Spring 5-15-2020

CSP-Completeness And Its Applications

Alexander Durgin
Washington University in St. Louis

Washington University in St. Louis

School of Engineering and Applied Science

Department of Computer Science and Engineering

Thesis Examination Committee:
Brendan Juba, Chair
Jeremy Buhler
William Yeoh

CSP-Completeness And Its Applications

by

Alexander Durgin

A thesis presented to the Graduate School of Arts and Sciences
of Washington University in partial fulfillment of the
requirements for the degree of

Master of Science

May 2020
Saint Louis, Missouri

Contents

Acknowledgments

First and foremost, I would like to thank Professor Brendan Juba for the time and energy he spent introducing me to exciting problems in learning theory and helping to guide me through them. I would also like to thank my thesis committee for taking the time out of their busy workloads to serve on my committee and help me move forward with my academic career. And lastly, I would like to thank the Boeing Company for providing the funding to pay for the entirety of my tuition.

Alexander Durgin

Washington University in Saint Louis
May 2020

ABSTRACT OF THE THESIS

CSP-Completeness And Its Applications

by

Alexander Durgin

Master of Science in Computer Science

Washington University in St. Louis, May 2020

Research Advisor: Brendan Juba

We build off of previous ideas used to study both reductions between CSP-refutation problems and improper learning and between CSP-refutation problems themselves to expand some hardness results that depend on the assumption that refuting random CSP instances are hard for certain choices of predicates (like k-SAT). First, we are able argue the hardness of the fundamental problem of learning conjunctions in a one-sided PAC-esque learning model that has appeared in several forms over the years. In this model we focus on producing a hypothesis that foremost guarantees a small false-positive rate while minimizing the false-negative rate for such hypotheses. Further, we formalize a notion of CSP-refutation reductions and CSP-refutation completeness that and use these, along with candidate CSP-refutatation complete predicates, to provide further evidence for the hardness of several problems.

Chapter 1

Introduction

1.1 Motivation and related work

Arguably, the main contribution of Computational Complexity to human knowledge has been the development of methods for establishing that problems are likely to be intractable. NP-completeness has found application well beyond computer science. In spite of its vast reach, there are limits to what can be established intractable using NP-completeness. First and foremost, NP-completeness, being only evidence of worst-case intractability, is no strong evidence that a problem is hard "in practice." Indeed, one need look no further than satisfiability to find a problem that is both NP-complete and yet widely held to be solvable "in practice." Second, moreover, there are large families of essentially average-case problems that are suspected to be intractable, specifically arising in cryptography (Bogdanov and Trevisan, 2006; Akavia et al., 2006; Haitner et al., 2010; Pass et al., 2011) and machine learning (Applebaum et al., 2008), but for which NP-completeness would violate other standard assumptions. NP-completeness is simply not an effective tool for studying these problems.

Famously, an analogue to the theory of NP-completeness for average-case problems was proposed by Levin (1986). In this framework, computational problems are paired with a distribution on inputs, called a "distributional problem," and it is this "distributional problem" pair that may be tractable or intractable. While this presents a natural candidate framework for the study of such problems, proving problems to be "average-case NP-complete" is much more difficult than for worst-case (standard) NP-completeness. It is difficult to construct reductions that enforce that typical instances of an arbitrary problem map to typical instances of a candidate complete problem. The current state of our knowledge can be summarized

as: all natural NP-complete problems are average-case hard for a somewhat unnatural distribution (Livne, 2010), and the existence of any average-case NP-complete problem for an efficiently sampleable distribution implies that there is an average-case complete problem with the uniform distribution on inputs, although the problem itself may be unnatural (Impagliazzo and Levin, 1990). Thus, although we do not have formal evidence against the possibility, it is still beyond the state of the art to establish the intractability of problems arising in cryptography and machine learning via average-case NP-completeness. For example, Goldwasser and Kalai (2016) describe such a hypothetical finding as a "triumph."

In both cryptography and machine learning, the standard for hardness has been largely reductions from a "famous" problem with no known efficient algorithms, such as factoring integers (equiv., computing modular square roots), discrete logarithms, or finding shortest vectors in lattices. In the best of circumstances, these problems have known worst-case to average-case (uniform random input) reductions, and so are based on the worst-case hardness of the corresponding famous problem. Even here, the reach of this criteria was sufficiently limited that alternatives have been sought. For example, relatively few assumptions have yielded public-key cryptography, and such reductions are not known for (improper) PAC-learning of DNF or approximate agnostic learning of standard hypothesis classes such as conjunctions or halfspaces. In the case of cryptography, this has left the field dangerously exposed to advances in quantum computers, for example, which could leave only a few related standard intractability assumptions (based on lattices and "learning with errors" (Regev, 2009; Brakerski et al., 2013)) valid in practice.

This has motivated the search for other, somewhat tested average-case assumptions, such as the planted clique problem (Jerrum, 1992; Kučera, 1995) and random constraint satisfaction problems such as refuting random 3-SAT (Feige, 2002). Indeed, for public-key cryptography, Applebaum et al. (2010) explored the use of variants of both of these assumptions. Likewise, in machine learning, given the intractability of a variant of refuting random k-SAT (for large k) recent work by Daniely and Shalev-Shwartz established the intractability of improper learning of DNF (Daniely and Shalev-Shwartz, 2016); under the intractability of k-XOR, further work by Daniely (2016) also established the intractability of "agnostic" learning of halfspaces with a constant-factor approximation to the optimal error rate. In comparison to problems such as integer factoring and shortest vectors in lattices, the specific average-case variants of these problems considered are relatively contemporary developments. As

discussed above, evidence for the hardness of these problems is primarily limited to the failure of specific algorithms, or specific classes of approaches. Thus, further evidence for the intractability of these problems is desirable.

1.2 Overview of the results

The goals of this work are two fold. First, it is to establish a general framework for formalizing the concepts of reductions between CSP-refutation problems and completeness over these reductions with respect to arbitrary classes of refutation problems, as well as providing examples of complete refutation problems for the class of all non-trivial refutation problems. And secondly, it is to provide evidence for the hardness of the fundamental problem of learning conjunctions in an often-studied one-sided PAC-esque learning model.

1.2.1 Formalizing CSP-Reductions and Associated Completeness

More to the first goal, we show new evidence for the intractability of *refuting random k-SAT*, the problem of distinguishing satisfiable k-CNFs from k-CNFs for which $m = f(n, k)$ clauses on n variables (for a specified function f) are chosen uniformly at random. It is known that beyond some constant clause-to-variable ratio, the formulas become unsatisfiable for each fixed $k \geq 2$ (Friedgut and Bourgain, 1999), and in particular that for sufficiently large, fixed k the limiting probability of satisfiability (as $n \to \infty$) drops from 1 to 0 at a fixed, constant density (Ding et al., 2015). (The same "sharp threshold" phenomenon is widely conjectured to hold more generally for all $k \geq 3$.) Thus, at least, for some $f(n, k) = \Omega(n)$, the problem is information-theoretically feasible in the sense that below this threshold, while the problem can be posed, no algorithm can solve it since the distributions are statistically indistinguishable. By contrast, the current best *efficient algorithms* for refuting random k-SAT require $\tilde{\Omega}(n^{k/2})$ clauses (Allen et al., 2015). Feige (2002) conjectured that the problem is intractable for some $f(n, k) = O(n)$ (sufficiently large that the problem is information-theoretically feasible), and recently Daniely and Shalev-Shwartz more boldly conjectured that the problem is hard for $f(n, k) = n^{\omega(1)}$ clauses. As we discussed in section 1.1, these

assumptions were used to show the intractability of certain problems, especially in machine learning, that had resisted previous efforts.

The only previous evidence for hardness of these problems is that broad classes of algorithms cannot solve them. Specifically, "statistical algorithms" (Feldman et al., 2018) and semidefinite programming hierarchies such as sum-of-squares (Schoenebeck, 2008) cannot detect the unsatisfiability of formulas with substantially fewer clauses. Indeed, more generally, for any random k-CSP for which a t-wise independent probability distribution can be constructed on the satisfying assignments of the defining constraint, neither statistical algorithms (Feldman et al., 2018) nor the sum-of-squares relaxation (Kothari et al., 2017) detect unsatisfiability with less than $\tilde{\Omega}(n^{t/2})$ constraints, and the best known polynomial-time algorithm (Allen et al., 2015, again) similarly requires this many constraints. Thus, we have essentially the same evidence for hardness of *every* k-CSP that supports $k - O(1)$-wise independence, and Barak et al. (2013) likewise conjecture that *every* such random CSP refutation problem is intractable.

Here, we show that refuting random k-SAT is *complete* for random k-CSP refutation problems for each fixed k under randomized, constraint-wise reductions. We give a simple family of reductions that work for any k and for the strong refutation variant in which the formula may only be ϵ-close to satisfiable. To our knowledge, this notion of completeness for random k-CSP refutation problems had not previously been considered. The closest analogue is the reductions between the sum-of-squares relaxations of such CSP problems, considered first by Schoenebeck (2008) and more generally by Tulsiani (2009) and finally Chan (2016), but our reduction applies to any algorithm for refuting random CSP instances, not just semidefinite programming relaxations (or other specific methods). Under this notion of completeness, we are able to vastly expand the classes of predicates which are candidates for producing hard refutation problems, a phenomenon that, to our knowledge, has not yet occurred in the study of hardness of improper learning. For example, the class of CSP-refutation problems for which k-SAT is complete includes all CSPs considered in the literature which tend to be nontrivial, monotone CSPs, as well as all exotic predicates such as not-all-equal-SAT, random predicates, and even the XOR$\oplus$MAJ predicate introduced by Applebaum and Lovett (2018) as a candidate hard predicate for PRGs in NC^0 achieving high stretch.

In particular, we now have this strengthened evidence for the hardness of problems in machine learning that had shown to be intractable under the assumption that improper learning of DNF is intractable. For example, in addition to the various classes that can express DNFs such as DFAs of size n^ϵ (Pitt and Valiant, 1988), this includes the hardness of agnostic learning of conjunctions (Kearns et al., 1994) and parities (Feldman et al., 2009) up to arbitrary accuracy, and the hardness of improper heuristic learning of conjunctions (Bshouty and Burroughs, 2005) (a.k.a. "positive-reliable" learning (Kalai et al., 2012)) and relatedly, learning abduction of conjunctions (Juba, 2016) and conditional linear regression for conjunctive conditions (Juba, 2017).

We also note that recent work by Brennan et al. (2018) established a variety of reductions among planted sparse graph problems, ultimately based on planted clique (similar to earlier work by Berthet and Rigollet (2013)). Although along the lines of Papadimitriou (1994) we could simply consider "the class of problems reducible to planted clique" (PPC?), it would be very interesting if the evidence for the hardness of planted clique could be strengthened to completeness for a more natural class of problems, similar to what we show here for random k-SAT.

It follows from the reductions of Feige (2002) that, for the special case of $k = 3$, several other predicates are also complete for strong 3-CSP refutation (under refutation reductions), specifically including 3-XOR, 3-AND, and 3-MAJ. While these reductions seem to rely on the limited space of assignments on three variables and it is not clear if they can be generalized to arbitrary k, we are able to show that a new variant of random XOR is also hard for strong k-CSP refutation. Specifically, when the sizes of the predicates are binomially distributed (conditioned on at least one success),[1] we find that strongly refuting such random XOR systems is strong k-CSP-hard, and hence this new XOR variant could serve as the strong foundation of a candidate family of predicates for something like Goldreich's proposed pseudorandom generator Goldreich (2011). We note that this equivalently gives a reduction to a standard random CSP problem with multiple predicates (where the predicate is also chosen uniformly at random) if we consider the predicates given by non-constant parities supported on strings of length k. Previously, it was essentially shown by Allen et al. (2015) (see in particular Raghavendra et al. (2017) for a more careful treatment) how to reduce

[1] Recall that the binomial distribution with parameters n and p is the number of heads ("successes") when a p-biased coin is tossed n times.

any k-CSP to a collection of *weighted* k-XOR refutation problems, where "strong refutation" now means bounding the maximum value of the weighted sum of the predicates, where the weights are given by sums of (uniformly) $\{-1, 0, +1\}$-valued random variables. Neither Allen et al. (2015) nor Raghavendra et al. (2017) analyzed this transformation as a reduction, but rather showed how it could be used to give polynomial-time algorithms for strong refutation for large formulas. The reduction produces instances with many different size predicates, and they obtain different bounds for different regimes of the output predicate sizes. Thus it's challenging to extract from their work one clean variant of even their weighted k-XOR refutation problem that is complete.

1.2.2 CSP-Reduction Framework Applied to One-Sided Learning

Toward the second goal, we begin by considering the following, closely related learning models:

(i) In Pitt and Valiant's *heuristic learning* model (Pitt and Valiant, 1988), one seeks a "rule of thumb" that commits (almost) no false-positive errors, and matches the best true-positive rate of members of a given class.

(ii) In Kalai, Kanade, and Mansour's *positive-reliable learning* model (Kalai et al., 2012), one seeks a classifier that, again, commits almost no false-positive errors, and almost matches the false-negative rate of the optimal classifier that makes no false-positive errors.

(iii) In Juba's *learning to abduce* model (Juba, 2016), one seeks a "hypothesis" condition with probability as large as possible such that in the corresponding conditional probability distribution over examples, the label is almost always true, thus "empirically entailed."

The only differences are that Kalai et al. formulate their problem as minimizing false-negatives as opposed to maximizing the positive classification rate, and Juba essentially formulates the problem in terms of precision rather than the raw false-positive rate. Thus, in the realizable setting (where a perfect rule exists) all three models are computationally

equivalent. We also remark briefly that these problems also arise as a special case of the *conditional linear regression* problem (Juba, 2017), which is a variant of robust linear regression for small (minority-fraction) subsets, in which we also ask for a formula describing the subset on which the linear predictor is intended to be used.

Conjunctions are among the simplest and least-expressive nontrivial representations. They present a natural starting point for studying the extent of learnability in any model. Moreover, conjunctions are of particular significance to the learning to abduce model. In the usual formulation of abduction as a reasoning task, conjunctions are widely considered to be the most natural hypothesis formulation. For example a typical application of abduction is in diagnosing faulty circuits, and the hypothesis explaining a given output is usually a conjunction of faults at various points in the circuit. Indeed, many classical formulations of the abduction task in AI only considered conjunctive hypotheses (e.g., ATMS Reiter and de Kleer (1987)). Thus, conjunctions can be considered to be the central representation class for abduction, and therefore the learnability of that class in Juba's model is particularly significant.

It is therefore somewhat surprising and unfortunate that all evidence to date suggests that conjunctions are not learnable in these models. The first results concerned the "proper learning" variant of the task, in which we seek the representation of a specific conjunction solving the task: Pitt and Valiant showed this problem to be NP-hard (Pitt and Valiant, 1988); indeed, Bshouty and Burroughs (2005) noted that it follows from results of Håstad (1996) that even getting a $n^{1-\gamma}$-approximation to the optimal positive classification rate for this problem (for any constant $\gamma > 0$) is NP-hard.

Bshouty and Burroughs (2005) showed furthermore that even for the "improper" variant in which any representation will do, we cannot obtain any polynomial approximation for the positive classification rate, or else we would obtain an algorithm for PAC-learning DNF in the usual, distribution-free model. This was the central problem in computational learning theory raised by Valiant (1984), and the state-of-the-art algorithm for this problem requires $2^{O(n^{1/3})}$ time and examples (Klivans and Servedio, 2004). Until recently, the only evidence for the hardness of learning DNF was its notoriety. But, Daniely and Shalev-Shwartz (2016), building on techniques pioneered by Daniely et al. (2014), show that learning DNF is hard given that it is hard to distinguish random k-CNFs on $n^{f(k)}$ constraints from satisfiable

k-CNFs for any $f(k) = \omega(1)$. This assumption is a slight strengthening of Feige's *R3SAT hypothesis,* Feige (2002) which (only) asserts that no polynomial-time algorithm can distinguish random 3-CNFs of size $O(n)$ from satisfiable 3-CNFs. These connections show, in turn that improper learning of conjunctions is intractable under the same assumptions. Thus, furthermore, learning of any representation that can *express* a conjunction is also intractable in these models. Since almost all natural representations can express conjunctions (except disjunctions and parities, which are learnable (Bshouty and Burroughs, 2005; Kanade and Thaler, 2014; Juba, 2016)), this essentially settles the extent of learnability in these models.

The pioneering work of Daniely et al. (2014) had earlier obtained a number of strong hardness of improper learning results, using assumptions about the hardness of random CSPs for unusual predicates, that were unfortunately subsequently falsified by Allen et al. (2015). While the current, random k-SAT variant used by Daniely and Shalev-Shwartz has yet to be falsified (and indeed seems plausible) and PAC-learning of DNF still seems formidable, it is still desirable to have stronger evidence for the hardness of these problems. But, the hardness of learning has almost always been based on the hardness of specific problems, such as the aforementioned hardness of specific random CSP refutation problems, the hardness of specific cryptographic problems, such as integer factoring (Kearns and Valiant, 1994) or shortest vector problems (Klivans and Sherstov, 2009), or the hardness of planted clique (Berthet and Rigollet, 2013). Applebaum et al. (2008) obtained evidence that such a result for improper learning based on NP-hardness is implausible: most simple kinds of reductions would imply the polynomial-time hierarchy collapses, and more complex reductions would yield a generic reduction to construct weak one-way functions from arbitrary problems in NP that are hard on average. Moreover, no natural problem with a natural distribution has yet been shown to be complete for the average-case analogues of NP[2] so hardness for the average-case analogues of NP seem beyond our current reach, at least.

We show that the task of improper learning of conjunctions in this one-sided learning model is hard unless all non-trivial k-CSPs can be weakly refuted: and while this same result can be derived via Bshouty and Burroughs (2005), Daniely and Shalev-Shwartz (2016), and the completeness of k-SAT. The reduction we derive is significantly simpler than the resulting

[2]In particular, this is in contrast to specific works where either the distribution is natural and the problem is not (Impagliazzo and Levin, 1990, e.g.,) or where the problem is natural and the distribution is not (Livne, 2010).

reduction from the composition of the former and is worth laying out and studying in its own right for what it reveals about what little structure an abductive learner of conjunctions requires from predicates to be able to refute them.

We note that Barak et al. (2013) had explicitly conjectured that a basic semidefinite program should be optimal for weak refutation of *all* predicates with a constant constraint-to-variable ratio. By contrast, we only require that refutation is hard for *some* predicate on $n^{f(k)}$ constraints for arbitrarily slowly growing $f(k)$. In particular, Kothari et al. (2017) show that the usual sum-of-squares formulation cannot efficiently refute such instances whenever there is a $2f(k) + 1$-wise independent distribution on the satisfying assignments of the predicate. (We say that such predicates *"support $2f(k) + 1$-wise independence."*) Therefore, unless we can improve upon the sum-of-squares algorithm for refutation on $n^{f(k)}$ constraints for *all* non-trivial CSPs that support $2f(k) + 1$-wise independence, there is no polynomial-time algorithm for learning conjunctions in these models. Kothari et al. note that for the same family of predicates, it furthermore follows from work by Lee et al. (2015) that no polynomial-size semidefinite programming extended formulation will succeed for the same family and same number of constraints. Therefore, again, a polynomial time algorithm for learning conjunctions in this model will establish that *no* semidefinite programming formulations are optimal for *any* non-trivial predicates (in stark contrast to the conjecture of Barak et al. (2013)).

Chapter 2

Preliminaries

In this chapter we will be formalizing the problem of CSP-refutations which is a fundamental and recurring element throughout the rest of this work, as well as some notation and concepts centered around CSP-refutation and reductions from these problems.

2.1 CSP-Refutations

Definition 1 (Constraint satisfaction problems)**.** *We call a boolean function* $P : \{\pm 1\}^k \to \{0,1\}$ *a* k-predicate *and say that* $C : \{\pm 1\}^n \to \{0,1\}$ *is a* P-constraint *if* $\exists \alpha_1, ..., \alpha_k \in \{\pm 1\}$ *and distinct choices* $i_1, \ldots, i_k \in [n]$ *are distinct such that* $C(x) = P(\alpha_1 x_{i_1}, ..., \alpha_k x_{i_k})$*. We denote the set of such* P*-constraints by* CSP_P

For k*-predicate* P *and* $\eta \in [0,1]$*, we say that a randomized algorithm* $\mathcal{A}$ *solves* the problem of distinguishing between $(1-\eta)$-satisfiable and random instances of P-constraints of size $m(n,k)$ *if on input* $S = \{C_1, ..., C_{m(n,k)}\} \subseteq \mathrm{CSP}_P$ *the following holds:*

(i) If there exists $x \in \{\pm 1\}^n$ *such that for at least* $(1-\eta)m(n,k)$ *many* $i \in [m(n,k)]$ *we have* $C_i(x) = 1$*, then* $\mathcal{A}(S)$ *outputs* "satisfiable" *with probability at least* $\frac{3}{4}$ *with respect to the internal randomness of* $\mathcal{A}$ *and*

(ii) If each $C_i \in S$ *is drawn uniformly at random from all of the* P*-constraints and independently from the other* P*-constraints, then for almost-all choices of* S*,* $\mathcal{A}(S)$ *outputs* "random" *with probability at least* $\frac{3}{4}$ *with respect to the internal randomness of* $\mathcal{A}$*.*

We denote this problem by $\mathrm{CSP}^{rand,1-\eta}_{m(n,k)}(P)$.

If there is no such $\mathcal{A}$ *that runs in time* $\mathrm{poly}(n)$*, then we say that this problem is* hard. *Otherwise, we say that this problem is* easy.

For example, we can let P be the familiar k-SAT predicate, taking $x \in \{\pm 1\}^k$ to 1 when at least one of the inputs is positive, ie $x \neq \{-1, \ldots, -1\}$ and to 0 when all of the inputs are negative. Then the corresponding distinguishing problem $\mathrm{CSP}^{rand,1-\eta}_{m(n,k)}(k\text{-SAT})$ is that of deciding whether a set of k-SAT clauses has at least ηm simultaneously satisfiable clauses.

Such random constraint satisfaction problems have been studied mostly as a starting point for the study of the intractability of other problems. A relatively minimal assumption in this direction, which we will show suffices for many applications as a consequence of our reductions, is the following:

Assumption 2. *There exists some* k*-predicate* P *and* $f(k) = \omega(1)$ *such that* $\mathrm{CSP}^{rand,1}_{n^{f(k)}}(P)$ *is hard.*

2.2 CSP-Refutation Reductions

It is also convenient to give a name to the notion of reducibility between random CSP refutation problems that we use in this paper and that has appeared in the literature prior (particularly in (Ding et al., 2015; Daniely and Shalev-Shwartz, 2016; Daniely, 2016; Feige, 2002)).

Definition 3 (Refutation Reductions)**.** *Let* P, P' *be* k*-predicates. Let* $\eta, \eta' \in [0, 1)$*, then we say a function random polynomial-time computable* $f : \mathrm{CSP}^{m(n,k)}_P \to \mathrm{CSP}^{m'(n,k)}_{P'}$ *is a* refutation reduction *from* $\mathrm{CSP}^{rand,1-\eta}_{m(n,k)}(P)$ *to* $\mathrm{CSP}^{rand,1-\eta'}_{m'(n,k)}(P')$ *if the following holds:*

Let $S = \{C_1, \ldots, C_{m(n,k)}\}$*, where each* $C_i \in \mathrm{CSP}_P$ *and let* $S' = f(S) = \{C'_1, C'_2, \ldots, C_{m'(n,k)}\}$.

(i) If there exists $x \in \{\pm 1\}^n$ *such that for at least* $(1-\eta)m(n,k)$ *many* $i \in [m(n,k)]$ *we have* $C_i(x) = 1$*, then with probability* $\geq 3/4$ *over the internal randomness of* f*, there*

exists some $x' \in \{\pm 1\}^n$ *such that* $C_i(x') = 1$ *on at least* $(1-\eta')m'(n,k)$ *many clauses of* S'.

(ii) If each $C_i \in S$ *is drawn uniformly at random from all of the* P*-constraints and independently from the other* P*-constraints, then the induced distribution on the* C'_i *is also the uniform distribution over* P' *constraints. That is,* S' *is a uniform random system of* $m'(n,k)$ P'*-constraints.*

We would like to highlight some of the more interesting structure that can be found in the refutation reductions we study in this paper. In particular, these transformations perform "constraint-wise" mappings between the input and the image formula.

Definition 4 (Constraintwise Refutation Reductions)**.** *Let* $f : \mathrm{CSP}_P^{m(n,k)} \to \mathrm{CSP}_{P'}^{m(n,k)}$ *be a refutation reduction from* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta}(P)$ *to* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta'}(P')$ *such that for every* P*-formula* $S = \{C_1, \ldots, C_{m(n,k)}\}$*, there is a* $g : \mathrm{CSP}_P \to \mathrm{CSP}_{P'}$ *where* $f(S) = \{g(C_1), \ldots, g(C_{m(n,k)})\}$. *We call* f *the* constraint-wise refutation reduction *from* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta}(P)$ *to* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta'}(P')$ *induced by* g.

And with an idea of what it means to reduce from one random CSP refutation problem to another given, we can naturally discuss the idea of completeness for classes of random refutation problems.

Definition 5 (Refutation Completeness)**.** *Let* $\mathcal{C}$ *be a set of CSP refutation problems, that is, a collection of* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta}(P)$ *for possibly many different choices of* P*,* η*, and* m*. Then we say* $\mathrm{CSP}_{m'(n,k)}^{rand,1-\eta'}(P')$ *is* $\mathcal{C}$ random refutation complete *for* $\mathcal{C}$ *if for each* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta}(P) \in \mathcal{C}$ *there exists a refutation reduction* f *from* $\mathrm{CSP}_{m(n,k)}^{rand,1-\eta}(P)$ *to* $\mathrm{CSP}_{m'(n,k)}^{rand,1-\eta'}(P')$.

Note that the above definitions give a description of reducibility and completeness that can apply to both "weak refutation" (where $\eta = 0$) and "strong refutation" (where $\eta > 0$).

Chapter 3

CSP Refutation Completeness

3.1 Random k-CSP Completeness of k-SAT

We now give our main reductions, showing that random k-SAT and a new variant of strong refutation of random XOR are at least as hard as all strong refutation problems for k-CSPs. This gives new evidence for the intractability of these specific problems, which in turn we can argue gives new evidence for the hardness of many other problems that had been previously established to be hard under the assumption that random k-SAT specifically was hard. As our variant of random XOR refutation is new, it currently does not establish such results, and indeed this suggests several directions for future work that we will review subsequently.

3.1.1 The Reduction

Our main result is that random k-SAT is complete for strong refutation of random k-CSPs: it is itself a random k-CSP problem, and every other (strong) refutation problem for random k-CSPs reduces to it. Intuitively, our reduction proceeds as follows: Observe that we can think of a clause as checking that a forbidden assignment does not appear. For any arbitrary predicate P, given an instance of a CSP on P, we can choose one of the falsifying assignments of P to check independently for each constraint of the instance. This will map a random constraint to a random clause, and we will catch each falsified constraint with probability $1/2^K(1 - \mathbb{E}[P])$ (where $\mathbb{E}[P]$ denotes the probability that P is satisfied on an assignment chosen uniformly at random). So, in expectation, if $\eta(1 - \mathbb{E}[P])m$ constraints are falsified

originally, we get a k-SAT instance in which $\frac{\eta}{2^k}m$ constraints are falsified. If the formula is large enough, this fraction is almost exact. More formally, now:

Theorem 6 (CSP Refutation Completeness of k-SAT)**.** *Let $\eta \in [0,1]$, $m(n,k) = \Omega\Big(\frac{2^k}{\mathbb{E}[P]^2}\Big)$, and $\mathcal{C} = \{\mathrm{CSP}^{rand,1-\eta(1-\mathbb{E}[P])}_{m(n,k)}(P) | P$ is a falsifiable predicate$\}$. Then $\mathrm{CSP}^{rand,1-\eta}_{m(n,k)}(k\text{-SAT})$ is random refutation complete for $\mathcal{C}$ under randomized constraint-wise reductions.*

In particular, this theorem shows completeness for both the weak and strong variants of the random refutation problems.

We will make use of the following simple bound on the tail of the binomial distribution given in Feller (1957) (found in chapter VI, section 3).

Lemma 7. *Let $X \sim Binomial(n,p)$, then $\Pr[X \leq r] \leq \frac{(n-r)p}{(np-r)^2}$ for $r < np$.*

Now our proof of theorem 6:

Proof. For the remainder of the proof, let m denote $m(n,k)$ for brevity. Let $J = \{C_1, ..., C_m\} \subseteq \mathrm{CSP}_P$ be an instance of $\mathrm{CSP}^{rand,1-\eta(1-\mathbb{E}[P])}_{m(n,k)}(P)$. For each $C_i \in J$ let $x^{(i)} \in \{-1,1\}^n$ be a falsifying assignment chosen randomly over all falsifying assignments of C_i, hence $C_i(x^{(i)}) = -1$. Define $C'_i \in \mathrm{CSP}_{k\text{-SAT}}$ to be the k-disjunction over the variables which appear in C_i that is falsified by $x^{(i)}$, and consider the k-SAT formula $J' = \{C'_1, ..., C'_m\}$. Clearly, J' can be produced in polynomial time with respect to n (for constant k). We will argue that this transformation induces a constraint-wise refutation reduction.

First, we observe that if J is drawn uniformly at random, then J' will also be produced uniformly at random. Let x^* be any falsifying assignment of $P(x_{i_1}, \ldots, x_{i_k})$. Note that every $C_i \in J$ is of the form $C_i(x) = P(\alpha_{i_1} x_{i_1}, \ldots, \alpha_{i_k} x_{i_k})$ where in particular each $\alpha_{i_j} \in \{\pm 1\}$ is either label with probability $\frac{1}{2}$. Hence, the falsifying assignment of C'_i induced by x^* (where an index i_j is negated iff $\alpha_{i_j} = -1$) is uniform random on the variables occurring in C_i, and so C'_i must also be uniform random over the clauses of size k.

Next, suppose that J is $(1 - \eta(1 - \mathbb{E}[P]))$-satisfiable. Let $x^* \in \{\pm 1\}^n$ be an assignment satisfying a $1 - \eta(1 - \mathbb{E}[P])$ fraction of J, ie $C_i(x^*) = 1$ for some $(1 - \eta(1 - \mathbb{E}[P]))m$ many i. Let $m' = \lfloor \eta(1 - \mathbb{E}[P]))m \rfloor$, and without loss of generality, let $C_1, \ldots, C_{m'}$ be the remaining

constraints of J that may or may not be satisfied x^*. Note that for each $i \in \{m'+1, \ldots, m\}$ we have that x^* must differ from $x^{(i)}$ on at least one variable appearing in C_i, and hence $C'_i(x^*) \neq 0$. Note that if $\eta = 0$, then J is entirely satisfiable and there are no extra constraints to consider (ie $m' = 0$) in our analysis.

Thus, if $\eta > 0$, then we only need that our randomized reduction produces at least $((1 - \eta(1 - \mathbb{E}[k\text{-SAT}])) - (1 - \eta(1 - \mathbb{E}[P])))m = m' - m\eta(1 - \mathbb{E}[k\text{-SAT}])$ many satisfiable clauses from these remaining m' constraints. We will show this occurs with high probability via a simple tail-bound. Let X_i be the indicator random variable such that $X_i = 1$ iff the falsifying assignment of C_i disagrees with x^* on a variable appearing in C_i. We note that the X_i are independent Bernoulli trials, since the $x^{(i)}$ are chosen independently of one another by the reduction. So the sum $\sum_{i=1}^{m'} X_i$ is a binomial distribution with $p = \mathbb{E}[k\text{-SAT}]$, hence we can bound the probability that fewer than $m' - m\eta(1 - \mathbb{E}[k\text{-SAT}])$ of the constraints are satisfied by x^* with lemma 7:

$$\begin{aligned}
\Pr[\sum_{i=1}^{m'} X_i \leq m' - m\eta(1 - \mathbb{E}[k\text{-SAT}])] &\leq \frac{(m' - (m' - m\eta(1 - \mathbb{E}[k\text{-SAT}])))\mathbb{E}[k\text{-SAT}]}{(m'\mathbb{E}[k\text{-SAT}] - (m' - m\eta(1 - \mathbb{E}[k\text{-SAT}])))^2} \\
&= \frac{m\eta 2^{-k}(1 - 2^{-k})}{2^{-2k}(m')^2 + 2^{-2k}m^2\eta^2 - 2 \cdot 2^{-2k}m'm\eta} \\
&= \frac{m\eta(2^k - 1)}{(m')^2 + m^2\eta^2 - 2m'm\eta}
\end{aligned}$$

And for this to be bounded above by $1/4$, it suffices that $m > \frac{4}{\eta}\frac{2^k - 1}{\mathbb{E}[P]^2}$ since

$$\begin{aligned}
m &> \frac{4}{\eta}\frac{2^k - 1}{\mathbb{E}[P]^2} \\
m\big((1 - \mathbb{E}[P])^2 + 1 - 2(1 - \mathbb{E}[P])\big) &> 4\frac{1}{\eta}(2^k - 1) \\
\frac{(m\eta)^2(1 - \mathbb{E}[P])^2 + (m\eta)^2 - 2(m\eta)^2(1 - \mathbb{E}[P])}{m} &> 4\eta(2^k - 1) \\
\frac{(m')^2 + m^2\eta^2 - 2m'm\eta}{m} &> 4\eta(2^k - 1) \\
\frac{m\eta(2^k - 1)}{(m')^2 + m^2\eta^2 - 2m'm\eta} &< 1/4
\end{aligned}$$

And so for $m = \Omega(\frac{2^k}{\mathbb{E}[P]^2})$, these m' constraints will fail with probability $< \frac{1}{4}$ to produce at least $((1 - \eta(1 - \mathbb{E}[k\text{-SAT}])) - (1 - \eta(1 - \mathbb{E}[P])))m$ k-SAT constraints satisfied by x^*. $\square$

In particular, theorem 6 tells us that as long as there is *some* k-predicate P and $m(n, k)$ for which it is hard to distinguish random from almost satisfiable instances, then it is hard to do so for k-SAT.

We obtain many corollaries, for the various problems shown to be hard based on the assumption that refuting random k-SAT is hard. For example, we have the following strengthening of the lower bound found by Daniely and Shalev-Shwartz (2016) for improperly PAC-learning DNF, since their hardness assumption was on the hardness of k-SAT:

Corollary 8. *If assumption 2 holds, then there is no polynomial-time algorithm for improperly PAC-learning DNF formulas.*

Remark. *For a somewhat more limited set of predicates (those which are "uniformly falsifiable"), we give similar evidence for the hardness of PAC-learning DNF in Chapter 4.*

And consequently, furthermore, any lower bound established given the hardness of learning DNF is similarly strengthened. For example, agnostically learning conjunctions up to additive error is at least as hard as PAC-learning DNF (Kearns et al., 1994), hence:

Corollary 9. *If assumption 2 holds, then there is no polynomial-time algorithm for agnostically PAC-learning conjunctions up to additive error.*

Remark on the satisfiability threshold. Again, we know that when there are asymptotically fewer than $(2^k \ln 2)n$ clauses, random k-SAT formulas are satisfiable with high probability; thus, for m smaller than this, the reduction produces instances of an information-theoretically impossible problem. There is no contradiction here, but the reduction is not useful since it produces unconditionally hard instances.

3.1.2 Beyond k-SAT

We observe that we can use this "random k-CSP completeness" of k-SAT to show random k-CSP hardness of other CSP refutation problems. As noted previously, Feige (2002) had

shown that strong refutation of random 3-SAT reduces to strong refutation of many other predicates, such as 3-AND, 3-XOR, and 3-MAJ. Hence, all of these are complete for strong 3-CSP refutation. But, Feige's reduction relies on the use of good algorithms for refuting NAE-SAT to handle cases where simply interpreting a clause as an XOR (for example) fails, and it is not clear how to generalize his algorithm beyond $k = 3$. We note that likewise, although later work has shown (for general k) how to use algorithms for strongly refuting sufficiently large *weighted* k-XOR formulas to refute arbitrary k-CSP instances (Allen et al., 2015; Raghavendra et al., 2017), the weighted k-XOR problem is not a standard CSP refutation problem, and furthermore the reduction requires solving a large collection of such weighted instances. Finally, these works do not analyze these transformations as a reduction per se, but rather use them in the specific large-formula ($\Omega(n^{k/2})$ constraint) regime to obtain efficient algorithms for refuting large formulas. By contrast here, we exhibit a new variant of random XOR that can be viewed as a standard CSP, which we can establish is hard for strong refutation of k-CSPs under constraint-mapping reductions.

For every non-empty $S \subseteq [k]$, let $\chi_S : \{\pm 1\}^k \mapsto \{\pm 1\}$ denote the parity function on the variables whose index appear in S, ie $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$. Then, for example, we can show k-CSP completeness of $(\leq k)$-XOR: Let $CSP_{(\leq k)\text{-XOR}}$ denote the set of constraints over all of the non-constant χ_S predicates. Define $\mathrm{CSP}^{rand,1-\eta(1-\mathbb{E}[(\leq k)\text{-XOR}])}_{m(n,k)}((\leq k)\text{-XOR})$ as before, but we let our formula's constraints be drawn from over any of the χ_S constraints, and the random formula are uniform not only over the literals but also over the choice of S. Equivalently, these are instances in which the constraints are independently chosen to have binomially distributed sizes, conditioned on a nonzero outcome ($k_i \sim \mathrm{Binomial}(k, 1/2) | k_i > 0$, independently).

Theorem 10. *Let P be any non-trivial k-CSP. Let $\eta \in [0, 1)$ and $m = \omega(1)$. Then there is a randomized constraint-wise refutation reduction from $\mathrm{CSP}^{rand,1-\eta(1-\mathbb{E}[P])}_{m(n,k)}(P)$ to $\mathrm{CSP}^{rand,1-\eta'(1-\mathbb{E}[(\leq k)\text{-XOR}])}_{m(n,k)}((\leq k)\text{-XOR})$ for $\eta' \in [0, 1]$ depending only on η and k.*

Proof. We will argue that $\mathrm{CSP}^{rand,1-\eta(1-\mathbb{E}[k\text{-SAT}])}_{m(n,k)}(k\text{-SAT})$ efficiently (constraint-wise refutation) reduces to $\mathrm{CSP}^{rand,1-\eta'(1-\mathbb{E}[(\leq k)\text{-XOR}])}_{m(n,k)}((\leq k)\text{-XOR})$, and the result will follow. First, we will observe that every parity function over some non-empty subset of the variables appearing in a clause of size k agrees with that clause on a $\frac{1}{2} + \frac{1}{2^k}$ fraction of all inputs (that is, the Fourier coefficient of k-SAT on every non-constant character is $\frac{1}{2^{k-1}}$).

Now consider the following efficient, refutation reduction. Let $J = \{C_1, \ldots, C_{m(n,k)}\} \subseteq \text{CSP}_{k\text{-SAT}}$ be an instance of $\text{CSP}^{rand,1-\eta(1-\mathbb{E}[k\text{-SAT}])}_{m(n,k)}(k\text{-SAT})$. For each $C_i \in J$, let $C'_i \in CSP_{(\leq k)\text{-XOR}}$ where C'_i is uniform randomly chosen among all ($\leq k$)-XOR constraints over all of the literals appearing in C_i. Let $J' = \{C'_1, \ldots, C'_{m(n,k)}\}$, an instance of $\text{CSP}^{rand,1-\eta'(1-\mathbb{E}[(\leq k)\text{-XOR}])}_{m(n,k)}((\leq k)\text{-XOR})$. It is easy to see that if J is a random k-SAT formula, then J' is a random ($\leq k$)-XOR formula over the same variable set, since the subset of literals appearing in each constraint is chosen uniformly at random over the literals of a uniformly drawn k-OR constraint.

Now, suppose instead that J is $(1 - \eta(1 - \mathbb{E}[k\text{-SAT}]))$-satisfiable. Let $x^* \in \{\pm 1\}$ be an assignment such that x^* satisfies, without loss of generality, $C_1, ..., C_{m'}$, where $m' = \lceil (1 - \eta(1 - \mathbb{E}[k\text{-SAT}]))\rceil$. Let X_i denote the indicator random variable that is 1 iff the randomly chosen ($\leq k$)-constraint C'_i is satisfied by x^*. Note that the X_i are iid, since the parity is sampled uniform randomly and independently of the other C'_i. Moreover, we can deduce that since $C_i(x^*) = 1 = \sum_{S \subseteq [k]} \widehat{C_i(S)}\chi_S(x^*)$, and that the Fourier coefficient corresponding to the empty set is $1 - \frac{1}{2^{k-1}}$, that the number of non-constant parities which agree with C_i on x^* (and so the number of possible choices of ($\leq k$)-constraints for C'_i which are satisfied by x^*) is greater by one than the number of non-constant parities that don't. Hence, $\mathbb{E}[X_i] = \frac{2^{k-1}}{2^k - 1} = \frac{1}{2} + \frac{1}{2^{k+1}-2}$.

So, we can bound the probability that the produced ($\leq k$)-XOR formula is not $1 - \frac{\eta'}{2}$ satisfiable via a Chernoff bound for some choice of η':

$$\Pr\left[\sum_{i=1}^{m'} X_i \leq \left(1 - \eta'\frac{1}{2}\right) m(n,k)\right] =$$

$$\Pr\left[\sum_{i=1}^{m'} X_i \leq (1-\delta)\left(\frac{1}{2} + \frac{1}{2^{k+1}-2}\right) m'\right] < \exp\left(-\frac{\delta^2\left(\frac{1}{2} + \frac{1}{2^{k+1}-2}\right) m'}{2}\right)$$

Now, take $0 < \delta < \frac{1}{2}\left(\frac{\eta - 1}{\eta - 2^k}\right)$, and we see that taking $\eta' = \frac{-\eta\delta + \eta + \delta 2^k + 2^k - 2}{2^k - 1}$ $(< \frac{\frac{1+\eta}{2} + 2^k - 2}{2^k - 1} < 1)$ is enough. And η' only depends on k and η. Hence, for fixed k and η, since $m(n,k) = \omega(1)$ we have for large enough n that J' is $(1 - \eta'\frac{1}{2})$-satisfiable with probability $\geq \frac{3}{4}$. □

Again, as this variant of random XOR refutation is new, it does not immediately strengthen any of the existing lower bounds based on strong refutation of random k-XOR.

Additional evidence of hardness of XOR for large k. We note an additional source of evidence that k-XOR (and consequently k-SAT and $(\leq k)$-XOR) strong refutation is hard. The reduction given by Alekhnovich (2011) for 3-sparse linear systems generalizes to arbitrary k as well as to other distributions, such as our binomially distributed systems. The problem considered by Alekhnovich is to distinguish ϵ-noisy 3-sparse linear systems from linear systems with the same distribution on constraints and a uniform random right-hand side, and he shows how to reduce this problem to strongly refuting 3-XOR. As it is conjectured that this variant of the problem is hard, it provides additional evidence for the hardness of such XOR refutation problems.

A small generalization. We can also observe that we can construct simpler (constraint-wise) refutation reductions from large classes of CSPs to CSPs other than k-SAT, which preserve solutions for families of constraints. Let $\mathrm{enc}_{P,P'} : \mathrm{CSP}_P \to \mathrm{CSP}_{P'}$, where $\mathrm{enc}(P(\ell_1, \ldots, \ell_k)) = P'(\ell_1, \ldots, \ell_k)$. Then it is easy to see that if the support of P is contained in the support of P', then $\mathrm{enc}_{P,P'}$ induces an efficient reduction from $\mathrm{CSP}_m^{rand,1-\eta}(P)$ to $\mathrm{CSP}_m^{rand,1-\eta}(P')$, sending instance $\{C_1, \ldots, C_m\}$ to $\{\mathrm{enc}_{P,P'}(C_1), \ldots, \mathrm{enc}P, P'(C_m)\}$.

This observation allows us to strengthen any hardness result based on hardness of refuting random CSP instances for a particular choice of predicate. Examples of this, as previously mentioned, include lower bounds on the hardness of agnostically learning halfspaces (Daniely, 2016):

Corollary 11. *If there exists* any *k-predicate P whose support consists only of assignments with an odd parity and any constants $c > 0$, $\frac{1}{2} > \eta > 0$ such that $\mathrm{CSP}_{n^{c \log(k)\sqrt{k}}}^{rand,1-\eta}(P)$ is hard, then it is hard to agnostically PAC-learn halfspaces even with a constant approximation ratio.*

Chapter 4

Hardness of Abductively Learning Conjunctions

In this chapter, we use the previously introduced infrastructure of CSP-completeness to give evidence for the hardness of abductively (even improperly) learning conjunctions.

4.1 Additional Preliminaries

To accomplish this task we will introduce some new concepts and associated notation in this section to help us describe the learning model and the reduction from the CSP refutation to this new learning model.

4.1.1 Abductive Learning Problem

First, we describe the one-sided learning model in which we are studying the learnability of conjunctions, as described by Juba (2016). (We relate this model to positive-reliable learning in the appendix.)

Definition 12. *For a class $\mathcal{H}$ of Boolean formulas over Boolean attributes $x_1, \ldots, x_n$, the* abduction *task is as follows. We are given as input m independent examples $x^{(1)}, \ldots, x^{(m)}$ from an arbitrary distribution D over $\{0,1\}^n$ (assignments to the n attributes), a* query *formula $c(x)$ over $x_1, \ldots, x_n$, and an* alphabet *$A \subseteq \{x_1, \ldots, x_n\}$, for which there exists*

$h^* \in \mathcal{H}$ only using attributes in A such that $\Pr[c(x) = 1 | h^*(x) = 1] = 1$ *and* $\Pr[h^*(x) = 1] \geq \mu$*. Then, with probability* $1 - \delta$*, in time polynomial in* $n, 1/\mu, 1/\epsilon$*, and* $1/\delta$*, we find an* explanation $h \in \mathcal{H}$ *only using attributes in* A *such that*

(i) $\Pr[c(x) = 1 | h(x) = 1] \geq 1 - \epsilon$ *and*
(ii) $\Pr[h(x) = 1] \geq 1/p(1/\mu, n, 1/(1 - \epsilon))$ *for some positive polynomial* p*.*

So, in the case of there being a good ("μ-plausible") explanation for the sample data (an $h \in \mathcal{H}$ with no error on its support), an efficient abductive learner in this model will probably output an approximately correct hypothesis (on its support) with the plausibility (size of the support) depending only polynomially on $n, \frac{1}{\mu}, \frac{1}{\epsilon}, \frac{1}{\delta}$.

4.1.2 Scattering and Explainability

Analogous to distinguishing between random and satisfiable instances of CSP problems, our reduction will make use of the following problem of distinguishing between scattered and explainable samples:

Definition 13 ($(\mathcal{H}, \mu)$-explainable and scattered samples)**.** *Let* $S = \{(x_1, y_1), ..., (x_{m(n)}, y_{m(n)})\} \subseteq \{0,1\}^n \times \{0,1\}$ *be a labeled sample.*

- *We say that* S *is* $(\mathcal{H}, \mu)$-explainable *for* $\mu > 0$ *if there exists* $h^* \in \mathcal{H}$ *such that*

  (i) $\frac{1}{m(n)} \sum_{i=1}^{m(n)} \mathbb{1}_{\{x | h^*(x) = 1\}}(x_i) \geq \mu$

  (ii) For each $i \in [m(n)]$, $h^*(x_i) = 1 \implies y_i = 1$

- *We say that a distribution over* $(\{0,1\}^n \times \{0,1\})^m$ *is* scattered *if for* $S \sim D$ *the examples* (x_i, y_i) *are independent and identically distributed, and the* y_i *in particular are Bernoulli*$(\frac{1}{2})$ *random variables that are independent of* x_i*.*

Definition 14 (Distinguishing Explainable From Scattered Samples)**.** *For hypothesis class* $\mathcal{H}$*, we say that a random algorithm* $\mathcal{A}$ *solves* the problem of distinguishing between $(\mathcal{H}, \mu)$-explainable samples and scattered samples of size $m(n)$ *if on input* $S = \{(x_1, y_1), ..., (x_{m(n)}, y_{m(n)})\} \subseteq \{0,1\}^n \times \{0,1\}$*,* $\mathcal{A}$ *has the following two behaviors:*

(i) if S is $(\mathcal{H}, \mu)$-explainable, then $A(S)$ outputs "explainable" *with probability at least $\frac{3}{4}$ with respect to the internal randomness of A.*

(ii) If S is drawn from a scattered distribution, then with probability $1 - o_n(1)$ with respect to the choice of S, $A(S)$ will output "scattered" *with probability at least $\frac{3}{4}$ with respect to the internal randomness of A.*

If there is no such $\mathcal{A}$ that runs in time poly(n)*, then we say that this problem is* hard. *Otherwise, we say that this problem is* easy.

In particular, our reduction will relate the problem of distinguishing $(\mathcal{H}, \mu)$-explainable from scattered examples to the abductive learnability of $\mathcal{H}$. The basic idea behind this relationship is similar to that of Daniely et al., in that since any efficient abductive learner can use at most a polynomial number of bits to describe an output hypothesis, it will almost certainly output a poorly performing explanation in the scattered case. Meanwhile, in the explainable case, it will do well by assumption. We will give a full proof in the next section.

4.2 Main Result: One-Sided Improper Learning of Conjunctions Refutes All Non-Trivial CSPs

Our main result concludes that under Assumption 2, abductively learning conjunctions is hard, even improperly:

Theorem 15 (Hardness of Improperly Abductively Learning Conjunctions)**.** *If assumption 2 holds, then there exists no random algorithm that efficiently abduces conjunctions, even improperly.*

Given the results above, an obvious place one might start would be trying to reduce $\mathrm{CSP}^{rand,1}_{n^{f(k)}}(k\text{-SAT})$ to abductively learning conjunctions, and this is what we will do. But to emphasise what little structure is required out of the predicate to make our reduction work, we will give the reduction from $\mathrm{CSP}^{rand,1}_{n^{f(k)}}(P)$ for any P that is uniformly falsifiable. That is, at least one of $P(0, 0, \ldots, 0) = 0$ or $P(1, 1, \ldots, 1) = 0$ holds.

Hence, we will show the following result:

Theorem 16. *If there exists some uniformly falsifiable k-predicate P for which* $\mathrm{CSP}^{rand,1}_{n^{f(k)}}(P)$ *is hard, then it is hard to improperly abduce conjunctions efficiently.*

And theorem 15 follows from the random refutation completeness of $\mathrm{CSP}^{rand,1}_{n^{f(k)}}(k\text{-SAT})$ by theorem 6.

The method behind our result is based on that of Daniely et al. Namely, we will reduce the (hard) problem of distinguishing between satisfiable and random CSP instances to the problem of distinguishing between conjunctively explainable and scattered samples. The idea will be to label the set of input constraints negatively, then to randomly (with probability $\frac{1}{2}$) replace constraints with (uniformly) random positively labeled constraints. We will then encode the constraints (as a collection of literals) into a larger set of Boolean attributes.

In the case that the input system is satisfiable, due to our mild assumption that P is false on either the all 0 or all 1 input, any satisfying assignment induces a mapping $h^* : \{0,1\}^{2n} \to \{0,1\}$ (computable by a conjunction) from the encoded constraints to $\{0,1\}$ such that for P-constraint C and encoding function $\mathrm{enc} : \mathrm{CSP}_P \to \{0,1\}^{2n}$ we have $h^*(\mathrm{enc}(C)) = 1$ only if $C(x^*) = 0$, i.e., there exists a conjunctive explanation for the encoded input system of constraints. In the random case, on the other hand, there is almost always no such explanation h^*. Hence, any efficient abductive learner for conjunctions will be able to solve our original CSP problem.

Note that for brevity the following analysis will be for the case that $P(0,0,\ldots,0) = 0$, but if we do not have falsification on all 0's and instead have falsification of P on all 1's ($P(1,1,\ldots,1) = 0$), that the analysis is much the same. Furthermore, we note that our argument extends directly to all constant size alphabets Γ given a suitable notion of "literals:" we only require that for any pair of symbols $\sigma, \tau \in \Gamma$, there is a literal function that takes σ to τ (the literals are "1-transitive"). In particular, the constant shift literals used by Georgiou et al. (2009) will suffice.

4.2.1 The Reduction

First, we will describe our encoding of CSP_P over n variables as elements of $\{0,1\}^{2n}$. Let $\mathrm{enc} : \mathrm{CSP}_P \to \{0,1\}^{2n}$, where for $C = P(\ell_{i_1}, \ldots \ell_{i_k})$, $\mathrm{enc}(C) = z$ such that when we identify $\{0,1\}^{2n}$ with $\{0,1\}^{n\times[2]}$, $z_{(i,1)} = 1 \iff$ either x_i does not appear in any literal ℓ_q of C or x_i appears as a literal itself in C, and similarly $z_{(i,2)} = 1 \iff$ either x_i does not appear in any literal ℓ_q of C or $\neg x_i$ appears as a literal itself in C. Hence, every index of $\mathrm{enc}(C)$ is 1 unless the negation of the associated literal appears in the constraint C.

We now describe the algorithm itself. Let $\mathcal{H}_{\mathrm{con}}$ denote the hypothesis class of conjunctions over n variables and let P denote a k-predicate over n variables. Let $\mathcal{A}$ denote an algorithm that efficiently distinguishes between $(\mathcal{H}_{\mathrm{con}}, \mu)$-explainable (with respect to $\mu = \mu(n,k)$) and scattered samples of size n^d. Then $\mathcal{A}'$, given as Algorithm 1 below, is a polynomial time algorithm for the problem of distinguishing between satisfiable and random instances of P-constraints of size n^d, by reducing to the problem of distinguishing between $(\mathcal{H}_{\mathrm{con}}, \mu)$-explainable and scattered samples of size n^d over $\{0,1\}^{2n}$.

Algorithm 1: $\mathcal{A}'$, Reduction of the CSP distinguishing problem to the distinguishing problem for scattered versus explainable samples.

Input : $S = \{C_1, \ldots, C_{n^d}\} \subseteq \mathrm{CSP}_P$
Output: Either "satisfiable" or "random"

Let $S' = \{(C_1, 0), \ldots, (C_{n^d}, 0)\}$ by labeling each input constraint 0
for $i \leftarrow 1$ **to** n^d **do**
  With probability $\frac{1}{2}$ replace, in S', labeled example $(C_i, 0)$ with labeled example $(C, 1)$ for C chosen uniformly at random.
end
Let $E = \{(\mathrm{enc}(C_1), b_1), \ldots, (\mathrm{enc}(C_{n^d}), b_{n^d})\}$, where b_i is the label of the ith example after the randomizing loop.
if $\mathcal{A}(E) =$ "explainable" **then**
  return "satisfiable"
else
  return "random"
end

4.2.2 Correctness of The Reduction

We will now establish the correctness of $\mathcal{A}'$, as shown in Algorithm 1.

Lemma 17. *On input a satisfiable system of constraints for some uniformly falsifiable k-constraint P, the algorithm $\mathcal{A}'$ will output "satisfiable" with probability at least $\frac{3}{4}$. And if S is random, then $\mathcal{A}'$ will return "random" with probability at least $\frac{3}{4}$ for a $1 - o_n(1)$-fraction of choices of S.*

We consider each of the two cases of the problem separately. In the case that S is structured (satisfiable) we will show that with probability at least $\frac{3}{4}$ we output "satisfiable". And, in the unstructured case (S is random), we will show that over a $1-o_n(1)$ fraction of all choices of S with probability at least $\frac{3}{4}$ $\mathcal{A}'$ outputs "random". And then afterwards, we will prove our main theorem, relating the problem to abduction.

Claim 18. *If $S = \{C_1, ..., C_{n^d}\}$ is satisfiable, then E is $(\mathcal{H}_{\text{con}}, \mu)$-explainable with respect to any $\mu < \frac{1}{2^{k+1}} - \sqrt{\frac{2^{k+1}}{n^d} \ln 4}$ with probability at least $\frac{3}{4}$.*

Proof. Let x^* be a satisfying assignment for $\{C_1, ..., C_{n^d}\}$ and consider the following $h^* \in \mathcal{H}_{\text{con}}$: $h^*(z) = \bigwedge_{i=1}^{n} \alpha_i$ where

$$\alpha_i = \begin{cases} z_{(i,1)}, & \text{if } x^*_i = 0 \\ z_{(i,2)}, & \text{if } x^*_i = 1 \end{cases}$$

In words, $h^*(\text{enc}(C)) = 0$ if and only if some literal belonging to C is satisfied by the assignment x^*. Hence, if we have $h^*(\text{enc}(C)) = 1$ on a μ fraction of E, then we are done. Note that this argument only depends on the value of the predicate on the all 0 string. If instead we had considered predicates that are falsified on the all 1 string, then only the construction of h^* changes in this proof (particularly, each $z_{(i,1)}$ appearing in h^* is swapped for a $z_{(i,2)}$ and vice versa).

Let X_i be the indicator random variable that is 1 if and only if $h^*(\text{enc}(C_i)) = 1$. We make two observations. First, we note that if $h^*(\text{enc}(C_i)) = 1$ then the label of the ith example is also 1. Second, we note that $\mathbb{E}[X_i] = (\frac{1}{2})(\frac{1}{2^k})$. And so, by application of the usual multiplicative Chernoff bounds:

$$\Pr\Big[\sum_{i=1}^{n^d} X_i < \mu n^d\Big] < \exp\Big(-\Big(\frac{1}{2}\Big)(\frac{1}{2^{k+1}} - \mu)^2\Big(\frac{1}{2^{k+1}}\Big)n^d\Big)$$

And so, in order to pick μ so that this probability h^* does a poor job at explaining the labeling is less than $\frac{1}{4}$ we need

$$\exp\Big(-\Big(\frac{1}{2}\Big)(\frac{1}{2^{k+1}} - \mu)^2\Big(\frac{1}{2^{k+1}}\Big)n^d\Big) < \frac{1}{4}$$

and continuing with elementary manipulations we find that happens when

$$\mu < \frac{1}{2^{k+1}} - \sqrt{\frac{2^{k+1}}{n^d}\ln 4}$$

and so we are done. □

Thus, our reduction maps satisfiable instances of P-constraints to conjunctively explainable instances over the boolean cube.

Next, observe that in the case of S being drawn uniformly at random, that for almost all choices of S, with probability at least $\frac{3}{4}$ with respect to the internal randomness of the algorithm, we will output "random". Indeed, we observe that the distribution induced by the algorithm on examples over $\{0,1\}^{2n} \times \{0,1\}$ is scattered in the random case, and so, by the assumption that $\mathcal{A}$ solves the problem of distinguishing between $(\mathcal{H}_{\text{con}}, \mu)$-explainable and scattered samples we have this guarantee:

Claim 19. *If S is drawn uniformly at random, then for a $1 - o_n(1)$ fraction of the possible S, with probability at least $\frac{3}{4}$ $\mathcal{A}'$ will return "random."*

Proof. As $\mathcal{A}'$ outputs "random" whenever $\mathcal{A}$ outputs "scattered" on its input E, and we are given that $\mathcal{A}$ outputs "scattered" with probability $\frac{3}{4}$ for a $1 - o_n(1)$-fraction of the possible samples produced by a scattered distribution, it suffices to argue that E is indeed scattered.

Indeed, since $\{C_1, \ldots, C_{n^d}\}$ is by assumption a collection of mutually independent and uniformly random constraints, in particular the set of literals appearing in each C_i is an independent and uniformly random set of k literals on distinct variables. Observe that this remains true even if we sample a new, independent constraint for C_i in the first loop of $\mathcal{A}'$. Hence, $\{\text{enc}(C_1), \ldots, \text{enc}(C_{n^d})\}$ is, by construction, a collection of mutually independent and identically distributed random variables. Moreover, our choice of whether or not to resample C_i and replace its label with 1 is an independent Bernoulli trial, and as the new constraint C_i was (once again) independently and uniformly sampled, we find that indeed the labels y_i are independent and unbiased Bernoulli random variables. Thus we see that E is a scattered sample as claimed. □

Thus, our reduction maps random CSP instances to scattered samples, and (by Claim 18) mapped satisfiable CSP instances to "explainable" samples. Since by hypothesis $\mathcal{A}$ is able to distinguish scattered from "explainable" samples, our reduction is correct.

4.2.3 Proof of Theorem 16

Proof. Toward a contradiction let $\mathcal{L}$ be an efficient abductive learner of $\mathcal{H}_{\text{con}}$. We take $g(n, \mu, \epsilon, \delta) \geq \Omega(\left((1-\epsilon)\frac{\mu}{n}\right)^d)$ (for some constant d) to be the lower bound of the plausibility of $\mathcal{L}$'s output explanation guaranteed by definition of $\mathcal{L}$ being an efficient abductive learner when there is a μ-plausible explanation.

Hence, for each fixed choice of $\delta, \epsilon \in (0, 1)$ there is some $d > 0$ such that (i) $\mathcal{L}$ reads and writes fewer than $(\frac{n}{\mu})^d$ bits over its execution, including those used to read the input examples and the bits required to describe an output explanation. In particular, the number of examples read and used by the algorithm is at most $(\frac{n}{\mu})^d$, and (ii) $(\frac{n}{\mu})^d \geq \frac{1}{g(n,\mu,\epsilon,\delta)}$.

Let $q = d+1$ and let k be large enough such that $f(k) \geq 3q$. Recall that for our reduction to distinguish between satisfiable and random samples our analysis requires that our subroutine $\mathcal{A}$ be able to efficiently distinguish between $(\mathcal{H}_{\text{con}}, \mu)$-explainable and scattered samples for some $\frac{1}{2^{k+2}} \leq \mu < \frac{1}{2^{k+1}}$. Notice that for sufficiently large n, i.e., $n > \frac{1}{\mu^d}$, we have $n^{d+1} > \left(\frac{n}{\mu}\right)^d$, that is, $n^q \geq (\frac{n}{\mu})^d$.

Now consider the algorithm $\mathcal{L}'$, given as Algorithm 2 below, that on input $S \subseteq \{0,1\}^n \times \{0,1\}$ of size n^{3q}, we claim distinguishes between the case that S is $(\mathcal{H}_{\text{con}}, \mu)$-explainable for such a μ and the case that S is scattered.

Algorithm 2: $\mathcal{L}'$, reduction of distinguishing scattered from explainable samples to abductive learning

Input : $S = \{(x_1, y_1), \ldots, (x_{n^{3q}}, y_{n^{3q}})\}$
Output: Either "explainable" or "random"

1 Run $\mathcal{L}$ with the following parameters: Examples drawn uniformly (with replacement) from S as the input example set, μ as above, $\delta = \frac{1}{8}$, and $\epsilon = \frac{1}{4}$ and let h be the output explanation hypothesis.
2 **if** If $\frac{1}{n^{3q}} \sum_{i=1}^{n^{3q}} \mathbb{1}_{\{(x,y) \in S | h(x)=1\}}(x_i) \geq g(n, \mu, \epsilon, \delta)$ and $\text{Err}_{\{(x,y) \in S | h(x)=1\}}(h) < \frac{1}{4}$ **then**
3 | **return** "explainable"
4 **else**
5 | **return** "random"
6 **end**

Where we define

$$\text{Err}_A(h) = \frac{1}{|A|} \sum_{(x,y) \in A} \mathbb{1}_{\{(x,y) \in A | y=0\}}(x, y)$$

i.e., the fraction of false positives h produces over S.

Suppose that S is $(\mathcal{H}_{\text{con}}, \mu)$-explainable, then by assumption of $\mathcal{L}$ being an abductive learner, for large enough n, with probability at least $1 - \delta > \frac{3}{4}$ it will return an h satisfying the condition in line 2 of the algorithm (since $\mathcal{L}$ must work with respect to the uniform distribution over examples we are using) and $\mathcal{L}'$ will return "explainable".

Now suppose that instead S is drawn from a scattered distribution. We bound the probability that h does well (that it satisfies the conditions of line 2). Let us fix an arbitrary h that may be output by $\mathcal{L}$. Note that when h passes the first condition, on the number of positive classifications, since we have chosen q so that $g(n, \mu, \epsilon, \delta) \geq n^{-q}$, h must be positive on at least n^{2q} examples from S. Since S is scattered, for this fixed h, half of the examples from S it classifies positively are false-positives in expectation. Note that since the labels y_i are independent of the examples x_i in a scattered distribution, conditioning on h passing the first check on the number of positive classifications leaves the distribution on labels uniform. Therefore, by a Chernoff bound, the probability that fewer than $1/4$ of the at least n^{2q}

examples h classifies positively are false positive errors is at most $e^{-n^{2q}/8}$. And since there are at most 2^{n^q} possible output hypotheses of $\mathcal{L}$, by applying the union bound over these the probability that the output of $\mathcal{L}$ has error lower than $\frac{1}{4}$ is at most $2^{n^q}e^{-n^{2q}/8} = e^{-n^q(n^q/8-\ln 2)}$. We conclude that with probability $1 - o_n(1)$ over choices of S, $\mathcal{L}'$ will return "random" with probability $7/8$. Hence, overall, for sufficiently large n $\mathcal{L}'$ returns "random" with probability greater than $3/4$ as needed.

And thus, if there were an efficient abductive learner $\mathcal{L}$ for $\mathcal{H}_{\text{con}}$ we would contradict the assumption on P as follows. We first take $\mathcal{L}'$ above as the $\mathcal{A}$ subroutine of $\mathcal{A}'$, which efficiently solves the problem of distinguishing between $(\mathcal{H}_{\text{con}}, \mu)$-explainable and scattered samples when we have such an efficient abductive learner $\mathcal{L}$, for all large enough k, and then use our algorithm $\mathcal{A}'$ to solve the problem of distinguishing between satisfiable and random P-instances for some choice of k, and $f(k)$. This contradicts the assumption that it is hard to solve $\text{CSP}^{rand,1}_{n^{f(k)}}(P)$. □

Chapter 5

Conclusion

5.1 Directions for future work

With regard to the k-CSP reduction framework developed above, we see five natural directions for future investigation. The first, obvious direction is to try to find additional examples of k-CSP hard average-case problems. As a starting point, it would be interesting to know if our result for $(\leq k)$-XOR can be strengthened to k-XOR, i.e., with all constraints the same size. Given that on the algorithms side, the "XOR principle" (Feige, 2002; Feige and Ofek, 2007; Feige et al., 2006; Allen et al., 2015) asserts that refuting k-XOR is sufficient to refute all CSPs, we should expect that k-XOR is complete. This might help provide additional evidence for the hardness of approximate agnostic learning of halfspaces, bolstering results of Daniely (2016) for example. Actually, noting that Feige (2002) showed that many natural 3-CSPs are complete, it is natural to ask if the same is true of the generalization to $k > 3$. Finally, as this notion of CSP completeness gives us a new way to study the complexity of refutation problems with unusual predicates, one interesting question in this direction (suggested to us by M. Tulsiani) is whether or not the linearity testing predicate used by Samorodnitsky and Trevisan (2000) can be shown to be complete for k-CSPs. Currently there is no evidence that such refutation problems are hard, but also there are no algorithms.

The second, related direction is to see if our new variant of random XOR refutation is useful for establishing hardness. Even if we cannot manage to establish that the standard k-XOR problem is k-CSP complete, we feel that this new variant is a reasonably natural problem, and it may be that reductions that had been based on k-XOR can be adapted to use this

problem instead. If so, then the problems based on the hardness of refuting k-XOR (such as approximate agnostic learning of halfspaces with constant approximation ratio) could be shown to be hard under the weaker assumption that some predicate is hard to strongly refute.

A third direction would be to tighten our bounds on m in the completeness analysis of $(\leq k)$-XOR to something like that in the k-SAT case, expressed in terms of $\mathbb{E}[(\leq k)\text{-XOR}]$ and $\mathbb{E}[k\text{-SAT}]$.

The fourth direction is to better understand the predicates on k variables for large k. Applebaum and Lovett (2018), for example, considered a slightly unusual combination predicate for cryptographic applications. We believe there is more scope for considering such unusual predicates, which would then translate into new understanding of the usual problems such as random k-SAT. For example, Impagliazzo and Levin (1990) considered somewhat unusual problems in which a predicate is evaluated on the preimage of a hash function. We could define a predicate along these lines, and perhaps it would allow us to connect to problems such as average-case NP-complete problems.

The final direction we propose here for making use of the CSP-refutation reduction framework is that of providing evidence for certain functions being pseudorandom generators. In particular, Goldreich (2011) provides a candidate pseudorandom generator based on the output of a system of CSP consraints. Loosely, Goldreich's proposed algorithm is a pseudorandom generator if it is hard to distinguish between random and satisfiable systems of constraints with respect to the underlying predicate(s). We believe there at least two opportunities here. The first is to explore whether popular choices for the input predicate (for example, the $\text{XORAND}_{t,k-t}$ of ODonnell and Witmer (2014)) are complete for k-CSPs. And the second is to modify Goldreich's candidate function slightly to allow the problem of distinguishing for our XOR variant to reduce to that of distinguishing between the generated and random strings.

And on the topic of this framework as applied to one-sided learning theory, we have positive results for the agnostic variant of this model: we have a $\tilde{O}(\sqrt{n})$-approximation to the optimal disjunction (Zhang et al., 2017) (based on an earlier algorithm by Peleg (2007)) and a simple $\tilde{O}(s \log\log n)$ approximation to the optimal size-s disjunction (Juba et al., 2018). We would like to know how close to optimal either of these algorithms are. We note that Daniely (2016)

has succeeded at using similar techniques (but based on strong assumptions about refuting random XOR) to show $\omega(1)$ lower bounds for the blow-up of agnostically learning halfspaces in the standard improper supervised learning model.[3] Noting the relative simplicity of the reductions for the one-sided error models we consider, we are optimistic that it might be possible to extend Daniely's techniques to analyze the blow-up needed for one-sided learning of disjunctions.

[3]Daniely also obtains a $2^{\log^{1-\epsilon} n}$ lower bound for very strong assumptions – for polylogarithmic k, Daniely requires refuting random k-XOR to remain hard with up to $n^{O(k)}$ constraints.

References

Akavia, A., Goldreich, O., Goldwasser, S., and Moshkovitz, D. (2006). On basing one-way functions on NP-hardness. In *Proc. 38th STOC*, pages 701–710.

Alekhnovich, M. (2011). More on average case vs approximation complexity. *Computational Complexity*, 20(4):755–786.

Allen, S. R., O'Donnell, R., and Witmer, D. (2015). How to refute a random CSP. In *Proc. 56th FOCS*, pages 689–708.

Alon, N., Krivelevich, M., and Sudakov, B. (1998). Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13(3-4):457–466.

Applebaum, B., Barak, B., and Wigderson, A. (2010). Public-key cryptography from different assumptions. In *Proc. 42nd STOC*, pages 171–180.

Applebaum, B., Barak, B., and Xiao, D. (2008). On basing lower-bounds for learning on worst-case assumptions. In *Proc. 49th FOCS*, pages 211–220.

Applebaum, B. and Lovett, S. (2018). Algebraic attacks against random local functions and their countermeasures. *SIAM J. Comput.*, 47(1):52–79.

Awasthi, P., Blum, A., and Sheffet, O. (2010). Improved guarantees for agnostic learning of disjunctions. In *Proc. 23rd COLT*, pages 359–367.

Barak, B., Hopkins, S., Kelner, J., Kothari, P., Moitra, A., and Potechin, A. (2016). A nearly tight sum-of-squares lower bound for the planted clique problem. In *Proc. 57th FOCS*, pages 428–437.

Barak, B., Kindler, G., and Steurer, D. (2013). On the optimality of semidefinite relaxations for average-case and generalized constraint satisfaction. In *Proc. 4th ITCS*, pages 197–214.

Ben-Sasson, E. and Wigderson, A. (2001). Short proofs are narrow – resolution made simple. *J. ACM*, 48(2):149–169.

Benabbas, S., Georgiou, K., Magen, A., and Tulsiani, M. (2012). SDP gaps from pairwise independence. *Theory of Computing*, 8(12):269–289.

Berthet, Q. and Rigollet, P. (2013). Complexity theoretic lower bounds for sparse principal component detection. *J. Mach. Learn. Res. (COLT)*, 30. Extended version, Conference On Learning Theory. arXiv:1304.0828.

Bogdanov, A. and Trevisan, L. (2006). On worst-case to average-case reductions for NP problems. *SIAM Journal on Computing*, 36(4):1119–1159.

Brakerski, Z., Langlois, A., Peikert, C., Regev, O., and Stehlé, D. (2013). Classical hardness of learning with errors. In *Proc. 45th STOC*, pages 575–584.

Braverman, M., Ko, Y. K., Rubinstein, A., and Weinstein, O. (2015). Eth hardness for densest-k-subgraph with perfect completeness. *arXiv preprint arXiv:1504.08352*.

Brennan, M., Bresler, G., and Huleihel, W. (2018). Reducibility and computational lower bounds for problems with planted sparse structure. In *Proc. 31st COLT*, volume 75 of *PMLR*, pages 48–166. JMLR.

Bshouty, N. H. and Burroughs, L. (2005). Maximizing agreements with one-sided error with applications to heuristic learning. *Machine Learning*, 59(1-2):99–123.

Chan, S. O. (2016). Approximation resistance from pairwise-independent subgroups. *J. ACM*, 63(3):27:1–27:32.

Coja-Oghlan, A., Cooper, C., and Frieze, A. (2010). An efficient sparse regularity concept. *SIAM Journal on Discrete Mathematics*, 23(4):2000–2034.

Daniely, A. (2016). Complexity theoretic limitations on learning halfspaces. In *Proc. 48th STOC*, pages 105–117.

Daniely, A., Linial, N., and Shalev-Shwartz, S. (2013). More data speeds up training time in learning halfspaces over sparse vectors. In *Advances in Neural Information Processing Systems*, pages 145–153.

Daniely, A., Linial, N., and Shalev-Shwartz, S. (2014). From average case complexity to improper learning complexity. In *Proc. 46th STOC*, pages 441–448.

Daniely, A. and Shalev-Shwartz, S. (2016). Complexity theoretic limtations on learning DNF's. In *Proc. 29th COLT*, volume 49 of *JMLR Workshops and Conference Proceedings*, pages 1–16. JMLR.

Ding, J., Sly, A., and Sun, N. (2015). Proof of the satisfiability conjecture for large k. In *Proc. 49th STOC*, pages 59–68. ACM.

Durgin, A. and Juba, B. (2019). Hardness of improper one-sided learning of conjunctions for all uniformly falsifiable CSPs. In *Proc. 30th ALT*, volume 98 of *PMLR*, pages 369–382.

Elkans, C. (2001). The foundations of cost-sensitive learning. In *Proc. IJCAI'01*, pages 973–978.

Feige, U. (2002). Relations between average case complexity and approximation complexity. In *Proc. 34th STOC*, pages 534–543.

Feige, U., Kim, J. H., and Ofek, E. (2006). Witnesses for non-satisfiability of dense random 3CNF formulas. In *Proc. 47th FOCS*, pages 497–508.

Feige, U. and Ofek, E. (2007). Easily refutable subformulas of large random 3CNF formulas. *Theory of Computing*, 3(1):25–43.

Feldman, V., Gopalan, P., Khot, S., and Ponnuswami, A. K. (2009). On agnostic learning of parities, monomials, and halfspaces. *SIAM J. Comput.*, 39(2):606–645. Preliminary versions appeared in 47th FOCS, 2006 and 21st CCC, 2006.

Feldman, V., Grigorescu, E., Reyzin, L., Vempala, S., and Xiao, Y. (2013). Statistical algorithms and a lower bound for detecting planted cliques. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 655–664. ACM. Full version: arXiv:1201.1214v6.

Feldman, V., Perkins, W., and Vempala, S. (2018). On the complexity of random satisfiability problems with planted solutions. *SIAM Journal on Computing*, 47(4):1294–1338.

Feller, W. (1957). *An Introduction to Probability Theory and Its Applications vol. 1.* John Wiley and Sons, Inc., Wiley.

Friedgut, E. and Bourgain, J. (1999). Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American Mathematical Society*, 12(4):1017–1054.

Georgiou, K., Magen, A., and Tulsiani, M. (2009). Optimal Sherali-Adams gaps from pairwise independence. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 5687 of *LNCS*, pages 125–139. Springer.

Goldreich, O. (2011). *Candidate One-Way Functions Based on Expander Graphs*, page 76–87. Springer-Verlag, Berlin, Heidelberg.

Goldwasser, S. and Kalai, Y. T. (2016). Cryptographic assumptions: A position paper. In *Theory of Cryptography Conference*, pages 505–522. Springer.

Haitner, I., Mahmoody, M., and Xiao, D. (2010). A new sampling protocol and applications to basing cryptographic primitives on the hardness of NP. In *Proc. 25th CCC*, pages 76–87.

Håstad, J. (1996). Clique is hard to approximate within $n^{1-\epsilon}$. In *Proc. 37th FOCS*, pages 627–636.

Impagliazzo, R. and Levin, L. A. (1990). No better ways to generate hard NP instances than picking uniformly at random. In *Proc. 31st FOCS*, pages 812–821.

Jerrum, M. (1992). Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359.

Juba, B. (2016). Learning abductive reasoning using random examples. In *Proc. 30th AAAI*, pages 999–1007.

Juba, B. (2017). Conditional sparse linear regression. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 67. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. Proc. 8th ITCS.

Juba, B., Li, Z., and Miller, E. (2018). Learning abduction under partial observability. In *Proc. 32nd AAAI*.

Kalai, A. T., Kanade, V., and Mansour, Y. (2012). Reliable agnostic learning. *Journal of Computer and System Sciences*, 78(5):1481–1495.

Kanade, V. and Thaler, J. (2014). Distribution-independent reliable learning. In *Proc. 27th COLT*, volume 35 of *JMLR Workshops and Conference Proceedings*, pages 3–24.

Kearns, M. and Valiant, L. (1994). Cryptographic limitations on learning boolean formulae and finite automata. *Journal of the ACM (JACM)*, 41(1):67–95.

Kearns, M. J., Schapire, R. E., and Sellie, L. M. (1994). Towards efficient agnostic learning. *Machine Learning*, 17(2-3):115–141.

Khot, S. (2002). On the power of unique 2-prover 1-round games. In *Proc. 34th STOC*, pages 767–775.

Klivans, A. R. and Servedio, R. A. (2004). Learning DNF in time $2^{O(n^{1/3})}$. *JCSS*, 68(2):303–318.

Klivans, A. R. and Sherstov, A. A. (2009). Cryptographic hardness for learning intersections of halfspaces. *JCSS*, 75(1):2–12.

Kothari, P. K., Mori, R., O'Donnell, R., and Witmer, D. (2017). Sum of squares lower bounds for refuting any CSP. In *Proc. 49th STOC*, pages 132–145.

Kučera, L. (1995). Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2):193–212.

Lee, J. R., Raghavendra, P., and Steurer, D. (2015). Lower bounds on the size of semidefinite programming relaxations. In *Proc. 47th STOC*, pages 567–576.

Levin, L. A. (1986). Average case complete problems. *SIAM Journal on Computing*, 15(1):285–286.

Livne, N. (2010). All natural NP-complete problems have average-case complete versions. *Computational Complexity*, 19(4):477–499.

McCarthy, J. and Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, Edinburgh. Available at http://www-formal.stanford.edu/jmc/mcchay69.html.

Moshkovitz, D. and Raz, R. (2010). Two-query PCP with subconstant error. *J. ACM*, 57(5):29.

ODonnell, R. and Witmer, D. (2014). Goldreich's prg: Evidence for near-optimal polynomial stretch. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 1–12.

Papadimitriou, C. H. (1994). On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Sciences*, 48(3):498–532.

Pass, R., Tseng, W.-L. D., and Venkitasubramaniam, M. (2011). Towards non-black-box lower bounds in cryptography. In *Theory of Cryptography Conference*, pages 579–596. Springer.

Peleg, D. (2007). Approximation algorithms for the label-covermax and red-blue set cover problems. *J. Discrete Algorithms*, 5:55–64.

Pitt, L. and Valiant, L. G. (1988). Computational limitations on learning from examples. *J. ACM*, 35(4):965–984.

Raghavendra, P. (2008). Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th STOC*, pages 245–254.

Raghavendra, P., Rao, S., and Schramm, T. (2017). Strongly refuting random CSPs below the spectral threshold. In *Proc. 49th STOC*, pages 121–131.

Regev, O. (2009). On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34.

Reiter, R. and de Kleer, J. (1987). Foundations for assumption-based truth maintenance systems: Preliminary report. In *Proc. AAAI-87*, pages 183–188.

Roth, D. (1995). Learning to reason: the non-monotonic case. In *Proc. 14th IJCAI*, volume 2, pages 1178–1184.

Samorodnitsky, A. and Trevisan, L. (2000). A pcp characterization of np with optimal amortized query complexity. In *Proc. 32nd STOC*, pages 191–199.

Schoenebeck, G. (2008). Linear level Lasserre lower bounds for certain k-CSPs. In *Proc. 49th FOCS*, pages 593–602.

Tulsiani, M. (2009). CSP gaps and reductions in the Lasserre hierarchy. In *Proc. 41st STOC*, pages 303–312.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, 18(11):1134–1142.

Valiant, L. G. (1994). *Circuits of the Mind*. Oxford University Press, Oxford.

Valiant, L. G. (1995). Rationality. In *Proc. 8th COLT*, pages 3–14.

Zhang, M., Mathew, T., and Juba, B. (2017). An improved algorithm for learning to perform exception-tolerant abduction. In *Proc. 31st AAAI*, pages 1257–1265.

CSP-Completeness And Applications, Durgin, M.S. 2020