with probability at least  $1/10$ . Then  $\text{rel-dist}(f, g_f) \leq \varepsilon/10$ .

**Lemma 19.** Suppose that when [Algorithm 2](#) is run on  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ , it passes Phase 1 (that is, it reaches line 4) with probability at least  $1/10$ . Then  $G$  is a linear subspace.

*Proof of Lemma 18.* Assume for a contradiction that  $\text{rel-dist}(f, g_f) \geq \varepsilon/10$ . First we show that

$$\Pr_{x, y \sim F} [f(x \oplus y) = 0] \geq \varepsilon/40. \quad (4)$$

Therefore, the probability we fail to reject  $f$  during Phase 1 is at most  $(1 - \varepsilon/40)^{c_1/\varepsilon}$ , which is at most  $1/10$  for a suitable choice of  $c_1$ .

It remains to prove [Equation \(4\)](#). Since  $\text{rel-dist}(f, g_f) \geq \varepsilon/10$ , we have  $|F \Delta G|/N \geq \varepsilon/10$  and there are two cases: at least half of the elements in  $F \Delta G$  are in  $F$  but not in  $G$ , or vice versa.

In the first case (at least half of the elements in  $F \Delta G$  are in  $F$  but not in  $G$ ), we have:

$$\Pr_{x, y \sim F} [f(x \oplus y) = 0] \geq \Pr_{x, y \sim F} [g_f(y) = 0] \cdot \Pr_{x, y \sim F} [f(x \oplus y) = 0 \mid g_f(y) = 0].$$

Since we are in the first case, the first term is at least  $\varepsilon/20$ , and for the second term, because we are conditioning on  $g_f(y) = 0$ , by the definition of  $g_f$ , this term is at least  $1/2$ . [Equation \(4\)](#) follows.

For the second case, we have  $|G \setminus F| \geq \varepsilon N/20$ . Fix any  $a \in G \setminus F$ , since  $g_f(a) = 1$ , we have  $\Pr_{x \sim F}[f(a \oplus x) = 1] \geq 1/2$ . For any pair  $(a, x)$  such that  $f(a \oplus x) = 1$ , we have  $a \oplus x = y$  for some  $y \in F$ . As a result, we have

$$\Pr_{x, y \sim F}[a \oplus x = y] \geq \frac{1}{2} \cdot \frac{1}{N} \implies \Pr_{x, y \sim F}[x \oplus y = a] \geq \frac{1}{2N}.$$

Since we are in the second case, this gives

$$\Pr_{x, y \sim F}[f(x \oplus y) = 0] \geq \Pr_{x, y \sim F}[x \oplus y \in G \setminus F] = \sum_{a \in G \setminus F} \Pr_{x, y \sim F}[x \oplus y = a] \geq \frac{\varepsilon N}{20} \cdot \frac{1}{2N} = \frac{\varepsilon}{40},$$

which completes the proof of the lemma.  $\square$

Recalling [Definition 11](#), we will prove [Lemma 19](#) by showing that under the stated condition on  $f$ , for any  $a, b \in G$ , we have  $g_f(a \oplus b) = 1$ .

Following [\[PRS02\]](#), we define the set of *witnesses* for  $a \in \{0, 1\}^n$  to be

$$W(a) := \{x \in F : f(a \oplus x) \neq f(a)\}.$$

Elements of the set  $W(a)$  are witnesses in the sense that they witness that  $F$  is not a linear subspace. Let

$$H = \{a \in \{0, 1\}^n : |W(a)| > \delta N\}$$

be the set of elements  $a \in \{0, 1\}^n$  for which the set of witnesses for  $a$  is “large,” i.e. at least a  $\delta$  fraction of  $N = |F|$ , where  $\delta := 1/36$ . The next lemma states that if the algorithm passes Phase 1 with probability at least  $1/10$ , then the set  $H$  is small.

**Lemma 20.** Assume that  $|H| \geq \delta N$ . Then  $\Pr[\text{Algorithm 2 on } f \text{ reaches line 7}] < 1/10$ .

*Proof.* Suppose that  $|H| \geq \delta N$ . Similar to the proof of [Lemma 18](#), we consider two cases: either (i)  $|H \cap F| \geq \delta N/2$ , or (ii)  $|H \cap f^{-1}(0)| \geq \delta N/2$ . We will argue that in both cases we have

$$\Pr_{x, y \sim F}[f(x \oplus y) = 0] \geq \delta^2/2 \tag{5}$$

and thus, the probability that we fail to reject  $f$  in line 3 is at most  $(1 - \delta^2/2)^{c_1/\varepsilon}$ , which is less than  $1/10$  for a suitably large  $c_1$ .

In the first case, we have for any  $y \in H \cap F$ :

$$\Pr_{x \sim F}[f(y \oplus x) = 0] \geq \delta.$$

On the other hand in this first case  $y \sim F$  belongs to  $H \cap F$  with probability at least  $\delta/2$ . So altogether we have

$$\Pr_{x, y \sim F}[f(x \oplus y) = 0] \geq \delta^2/2,$$

giving [Equation \(5\)](#) as desired in the first case.

In the second case, there exist at least  $\delta N/2$  points  $a$  with  $f(a) = 0$  and  $a \in H$ . For any such point  $a \in H \cap f^{-1}(0)$  we have

$$\Pr_{x \sim F}[f(a \oplus x) = 1] \geq \delta.$$

But if  $f(a \oplus x) = 1$  then  $a \oplus x = y$  for some  $y \in F$ . Hence

$$\Pr_{x, y \sim F}[a \oplus x = y] \geq \delta/N, \text{ which we can rewrite as } \Pr_{x, y \sim F}[x \oplus y = a] \geq \delta/N.$$

Summing over all  $a \in H \cap f^{-1}(0)$ , we have

$$\Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(\mathbf{x} \oplus \mathbf{y}) = 0] \geq \Pr_{\mathbf{x}, \mathbf{y} \sim F} [\mathbf{x} \oplus \mathbf{y} \in H \cap f^{-1}(0)] \geq \frac{\delta N}{2} \cdot \frac{\delta}{N} \geq \frac{\delta^2}{2},$$

so indeed Equation (5) holds in the second case as well.  $\square$

We know from the definition of  $g_f$  that for every  $a \in \{0, 1\}^n$ ,

$$\Pr_{\mathbf{x} \sim F} [g_f(a) = f(a \oplus \mathbf{x})] \geq 1/2.$$

The next Lemma states that whenever  $|H|$  is small, the agreement probability is a lot higher. This in turn will allow us to prove Lemma 19.

**Lemma 21.** Assume that  $|H| \leq \delta N$ . Then for every  $a \in \{0, 1\}^n$ , we have

$$\Pr_{\mathbf{x} \sim F} [g_f(a) = f(a \oplus \mathbf{x})] \geq 1 - 4\delta.$$

Armed with the above Lemma (which we defer to the end of this section), we can now complete the proof of Lemma 19.

*Proof of Lemma 19.* By Lemma 20 if  $|H| > \delta N$ , then with probability at least 9/10,  $f$  is rejected in Phase 1. Thus it suffices to show that if  $|H| \leq \delta N$ , then  $G$  is a linear subspace. Towards this goal, assume for sake of contradiction that  $G$  is not a linear subspace, and thus there exist  $a, b \in \{0, 1\}^n$  such that  $g_f(a) = g_f(b) = 1$  but  $g_f(a \oplus b) = 0$ . At a high level, we obtain a contradiction as follows:

- (1) First, using the assumption that  $g_f(a) = g_f(b) = 1$  plus Lemma 21 we show that

$$\Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(a \oplus \mathbf{x} \oplus b \oplus \mathbf{y}) = 1] \geq 1 - 10\delta.$$

- (2) On the other hand using the assumption that  $g_f(a \oplus b) = 0$  plus Lemma 21 we show that

$$\Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(a \oplus \mathbf{x} \oplus b \oplus \mathbf{y}) = 1] < 6\delta.$$

Together this gives a contradiction by our choice of  $\delta = 1/36$ .

*Proving (1):* First, by definition of  $W(\cdot)$  we have

$$\begin{aligned} \Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(a \oplus \mathbf{x} \oplus b \oplus \mathbf{y}) = 1] &\geq \Pr_{\mathbf{x}, \mathbf{y} \sim F} \left[ (a \oplus \mathbf{x} \in F \setminus H) \wedge (b \oplus \mathbf{y} \in F \setminus W(a \oplus \mathbf{x})) \right] \\ &\geq \Pr_{\mathbf{x} \sim F} [a \oplus \mathbf{x} \in F \setminus H] \times \min_{a' \in F \setminus H} \Pr_{\mathbf{y} \sim F} [b \oplus \mathbf{y} \in F \setminus W(a')]. \end{aligned}$$

We claim that the first term above is at least  $1 - 5\delta$ . This is because

1.  $\Pr_{\mathbf{x} \sim F} [a \oplus \mathbf{x} \in F] \geq 1 - 4\delta$  using  $g_f(a) = 1$  and Lemma 21, and
2.  $\Pr_{\mathbf{x} \sim F} [a \oplus \mathbf{x} \in H] \leq \delta$  since  $|H| \leq \delta N$  and for a fixed  $a$ , by linearity every  $z \in H$  has at most one unique representation as  $a \oplus \mathbf{x}$  with  $\mathbf{x} \in F$ .

It follows that the first term is at least  $1 - 5\delta$ .

Next we claim that the second term is at least  $1 - 5\delta$ . This is because



1.  $\Pr_{\mathbf{y} \sim F}[b \oplus \mathbf{y} \in F] \geq 1 - 4\delta$  using  $g_f(b) = 1$  and [Lemma 21](#), and

2. Given that  $a' \in F \setminus H$ , we have  $|W(a')| \leq \delta N$  and thus,  $\Pr_{\mathbf{y} \sim F}[b \oplus \mathbf{y} \in W(a')] \leq \delta$ .

So the second term is at least  $1 - 5\delta$ , and the product is at least  $(1 - 5\delta)(1 - 5\delta) \geq 1 - 10\delta$ .

*Proving (2):* We bound the probability of interest as follows:

$$\begin{aligned} & \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x} \oplus b \oplus \mathbf{y}) = 1] \\ & \leq \Pr_{\mathbf{x} \sim F}[a \oplus b \oplus \mathbf{x} \in H] + \Pr_{\mathbf{x} \sim F}[f(a \oplus b \oplus \mathbf{x}) = 1] \\ & \quad + \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x} \oplus b \oplus \mathbf{y}) = 1 \mid a \oplus b \oplus \mathbf{x} \notin H \wedge f(a \oplus b \oplus \mathbf{x}) = 0]. \end{aligned}$$

The first term is at most  $\delta$  by linearity; the second term is at most  $4\delta$  since  $g_f(a \oplus b) = 0$  and by [Lemma 21](#), and the last term is also bounded by  $\delta$  by the definition of  $H$ . Therefore the probability that  $f(a \oplus \mathbf{x} \oplus b \oplus \mathbf{y}) = 1$  is at most  $6\delta$ .  $\square$

It is left to prove [Lemma 21](#).

*Proof of Lemma 21.* We want to prove that if  $H$  is small ( $|H| \leq \delta N$ ), then the agreement between  $g_f$  and  $f$  is amplified in the sense that for any  $a \in \{0, 1\}^n$ ,  $\Pr_{\mathbf{x} \sim F}[g_f(a) = f(a \oplus \mathbf{x})] \geq 1 - 4\delta$ . By the definition of  $g_f$ , this probability is at least  $1/2$ . Clearly any  $a \notin H$  has at most  $\delta N$  witnesses  $x$  that violate the equality, so for these  $a$ 's, the probability is easily seen to get amplified to  $1 - \delta \geq 1 - 4\delta$ . However, we want to show that small  $H$  implies that the agreement between  $g_f$  and  $f$  is amplified for *every*  $a$ . To show this, fix  $a$  and let  $\gamma = \Pr_{\mathbf{x} \sim F}[f(a \oplus \mathbf{x}) = g_f(a)]$ . We will consider the quantity

$$\Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x}) = f(a \oplus \mathbf{y})]. \quad (6)$$

On the one hand, we will relate it to  $\gamma$  and on the other hand derive an expression in terms of  $\delta$ .

$$\begin{aligned} & \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x}) = f(a \oplus \mathbf{y})] \\ & \geq \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x}) = f(a \oplus \mathbf{x} \oplus \mathbf{y}) \wedge f(a \oplus \mathbf{y}) = f(a \oplus \mathbf{x} \oplus \mathbf{y})] \\ & = 1 - \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x}) \neq f(a \oplus \mathbf{x} \oplus \mathbf{y}) \vee f(a \oplus \mathbf{y}) \neq f(a \oplus \mathbf{x} \oplus \mathbf{y})] \\ & \geq 1 - 2 \times \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x}) \neq f(a \oplus \mathbf{x} \oplus \mathbf{y})]. \end{aligned}$$

Now we will show that the last term is at most  $2\delta$ :

$$\begin{aligned} & \Pr_{\mathbf{x}, \mathbf{y} \sim F}[f(a \oplus \mathbf{x}) \neq f(a \oplus \mathbf{x} \oplus \mathbf{y})] \\ & \leq \Pr_{\mathbf{x} \sim F}[a \oplus \mathbf{x} \in H] \times \max_{a' \in H} \left\{ \Pr_{\mathbf{y} \sim F}[f(a') \neq f(a' \oplus \mathbf{y})] \right\} \\ & \quad + \Pr_{\mathbf{x} \sim F}[a \oplus \mathbf{x} \notin H] \times \max_{a' \notin H} \left\{ \Pr_{\mathbf{y} \sim F}[f(a') \neq f(a' \oplus \mathbf{y})] \right\}. \end{aligned}$$

The second term is at most  $\delta$  since for every  $a' \notin H$ ,  $\Pr_{\mathbf{y} \sim F}[f(a') \neq f(a' \oplus \mathbf{y})] \leq \delta$  and trivially we have  $\Pr_{\mathbf{x} \sim F}[a \oplus \mathbf{x} \notin H] \leq 1$ . The first term is also at most  $\delta$ , since we have  $\Pr_{\mathbf{x} \sim F}[a \oplus \mathbf{x} \in H] \leq \delta$  by linearity and trivially  $\Pr_{\mathbf{y} \sim F}[f(a') \neq f(a' \oplus \mathbf{y})] \leq 1$ . Therefore, [Equation \(6\)](#) is at least  $1 - 4\delta$ .

Now we will express the same quantity in terms of the agreement probability,  $\gamma$ :

$$\begin{aligned}
& \Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(a \oplus \mathbf{x}) = f(a \oplus \mathbf{y})] \\
&= \Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(a \oplus \mathbf{x}) = g_f(a) \wedge f(a \oplus \mathbf{y}) = g_f(a)] \\
&\quad + \Pr_{\mathbf{x}, \mathbf{y} \sim F} [f(a \oplus \mathbf{x}) \neq g_f(a) \wedge f(a \oplus \mathbf{y}) \neq g_f(a)] \\
&= \gamma^2 + (1 - \gamma)^2,
\end{aligned}$$

using independence of  $\mathbf{x}$  and  $\mathbf{y}$ . By combining our two inequalities, we have  $\gamma^2 + (1 - \gamma)^2 \geq 1 - 4\delta$ . Rearranging we have  $\gamma(1 - \gamma) \leq 2\delta$ . But  $\gamma \geq 1/2$ , so  $1 - \gamma \leq 4\delta$ , as we wanted to show.  $\square$

### 3.5.2 Analysis of phase 2 (lines 4-9)

For the second phase, we will establish the following lemma that states that any  $f$  that has passed Phase 1 but is far from an anti-monotone conjunction, will be rejected with high probability.

**Lemma 22.** Consider an execution of [Algorithm 2](#) up through line 8 on a function  $f$  that satisfies the following conditions:

- $f$  is  $\varepsilon$ -far in relative distance from every anti-monotone conjunction;
- $\text{rel-dist}(f, g_f) \leq \varepsilon/10$ ;
- and  $G := g_f^{-1}(1)$  is a linear subspace.

Then  $\Pr[\text{Algorithm 2 rejects } f \text{ in lines 4-8}] \geq 0.9$ .

We will need the following two definitions. For any  $x \in G$ , we define

$$\text{LT}(x) = \{y \in \{0, 1\}^n : y \preceq x\} \quad \text{and} \quad \mathcal{X} = \{x \in G : \text{LT}(x) \subseteq G\}.$$

We recall the following useful claim that is analogous to Claim 21 from [\[PRS02\]](#):

**Claim 23.** If  $G$  is a linear subspace of  $\{0, 1\}^n$  then the set  $\mathcal{X}$  is a linear subspace of  $G$ . Moreover, if  $g_f$  is not an anti-monotone conjunction then  $|\mathcal{X}| \leq |G|/2$ .

*Proof.* We first show  $\mathcal{X}$  is a linear subspace of  $G$ . To do this, we need to show that if  $x^1, x^2 \in \mathcal{X}$  then  $z = x^1 \oplus x^2 \in \mathcal{X}$ , meaning  $y \in G$  for every  $y \preceq z$ . Fix  $y \preceq z$ , we want to write  $y = y^1 \oplus y^2$  where  $y^1 \preceq x^1, y^2 \preceq x^2$ . For a coordinate  $i$  if  $y_i = 0$ , then we set  $y_i^1 = y_i^2 = 0$ . Otherwise if  $y_i = 1$ , then  $z_i = 1$  so at least one of  $x_i^1$  or  $x_i^2$  is 1. If  $x_i^1 = 1$ , we set  $y_i^1 = 1$  and  $y_i^2 = 0$  and otherwise  $y_i^2 = 1$  and  $y_i^1 = 0$ . It's easy to check  $y^1 \preceq x^1$  and  $y^2 \preceq x^2$ . Since  $x^1, x^2 \in \mathcal{X}$  this implies  $y^1, y^2 \in G$  and since  $G$  is a linear subspace we also have  $y \in G$ . So  $\mathcal{X} \subseteq G$ .

If  $\mathcal{X} = G$ , then  $G$  is an anti-monotone subspace. By [Lemma 14](#), this implies  $g_f$  is an anti-monotone conjunction. So, if  $g_f$  isn't an anti-monotone conjunction, we must have  $G \not\subseteq \mathcal{X}$ , by [Fact 12](#) this implies  $|\mathcal{X}| \leq |G|/2$ .  $\square$

We will also use the following simple claim:

**Claim 24.** If  $G$  is a linear subspace of  $\{0, 1\}^n$  and  $g_f$  is not anti-monotone, then

$$\Pr_{\substack{\mathbf{u} \sim F \\ \mathbf{y} \preceq \mathbf{x}}} [f(\mathbf{y} \oplus \mathbf{u}) = 0] \geq 1/4, \quad \text{for any } x \in G \setminus \mathcal{X}.$$

*Proof.* Note that  $\text{LT}(x)$  is a linear subspace, and since  $x \notin \mathcal{X}$  we have that  $\text{LT}(x) \not\subseteq G$ . By [Fact 12](#):

$$\Pr_{\mathbf{y} \preceq x} [\mathbf{y} \in G] \leq 1/2.$$

Thus, if we sample  $\mathbf{y} \preceq x$  uniformly at random, then with probability at least  $1/2$  we have  $\mathbf{y} \notin G$  and  $g_f(\mathbf{y}) = 0$  which means  $\Pr_{\mathbf{u} \sim F}[f(\mathbf{y} \oplus \mathbf{u}) = 0] \geq 1/2$ . This finishes the proof of the claim.  $\square$

**Lemma 25.** Assume that  $G$  is a linear subspace and  $g_f$  is not an anti-monotone conjunction, and  $\text{rel-dist}(f, g_f) \leq 0.1$ . Then we have

$$|(G \setminus \mathcal{X}) \cap F| \geq 0.1N.$$

*Proof.* Since  $g_f$  is not an anti-monotone conjunction, by [Claim 23](#) we have  $|G \setminus \mathcal{X}| \geq |G|/2$ . But since  $\text{rel-dist}(f, g_f) \leq 0.1$ , we have  $|G \setminus F| \leq 0.1N$ , and we also must have  $|G| \geq 0.9N$ . Hence

$$|(G \setminus \mathcal{X}) \cap F| \geq 0.9N/2 - 0.1N \geq 0.1N.$$

This finishes the proof of the lemma.  $\square$

We can now prove [Lemma 22](#):

*Proof of Lemma 22.* Assume that (i)  $\text{rel-dist}(f, f') > \varepsilon$  for every anti-monotone conjunction  $f'$ ; (ii)  $\text{rel-dist}(f, g_f) \leq \varepsilon/10$ ; and (iii)  $G$  is a linear subspace. Then  $G$  cannot be an anti-monotone linear subspace, as otherwise by [Lemma 14](#)  $g_f$  would be an anti-monotone conjunction and violate (i) and (ii). Hence we have  $|(G \setminus \mathcal{X}) \cap F| \geq N/10$  by [Lemma 25](#); this in turn means that for  $\mathbf{x} \sim F$ , with probability at least  $1/10$  we have  $\mathbf{x} \in G \setminus \mathcal{X}$ .

By [Claim 24](#), for any  $x \in G \setminus \mathcal{X}$  we have

$$\Pr_{\substack{\mathbf{u} \sim F \\ \mathbf{y} \preceq x}} [f(\mathbf{y} \oplus \mathbf{u}) = 0] \geq 1/4.$$

Thus by [Lemma 25](#) we have that

$$\Pr_{\substack{\mathbf{x}, \mathbf{u} \sim F \\ \mathbf{y} \preceq \mathbf{x}}} [f(\mathbf{y} \oplus \mathbf{u}) = 0] \geq 1/40.$$

Hence, each iteration of lines 4–8 has at least a  $1/40$  probability to reject  $f$ . So the probability we fail to reject  $f$  during lines 4–8 is  $(1 - 1/40)^{c_2}$ , which is at most  $0.1$  for a suitable choice of  $c_2$ .  $\square$

### 3.5.3 Putting the pieces together

We have all the ingredients we need to establish the soundness of [Algorithm 2](#):

**Theorem 26.** Suppose that  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  satisfies  $\text{rel-dist}(f, f') > \varepsilon$  for every anti-monotone conjunction  $f'$ . Then [Algorithm 2](#) rejects  $f$  with probability at least  $0.9$ .

*Proof.* We prove the contrapositive, i.e. that if [Algorithm 2](#) accepts  $f$  with probability at least  $0.1$  then  $f$  has  $\text{rel-dist}(f, f') \leq \varepsilon$  for some anti-monotone conjunction  $f'$ .

By [Lemma 18](#), if  $f$  passes Phase 1 with probability at least  $1/10$ , then  $\text{rel-dist}(f, g_f) \leq \varepsilon/10$ . By [Lemma 19](#), if  $f$  passes Phase 1 with probability at least  $1/10$ , then  $G$  is a linear subspace. Finally, by [Lemma 22](#) if  $f$  passes Phase 2 with probability at least  $0.1$ , then we must have  $\text{rel-dist}(f, f') \leq \varepsilon$  for some anti-monotone conjunction  $f'$ .  $\square$

We can now prove [Theorem 1](#) about [CONJ-TEST](#):

*Proof of [Theorem 1](#).* Recall that [CONJ-TEST](#) first samples  $\mathbf{y} \sim F$  and then runs [Algorithm 2](#) on  $f_{\mathbf{y}}$ . When  $f$  is a conjunction,  $f_{\mathbf{y}}$  is an anti-monotone conjunction and thus, by [Theorem 17](#) [CONJ-TEST](#) accepts with probability 1. When  $f$  is far from any conjunction,  $f_{\mathbf{y}}$  must be far from any anti-monotone conjunction. So by [Theorem 26](#), [CONJ-TEST](#) rejects with probability at least 0.9.  $\square$

### 3.6 Robustness of [CONJ-TEST](#) to faulty oracles

[CONJ-TEST](#) will play a crucial role in the relative-error tester for decision lists in the next section. As will become clear there, [CONJ-TEST](#) will be run on a function  $f$  to test whether it is relative-error close to a conjunction when it can only access a “faulty” version of the  $\text{Samp}(f)$  oracle. For later reference we record the following two explicit statements about the performance of [CONJ-TEST](#); we show that as long as the faulty version of  $\text{Samp}(f)$  satisfies some mild conditions, [CONJ-TEST](#) still outputs the correct answer with high probability. In both statements, we write  $\mathcal{D}$  to denote the distribution over  $\{0, 1\}^n$  underlying the faulty sampling oracle  $\text{Samp}^*(f)$ , which is not necessarily uniform over  $F = f^{-1}(1)$  (and in [Theorem 28](#) need not even necessarily be supported over  $F$ ).

**Theorem 27.** Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  be a conjunction. Assume that the oracle  $\text{Samp}^*(f)$  always returns some string in  $F$  (i.e., the distribution  $\mathcal{D}$  is supported over  $F$ ). Then [CONJ-TEST](#), running on  $\text{MQ}(f)$  and  $\text{Samp}^*(f)$ , accepts  $f$  with probability 1.

*Proof.* When run on an anti-monotone conjunction  $f$ , [Algorithm 2](#) will never reject as long as all its samples are from  $F$ ; therefore [Algorithm 2](#) will accept  $f$  with probability 1. Since [CONJ-TEST](#) works by drawing a single sample  $\mathbf{y} \sim \text{Samp}^*(f)$ , which by assumption is in  $F$ , and running [Algorithm 2](#) on  $f_{\mathbf{y}}$ , even that  $\mathbf{y} \in F$ , the function  $f_{\mathbf{y}}$  is guaranteed to be an anti-monotone conjunction when  $f$  is a conjunction, and [CONJ-TEST](#) is guaranteed to accept.  $\square$

**Theorem 28.** Suppose that  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  satisfies  $\text{rel-dist}(f, f') > \varepsilon$  for any conjunction  $f'$ . Let  $\text{Samp}^*(f)$  be a sampling oracle for  $f$  such that the underlying distribution  $\mathcal{D}$  satisfies

$$\Pr_{\mathbf{x} \sim \mathcal{D}} [\mathbf{x} = z] \geq \frac{1}{20|F|}, \quad \text{for any } z \in F.$$

Then [CONJ-TEST](#), running on  $\text{MQ}(f)$  and  $\text{Samp}^*(f)$ , rejects  $f$  with probability at least 0.9.

*Proof.* The proof of soundness ([Theorem 26](#)) of [Algorithm 2](#) shows that, given any  $f$  that satisfies  $\text{rel-dist}(f, f') > \varepsilon$  for any anti-monotone conjunction  $f'$ , either each round of Phase 1 rejects  $f$  with probability  $\Omega(1/\varepsilon)$ , or each round of Phase 2 rejects  $f$  with probability  $\Omega(1)$ , under  $\text{Samp}(f)$ . It is easy to show that under our assumption of  $\text{Samp}^*$ , both probabilities basically go down by at most a factor of  $(1/20)^2$ . So by making  $c_1$  and  $c_2$  sufficiently large, we have that [Algorithm 2](#) rejects  $f$  with probability at least 0.9 even when it is given access to  $\text{Samp}^*(f)$  instead of  $\text{Samp}(f)$ .

Now consider running [CONJ-TEST](#) on a function  $f$  with  $\text{rel-dist}(f, f') > \varepsilon$  for any conjunction  $f'$ , with access to a faulty oracle  $\text{Samp}^*(f)$ . [CONJ-TEST](#) works by first drawing a single sample  $\mathbf{y} \sim \text{Samp}^*(f)$  and running [Algorithm 2](#) on  $f_{\mathbf{y}}$ . Since  $f$  is far from any conjunction, we must have  $\text{rel-dist}(f_{\mathbf{y}}, f') > \varepsilon$  for any anti-monotone conjunction  $f'$  (no matter whether  $\mathbf{y} \in F$  or not). Hence, [CONJ-TEST](#) is guaranteed to reject  $f$  with probability at least 0.9.  $\square$

## 4 A relative-error testing algorithm for decision lists

Let’s first recall how a decision list is represented and some basic facts about decision lists.

## 4.1 Basics about decision lists

A decision list  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  can be represented as a list of rules:

$$(x_{i_1}, b_1, v_1), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}},$$

where  $i_1, \dots, i_k \in [n]$  are indices of variables and  $b_1, \dots, b_k, v_1, \dots, v_k, v_{\text{default}} \in \{0, 1\}$ . For each  $x \in \{0, 1\}^n$ ,  $f(x)$  is set to be  $v_j$  with the smallest  $j$  such that  $x_{i_j} = b_j$ , and is set to be  $v_{\text{default}}$  if  $x_{i_j} \neq b_j$  for all  $j \in [k]$ . Without loss of generality, we will make the following assumptions:

- *Each variable appears at most once in the list of rules (i.e., no two  $i_j$ 's are the same).*  
Indeed if two rules of the form  $(x_{i_j}, b_j, v_j)$  and  $(x_{i_\ell}, b_\ell, v_\ell)$  with  $j < \ell$  and  $i_j = i_\ell$  are present, then either  $b_j = b_\ell$ , in which case the latter rule can be trivially deleted, or  $b_\ell = 1 - b_j$ , in which case the latter rule and all rules after it can be replaced by  $v_{\text{default}} = v_\ell$ , without changing the function  $f$  represented.
- *There exists at least one rule  $(x_{i_j}, b_j, v_j)$  with  $v_j = 1$ .* This is because if  $v_1 = \dots = v_k = 0$ , then either (1)  $v_{\text{default}} = 0$ , in which case  $f$  is the all-0 function; (2)  $v_{\text{default}} = 1$  and at least one variable  $x_i$  does not appear in the list, in which case we can insert  $(x_i, 1, 1)$  before  $v_{\text{default}}$ , which does not change the function and makes sure at least one  $v_j$  is 1; or (3)  $v_{\text{default}} = 1$  but all variables appear in the list, in which case we have  $|f^{-1}(1)| = 1$ . Functions  $f$  with  $|f^{-1}(1)| \leq 1$  are easy to recognize with the sampling oracle.

In the rest of the section, whenever we talk about decision lists, the function  $f$  is assumed to satisfy  $|f^{-1}(1)| > 1$  so it can be represented by a list of rules that satisfies the two conditions above.

Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  be a decision list represented by  $(x_{i_1}, b_1, v_1), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}}$ . The following decomposition of the list plays a crucial role both in the algorithm and its analysis:

- Let  $p \in [k]$  be the smallest index with  $v_p = 1$ . We will call  $p$  the **pivot**.
- The prefix before  $(x_{i_p}, b_p, v_p)$ , i.e.,  $(x_{i_1}, b_1, 0), \dots, (x_{i_{p-1}}, b_{p-1}, 0)$  will be called the **head**.  
The key insight is that we can view the head as a conjunction  $C : \{0, 1\}^n \rightarrow \{0, 1\}$ , where

$$C(x) = (x_{i_1} = 1 - b_1) \wedge \dots \wedge (x_{i_{p-1}} = 1 - b_{p-1}).$$

And we write  $\mathcal{H}_C := \{i_1, \dots, i_{p-1}\}$  to denote the set of  $p-1$  variables in the conjunction.

- The part after  $(x_{i_p}, b_p, v_p)$ , i.e.,  $(x_{i_{p+1}}, b_{p+1}, v_{p+1}), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}}$  will be called the **tail**. Let  $L : \{0, 1\}^n \rightarrow \{0, 1\}$  denote the decision list represented by the pivot and tail:

$$(x_{i_p}, b_p, v_p = 1), (x_{i_{p+1}}, b_{p+1}, v_{p+1}), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}}.$$

Then it is easy to verify that  $f(x) = C(x) \wedge L(x)$  for all  $x \in \{0, 1\}^n$ .

We record the following simple fact about the decomposition  $f = C \wedge L$ :

**Observation 29.** All  $z \in f^{-1}(1)$  satisfy  $C(z) = 1$ , i.e.,  $z_{i_j} = 1 - b_j$  for all  $i_j \in \mathcal{H}_C$ .

We will need the following lemma about the decomposition  $f = C \wedge L$  of a decision list, which says that coordinates in the tail are fairly unbiased. The intuition is that at least half the points in  $f^{-1}(1)$  are accepted by the rule  $(x_{i_p}, b_p, 1)$ , and for these points the coordinates  $x_{i_j}$  with  $j > p$ , do not matter, and thus are distributed uniformly.

**Lemma 30.** Let  $f$  be a decision list represented by  $(x_{i_1}, b_1, v_1), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}}$ , and let  $p$  be its pivot. Then we have

$$\Pr_{z \sim f^{-1}(1)} [z_{i_p} = b_p] \geq 1/2 \quad \text{and} \quad \Pr_{z \sim f^{-1}(1)} [z_{i_j} = 1] \in [1/4, 3/4], \quad \text{for all } j : p < j \leq k.$$

*Proof.* Using the decomposition  $f = C \wedge L$ , we have

$$\{z \in \{0, 1\}^n : C(z) = 1 \text{ and } z_{i_p} = b_p\} \subseteq f^{-1}(1) \subseteq \{z \in \{0, 1\}^n : C(z) = 1\}. \quad (7)$$

Given that the LHS is exactly half of the size of the set on the RHS, the first part follows.

Furthermore, consider any  $i_j$  with  $p < j \leq k$ . Exactly half of the strings  $z$  in the set on the LHS of Equation (7) have  $z_{i_j} = 1$  and exactly half of them have  $z_{i_j} = 0$ . Thus, we have for any  $b \in \{0, 1\}$ ,

$$\Pr_{z \sim f^{-1}(1)} [z_{i_j} = b] \geq \frac{|\{z : C(z) = 1, z_{i_p} = b_p \text{ and } z_{i_j} = b\}|}{|f^{-1}(1)|} = \frac{|\{z : C(z) = 1\}|}{4|f^{-1}(1)|} \geq \frac{1}{4},$$

where the last inequality used the second inclusion of Equation (7). This finishes the proof.  $\square$

## 4.2 The Algorithm

Our main algorithm for testing decision lists in relative distance is presented in Algorithm 3, which uses two subroutines described in Algorithm 4 and Algorithm 5, respectively. Before diving into the analysis of Algorithm 3, we start with an overview of the algorithm in this subsection, to give some intuition behind it about why it accepts functions that are decision lists and rejects functions that are far from decision lists in relative distance with high probability.

We will start with the case when  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  is a decision list, and explain how the algorithm extracts information about the underlying decomposition  $C \wedge L$  of  $f$  so that  $f$  passes all the tests in Algorithm 3 with high probability. Towards the end of this subsection, we will switch to the case when  $f$  is far from any decision list in relative distance, and discuss why passing all the tests in Algorithm 3 gives evidence that  $f$  is close in relative distance to some decomposition  $C \wedge L$  of a decision list, which should not happen given the assumption that  $f$  is far from every decision list in relative distance. The formal proof is given in Sections 4.3 to 4.5.

**The case when  $f$  is a decision list.** Let  $f$  be a decision list represented by  $(x_{i_1}, b_1, v_1), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}}$ . Let  $p$  be the pivot and  $f = C \wedge L$  be the decomposition, where  $C$  is a conjunction over variables in  $\mathcal{H}_C := \{i_1, \dots, i_{p-1}\}$  and  $L$  is a decision list formed by the pivot and the tail of the list.

The first step of Algorithm 3 draws samples from  $\text{Samp}(f)$  to partition variables into  $\mathbf{U} \cup \mathbf{R}$ :

2: Draw  $O(1/\varepsilon)$  samples  $\mathbf{z} \sim \text{Samp}(f)$ , and call the set of samples  $\mathbf{S}$ . Let

$$\mathbf{U} = \{i \in [n] : \exists b \in \{0, 1\} \text{ s.t. } \mathbf{z}_i = b \text{ for all } \mathbf{z} \in \mathbf{S}\} \quad \text{and} \quad \mathbf{R} = [n] \setminus \mathbf{U}$$

and  $\mathbf{u} \in \{0, 1\}^{\mathbf{U}}$  be the unique string such that  $\mathbf{u}_i = \mathbf{z}_i$  for all  $i \in \mathbf{U}$  and  $\mathbf{z} \in \mathbf{S}$ .

So  $\mathbf{U}$  is the set of *unanimous coordinates* of samples in  $\mathbf{S}$  and thus, we always have  $\mathcal{H}_C \subseteq \mathbf{U}$  and  $\mathbf{u}_{i_j} = 1 - b_j$  for all  $i_j \in \mathcal{H}_C$  by Observation 29. Ideally, we would hope for  $\mathbf{R}$  to contain all variables in the tail, i.e.,  $i_{p+1}, \dots, i_k$ ; however, this won't be the case in general, unfortunately, since  $k$  can

be much bigger than  $p$  (e.g.,  $k - p$  could be polynomial in  $n$ ). But by [Lemma 30](#) and given that  $\mathbf{S}$  contains  $O(1/\varepsilon)$  samples from  $\text{Samp}(f)$ , it is easy to show that with high probability, the variables occurring in the beginning part of the tail (excluding those variables that are hidden too deep in the tail) all lie in  $\mathbf{R}$  with high probability, and it turns out that this is all we will need for the proof of completeness to go through.

Now, the pivot variable  $x_{i_p}$  can go either way, since it may be highly biased towards  $b_p$ . However, if it is too biased towards  $b_p$  so that none of the samples  $\mathbf{z}$  in  $\mathbf{S}$  have  $z_p = 1 - b_p$ , then one can argue that  $f$  must be very close to a conjunction over  $\{0, 1\}^n$  in relative distance. Because of this, Step 0 of [Algorithm 3](#) will accept  $f$  with high probability.

Proceeding to Step 2, assume without loss of generality that [Algorithm 3](#) has found  $\mathbf{U}$  and  $\mathbf{R}$  such that  $\mathcal{H}_C \subseteq \mathbf{U}$  and  $\mathbf{R}$  contains variables in the beginning part of  $L$ , including  $i_p$ , and the string  $\mathbf{u} \in \{0, 1\}^{\mathbf{U}}$  satisfies  $u_{i_j} = 1 - b_j$  for all  $i_j \in \mathcal{H}_C$ . This implies the following properties:

1. The function  $f|_{\mathbf{u}}$  over  $\{0, 1\}^{\mathbf{R}}$  is a “dense” decision list, where “dense” means the number of its satisfying assignments is at least half of  $2^{|\mathbf{R}|}$ , and it is a decision list because the restriction of any decision list remains a decision list.
2. For any  $\alpha \in \{0, 1\}^{\mathbf{U}}$ ,  $f|_{\alpha}$  is either the all-0 function (because  $\alpha$  does not satisfy the conjunction  $C$ ), or very close to  $f|_{\mathbf{u}}$  (because  $\mathbf{R}$  contains the beginning part of the tail) and thus, both  $f|_{\mathbf{u}}$  and  $f|_{\alpha}$  are close to  $L$  (when viewed as a decision list over  $\mathbf{R}$ ) under the uniform distribution.

Given these properties of  $\mathbf{U}$ ,  $\mathbf{R}$  and  $\mathbf{u}$ , it is easy to see that  $f$  passes both tests in Step 2 with high probability:

- 3: Draw  $O(1)$  points  $\mathbf{w} \sim \{0, 1\}^{\mathbf{R}}$ ; halt and reject  $f$  if less than  $1/4$  of the sampled  $\mathbf{w}$ 's have  $f(\mathbf{u} \circ \mathbf{w}) = 1$ .
- 4: Run the algorithm of [Theorem 6](#) on  $f|_{\mathbf{u}}$  with  $\varepsilon/100$ ; halt and reject  $f$  if it rejects.

since line 3 checks whether  $f|_{\mathbf{u}}$  is “dense,” and line 4 checks if  $f|_{\mathbf{u}}$  is close to a decision list under the uniform distribution. The function  $f$  would also pass the tests in Step 3 with high probability:

- 5: Repeat  $O(1/\varepsilon)$  times:
- 6: Draw  $\mathbf{z} \sim \text{Samp}(f)$  and  $\mathbf{w} \sim \{0, 1\}^{\mathbf{R}}$ ; if  $f(\mathbf{z}_{\mathbf{U}} \circ \mathbf{w}) \neq f(\mathbf{u} \circ \mathbf{w})$ , halt and reject  $f$ .

as for any  $z \in f^{-1}(1)$ ,  $\mathbf{z}_{\mathbf{U}}$  satisfies  $C$  and thus,  $f|_{\mathbf{z}_{\mathbf{U}}}$  is close to  $f|_{\mathbf{u}}$  under the uniform distribution.

For step 4, let  $\mathbf{\Gamma} : \{0, 1\}^{\mathbf{U}} \rightarrow \{0, 1\}$  be defined as follows:

$$\mathbf{\Gamma}(\alpha) = \begin{cases} 1 & \text{if } \Pr_{\mathbf{w} \sim \{0, 1\}^{\mathbf{R}}} [f(\alpha \circ \mathbf{w}) = 1] \geq 1/16 \\ 0 & \text{otherwise} \end{cases}$$

Now, observe that for any  $\alpha \in \{0, 1\}^{\mathbf{U}}$ , either  $\alpha$  does not satisfy  $C$ , in which case  $f|_{\alpha}$  is the all-0 function, or  $\alpha$  satisfies  $C$ , in which case  $f|_{\alpha}$  is close to  $f|_{\mathbf{u}}$  and thus  $\mathbf{\Gamma}(\alpha) = 1$  (because  $f|_{\mathbf{u}}$  is “dense”). As a result,  $\mathbf{\Gamma}$  is indeed the same as  $C$  and thus it is a conjunction over  $\{0, 1\}^{\mathbf{U}}$ . From this, it may seem straightforward to conclude that  $\mathbf{\Gamma}$  always passes the test in Step 4 below:



7: Run **CONJ-TEST** with  $\text{Samp}(\Gamma) \leftarrow (\text{Algorithm 5})$ ,  $\text{MQ}(\Gamma) \leftarrow (\text{Algorithm 4})$  to test if  $\Gamma$  is  $(\varepsilon/100)$ -close to a conjunction in relative distance; halt and return the same answer.

except that we don't actually have direct access to sampling and query oracles of  $\Gamma$ . They are simulated by **Algorithm 5** and **Algorithm 4**, respectively. We show in the proof that (1) **Algorithm 5** (for the sampling oracle of  $\Gamma$ ) only returns samples in  $\Gamma^{-1}(1)$ ; and (2) with high probability, **Algorithm 4** returns the correct answer to all  $O(1/\varepsilon)$  membership queries made by **CONJ-TEST**. From this we can apply **Theorem 27** to conclude that **CONJ-TEST** accepts with high probability.

**The case when  $f$  is far from every decision list.** Let's now switch to the case when  $f$  is  $\varepsilon$ -far from any decision list in relative distance. Because conjunctions are special cases of decision lists, Step 0 does not accept  $f$  with high probability and **Algorithm 3** goes through Step 1 to obtain  $U, R$  and  $u \in \{0, 1\}^U$ . Let's consider what conditions  $f$  needs to satisfy in order for it to pass tests in Steps 2 and 3 with high probability:

1. To pass Step 2,  $f|_u$  over  $\{0, 1\}^R$  needs to be "dense:"  $|f|_u^{(-1)}(1)| \geq 2^{|R|}/8$ , and must also be  $(\varepsilon/100)$ -close to a decision list under the uniform distribution.
2. To understand the necessary condition posed on  $f$  to pass Step 3, let's introduce the following function  $g : \{0, 1\}^n \rightarrow \{0, 1\}$ , which is closely related to  $\Gamma$  defined earlier and will play a crucial role in the analysis:

$$g(x) = \Gamma(x_U) \wedge f(u \circ x_R).$$

We show that to pass Step 3 with high probability, it must be the case that  $\text{rel-dist}(f, g)$  is small. This means that for each  $\alpha \in \{0, 1\}^U$ , one can replace  $f|_\alpha$  by the all-0 function if  $\Gamma(\alpha) = 0$  (which means that  $f|_\alpha$  is not dense enough), and replace  $f|_\alpha$  by  $f|_u$  otherwise, and this will only change a small number of bits of  $f$  (relative to  $|f^{-1}(1)|$ ). Note that  $g$  is getting closer to the decomposition of a decision list since we know  $f|_u$  is close to a decision list over  $\{0, 1\}^R$ .

Assuming all the necessary conditions summarized above for  $f$  to pass Steps 2 and 3 and using the assumption that  $f$  is far from decision lists in relative distance, we show that  $\Gamma$  must be far from conjunctions in relative distance. Finally, given the latter, it can be shown that **CONJ-TEST** in Step 4 rejects with high probability, again by analyzing performance guarantees of **Algorithm 5** and **Algorithm 4** as simulators of sampling and query oracles of  $\Gamma$ , respectively, so that **Theorem 28** for **CONJ-TEST** can be applied.

### 4.3 Query Complexity

Let  $c_0$  be the absolute constant such that running **CONJ-TEST** with relative-error distance parameter  $\varepsilon/100$  uses no more than  $c_0/\varepsilon$  many samples and membership queries. Then the two constants  $c_1$  and  $c_2$  are chosen to satisfy the following conditions: 1)  $c_2$  is both sufficiently large and sufficiently larger than  $c_0$ ; 2)  $c_1$  is chosen to be sufficiently larger than  $c_2$  (and thus,  $c_1$  is also sufficiently large and sufficiently larger than  $c_0$ ).

The following lemma gives the easy efficiency analysis of **Algorithm 3**:

**Lemma 31.** **Algorithm 3** makes no more than  $\tilde{O}(1/\varepsilon)$  calls to  $\text{Samp}(f)$  and  $\text{MQ}(f)$ .

**Input:** Oracle access to  $\text{MQ}(f)$ ,  $\text{Samp}(f)$  of  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  and a parameter  $\varepsilon$ .

**Output:** “Reject” or “Accept.”

---

**Step 0: Special case when  $f$  is close to a conjunction**

- 1: Run **CONJ-TEST** on  $f$  with  $\varepsilon$ ; halt and accept  $f$  if it accepts.

**Step 1: Draw samples to obtain  $U$  and  $R$**

- 2: Draw  $c_1/\varepsilon$  samples  $z \sim \text{Samp}(f)$ , and call the set of samples  $S$ . Let

$$U = \{i \in [n] : \exists b \in \{0, 1\} \text{ s.t. } z_i = b \text{ for all } z \in S\} \quad \text{and} \quad R = [n] \setminus U$$

and  $u \in \{0, 1\}^U$  be the unique string such that  $u_i = z_i$  for all  $i \in U$  and  $z \in S$ .

**Step 2: Check that  $f|_u$  is close to a “dense” decision list**

- 3: Draw  $c_2$  points  $w \sim \{0, 1\}^R$ ; halt and reject  $f$  if less than  $1/4$  of the sampled  $w$ ’s have  $f(u \circ w) = 1$ .
- 4: Run the algorithm of **Theorem 6** on  $f|_u$  with  $\varepsilon/100$ ; halt and reject  $f$  if it rejects.

**Step 3: Check that  $f|_w$  is close to  $f|_u$  on “most”  $w \in \{0, 1\}^U$**

- 5: Repeat  $c_2/\varepsilon$  times:
- 6: Draw  $z \sim \text{Samp}(f)$  and  $w \sim \{0, 1\}^R$ ; if  $f(z_U \circ w) \neq f(u \circ w)$ , halt and reject  $f$ .

**Step 4: Is  $\Gamma$  a conjunction?**

Let  $\Gamma : \{0, 1\}^U \rightarrow \{0, 1\}$  be defined as follows:

$$\Gamma(\alpha) = \begin{cases} 1 & \text{if } \Pr_{w \sim \{0, 1\}^R} [f(\alpha \circ w) = 1] \geq 1/16 \\ 0 & \text{otherwise.} \end{cases}$$

- 7: Run **CONJ-TEST** with  $\text{Samp}(\Gamma) \leftarrow (\text{Algorithm 5})$ ,  $\text{MQ}(\Gamma) \leftarrow (\text{Algorithm 4})$  to test if  $\Gamma$  is  $(\varepsilon/100)$ -close to a conjunction in relative distance; halt and return the same answer.

**Algorithm 3:** Relative-error decision list tester. Here  $c_1$  and  $c_2$  are absolute constants chosen at the beginning of **Section 4.3**.

**Input:** Oracle access to  $\text{MQ}(f)$  and a query  $\alpha \in \{0, 1\}^U$ .

---

- 1: Draw  $w \in \{0, 1\}^R$  for  $c_2 \log(1/\varepsilon)$  many times and return 0 if  $f(\alpha \circ w) = 0$  for all  $w$ .
- 2: Halt and reject  $f$  if  $f(\alpha \circ w) \neq f(u \circ w)$  for some  $w$ ; return 1 otherwise.

**Algorithm 4:** Simulator for query access to  $\Gamma$ .

**Input:** Oracle access to  $\text{Samp}(f)$ .

1: Draw  $\mathbf{z} \sim \text{Samp}(f)$  and return  $\mathbf{z}_U$ .

**Algorithm 5:** Simulator for sample access to  $\Gamma$ .

*Proof.* Running **CONJ-TEST** in line 1 takes  $O(1/\varepsilon)$  samples and non-adaptive queries. Line 2 uses  $O(1/\varepsilon)$  samples. Line 3 uses  $O(1)$  queries. Line 4 uses  $\tilde{O}(1/\varepsilon)$  queries by **Theorem 6**. Line 6 uses  $O(1/\varepsilon)$  samples and queries. **CONJ-TEST** on line 7 runs **Algorithm 4** and **Algorithm 5** each  $O(1/\varepsilon)$  times, which in total uses  $\tilde{O}(1/\varepsilon)$  samples and queries to  $\text{Samp}(f)$  and  $\text{MQ}(f)$ .  $\square$

**Remark 32.** An alternative version of **Algorithm 3** is obtained simply by replacing the algorithm of **Theorem 6** with the algorithm of **Theorem 7**. As discussed in **Section 1.1**, this yields a version of **Theorem 3** which is nonadaptive and uses  $\tilde{O}(1/\varepsilon)$  samples and makes  $\tilde{O}(1/\varepsilon^2)$  calls to  $\text{MQ}(f)$ .

#### 4.4 Completeness: Correctness of **Algorithm 3** when $f$ is a decision list

Throughout this section, we assume  $f$  to be a decision list with pivot  $p$  and decomposition  $C \wedge L$ , where  $C : \{0, 1\}^n \rightarrow \{0, 1\}$  is a conjunction over variables in  $\mathcal{H}_C = \{i_1, \dots, i_{p-1}\}$  and  $L : \{0, 1\}^n \rightarrow \{0, 1\}$  is a decision list over other variables, represented by the following list of rules:

$$(x_{i_p}, b_p, v_p), \dots, (x_{i_k}, b_k, v_k), v_{\text{default}}$$

with  $v_p = 1$  given that  $p$  is the pivot. Our goal is to show the following:

**Theorem 33.** If  $f$  is a decision list, then **Algorithm 3** accepts  $f$  with probability at least 0.7.

To prove **Theorem 33**, we will consider two cases (recall that  $c_0$  is the constant such that  $c_0/\varepsilon$  is the number of queries and samples asked by **CONJ-TEST** with relative error parameter  $\varepsilon/100$ ):

$$\text{Case 1: } \Pr_{\mathbf{z} \sim f^{-1}(1)} [\mathbf{z}_{i_p} \neq b_p] \leq \frac{\varepsilon}{10c_0}; \quad \text{Case 2: } \Pr_{\mathbf{z} \sim f^{-1}(1)} [\mathbf{z}_{i_p} \neq b_p] > \frac{\varepsilon}{10c_0}.$$

The key idea, is that in the first case,  $f$  is so close to a conjunction that **CONJ-TEST** will accept in line 1 with high probability. In the second case, with high probability we will have  $i_p \in U$ .

##### 4.4.1 Case 1

We will first prove a lemma regarding our conjunction tester.

**Lemma 34.** Let  $h : \{0, 1\}^n \rightarrow \{0, 1\}$  and let  $D$  be an anti-monotone conjunction with  $D^{-1}(1) \subseteq h^{-1}(1)$ . Consider running **Algorithm 2** on  $h$  using  $\text{MQ}(h)$  and  $\text{Samp}(h)$ . If all samples **Algorithm 2** received lie in  $D^{-1}(1)$ , then it always accepts  $h$ .

*Proof.* The algorithm can only reject on lines 3 and 8. On lines 3 assumption on samples we have  $\mathbf{x}, \mathbf{y} \in D^{-1}(1)$ . Thus  $\mathbf{x} \oplus \mathbf{y} \in D^{-1}(1)$  so  $h(\mathbf{x} \oplus \mathbf{y}) = 1$ . For line 11, we have  $\mathbf{x}, \mathbf{u} \in D^{-1}(1)$ , and since  $\mathbf{y} \preceq \mathbf{x}$  we have that  $\mathbf{y} \in D^{-1}(1)$ . So again,  $h(\mathbf{y} \oplus \mathbf{u}) = 1$ .  $\square$

**Corollary 35.** Let  $h : \{0, 1\}^n \rightarrow \{0, 1\}$ , and  $D$  be a conjunction with  $D^{-1}(1) \subseteq h^{-1}(1)$ . Consider running **CONJ-TEST** on  $h$  using  $\text{MQ}(h)$  and  $\text{Samp}(h)$ . If all samples that **CONJ-TEST** received lie in  $D^{-1}(1)$ , then it always accepts  $h$ .

*Proof.* Let  $D'$  be the anti-monotone version of  $D$ , i.e.  $D'$  is obtained from  $D$  by replacing every positive literal  $x_i$  with the negated version  $\bar{x}_i$ . Recall that **CONJ-TEST** first draws  $\mathbf{y} \sim h^{-1}(1)$ , which by assumption satisfies  $\mathbf{y} \in D^{-1}(1)$ , and then runs **Algorithm 2** on  $h_{\mathbf{y}}$  (cf. **Remark 10**).  $D^{-1}(1) \subseteq h^{-1}(1)$  implies that  $D'^{-1}(1) \subseteq h_{\mathbf{y}}^{-1}(1)$ . Moreover, since all samples lie in  $D^{-1}(1)$ , all samples to **Algorithm 2** must lie in  $D'^{-1}(1)$ . It follows from **Lemma 34** that **CONJ-TEST** accepts.  $\square$

We are ready to prove **Theorem 33** for Case 1:

**Lemma 36.** Let  $f = C \wedge L$  be a decision list as described at the start of **Section 4.4** with

$$\Pr_{\mathbf{z} \sim f^{-1}(1)} [\mathbf{z}_{i_p} \neq b_p] \leq \frac{\varepsilon}{10c_0}.$$

Then **Algorithm 3** accepts  $f$  on line 1 with probability at least 0.9.

*Proof.* Consider the conjunction  $D = C \wedge (x_{i_p} = b_p)$ . We have  $D^{-1}(1) \subseteq f^{-1}(1)$ .

It follows from **Corollary 35** that **CONJ-TEST** accepts  $f$  on line 1 if all the  $c_0/\varepsilon$  samples drawn from  $f^{-1}(1)$  lie in  $D^{-1}(1)$ . (Recall that  $c_0/\varepsilon$  is the number of samples and queries asked by **CONJ-TEST** with parameter  $\varepsilon/100$ . So  $c_0/\varepsilon$  is also an upper bound for the number of samples **CONJ-TEST** draws in line 1 with  $\varepsilon$ .) Note that every  $\mathbf{z} \in f^{-1}(1)$  satisfies  $\mathbf{z} \in D^{-1}(1)$  as long as  $\mathbf{z}_{i_p} = b_p$ . So all  $c_0/\varepsilon$  samples lie in  $D^{-1}(1)$  with probability at least 0.9 by a union bound.  $\square$

#### 4.4.2 Case 2

In the following analysis we set  $\ell$  to be the following parameter:

$$\ell := \min \left( k - p, 3 \log \left( \frac{c_1}{\varepsilon} \right) \right) \leq 3 \log \left( \frac{c_1}{\varepsilon} \right).$$

**Lemma 37.** Let  $f = C \wedge L$  be a decision list as described at the start of **Section 4.4**. Suppose that

$$\Pr_{\mathbf{z} \sim f^{-1}(1)} [\mathbf{z}_{i_p} \neq b_p] \geq \frac{\varepsilon}{10c_0}.$$

If the algorithm reaches line 3, then  $i_{p+j} \in \mathbf{R}$  for all  $j : 0 \leq j \leq \ell$  with probability at least 0.9.

*Proof.* The event follows if for all  $i \in \{i_p, \dots, i_{p+\ell}\}$  and  $b \in \{0, 1\}$ , at least one  $\mathbf{z} \in \mathbf{S}$  has  $\mathbf{z}_i = b$ .

On line 2, we sample  $c_1/\varepsilon$  many  $\mathbf{z} \sim f^{-1}(1)$  with  $c_1$  sufficiently larger than  $c_0 \geq 1$ . First we handle  $i = i_p$ : because we are in Case 2, the probability of  $\mathbf{z}_{i_p} = b_p$  for all  $\mathbf{z} \in \mathbf{S}$  is at most

$$\left( 1 - \frac{\varepsilon}{10c_0} \right)^{c_1/\varepsilon}.$$

By **Lemma 30**,  $\Pr_{\mathbf{z} \sim f^{-1}(1)} [\mathbf{z}_{i_p} = b_p] \geq 1/2$ . Hence the probability of  $\mathbf{z}_{i_p} = 1 - b_p$  for all  $\mathbf{z}$  is at most  $(1/2)^{c_1/\varepsilon}$ . Setting  $c_1$  to be sufficiently larger than  $c_0$ , we have  $i_p \notin \mathbf{R}$  with probability at most 0.05.

For each  $i_{p+j}$  with  $j > 0$ , by **Lemma 30** we have for either  $b \in \{0, 1\}$ ,

$$\Pr_{\mathbf{z} \sim f^{-1}(1)} [\mathbf{z}_{i_{p+j}} = b \text{ for all } \mathbf{z} \in \mathbf{S}] \leq \left( \frac{3}{4} \right)^{c_1/\varepsilon}.$$

The probability of each of these  $2\ell$  events can be made sufficiently smaller than  $1/\ell$  (given that  $\ell$  is only logarithmic in  $c_1/\varepsilon$ ) by taking  $c_1$  sufficiently large. The lemma follows from a union bound.  $\square$

We will also need the following simple observation about  $\mathbf{U}$ :

**Lemma 38.** If **Algorithm 3** reaches line 3, then  $\mathcal{H}_C \subseteq \mathbf{U}$  and  $u_{i_j} = 1 - b_j$  for all  $i_j \in \mathcal{H}_C$ .

*Proof.* Given that  $f = C \wedge L$ , by definition we must have that every  $z \in f^{-1}(1)$  satisfies  $z_{i_j} = 1 - b_j$  for all  $i_j \in \mathcal{H}_C$ . So this must also be the case for points in  $\mathbf{S}$ .  $\square$

For all lemmas in the rest of this section we will assume that the conclusions of **Lemma 37** and **Lemma 38** hold, i.e.:

**Assumption 1.** **Algorithm 3** reaches line 3 with  $\mathbf{U} = U$ ,  $\mathbf{R} = R$  and  $\mathbf{u} = u$  satisfying

- $\mathcal{H}_C \subseteq U$  and  $u_{i_j} = 1 - b_j$  for all  $i_j \in \mathcal{H}_C$ .
- For any  $j : p \leq j \leq p + \ell$ ,  $i_j \in R$ .

Note that once  $U, R$  and  $u$  are fixed, the function  $\mathbf{\Gamma} = \Gamma$  over  $\{0, 1\}^U$  is also fixed.

The next lemma will play a key role in the rest of the proof; it says that under **Assumption 1**, one can ignore the rules of  $L$  that are too deep in the list.

**Lemma 39.** Under **Assumption 1**, for any  $\alpha \in \{0, 1\}^U$  with  $\alpha_i = 1 - b_i$  for all  $i \in \mathcal{H}_C$ , we have

$$\Pr_{w \sim \{0, 1\}^R} [f(\alpha \circ w) \neq f(u \circ w)] \leq \left( \frac{\varepsilon}{c_1} \right)^3.$$

*Proof.* Notice that both  $u$  and  $\alpha$  satisfy the conjunction given by  $C$ . Hence we have  $f(u \circ w) = L(u \circ w)$  and  $f(\alpha \circ w) = L(\alpha \circ w)$  for every  $w \in \{0, 1\}^R$ . By **Assumption 1**, we have  $i_p, \dots, i_{p+\ell} \in R$ . On the one hand, if  $\ell < 3 \log(c_1/\varepsilon)$ , then we have  $\ell = k - p$  and thus,  $L(\alpha \circ w) = L(u \circ w)$  for all  $w \in \{0, 1\}^R$  in which case the probability we are interested in is trivially 0. On the other hand, if  $\ell \geq 3 \log(c_1/\varepsilon)$  then if  $w_{i_j} = b_{i_j}$  for some  $j : p \leq j \leq p + \ell$  we have  $L(\alpha \circ w) = L(u \circ w)$ . Thus,

$$\Pr_{w \sim \{0, 1\}^R} [f(\alpha \circ w) \neq f(u \circ w)] \leq \Pr_{w \sim \{0, 1\}^R} [w_{i_j} \neq b_{i_j} \text{ for all } j : p \leq j \leq p + \ell] = 2^{-(\ell+1)}.$$

Plugging in  $\ell = 3 \log(c_1/\varepsilon)$  finishes the proof of the lemma in this case.  $\square$

Next we show that, under **Assumption 1**, **Algorithm 3** is unlikely to reject during lines 3–6:

**Lemma 40.** Under **Assumption 1**, **Algorithm 3** rejects  $f$  on lines 3–6 with probability at most 0.1.

*Proof.* Since  $i_p \in R$  and  $u$  satisfies the conjunction in  $C$ , we have that

$$\Pr_{w \sim \{0, 1\}^R} [f(u \circ w) = 1] \geq \Pr_{w \sim \{0, 1\}^R} [w_{i_p} = b_p] = 1/2.$$

Hence, by a simple Chernoff bound and by making the constant  $c_2$  sufficiently large, we have that the probability that less than  $1/4$  of the points queried have  $f(z) = 1$  is at most 0.01.

Next for line 4, note that by **Assumption 1**, we have that  $f|_u = (C \wedge L)|_u = L|_u$ . But  $L|_u$  is a decision list. Hence line 4 will reject with probability at most 0.05 by **Theorem 6**.

Recall that line 6 is repeated  $c_2/\varepsilon$  times. Fix any  $z \in f^{-1}(1)$  as the point sampled in line 6. By **Lemma 39**, we have that

$$\Pr_{w \sim \{0, 1\}^R} [f(u \circ w) \neq f(z_U \circ w)] \leq \left( \frac{\varepsilon}{c_1} \right)^3.$$

By a union bound over the  $c_2/\varepsilon$  rounds on line 11 and noting that  $c_1$  is sufficiently larger than  $c_2$ , the algorithm rejects with probability at most 0.01. The lemma follows by a union bound.  $\square$

Finally it remains to examine line 7. Let  $C^* : \{0, 1\}^U \rightarrow \{0, 1\}$  be the restriction of  $C$ , which has domain  $\{0, 1\}^n$  but is just a conjunction over  $\mathcal{H}_C \subseteq U$ , onto the domain  $\{0, 1\}^U$ :

$$C^*(\alpha) = (\alpha_{i_1} = 1 - b_1) \wedge \cdots \wedge (\alpha_{i_{p-1}} = 1 - b_{p-1}),$$

for any  $\alpha \in \{0, 1\}^U$ . Our first goal will be to show that  $\Gamma = C^*$ .

**Lemma 41.** Under [Assumption 1](#), we have that  $\Gamma = C^*$  (so in particular,  $\Gamma$  is a conjunction).

*Proof.* If  $C^*(\alpha) = 0$ , then  $f(\alpha \circ w) = 0$  for all  $w \in \{0, 1\}^R$  and by definition  $\Gamma(\alpha) = 0$ .

On the other hand, if  $C^*(\alpha) = 1$ , it follows from [Assumption 1](#) and [Lemma 39](#) that

$$\Pr_{w \in \{0, 1\}^R} [f(\alpha \circ w) = 1] \geq \Pr_{w \in \{0, 1\}^R} [f(u \circ w) = 1] - \left(\frac{\varepsilon}{c_1}\right)^3 \geq \frac{1}{2} - \left(\frac{\varepsilon}{c_1}\right)^3 \geq \frac{1}{16}$$

when  $c_1$  is sufficiently large and thus,  $\Gamma(\alpha) = 1$ .  $\square$

By [Lemma 41](#), if we were to run [CONJ-TEST](#) on  $\Gamma$  with the correct membership and sampling oracles, we know it would always accept (recall that [CONJ-TEST](#) has one-sided error). However, we need to account for the fact we are only able to simulate access to oracles  $\text{Samp}(\Gamma)$  and  $\text{MQ}(\Gamma)$  using [Algorithm 5](#) and [Algorithm 4](#), respectively.

**Lemma 42.** Suppose that  $f$  reaches line 8 of [Algorithm 3](#) under [Assumption 1](#). Then [CONJ-TEST](#) accepts with probability at least 0.9.

*Proof.* By [Lemma 41](#), all samples [CONJ-TEST](#) receives using [Algorithm 5](#) are from  $\Gamma^{-1}(1)$  because they are from  $C^{*-1}(1)$  and  $\Gamma = C^*$ . We prove below that with probability at least 9/10, all the  $c_0/\varepsilon$  membership queries made by [CONJ-TEST](#) are correctly answered by [Algorithm 4](#). When this happens, it follows directly from [Theorem 27](#) that [CONJ-TEST](#) accepts and the lemma follows.

For this purpose, it suffices to show that given any  $\alpha \in \{0, 1\}^U$ , the probability that [Algorithm 4](#) rejects or returns the wrong value for  $\Gamma(\alpha)$  is at most  $\varepsilon/(10c_0)$ . The lemma then follows by a union bound over all  $c_0/\varepsilon$  membership queries made by [CONJ-TEST](#).

If  $\Gamma(\alpha) = 0$ , then it follows from [Lemma 41](#) that we have  $f(\alpha \circ w) = 0$  for all  $w \in \{0, 1\}^R$  and thus [Algorithm 4](#) always returns 0.

If  $\Gamma(\alpha) = 1$ , then  $C^*(\alpha) = 1$  and thus, the probability of  $w \sim \{0, 1\}^R$  having  $f(\alpha \circ w) = 1$  is at least  $1/2$ . So the probability of [Algorithm 4](#) returning 0 is at most  $2^{-c_2 \log(1/\varepsilon)} \leq \varepsilon/(20c_0)$  given that  $c_2$  is sufficiently larger than  $c_0$ . On the other hand, by [Lemma 39](#) we have

$$\Pr_{w \sim \{0, 1\}^n} [f(\alpha \circ w) \neq f(u \circ w)] \leq \left(\frac{\varepsilon}{c_1}\right)^3.$$

By a union bound over all  $c_2 \log(1/\varepsilon)$  samples  $w$ , the probability that [Algorithm 4](#) rejects is at most  $\varepsilon/(20c_0)$  given that  $c_1$  is sufficiently larger than  $c_0$  and  $c_2$ . So the probability of either rejecting or returning 0 is at most  $\varepsilon/(10c_0)$ . This finishes the proof of the lemma.  $\square$

Finally, we prove [Theorem 33](#):

*Proof.* Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  be a decision list. If Case 1 holds, then by [Lemma 36](#) the algorithm accepts  $f$  with probability at least 0.9. Otherwise,  $f$  falls in Case 2. By [Lemma 37](#) and [Lemma 38](#), with probability at least 0.9, the algorithm either accepts  $f$  in line 1 (in which case we are done) or reaches line 4 with  $U, R$  and  $u$  that satisfy [Assumption 1](#).

Under [Assumption 1](#), by [Lemma 40](#)  $f$  gets rejected in lines 3–6 with probability at most 0.1, and gets rejected in line 7 by [CONJ-TEST](#) (including the case when it is rejected by [Algorithm 4](#)) with probability at most 0.1. So altogether,  $f$  is rejected with probability at most 0.3.  $\square$

#### 4.5 Soundness: Correctness of Algorithm 3 when $f$ is far from decision lists

Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  be a function that is  $\varepsilon$ -far from decision lists in relative distance. Given that conjunctions are special cases of decision lists, by Theorem 1, CONJ-TEST on line 1 rejects with probability at least 0.9, in which case the algorithm continues and reaches line 3. Let  $U = U$ ,  $R = R$  and  $u = u$  when it reaches line 3. After  $U, R$  and  $u$  are fixed, the function  $\Gamma = \Gamma$  is also fixed.

If  $f$  gets rejected on line 3 or 4 we are done; the next lemma gives a necessary condition for  $f$  to get through lines 3–4 without being rejected with high probability:

**Lemma 43.** Assume that  $f$  reaches line 3 with  $U, R$  and  $u$ . If either

$$|f|_u^{-1}(1)| < \frac{2^{|R|}}{8} \quad (8)$$

or  $\text{rel-dist}(f|_u, \text{DL}) \geq \varepsilon/10$ , then  $f$  gets rejected during lines 3–4 with probability at least 0.9.

*Proof.* If Equation (8) holds then the probability of  $f(u \circ \mathbf{w}) = 1$  when  $\mathbf{w} \sim \{0, 1\}^R$  is at most  $1/8$ . Given the  $1/4$  vs  $1/8$  gap and that  $c_2$  is sufficiently large, Algorithm 3 rejects on line 3 with probability at least 0.9.

Now assume that Equation (8) does not hold but  $\text{rel-dist}(f|_u, \text{DL}) \geq \varepsilon/10$ . We show that these two conditions together imply  $\text{dist}(f|_u, \text{DL}) \geq \varepsilon/80$  and thus, by Theorem 6 the algorithm rejects on line 4 with probability at least 0.9. To lower bound  $\text{dist}(f|_u, \text{DL})$ , we use

$$\frac{\varepsilon}{10} \leq \text{rel-dist}(f|_u, \text{DL}) = \text{dist}(f|_u, \text{DL}) \cdot \frac{2^{|R|}}{|f|_u^{-1}(1)|} \leq 8 \cdot \text{dist}(f|_u, \text{DL}).$$

This finishes the proof of the lemma.  $\square$

Now we look at lines 5–6. The next lemma gives a necessary condition for  $f$  not getting rejected on line 6 with high probability. To this end, we introduce a new function  $g : \{0, 1\}^n \rightarrow \{0, 1\}$ :

$$g(x) = \Gamma(x_U) \wedge f(u \circ x_R).$$

By definition, for each  $\alpha \in \{0, 1\}^U$ ,  $g|_\alpha$  is all-0 if  $\Gamma(\alpha) = 0$ , and  $g|_\alpha = f|_u$  if  $\Gamma(\alpha) = 1$ .

**Lemma 44.** Assume that  $f$  reaches line 4 with  $U, R$  and  $u$  and satisfies  $|f|_u^{-1}(1)| \geq 2^{|R|}/8$ . If we have  $\text{rel-dist}(f, g) \geq \varepsilon/10$ , then  $f$  gets rejected on line 6 with probability at least 0.9.

*Proof.* Assume that  $\text{rel-dist}(f, g) \geq \varepsilon/10$ . We partition  $f^{-1}(1) \triangle g^{-1}(1)$  into

$$\begin{aligned} P &:= \{y \in \{0, 1\}^n : \Gamma(y_U) = 0 \text{ and } f(y) = 1\} \quad \text{and} \\ Q &:= \{y \in \{0, 1\}^n : \Gamma(y_U) = 1 \text{ and } f(y) \neq g(y)\}. \end{aligned}$$

So we either have  $|P| \geq (\varepsilon/20) \cdot |f^{-1}(1)|$  or  $|Q| \geq (\varepsilon/20) \cdot |f^{-1}(1)|$ .

First consider the case when  $P$  is large. Each round of line 6 rejects with probability at least

$$\frac{\varepsilon}{20} \cdot \frac{1}{16},$$

where  $\varepsilon/20$  is the probability of  $z$  landing in  $P$ . To see where  $1/16$  comes from, note that whenever  $z \in P$ , we have  $\Gamma(z_U) = 0$  and thus  $|f|_{z_U}^{-1}(1)| \leq 2^{|R|}/16$  while  $|f|_u^{-1}(1)| \geq 2^{|R|}/8$  by assumption.



The volume difference between  $f|_u$  and  $f|_{z_U}$  shows that  $f(u \circ \mathbf{w}) \neq f(z_U \circ \mathbf{w})$  with probability at least  $1/16$ . By setting  $c_2$  sufficiently large, line 6 rejects with probability at least 0.9.

Next for the case when  $Q$  is large, the probability that each round of line 6 rejects is at least

$$\sum_{y \in Q} \frac{|f|_{y_U}^{-1}(1)|}{|f^{-1}(1)|} \cdot \frac{1}{2^{|R|}}.$$

Here for each  $y \in Q$ , the term above is exactly the probability that  $\mathbf{z} \sim \text{Samp}(f)$  and  $\mathbf{w} \sim \{0, 1\}^R$  of line 6 together satisfy  $\mathbf{z}_U \circ \mathbf{w} = y$ , in which case line 6 rejects. The numerator is at least  $2^{|R|}/16$  given that  $\Gamma(z_U) = 1$ . Plugging it in, each round of line 6 rejects with probability at least

$$|Q| \cdot \frac{2^{|R|}/16}{|f^{-1}(1)|} \cdot \frac{1}{2^{|R|}} \geq \frac{\varepsilon}{20} \cdot |f^{-1}(1)| \cdot \frac{2^{|R|}/16}{|f^{-1}(1)|} \cdot \frac{1}{2^{|R|}} \geq \frac{\varepsilon}{320}.$$

We again finish the proof by setting the constant  $c_2$  to be sufficiently large.  $\square$

Finally, we look at Step 4 of the tester. Given the lemmas established so far, we assume that  $f$  reaches line 7 with  $U, R$  and  $u$  (as well as  $g$  and  $\Gamma$ ) that satisfy the following assumption:

**Assumption 2.**  $f$  reaches line 7 with  $U, R, u$  and  $g$  that satisfy the following conditions:

- $\text{rel-dist}(f, \text{DL}) \geq \varepsilon$ ;
- $|f|_u^{-1}(1)| \geq 2^{|R|}/8$ ;
- $\text{rel-dist}(f|_u, \text{DL}) \leq \varepsilon/10$ ; and
- $\text{rel-dist}(f, g) \leq \varepsilon/10$ .

We need to show that under **Assumption 2**, **CONJ-TEST** on line 7 rejects with high probability. We proceed in two steps.

1. First we show in **Lemma 45** that if  $f$  satisfies **Assumption 2**, then  $\Gamma$  must be  $(\varepsilon/100)$ -far from conjunctions in relative distance.
2. Next we show in **Lemma 48** that **CONJ-TEST** on line 7 rejects with high probability. Note that this would follow trivially if **CONJ-TEST** were given access to sampling and membership oracles of  $\Gamma$ . However, **CONJ-TEST** is only given access simulated by **Algorithm 4** and **Algorithm 5**, respectively. So we analyze the their performance guarantees in **Lemma 47** and **Lemma 46**, respectively, and then apply **Theorem 28** to prove **Lemma 48**.

We start with the first step:

**Lemma 45.** Under **Assumption 2**,  $\Gamma$  is  $(\varepsilon/100)$ -far from conjunctions in relative distance.

*Proof.* Assume for a contradiction that  $\text{rel-dist}(\Gamma, C) \leq \varepsilon/100$  for some conjunction  $C$  over  $\{0, 1\}^U$ . We prove in the rest of the proof that  $\text{rel-dist}(f, \text{DL}) < \varepsilon$ , a contradiction.

The proof proceeds in two steps, where we define two Boolean functions  $\zeta$  and  $\lambda$  over  $\{0, 1\}^n$ . In particular, the last function  $\lambda$  will be a decision list. We prove the following inequalities:

$$\text{rel-dist}(g, \zeta) \leq \varepsilon/10 \quad \text{and} \quad \text{rel-dist}(\zeta, \lambda) \leq \varepsilon/100.$$

Combining with  $\text{rel-dist}(f, g) \leq \varepsilon/10$  from **Assumption 2**, it follows from the approximate triangle inequality for relative distance (**Lemma 5**) that  $\text{rel-dist}(f, \lambda) < \varepsilon$ , a contradiction.

First, given  $\text{rel-dist}(f|_u, \text{DL}) \leq \varepsilon/10$  from [Assumption 2](#), we let  $L$  be a decision list over  $\{0, 1\}^R$  that satisfies  $\text{rel-dist}(f|_u, L) \leq \varepsilon/10$ . Then our first function  $\zeta$  over  $\{0, 1\}^n$  is defined as

$$\zeta(z) := \Gamma(z_U) \wedge L(z_R).$$

Observe that for any  $\alpha \in \{0, 1\}^U$ , if  $\Gamma(\alpha) = 0$  then  $g|_\alpha = \zeta|_\alpha = 0$ . And if  $\Gamma(\alpha) = 1$ , then  $g|_\alpha = f|_u$  while  $\zeta|_\alpha = L$ , in which case,  $\text{rel-dist}(g|_\alpha, \zeta|_\alpha) = \text{rel-dist}(f|_u, L) \leq \varepsilon/10$ . We thus have that:

$$\begin{aligned} \text{rel-dist}(g, \zeta) &= \frac{1}{|g^{-1}(1)|} \sum_{\alpha: \Gamma(\alpha)=1} \left| \{w \in \{0, 1\}^R : g(\alpha \circ w) \neq \zeta(\alpha \circ w)\} \right| \\ &= \frac{1}{|g^{-1}(1)|} \sum_{\alpha: \Gamma(\alpha)=1} \text{rel-dist}(g|_\alpha, \zeta|_\alpha) \cdot |g|_\alpha^{-1}(1)| \\ &\leq \frac{\varepsilon}{10} \cdot \frac{1}{|g^{-1}(1)|} \sum_{\alpha: \Gamma(\alpha)=1} |g|_\alpha^{-1}(1)|. \end{aligned}$$

Since  $|g^{-1}(1)| = \sum_{\alpha} |g|_\alpha^{-1}(1)|$ , we get that  $\text{rel-dist}(g, \zeta) \leq \varepsilon/10$ .

Next recall that  $C$  is a conjunction over  $\{0, 1\}^U$  that satisfies  $\text{rel-dist}(\Gamma, C) \leq \varepsilon/100$ . The last function  $\lambda : \{0, 1\}^n \rightarrow \{0, 1\}$  is obtained as follows:

$$\lambda(z) := C(z_U) \wedge L(z_R)$$

(i.e. we are swapping out  $\Gamma(z_U)$  for  $C(z_U)$  in the definition of  $\zeta$ ). In particular,  $\lambda$  is a decision list as promised earlier in the proof. Below we bound  $\text{rel-dist}(\zeta, \lambda)$ .

Observe that for any  $\alpha \in \{0, 1\}^U$ , if  $\Gamma(\alpha) = C(\alpha)$ , then  $\lambda|_\alpha = \zeta|_\alpha$ . Otherwise, we have that one of  $\lambda|_\alpha$  or  $\zeta|_\alpha$  is  $L$  while the other is the all-0 function. So we have

$$\text{rel-dist}(\zeta, \lambda) = \frac{1}{|\zeta^{-1}(1)|} \cdot \left| \{ \alpha \in \{0, 1\}^U : \Gamma(\alpha) \neq C(\alpha) \} \right| \cdot |L^{-1}(1)|.$$

On the other hand,  $|\zeta^{-1}(1)| = |\Gamma^{-1}(1)| \times |L^{-1}(1)|$ . Thus, we have

$$\text{rel-dist}(\zeta, \lambda) = \frac{|\{ \alpha \in \{0, 1\}^U : \Gamma(\alpha) \neq C(\alpha) \}|}{|\Gamma^{-1}(1)|} = \text{rel-dist}(\Gamma, C) \leq \frac{\varepsilon}{100}.$$

This finishes the proof of the lemma.  $\square$

Next we analyze performance guarantees of [Algorithm 5](#) and [Algorithm 4](#) to simulate sampling and membership oracles of  $\Gamma$ , respectively:

**Lemma 46.** Assume that  $f$  respects [Assumption 2](#). For any  $\alpha \in \{0, 1\}^U$  with  $\alpha \in \Gamma^{-1}(1)$ , we have

$$\Pr[\text{Algorithm 5 returns } \alpha] \geq \frac{1}{20|\Gamma^{-1}(1)|}.$$

*Proof.* Let  $\alpha \in \{0, 1\}^U$  with  $\Gamma(\alpha) = 1$ . By definition of  $\Gamma$ , we have

$$\Pr_{z \sim f^{-1}(1)}[z_U = \alpha] \geq \frac{2^{|R|}/16}{|f^{-1}(1)|}. \quad (9)$$

On the other hand, we know that

$$\frac{\varepsilon}{10} \geq \text{rel-dist}(f, g) \geq \frac{|f^{-1}(1)| - |g^{-1}(1)|}{|f^{-1}(1)|}$$

and  $|g^{-1}(1)| \leq |\Gamma^{-1}(1)| \cdot 2^{|R|}$ . Combining these two we have

$$|f^{-1}(1)| \leq \frac{|g^{-1}(1)|}{1 - (\varepsilon/10)} \leq \frac{|\Gamma^{-1}(1)| \cdot 2^{|R|}}{1 - (\varepsilon/10)} \leq 1.2 \cdot |\Gamma^{-1}(1)| \cdot 2^{|R|}. \quad (10)$$

The lemma follows by combining Equation (9) and Equation (10).  $\square$

**Lemma 47.** Assume  $f$  respects Assumption 2. For any  $\alpha \in \{0, 1\}^U$ , Algorithm 4 run on  $\alpha$  satisfies

$$\Pr \left[ \text{Algorithm 4 does not reject on line 2 and returns } 1 - \Gamma(\alpha) \right] \leq \frac{\varepsilon}{40c_0}.$$

*Proof.* If  $\Gamma(\alpha) = 0$ , then by the definition of  $\Gamma$  we have  $|f|_{\alpha}^{-1}(1)| \leq 2^{|R|}/16$ . So by the second bullet of Assumption 2, we have

$$\Pr_{\mathbf{w} \sim \{0,1\}^R} [f(\alpha \circ \mathbf{w}) \neq f(u \circ \mathbf{w})] \geq \frac{1}{16}.$$

The probability of Algorithm 4 returning 1 without rejecting  $f$  is at most

$$\left( \frac{15}{16} \right)^{c_2 \log(1/\varepsilon)} \leq \frac{\varepsilon}{40c_0}$$

by making  $c_2$  sufficiently large and also sufficiently larger than  $c_0$ .

If  $\Gamma(\alpha) = 1$ , the algorithm returns 0 if all samples  $\mathbf{w}$  had  $f(\alpha \circ \mathbf{w}) = 0$ . By the definition of  $\Gamma$ , this happens with probability at most  $(15/16)^{c_2 \log(1/\varepsilon)}$ .  $\square$

We are now ready to show that  $f$  is rejected on line 7 with high probability:

**Lemma 48.** Assume  $f$  respects Assumption 2. Then CONJ-TEST rejects on line 7 of Algorithm 3 with probability at least 0.85.

*Proof.* By Lemma 45,  $\Gamma$  is  $(\varepsilon/100)$ -far from conjunctions in relative distance. By Lemma 46, every  $\alpha \in \Gamma^{-1}(1)$  is returned by Algorithm 5 with probability at least  $1/(20|\Gamma^{-1}(1)|)$ . If all the  $c_0/\varepsilon$  many membership queries made by CONJ-TEST are answered correctly, then by Theorem 28 CONJ-TEST rejects  $\Gamma$  with probability at least 0.9.

Now consider an execution of CONJ-TEST under the sampling oracle simulated by Algorithm 5 together with the correct membership query oracle, and the  $c_0/\varepsilon$  membership queries it makes. If all these  $c_0/\varepsilon$  membership queries are sent to Algorithm 4, then by Lemma 47 and a union bound, with probability at least 39/40, either  $f$  is rejected on line 2 of Algorithm 4 or all  $c_0/\varepsilon$  membership queries are answered correctly. This shows that with probability at least 39/40, either  $f$  is rejected or CONJ-TEST returns the same answer as it would with the correct membership oracle.

As a result, line 7 of Algorithm 3 rejects with probability at least  $0.9 \cdot (39/40) \geq 0.85$ .  $\square$

Finally we can prove our result establishing soundness:

**Theorem 49.** If  $f$  is  $\varepsilon$ -far from any decision list in relative distance, then Algorithm 3 rejects it with probability at least 0.7.

*Proof.* Given that  $f$  is  $\varepsilon$ -far from decision lists, line 1 accepts with probability at most 0.1 by Theorem 1. When this does not happen, the algorithm reaches line 3 and we consider two cases.

If  $f$  does not satisfy Assumption 2, then by Lemma 43 and Lemma 44 it is rejected during lines 3–6 with probability at least 0.9. If  $f$  satisfies Assumption 2, then it follows from Lemma 48 that it is rejected on line 7 with probability at least 0.85.

As a result, the algorithm rejects  $f$  with probability at least  $0.9 \cdot 0.85 \geq 0.7$ .  $\square$

## References

- [BB16] A. Belovs and E. Blais. A polynomial lower bound for testing monotonicity. In *Proceedings of the 48th ACM Symposium on Theory of Computing*, pages 1021–1032, 2016. [1](#)
- [BG22] Nader H. Bshouty and Oded Goldreich. On properties that are non-trivial to test. *Electron. Colloquium Comput. Complex.*, TR22-013, 2022. [8](#), [33](#)
- [BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47:549–595, 1993. Earlier version in STOC’90. [1](#)
- [Bsh20] Nader H. Bshouty. Almost optimal testers for concise representations. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, pages 5:1–5:20, 2020. , [1](#), [2](#), [3](#), [4](#), [6](#)
- [BY22] Arnab Bhattacharyya and Yuichi Yoshida. *Property Testing - Problems and Techniques*. Springer, 2022. [1](#)
- [CDH<sup>+</sup>25] X. Chen, A. De, Y. Huang, S. Nadimpalli, R. Servedio, and T. Yang. Relative error monotonicity testing. In *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2025. , [1](#), [2](#), [5](#), [6](#), [9](#)
- [CDST15] Xi Chen, Anindya De, Rocco A. Servedio, and Li-Yang Tan. Boolean Function Monotonicity Testing Requires (Almost)  $n^{1/2}$  Non-adaptive Queries. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015*, pages 519–528, 2015. [1](#)
- [CFP24] Xi Chen, Yumou Fei, and Shyamal Patel. Distribution-free testing of decision lists with a sublinear number of queries. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1051–1062, 2024. [1](#), [2](#)
- [CGM11] Sourav Chakraborty, David García-Soriano, and Arie Matsliah. Efficient sample extractors for juntas with applications. In *Automata, Languages and Programming - 38th International Colloquium, ICALP*, pages 545–556, 2011. [1](#), [4](#)
- [CP22] Xi Chen and Shyamal Patel. Distribution-free testing for halfspaces (almost) requires PAC learning. In *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1715–1743, 2022. [1](#)
- [CS13a] Deeparnab Chakrabarty and C. Seshadhri. A  $o(n)$  monotonicity tester for boolean functions over the hypercube. In *Proceedings of the 45th ACM Symposium on Theory of Computing*, pages 411–418, 2013. [1](#)
- [CS13b] Deeparnab Chakrabarty and C. Seshadhri. Optimal bounds for monotonicity and Lipschitz testing over hypercubes and hypergrids. In *Symposium on Theory of Computing Conference, STOC*, pages 419–428, 2013. [1](#)
- [CS16] Deeparnab Chakrabarty and C. Seshadhri. A  $\tilde{O}(n)$  non-adaptive tester for unateness. *CoRR*, abs/1608.06980, 2016. [1](#), [4](#)

- [CST14] Xi Chen, Rocco A. Servedio, and Li-Yang Tan. New algorithms and lower bounds for testing monotonicity. In *Proceedings of the 55th IEEE Symposium on Foundations of Computer Science*, pages 286–295, 2014. [1](#)
- [CW19] Xi Chen and Erik Waingarten. Testing unateness nearly optimally. In *Proceedings of the 51th ACM Symposium on Theory of Computing*, 2019. [1](#), [4](#)
- [CWX17a] Xi Chen, Erik Waingarten, and Jinyu Xie. Beyond Talagrand functions: new lower bounds for testing monotonicity and unateness. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 523–536, 2017. [1](#), [4](#)
- [CWX17b] Xi Chen, Erik Waingarten, and Jinyu Xie. Boolean unateness testing with  $\tilde{O}(n^{3/4})$  adaptive queries. In *Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 868–879, 2017. [1](#), [4](#)
- [CX16] Xi Chen and Jinyu Xie. Tight bounds for the distribution-free testing of monotone conjunctions. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 54–71, 2016. [1](#), [2](#)
- [DLM<sup>+</sup>07] I. Diakonikolas, H. Lee, K. Matulef, K. Onak, R. Rubinfeld, R. Servedio, and A. Wan. Testing for concise representations. In *Proc. 48th Ann. Symposium on Computer Science (FOCS)*, pages 549–558, 2007. [1](#), [2](#), [4](#), [6](#)
- [FLN<sup>+</sup>02] E. Fischer, E. Lehman, I. Newman, S. Raskhodnikova, R. Rubinfeld, and A. Samorodnitsky. Monotonicity testing over general poset domains. In *Proc. 34th Annual ACM Symposium on the Theory of Computing*, pages 474–483, 2002. [1](#)
- [GGL<sup>+</sup>00] Oded Goldreich, Shafi Goldwasser, Eric Lehman, Dana Ron, and Alex Samordinsky. Testing monotonicity. *Combinatorica*, 20(3):301–337, 2000. [1](#)
- [GGR98] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45:653–750, 1998. [1](#)
- [GLR<sup>+</sup>91] Peter Gemmell, Richard J. Lipton, Ronitt Rubinfeld, Madhu Sudan, and Avi Wigderson. Self-testing/correcting for polynomials and for approximate functions. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 32–42. ACM, 1991. [1](#)
- [Gol17] O. Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017. [1](#)
- [Gol20] Oded Goldreich, editor. *Computational Complexity and Property Testing - On the Interplay Between Randomness and Computation*, volume 12050 of *Lecture Notes in Computer Science*. Springer, 2020. [1](#)
- [GR20] Oded Goldreich and Dana Ron. One-sided error testing of monomials and affine subspaces. *Electron. Colloquium Comput. Complex.*, TR20-068, 2020. [1](#), [2](#), [3](#), [8](#)
- [GS09] Dana Glasner and Rocco A. Servedio. Distribution-free testing lower bound for basic boolean functions. *Theory of Computing*, 5(1):191–216, 2009. [1](#), [2](#)
- [HK03] S. Halevy and E. Kushilevitz. Distribution-Free Property Testing. In *Proceedings of the Seventh International Workshop on Randomization and Computation*, pages 302–317, 2003. [1](#)

- [HK05] S. Halevy and E. Kushilevitz. A lower bound for distribution-free monotonicity testing. In *Proceedings of the Ninth International Workshop on Randomization and Computation*, pages 330–341, 2005. [1](#)
- [KMS18] Subhash Khot, Dor Minzer, and Muli Safra. On monotonicity testing and boolean isoperimetric-type theorems. *SIAM J. Comput.*, 47(6):2238–2276, 2018. [1](#)
- [KS16] Subhash Khot and Igor Shinkar. An  $\tilde{o}(n)$  queries adaptive tester for unateness. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM)*, pages 37:1–37:7, 2016. [1](#), [4](#)
- [LW19] Amit Levi and Erik Waingarten. Lower bounds for tolerant junta and unateness testing via rejection sampling of graphs. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 52:1–52:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. [1](#), [4](#)
- [MORS10] Kevin Matulef, Ryan O’Donnell, Ronitt Rubinfeld, and Rocco A. Servedio. Testing halfspaces. *SIAM Journal on Computing*, 39(5):2004–2047, 2010. [1](#)
- [O’D14] R. O’Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. [1](#)
- [PRS02] M. Parnas, D. Ron, and A. Samorodnitsky. Testing Basic Boolean Formulae. *SIAM J. Disc. Math.*, 16:20–46, 2002. , [1](#), [2](#), [3](#), [4](#), [8](#), [9](#), [10](#), [12](#), [15](#)
- [Ron08] D. Ron. Property Testing: A Learning Theory Perspective. *Foundations and Trends in Machine Learning*, 1(3):307–402, 2008. [1](#)
- [Val84] L. Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, 1984. [1](#)

## A Lower Bounds

We follow ideas of [BG22] to give an  $\Omega(1/\varepsilon)$  lower bound for any two-sided adaptive relative-error testing algorithm for conjunctions and decision lists.

To this end we define the following two distributions  $\mathcal{D}_{\text{yes}}$  and  $\mathcal{D}_{\text{no}}$ :

1.  $\mathcal{D}_{\text{yes}}$  is supported on the constant-1 function only; and
2.  $\mathbf{f} \sim \mathcal{D}_{\text{no}}$  is drawn by setting independently every entry  $\mathbf{f}(x)$  of  $\mathbf{f}$  to be 0 with probability  $3\varepsilon$  and 1 with probability  $1 - 3\varepsilon$ .

The constant-1 function is trivially both a conjunction and a decision list. On the other hand, the following lemma shows that, as long as  $\varepsilon$  satisfies  $C \cdot \log n/2^n \leq \varepsilon \leq 0.1$  for some sufficiently large constant  $C$ ,  $\mathbf{f} \sim \mathcal{D}_{\text{no}}$  is far from decision lists (and thus, conjunctions) with high probability.

**Lemma 50.** With probability at least  $1 - o_n(1)$ ,  $\mathbf{f} \sim \mathcal{D}_{\text{no}}$  satisfies  $\text{rel-dist}(\mathbf{f}, \text{DL}) \geq \varepsilon$ .

*Proof.* By setting  $C$  suitably large and a Chernoff bound, with probability  $1 - o_n(1)$ , we have

1.  $|\mathbf{f}^{-1}(0)|$  lies in  $(1 \pm 0.1)(3\varepsilon 2^n)$ ; and
2. For any  $i \in [n]$  and  $b \in \{0, 1\}$ , we have  $|\{x : x_i = b \text{ and } \mathbf{f}(x) = 0\}| \geq 0.9(3\varepsilon 2^{n-1})$ .

We show that  $\text{rel-dist}(\mathbf{f}, \text{DL}) \geq \varepsilon$  whenever these conditions hold. Consider any decision list  $g$ :

1. If  $g$  starts with  $v_{\text{default}} = 1$  then  $g$  is the constant-1 function and we have  $\text{dist}(\mathbf{f}, g) \geq 2.7\varepsilon$ ;
2. If  $g$  starts with  $v_{\text{default}} = 0$  or  $(x_{i_1}, b_1, 0)$ , then  $|g^{-1}(1)| \leq 2^{n-1}$  and  $\text{dist}(\mathbf{f}, g) \geq 0.5 - 3.3\varepsilon$ ;
3. If  $g$  starts with  $(x_{i_1}, b_1, 1)$ , then by the second condition, we have  $\text{dist}(\mathbf{f}, g) \geq 1.35\varepsilon$ .

The lemma then follows from  $\text{rel-dist}(\mathbf{f}, g) \geq \text{dist}(\mathbf{f}, g)$  and  $0.5 - 3.3\varepsilon \geq \varepsilon$ .  $\square$

Let  $q = c/\varepsilon$  for some constant  $c$ . Assume for a contradiction that  $\mathcal{A}$  is a randomized, two-sided, adaptive relative-error testing algorithm for conjunctions / decision lists, which receives  $q$  samples from  $\text{Samp}(f)$  and makes  $q$  queries on  $\text{MQ}(f)$ . Then by the correctness of the algorithm, we have

$$\Pr[\mathcal{A} \text{ accepts the constant-1 function}] \geq 0.9 \quad \text{and} \quad \Pr_{\mathbf{f} \sim \mathcal{D}_{\text{no}}}[\mathcal{A} \text{ accepts } \mathbf{f}] \leq 0.1$$

and thus, we have

$$\Pr[\mathcal{A} \text{ accepts the constant-1 function}] - \Pr_{\mathbf{f} \sim \mathcal{D}_{\text{no}}}[\mathcal{A} \text{ accepts } \mathbf{f}] \geq 0.8.$$

Given that  $\mathcal{A}$  is a distribution over deterministic algorithms, there is a deterministic algorithm  $\mathcal{B}$  (i.e., what it does after receiving the samples is deterministic) with the same complexity such that

$$\Pr[\mathcal{B} \text{ accepts the constant-1 function}] - \Pr_{\mathbf{f} \sim \mathcal{D}_{\text{no}}}[\mathcal{B} \text{ accepts } \mathbf{f}] \geq 0.8.$$

We show below that when  $c$  is sufficiently small, the LHS is at most 0.1, a contradiction.

To this end, consider an execution of  $\mathcal{B}$  on the constant-1 function. First  $\mathcal{B}$  draws  $q$  samples from  $\{0, 1\}^n$ , which we denote by  $\mathbf{X} = (\mathbf{x}^1, \dots, \mathbf{x}^q)$  and is uniform over  $(\{0, 1\}^n)^q$ . After receiving  $\mathbf{X}$ , the deterministic algorithm  $\mathcal{B}$  makes a sequence of  $q$  queries. While  $\mathcal{B}$  is adaptive, given that it is run on the constant-1 function, all queries return 1 and thus, the  $q$  queries made by  $\mathcal{B}$  is determined by  $\mathbf{X}$  and we denote them by  $\mathbf{Y} = (\mathbf{y}^1, \dots, \mathbf{y}^q)$ . We refer to a pair  $(X, Y)$  as a transcript of  $\mathcal{B}$  on the constant-1 function if  $Y$  consists of queries that  $\mathcal{B}$  makes after receiving  $X$  as samples. There are  $2^{nq}$  transcripts  $(X, Y)$ , one for each  $X \in (\{0, 1\}^n)^q$ , and  $(\mathbf{X}, \mathbf{Y})$  is uniformly distributed among all the  $2^{nq}$  transcripts.

Let  $\mathcal{T}$  denote the set of all transcripts that lead  $\mathcal{B}$  to accept the constant-1 function. Then

$$\Pr[\mathcal{B} \text{ accepts the constant-1 function}] = |\mathcal{T}|/2^{nq}. \tag{11}$$

On the other hand, fix any  $(X, Y)$  in  $\mathcal{T}$ , where  $X = (x^1, \dots, x^q)$  and  $Y = (y^1, \dots, y^q)$ , and we consider the probability of the following event:

$\mathcal{E}_{X,Y}$ : Draw  $\mathbf{f} \sim \mathcal{D}_{\text{no}}$ ; when running  $\mathcal{B}$  on  $\mathbf{f}$ ,  $\mathcal{B}$  receives  $X$  as the sequence of  $q$  samples and the sequence of  $q$  queries it makes are  $Y = (y^1, \dots, y^q)$  and all of them return 1.

We note that when  $\mathcal{E}_{X,Y}$  occurs,  $\mathcal{B}$  must accept  $\mathbf{f}$ . As a result we have

$$\Pr_{\mathbf{f} \sim \mathcal{D}_{\text{no}}}[\mathcal{B} \text{ accepts } \mathbf{f}] \geq \sum_{(X,Y) \in \mathcal{T}} \Pr_{\mathbf{f} \sim \mathcal{D}_{\text{no}}}[\mathcal{E}_{X,Y}].$$

Finally, for any  $(X, Y) \in \mathcal{T}$ , by setting  $c$  sufficiently small, the probability of  $\mathcal{E}_{X,Y}$  is at least

$$(1 - 3\varepsilon)^{2q} \cdot \left(\frac{1}{2^n}\right)^q = (1 - 3\varepsilon)^{2c/\varepsilon} \cdot \frac{1}{2^{nq}} \geq 0.9 \cdot \frac{1}{2^{nq}},$$



where the  $(1 - 3\varepsilon)^{2q}$  on the LHS is the probability that every point  $z \in X \cup Y$  has  $\mathbf{f}(z) = 1$ ; when this happens, the probability of getting  $X$  as the samples is

$$\left( \frac{1}{|\mathbf{f}^{-1}(1)|} \right)^q \geq \left( \frac{1}{2^n} \right)^q,$$

and when this also happens,  $Y$  must be the queries that  $\mathcal{B}$  makes and all queries return 1.

Putting all pieces together, we have

$$\Pr[\mathcal{B} \text{ accepts the constant-1 function}] - \Pr_{\mathbf{f} \sim \mathcal{D}_{\text{no}}}[\mathcal{B} \text{ accepts } \mathbf{f}] \leq \frac{|\mathcal{T}|}{2^{nq}} - 0.9 \cdot \frac{|\mathcal{T}|}{2^{nq}} = 0.1 \cdot \frac{|\mathcal{T}|}{2^{nq}}$$

and using  $|\mathcal{T}| \leq 2^{nq}$ , the LHS is at most 0.1, a contradiction.