Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)

General Bounds on Satisfiability Thresholds for Random CSPs via Fourier Analysis

Colin Wei and **Stefano Ermon**
Computer Science Department
Stanford University
{colinwei,ermon}@cs.stanford.edu

Abstract

Random constraint satisfaction problems (CSPs) have been widely studied both in AI and complexity theory. Empirically and theoretically, many random CSPs have been shown to exhibit a phase transition. As the ratio of constraints to variables passes certain thresholds, they transition from being almost certainly satisfiable to unsatisfiable. The exact location of this threshold has been thoroughly investigated, but only for certain common classes of constraints.

In this paper, we present new bounds for the location of these thresholds in boolean CSPs. Our main contribution is that our bounds are fully general, and apply to any fixed constraint function that could be used to generate an ensemble of random CSPs. These bounds rely on a novel Fourier analysis and can be easily computed from the Fourier spectrum of a constraint function. Our bounds are within a constant factor of the exact threshold location for many well-studied random CSPs. We demonstrate that our bounds can be easily instantiated to obtain thresholds for many constraint functions that had not been previously studied, and evaluate them experimentally.

1 Introduction

Constraint satisfaction problems (CSPs) are widely used in AI, with applications in optimization, control, and planning (Russell et al. 2003). While many classes of CSPs are intractable in the worst case (Cook 1971), many real-world CSP instances are easy to solve in practice (Vardi 2014). As a result, there has been significant interest in understanding the average-case complexity of CSPs from multiple communities, such as AI, theoretical computer science, physics, and combinatorics (Biere, Heule, and van Maaren 2009).

Random CSPs are an important model for studying the average-case complexity of CSPs. Past works have proposed several distributional models for random CSPs (Molloy 2003; Creignou and Daudé 2003). An interesting feature arising from many models is a phase transition phenomenon that occurs as one changes the ratio of number of constraints, m, to the number of variables, n. Empirical results (Mitchell, Selman, and Levesque 1992) show that for many classes of CSPs, randomly generated instances are satisfiable with probability near 1 when m/n is below a certain threshold. For m/n larger than a threshold, the probability of satisfiability is close to 0. The statistical physics, computer science, and mathematics communities have focused much attention on identifying these threshold locations (Achlioptas, Naor, and Peres 2005; Biere, Heule, and van Maaren 2009).

Phase transitions for common classes of CSPs such as k-SAT and k-XORSAT are very well-studied. For k-SAT, researchers struggled to find tight lower bounds on the satisfiability threshold until the breakthrough work of Achlioptas and Moore, which provided lower bounds a constant factor away from the upper bounds. Later works closed this gap for k-SAT (Coja-Oghlan and Panagiotou 2013; 2016). More recently, Dudek, Meel, and Vardi (2016) also studied the satisfiability threshold for a more general CSP class, namely k-CNF-XOR, where both k-SAT and k-XORSAT constraints can be used. The results and analyses from these works, however, are all specific to the constraint classes studied.

In this paper, we provide new lower bounds on the location of the satisfiability threshold that hold for general boolean CSP classes. We focus on the setting where CSPs are generated by a single constraint type, though our analysis can extend to the setting with uniform mixtures of different constraint functions. We extend techniques from (Achlioptas and Peres 2004) and build on (Creignou and Daudé 2003), which proposes a distributional model for generating CSPs and provides lower bounds on the satisfiability threshold for these models. The significance of our work is that *our bounds hold for all functions that could be used to generate random CSP instances*. The lower bounds from (Creignou and Daudé 2003) are also broadly applicable, but they are looser than ours because they do not depend on constraint-specific properties. Our lower bounds are often tight (within a constant factor of upper bounds for many CSP classes) because they depend on specific properties of the Fourier spectrum of the function used to generate the random CSPs. Since these properties are simple to compute for any constraint function, our lower bounds are broadly applicable too.

The Fourier analysis of boolean functions (O'Donnell 2014) will be vital for obtaining our main results. Expressing functions in the Fourier basis allows for clean analyses of random constraints (Friedgut and Bourgain 1999; Barak et al. 2015; Achim, Sabharwal, and Ermon 2016). Our use of Fourier analysis is inspired by the work of Achim, Sabharwal, and Ermon (2016), who analyze the Fourier spectra of random hash functions used as constraints in CSP-based model counting. We show that the Fourier spectrum of our constraint-generating function controls the level of spatial correlation in the set of satisfying assignments to the random CSP. If the Fourier spectrum is concentrated on first and second order coefficients (corresponding to "low frequencies"), this correlation will be very high, roughly increasing the variance of the number of solutions to a random CSP and decreasing the probability of satisfiability. In related work, Montanari, Restrepo, and Tetali (2011) also use Fourier analysis to provide tight thresholds in the case where odd Fourier coefficients are all zero.

2 Notation and Preliminaries

In this section, we will introduce the preliminaries necessary for presenting our main theorem. First, we formally define our distribution for generating random CSPs, inspired from (Creignou and Daudé 2003).

We will use n and m to denote the number of variables and number of constraints in our CSPs, respectively. We will also let $f : \{-1,1\}^k \to \{0,1\}$ denote a binary function and refer to f as our constraint function. Often we use the term "solution set of f" to refer to the set $\{u : f(u) = 1\}$. Using the constraint function f, we create constraints by applying f to a signed subset of k variables.

Definition 1 (Constraint)**.** *Let $I = (i_1, \ldots, i_k)$ be an ordered tuple of k indices in $[n]$, and let s be a sign vector from $\{-1,1\}^k$. Given a vector $\sigma \in \{-1,1\}^n$, we will define the vector $\sigma_{I,s}$ of size k as follows:*

$$\sigma_{I,s} = (s_1\sigma_{i_1}, \ldots, s_k\sigma_{i_k}) \tag{1}$$

Now we can denote the application of f to these indices by $f_{I,s}(\sigma) = f(\sigma_{I,s})$. We call $f_{I,s}$ a constraint, and we say that $\sigma \in \{-1,1\}^n$ satisfies the constraint $f_{I,s}$ if $f_{I,s}(\sigma) = 1$.

The definition of a CSP generated from f follows.

Definition 2 (CSP generated from f)**.** *We will represent a CSP with m constraints and n variables generated from f as a collection of constraints $C_f(n,m) = \{f_{I_1,s_1}, \ldots, f_{I_m,s_m}\}$. Then $\sigma \in \{-1,1\}^n$ satisfies $C_f(n,m)$ if σ satisfies f_{I_j,s_j} for $j = 1, \ldots, m$.*

Example 1 (3-SAT)**.** *Let $f : \{-1,1\}^3 \to \{1,0\}$ where $f(u) = 0$ for $u = (-1,-1,-1)$ and $f(u) = 1$ for all other u. Then f is the constraint function for the 3-SAT problem.*

With these basic definitions in place, we are ready to introduce the model for random CSPs.

2.1 Random CSPs

We discuss our model for randomly generating $C_f(n,m)$, and formally define a "satisfiability threshold."

To generate instances of $C_f(n,m)$, we simply choose $I_1, \ldots, I_m$ and $s_1, \ldots, s_m$ uniformly at random. For completeness, we sample without replacement, i.e. there are no repeated variables in a constraint, and no duplicate constraints. However, sampling with replacement does not affect final results. For the rest of this paper we abuse notation and let $C_f(n,m)$ denote a randomly generated CSP instance following this model.

Now we formally discuss satisfiability thresholds. We let $r = m/n$. For many constraint functions f, there exist thresholds $r_{f,\text{sat}}$ and $r_{f,\text{unsat}}$ such that

$$\lim_{n\to\infty} \Pr[C_f(n,rn) \text{ is satisfiable}] = \begin{cases} 1 \text{ if} & r < r_{f,\text{sat}} \\ 0 \text{ if} & r > r_{f,\text{unsat}} \end{cases}$$

In general, it is unknown whether $r_{f,\text{sat}} = r_{f,\text{unsat}}$, but for some problems such as k-SAT (for large k) and k-XORSAT, affirmative results exist (Ding, Sly, and Sun 2015; Pittel and Sorkin 2016). If $r_{f,\text{sat}} = r_{f,\text{unsat}}$, then we say that the random CSP $C_f(n,m)$ exhibits a sharp threshold in m/n.

We are concerned with finding lower bounds $r_{f,\text{low}}$ on $r_{f,\text{unsat}}$ such that there exists a constant $C > 0$ independent of n so that for sufficiently large n,

$$\Pr[C_f(n,rn) \text{ is satisfiable}] > C \text{ for all } r < r_{f,\text{low}} \tag{2}$$

For CSP classes with a sharp threshold, (2) implies that

$$\lim_{n\to\infty} \Pr[C_f(n,rn) \text{ is satisfiable}] = 1 \text{ for all } r < r_{f,\text{low}}$$

We also wish to find upper bounds $r_{f,\text{up}}$ such that

$$\lim_{n\to\infty} \Pr[C_f(n,rn) \text{ is satisfiable}] = 0 \text{ for all } r > r_{f,\text{up}}$$

For an example of these quantities instantiated on a concrete example, refer to the experiments in Section 5.

We provide a value for $r_{f,\text{up}}$ which was derived earlier in (Creignou and Daudé 2003). Dubois (2001) and Creignou, Daudé, and Dubois (2007) provide methods for obtaining tighter upper bounds, but the looser values that we use are sufficient for showing that $r_{f,\text{low}}$ is on the same asymptotic order as $r_{f,\text{unsat}}$ for many choices of f.

The bound $r_{f,\text{low}}$ depends on both the symmetry and size of the solution set of f. The more assignments $u \in \{-1,1\}^k$ such that $f(u) = 1$, the more likely it is that each constraint is satisfied. Increased symmetry reduces the variance in the number of solutions to $C_f(n,rn)$, so solutions are more spread out among possible CSPs in our class and the probability that $C_f(n,rn)$ will have a solution is higher. We will formally quantify this symmetry in terms of the Fourier spectrum of f, which we introduce next.

2.2 Fourier Expansion of Boolean Functions

We discuss basics of Fourier analysis of boolean functions. For a detailed review, refer to (O'Donnell 2014). We define the vector space $\mathcal{F}_k$ of all functions mapping $\{-1,1\}^k$ to $\mathbb{R}$. The set $\mathcal{F}_k$ has the inner product $\langle f_1, f_2 \rangle = \sum_{u\in\{-1,1\}^k} f_1(u) f_2(u)/2^k$ for any $f_1, f_2 \in \mathcal{F}_k$.

This inner product space has orthonormal basis vectors χ_S, where the parity functions χ_S follow $\chi_S(u) = \prod_{i\in S} u_i$ for all $S \subseteq [k]$, subsets of the k indices. Because $(\chi_S)_{S\subseteq[k]}$ forms an orthonormal basis, if we write

$$\hat{f}(S) = \langle f, \chi_S \rangle = \frac{1}{2^k} \sum_{u\in\{-1,1\}^k} f(u)\chi_S(u) \tag{3}$$

then we can write f as a linear combination of these vectors: $f = \sum_{S\subseteq[k]} \hat{f}(S)\chi_S$. We note that when $S = \emptyset$, the empty set, $\hat{f}(\emptyset)$ is simply the average of f over $\{-1,1\}^k$. We will refer to the coefficients $(\hat{f}(S))_{S\subseteq[k]}$ as the Fourier spectrum of f. Since these coefficients are well-studied in theoretical computer science (O'Donnell 2014), the Fourier coefficients of many boolean functions are easily obtained.

Example 2 (3-SAT). *For 3-SAT,* $\hat{f}(\emptyset) = 7/8$. $\hat{f}(\{1\}) = \hat{f}(\{2\}) = \hat{f}(\{3\}) = 1/8$, *and* $\hat{f}(\{1,2\}) = \hat{f}(\{2,3\}) = \hat{f}(\{1,3\}) = -1/8$.

Representing f in the Fourier bases will facilitate our proofs, providing a simple way to express expectations over our random CSPs. The Fourier spectrum can also provide a measure of "symmetry" in f - if some values of $\hat{f}(S)$ are high where $|S| = 1$, then satisfying assignments to f are more skewed in the variable corresponding to S. We will show how this impacts satisfiability in Section 3.2.

3 Main Results

We provide simple formula for $r_{f,\text{up}}$. A similar result is in (Creignou and Daudé 2003), and the full proof is in the companion technical report.

Proposition 1. *For all constraint functions f, let*

$$r_{f,up} = \frac{\log 2}{\log 1/\hat{f}(\emptyset)} < \frac{\log 2}{1-\hat{f}(\emptyset)}$$

If $r \geq r_{f,up}$, $\lim_{n\to\infty} \Pr[C_f(n,rn)$ *is satisfiable*$] = 0$.

Proof Sketch. We compute the expected solution count for $C_f(n,rn)$. The expected solution count will scale with $\hat{f}(\emptyset)$, since $2^k\hat{f}(\emptyset)$ is simply the number of $u \in \{-1,1\}^k$ where $f(u) = 1$ and therefore governs how easily each constraint will be satisfied. If $r > r_{f,\text{up}}$, the expected solution count converges to 0 as $n \to \infty$, so Markov's inequality implies that the probability that a solution exists goes to 0. □

Next, we will present our value for $r_{f,\text{low}}$. First, some notation: let $U = \{u \in \{-1,1\}^k : f(u) = 1\}$, and let A be the $k \times |U|$ matrix whose columns are the elements of u. We will use A^+ to denote the Moore-Penrose pseudoinverse of A. For a reference on this, see (Barata and Hussein 2012). Finally, let $\mathbf{1}$ be the $|U|$-dimensional vector of 1's.

Example 3 (3-SAT). *For 3-SAT, $k = 3$ and $|U| = 7$, and we can write A as follows (up to permutation of its columns):*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix}$$

where columns of A satisfy the 3-SAT constraint function.

The following main theorem provides the *first computable equation for obtaining lower bounds that are specific to the constraint-generating function f.*

Theorem 1. *For all constraint functions f, let*

$$r_{f,low} = \frac{1}{2}\frac{c}{1-c} \text{ where } c = \hat{f}(\emptyset) - \frac{\mathbf{1}^T A^+ A \mathbf{1}}{2^k}$$

If $r < r_{f,low}$, *then there exists a constant* $C > 0$ *such that* $\lim_{n\to\infty} \Pr[C_f(n,rn)$ *is satisfiable*$] > C$.

The lower bound $r_{f,\text{low}}$ is an increasing function of c which is dependent on two quantities. First, with higher values of $\hat{f}(\emptyset)$, $C_f(n,rn)$ will have more satisfying assignments on average, so c and the threshold value will be higher. Second, c depends on the level of symmetry in the solution set of f, which we will show is connected to the Fourier spectrum of f. We explain this dependence in Section 3.2.

In comparison, Creignou and Daudé (2003) obtain lower bounds which depend only on the arity of f. While we cannot make an exact comparison because Creignou and Daudé use a different random CSP ensemble, for reference, they provide the general lower bound of $1/(ke^k - k)$ expected constraints per variable for functions of arity k. Our bounds are much tighter because of their specificity while remaining simple to compute. To demonstrate, we instantiate our bounds for some example constraint functions in Figure 1. Whereas their bounds are exponentially decreasing in k, our bounds are constant or increasing in k for the functions shown.

3.1 Constraint Functions in Figure 1

We define the constraint functions in Figure 1. Unless specified otherwise, they will be in the form $f : \{-1,1\}^k \to \{0,1\}$.

1. k-SAT: $f(u) = 0$ if u is the all negative ones vector, and $f(u) = 1$ otherwise.
2. k-XORSAT: $f(u) = \mathbb{1}(\chi_{[k]}(u) = -1)$
3. k-NAESAT: $f(u) = 0$ if u is the all negative ones or all ones vector, and $f(u) = 1$ otherwise.
4. k-MAJORITY: Defined when k is odd, $f(u) = 1$ if more than half of the variables of u are 1.
5. a-MAJ $\otimes$ 3-MAJ: Defined when a is odd, where $f : \{-1,1\}^{3a} \to \{0,1\}$. Defined as the composition of a-MAJORITY on a groups of 3-MAJORITY, as follows:

$$f(u_1,\ldots,u_{3a}) = f_{a-\text{MAJ}}(f_{3-\text{MAJ}}(u_1,u_2,u_3),\ldots, f_{3-\text{MAJ}}(u_{3a-2},u_{3a-1},u_{3a}))$$

6. k-MOD-3: $f(u) = 1$ when the number of 1's in u is divisible by 3, and 0 otherwise.
7. $\text{OR}_b\otimes$ XOR_a: In this case, $f : \{-1,1\}^{ab} \to \{0,1\}$, and f is the composition of a OR over b groups of XORs over a variables, as follows:

$$f(u_1,\ldots,u_{ab}) = f_{\text{OR}_b}(f_{\text{XOR}_a}(u_1,\ldots,u_a),\ldots, f_{\text{XOR}_a}(u_{ab-a+1},\ldots,u_{ab}))$$

While the last four constraint functions have not been analyzed much in the existing CSP literature, these types of general constraints are of practical interest because of (Achim, Sabharwal, and Ermon 2016), which performs probabilistic inference by solving CSPs based on arbitrary hash functions. For example, Achim, Sabharwal, and Ermon (2016) show that MAJORITY constraints are effective in practice for solving probabilistic inference problems.

CSP class (f)	Best lower bound on $r_{f,\text{unsat}}$	Our bound $r_{f,\text{low}}$	Upper bound $r_{f,\text{up}}$
k-XORSAT	1	$\frac{1}{2}$	1
k-SAT	$2^k \ln 2 - \frac{1+\ln 2}{2} - o_k(1)$	$2^{k-1} - O(k)$	$2^k \ln 2$
k-NAESAT	$2^{k-1} \ln 2 - \frac{\ln 2}{2} - \frac{1}{4} - o_k(1)$	$2^{k-2} - \frac{1}{2}$	$2^{k-1} \ln 2$
k-MAJORITY	?	$\frac{\frac{1}{2} - k\binom{k-1}{\frac{k-1}{2}}^2 2^{-2k+1}}{1+k\binom{k-1}{\frac{k-1}{2}}^2 2^{-2k+2}} = 0.111 - o_a(1)$	1
a-MAJ $\otimes$3-MAJ	?	$\frac{\frac{1}{2} - 3a\binom{a-1}{\frac{a-1}{2}}^2 2^{-2a-1}}{1+3a\binom{a-1}{\frac{a-1}{2}}^2 2^{-2a-2}} = 0.177 - o_a(1)$	1
k-MOD-3	?	$\frac{1}{4} - o_k(1)$	$\frac{\ln 2}{\ln 3} + o_k(1)$
$\text{OR}_b \otimes \text{XOR}_a$	?	$2^{b-1} - 1/2$	$2^{b-1} \ln 2$

Figure 1: We compare the best known lower bounds on the satisfiability threshold to our lower and upper bounds. For k-XORSAT (Pittel and Sorkin 2016), k-SAT (Ding, Sly, and Sun 2015), and k-NAESAT (Coja-Oglan and Panagiotou 2012), the numbers listed are known as exact sharp threshold locations. For the last four, we do not know of existing lower bounds. $\otimes$ is the composition operator for boolean functions, and we define these functions in Section 3.1.

3.2 Connecting Bounds with Fourier Spectrum

We explain how the Fourier spectrum can help us interpret Theorem 1. We first show the connection between c and the Fourier spectrum. Let $\hat{f}_{S:|S|=1}$ be the k-dimensional vector whose entries are Fourier coefficients of f for size 1 sets. Let B be the $k \times k$ matrix with diagonal entries $B_{ii} = \hat{f}(\emptyset)$ and off-diagonal entries $B_{ij} = \hat{f}(\{i, j\})$ for $i \neq j$.

Example 4 (3-SAT). *Following the coefficients in Example 2, for* 3*-SAT,* $\hat{f}_{S:|S|=1} = (1/8, 1/8, 1/8)$ *and*

$$B = \begin{bmatrix} 7/8 & -1/8 & -1/8 \\ -1/8 & 7/8 & -1/8 \\ -1/8 & -1/8 & 7/8 \end{bmatrix}$$

Lemma 1. *When the rows of A are linearly independent,* $c = \hat{f}(\emptyset) - \hat{f}^T_{S:|S|=1} B^{-1} \hat{f}_{S:|S|=1}$.

From this lemma, we see that larger values of c correspond to smaller $\hat{f}_{S:|S|=1}$. These terms will measure the amount of "symmetry" in the solution set for f. The matrix B and vectors $\hat{f}_{S:|S|=1}$ are easily obtained for many f since Fourier coefficients are well-studied (O'Donnell 2014).

Figure 1 shows how $\hat{f}(\emptyset) - c$ and f relate. Since k-SAT has a mostly symmetric solution set, $\hat{f}_{k-\text{SAT}}(\emptyset) - c = O(k/2^{2k})$ since $c = 1 - 2^{-k} - O(k/2^{2k})$, which is small compared to $\hat{f}_{k-\text{SAT}}(\emptyset)$. The solution set of k-NAESAT is completely symmetric as if $f_{k-\text{NAESAT}}(x) = 1$, then $f_{k-\text{NAESAT}}(-x) = 1$. Thus, $f_{k-\text{NAESAT}}$ has 0 weight on Fourier coefficients for sets with odd size so we can compute that $\hat{f}_{k-\text{NAESAT}}(\emptyset) - c = 0$. k-MAJORITY, however, is less symmetric, as shown by larger first order coefficients. Here, the bound in Figure 1 gives $\lim_{k\to\infty} \hat{f}_{k-\text{MAJORITY}}(\emptyset) - c = 1/\pi$, which is large compared to $\hat{f}_{k-\text{MAJORITY}}(\emptyset) \approx 1/2$.

4 Proof Strategy

Our proof relies on the second moment method, which has been applied with great success to achieve lower bounds for problems such as k-SAT (Achlioptas and Peres 2004) and k-XORSAT (Dubois and Mandler 2002). The second moment method is based on the following lemma, which can be derived using the Cauchy-Schwarz inequality:

Lemma 2. *Let X be any real-valued random variable. Then*

$$Pr[X \neq 0] = \Pr[|X| \neq 0] \geq \frac{E[|X|]^2}{E[X^2]} \geq \frac{E[X]^2}{E[X^2]} \quad (4)$$

If X is only nonzero when $C_f(n, rn)$ has a solution, we obtain lower bounds on the probability that a solution exists by upper bounding $\text{E}[X^2]$. For example, we could let X be the number of solutions to $C_f(n, rn)$. However, as shown in (Achlioptas and Peres 2004), this choice of X fails in most cases. Whether two different assignments satisfy $C_f(n, rn)$ is correlated: if the assignments are close in Hamming distance and one assignment is satisfying, it is more likely that the other is satisfying as well. This will make $\text{E}[X^2]$ much larger than $\text{E}[X]^2$, so (4) will not provide useful information. Figure 2a demonstrates this failure for k-SAT. Achlioptas and Peres (2004) show formally that the ratio $\text{E}[X]^2/\text{E}[X^2]$ will decrease exponentially (albeit at a slow rate). On the other hand, k-NAESAT is "symmetric", so the second moment method works directly here. In the plot, $\text{E}[X]^2/\text{E}[X^2]$ for 3-NAESAT stays above a constant. This also follows formally from our main theorem as well as (Achlioptas and Moore 2002). We formally define our requirements on symmetry in (13).

We circumvent this issue by weighting solutions to reduce correlations before applying the second moment method. As in (Achlioptas and Peres 2004), we use a weighting which factors over constraints in $C_f(n, rn)$ and apply the second moment method to the random variable

$$X = \sum_{\sigma \in \{-1,1\}^n} \prod_{c \in C_f(n,rn)} w(\sigma, c) \quad (5)$$

where $C_f(n, rn)$ is a collection of constraints $\{f_{I_1,s_1}, \ldots, f_{I_m,s_m}\}$ and the randomness in X comes over the choices of I_j, s_j. Now we can restrict our attention to constraint weightings of the form $w(\sigma, f_{I,s}) = w(\sigma_{I,s})$. In the special case where $w(\sigma_{I,s}) = f(\sigma_{I,s})$, X will simply represent the number of solutions to $C_f(n, rn)$. In general, we require $w(\sigma_{I,s}) = 0$ whenever $f(\sigma_{I,s}) = 0$. This way, if $X \neq 0$, then $C_f(n, rn)$ must have a solution.

For convenience, we assume that the index sets $I_1, \dots, I_m$ are sampled with replacement. They are chosen uniformly from $[n]^k$. We also allow constraints to be identical. In the companion technical report, we justify why proofs in this setting carry over to the without-replacement setting in Section 2.1 and also provide full proofs to the lemmas presented below.

In this setting, we will compute the first and second moments of the X chosen in (5) in terms of the Fourier spectrum of w.

Lemma 3. *The squared first moment of X is given by*

$$E[X]^2 = 2^{2n}(\hat{w}(\emptyset)^2)^{rn} \tag{6}$$

Proof. We can expand $\mathrm{E}[X]$ as follows:

$$\begin{aligned} \mathrm{E}[X] &= \sum_{\sigma \in \{-1,1\}^n} \mathrm{E}\left[\prod_{j=1}^{rn} w(\sigma_{I_j,s_j})\right] \\ &= \sum_{\sigma \in \{-1,1\}^n} \mathrm{E}[w(\sigma_{I,s})]^{rn} \end{aligned} \tag{7}$$

where we used the fact that constraints are chosen independently. Now we claim that for any $u \in \{-1,1\}^k$,

$$\Pr[\sigma_{I,s} = u] = \frac{1}{2^k}$$

This follows from the fact that we choose s uniformly over $\{-1,1\}^k$ and our definition of $\sigma_{I,s}$ in (1). Thus,

$$\begin{aligned} \mathrm{E}[w(\sigma_{I,s})] &= \sum_{u \in \{-1,1\}^k} w(u) \Pr[\sigma_{I,s} = u] \\ &= \frac{1}{2^k} \sum_{u \in \{-1,1\}^k} w(u) \\ &= \hat{w}(\emptyset) \end{aligned}$$

Plugging back into (7) gives the desired result. □

Next, we will compute the second moment $\mathrm{E}[X^2]$.

Lemma 4. *Let $g_w(\alpha) = \sum_{S \subseteq [k]} (2\alpha - 1)^{|S|} \hat{w}(S)^2$. The second moment of X is given by*

$$E[X^2] = 2^n \sum_{j=0}^{n} \binom{n}{j} g_w(j/n)^{rn} \tag{8}$$

The function $g_w(\alpha)$ is similar to the noise sensitivity of a boolean function (O'Donnell 2003) and measures the correlation in the value of w between two assignments σ, τ which overlap at $\alpha(\sigma, \tau)n$ locations. As a visual example, Figure 2b shows how $g_w(\alpha)$ changes for k-XORSAT with varying k. The key of our proof is showing that $\mathrm{E}[w(\sigma_{I,s})w(\tau_{I,s})] = g_w(\alpha(\sigma, \tau))$ for a random constraint $f_{I,s}$.

We will now write $\mathrm{E}[X]^2$ in terms of g_w. Since $g_w(1/2) = \hat{w}(\emptyset)^2$, plugging this into (6) gives us

$$\mathrm{E}[X]^2 = 2^{2n} g_w(1/2)^{rn} \tag{9}$$

This motivates us to apply the following lemma from (Achlioptas and Peres 2004), which will allow us to translate bounds on $g_w(\alpha)$ into bounds on $\mathrm{E}[X]^2/\mathrm{E}[X^2]$:

Lemma 5. *Let ϕ be any real, positive, twice-differentiable function on $[0,1]$ and let*

$$S_n = \sum_{j=0}^{n} \binom{n}{j} \phi(j/n)^n$$

Define ψ on $[0,1]$ as $\psi(\alpha) = \frac{\phi(\alpha)}{\alpha^\alpha (1-\alpha)^{1-\alpha}}$. If there exists $\alpha_{\max} \in (0,1)$ such that $\psi(\alpha_{\max}) > \psi(\alpha)$ for all $\alpha \neq \alpha_{\max}$, and $\psi''(\alpha_{\max}) < 0$, then there exist constants $B, C > 0$ such that for all sufficiently large n,

$$B\psi(\alpha_{\max})^n \leq S_n \leq C\psi(\alpha_{\max})^n \tag{10}$$

To apply the lemma, we can define $\phi_r(\alpha) = g_w(\alpha)^r$ and $\psi_r(\alpha) = \frac{\phi_r(\alpha)}{\alpha^\alpha (1-\alpha)^{1-\alpha}}$. Then from (9), we note that

$$\psi_r(1/2)^n = 2^n (g_w(1/2)^r)^n = \mathrm{E}[X]^2/2^n \tag{11}$$

On the other hand, from (8),

$$\sum_{j=0}^{n} \binom{n}{j} \phi_r(j/n)^n = \mathrm{E}[X^2]/2^n \tag{12}$$

so if the conditions of Lemma 5 hold for $\alpha_{\max} = 1/2$, we recover that $\mathrm{E}[X]^2/\mathrm{E}[X^2] \geq C$ for some constant $C > 0$.

One requirement for $\psi_r(\alpha)$ to be maximized at $\alpha = 1/2$ is that $\psi_r'(1/2) = 0$. Expanding $\psi_r'(1/2)$ gives $2g_w(1/2)^{r-1}(rg_w'(1/2)) = 0$. Since $g_w(1/2) = \hat{w}(\emptyset)^2 > 0$, we thus require

$$g_w'(1/2) = 2 \sum_{S \subseteq [k] : |S| = 1} \hat{w}(S)^2 = 0 \tag{13}$$

In order to satisfy (13), we need $\hat{w}(S) = 0$ for all $S \subseteq [k]$ where $|S| = 1$. To use Lemma 5, we would like to choose w such that (13) holds. We discuss how to choose w to optimize our lower bounds in the companion technical report. In the next section, we will provide r so that the conditions of Lemma 5 hold at $\alpha = 1/2$ for arbitrary w when (13) is satisfied.

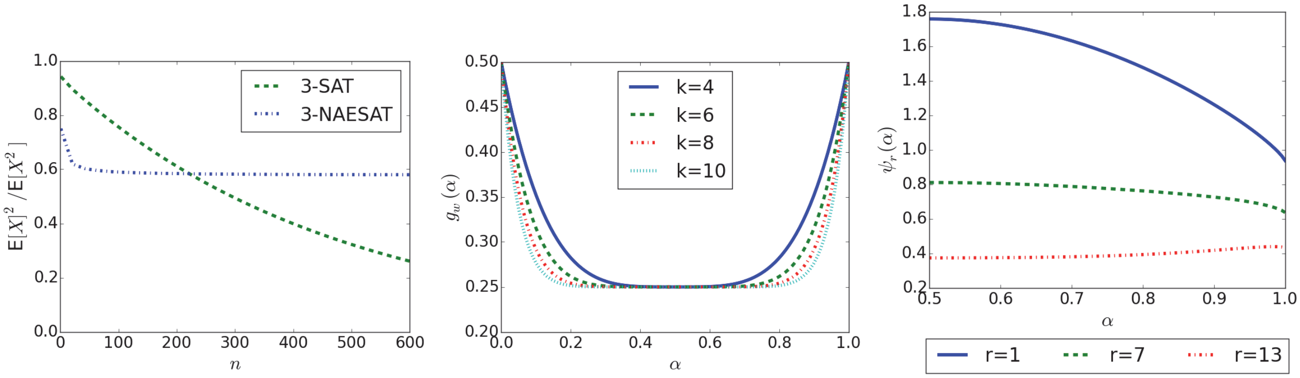
4.1 Bounding the Second Moment For Fixed w

We give a general bound on r in terms of our weight function w so that the conditions of Lemma 5 are satisfied for $\alpha = 1/2$. For now, the only constraint we place on w is that (13) holds. The next lemma lets us consider only $\alpha \in [1/2, 1]$.

Lemma 6. *Let $\alpha \geq 1/2$. Then $g_w(\alpha) \geq g_w(1-\alpha)$.*

This lemma follows because $g_w(\alpha)$ is a polynomial in $(2\alpha - 1)$ with nonnegative coefficients, and $(2\alpha - 1) > 0$ for $\alpha > 1/2$.

Now we can bound $\psi_r(\alpha)$ for $\alpha \in [1/2, 1]$. Combined with Lemma 6, the next lemma will give conditions on r such that $\psi_r(1/2) > \psi_r(\alpha)$ for all $\alpha \in [0,1]$.



(a) $\mathrm{E}[X]^2/\mathrm{E}[X^2]$ vs. n for 3-SAT and 3-NAESAT. X is the solution count, $r = 1$.

(b) $g_w(\alpha)$ when f is k-XORSAT and $w = f$.

(c) $\psi_r(\alpha)$ with f as 5-NAESAT and $w = f$.

Figure 2: For concreteness, we provide sample plots of the relevant quantities in our proofs.

Lemma 7. *Let the weight function* w *satisfy* (13). *If*

$$r \leq \frac{1}{2} \frac{\hat{w}(\emptyset)^2}{\sum_{S:|S|\geq 2} \hat{w}(S)^2} \tag{14}$$

$\psi_r(1/2) > \psi_r(\alpha)$ *for* $\alpha \in [0, 1]$ *and* $\psi_r''(1/2) < 0$.

Figure 2c shows how r controls the shape of the function $\psi_r(\alpha)$. As r increases, $\psi_r''(1/2)$ becomes positive and $\psi_r(\alpha)$ will no longer attain a local maximum in that region. The key step in proving Lemma 7 is rearranging $\psi_r(1/2) > \psi_r(\alpha)$ and simplify calculations by using approximations for the logarithmic terms that appear.

Our bound on r compares the average of w over $\{-1, 1\}^k$ with the correlations between w and the Fourier basis functions. If w has strong correlations with the other Fourier basis functions, two assignments which are equal at αn variables will likely either be both satisfying or both not satisfying as α approaches 1. This increases $\mathrm{E}[X^2]$ but not $\mathrm{E}[X]^2$ and makes Lemma 2 provide a trivial bound if r is too large. Thus, if w has strong correlations with the Fourier basis functions, we must choose smaller r as reflected by (14).

To get the tightest bounds, we wish to maximize the expression in (14). Although we prove our lemma for general w requiring only (13), we also need $w(u) = 0$ whenever $f(u) = 0$ to apply our lemma to satisfiability. Recalling our definition of X in (5), this condition ensures that $C_f(n, rn)$ has a solution whenever $X \neq 0$. Thus,

$$w(u) = \lambda(u) f(u) \tag{15}$$

for some $\lambda : \{-1, 1\}^k \to \mathbb{R}$. If we disregard (13), choosing $\lambda(u) = 1$ would maximize the bound on r in (14). The additional requirement of (13) for the second moment method to succeed can be viewed as a "symmetrization penalty" on r. In the companion technical report, we discuss how to choose w to optimize our bound on r while satisfying (13) and (15).

4.2 Proving the Main Theorem

We will combine our lemmas to prove Theorem 1.

Proof of Theorem 1. We wish to apply the second moment method on X defined in (5), where w is a function we use to weigh assignments to individual constraints. We choose w as described in the full version of the paper, which satisfies both (13) and (15). Since w satisfies (15), $\Pr[C_f(n, rn) \text{ is satisfiable}] \geq \Pr[X \neq 0] \geq \mathrm{E}[X]^2/\mathrm{E}[X^2]$ by Lemma 2. Now we will use Lemma 7 to show that the conditions for Lemma 5 are satisfied for $\phi_r = g_w(\alpha)^r$ and r satisfying (14). For our choice of w, it follows from the derivations in the full technical report that the RHS of (14) becomes

$$r < r_{f,\text{low}} = \frac{1}{2} \frac{c}{1 - c} \text{ where } c = \hat{f}(\emptyset) - \frac{\mathbf{1}^T A^+ A \mathbf{1}}{2^k}$$

where A is defined in Section 3. There is a slight technicality in directly applying Lemma 5 because ϕ_r might not be nonnegative for $\alpha < 1/2$; we discuss this in the companion technical report. Now using (11) and (12), and applying Lemma 5, we can conclude that there exists $C > 0$ such that

$$\Pr[C_f(n, rn) \text{ is satisfiable}] \geq \mathrm{E}[X]^2/\mathrm{E}[X^2] \geq C$$

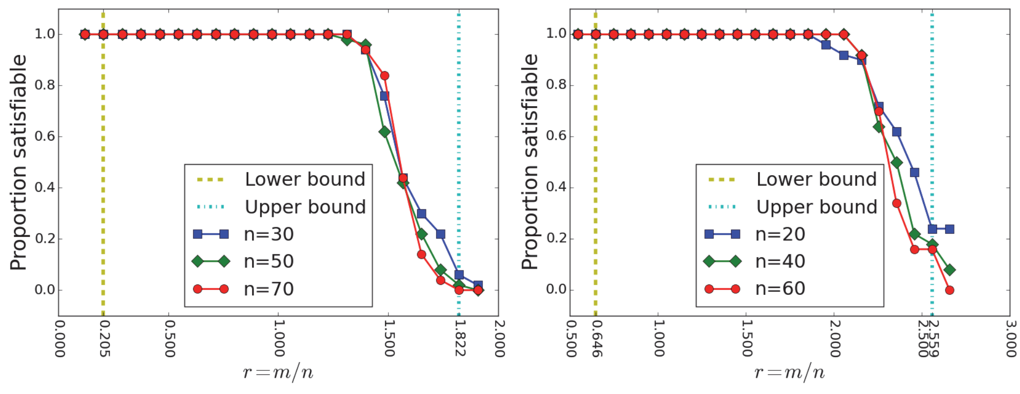
for sufficiently large n and all $r < r_{f,\text{low}}$. □

There remains a question of what $r_{f,\text{low}}$ we can hope achieve using a second moment method proof where X is defined as in (5). The following lemma provides some intuition for this:

Lemma 8. *In order for the conditions of Lemma 5 to hold at* $\alpha_{\max} = 1/2$ *for* X *in the form of* (5) *and any choice of* w *satisfying* (15)*, we require*

$$r < \log 2 / \log \frac{1}{\hat{f}(\emptyset) - \frac{\mathbf{1}^T A^+ A \mathbf{1}}{2^k}} \leq r_{f,up} = \frac{\log 2}{\log \frac{1}{\hat{f}(\emptyset)}} \tag{16}$$

The difference of $\mathbf{1}^T A^+ A \mathbf{1}/2^k$ in the lower logarithm compared to $r_{f,\text{up}}$ can be viewed as a "symmetrization penalty" necessary for our proof to work. While Lemma 8 does not preclude applications of the second moment



(a) $a = 2, b = 4$ (b) $a = 2, b = 5$

Figure 3: Proportion of CSPs satisfiable out of 50 trials vs. r for tribes functions. We show our bounds for reference.

method that do not rely on Lemma 5, consider what happens for r that do not satisfy (16). For these r, the function $\psi_r(\alpha)$ must obtain a maximum at some $\alpha^* \in [0,1], \alpha^* \neq 1/2$. If it also happens that $\phi_r(\alpha)$ is nonnegative and twice differentiable on $[0,1]$, and $\psi_r''(\alpha^*) < 0$, then conditions of Lemma 5 hold, and applying it to $\alpha_{\max} = \alpha^*$ along with (10), (11), and (12) will actually imply that

$$\frac{\mathrm{E}[X]^2}{\mathrm{E}[X^2]} \leq \frac{1}{B}\left(\frac{\psi_r(1/2)}{\psi_r(\alpha^*)}\right)^n$$

for some constant $B > 0$, which gives us an exponentially decreasing, and therefore trivial lower bound for the second moment method. Therefore, we believe that (16) is near the best lower bound on $r_{f,\text{unsat}}$ that we can achieve by applying the second moment method on X in the form of (5).

5 Experimental Verification of Bounds

We empirically test our bounds with the goal of examining their tightness. For our constraint functions, we will use tribes functions. The tribes function takes the disjunction of b groups of a variables and evaluates to 1 or 0 based on whether the following formula is true:

$$\text{TRIBES}_{a,b}(x_1, \ldots, x_{ab}) = \vee_{i=0}^{b-1}\left(\wedge_{j=1}^{a} x_{ia+j}\right)$$

where $+1$ denotes true and -1 denotes false. For our experiments, we randomly generate CSP formulas based on $\text{TRIBES}_{a,b}$. We use the Dimetheus[1] random CSP solver to solve these formulas, or report if no solution exists. We show our results in Figure 3. As expected, our values for lower bounds $r_{f,\text{low}}$ are looser than the upper bounds $r_{f,\text{up}}$.

6 Conclusion

Using Fourier analysis and the second moment method, we have shown general bounds on m/n, the ratio of constraints to variables; for m/n below these bounds, there is constant probability that a random CSP is satisfiable. We demonstrate that our bounds are easily instantiated and can be applied to obtain novel estimates of the satisfiability threshold for many classes of CSPs. Our bounds depend on how easy it is to symmetrize solutions to the constraint function. We provide a heuristic argument to approximate the best possible lower bounds that our application of the second moment method can achieve; these bounds differ from upper bounds on the satisfiability threshold by a "symmetrization penalty." Thus, an interesting direction of future research is to determine whether we can provide tighter upper bounds that account for symmetrization, or whether symmetrization terms are an artificial product of the second moment method.

[1]https://www.gableske.net/dimetheus

7 Acknowledgements

This work was supported by the Future of Life Institute (grant 2016-158687) and by the National Science Foundation (grant 1649208).

References

Achim, T.; Sabharwal, A.; and Ermon, S. 2016. Beyond parity constraints: Fourier analysis of hash functions for inference. In *Proceedings of The 33rd International Conference on Machine Learning*, 2254–2262.

Achlioptas, D., and Moore, C. 2002. The asymptotic order of the random k-SAT threshold. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, 779–788. IEEE.

Achlioptas, D., and Peres, Y. 2004. The threshold for random-SAT is $2^k \log 2 - \mathrm{O}(k)$. *Journal of the American Mathematical Society* 17(4):947–973.

Achlioptas, D.; Naor, A.; and Peres, Y. 2005. Rigorous location of phase transitions in hard optimization problems. *Nature* 435(7043):759–764.

Barak, B.; Moitra, A.; O'Donnell, R.; Raghavendra, P.; Regev, O.; Steurer, D.; Trevisan, L.; Vijayaraghavan, A.; Witmer, D.; and Wright, J. 2015. Beating the random assignment on constraint satisfaction problems of bounded degree. *arXiv preprint arXiv:1505.03424*.

Barata, J. C. A., and Hussein, M. S. 2012. The moore–penrose pseudoinverse: A tutorial review of the theory. *Brazilian Journal of Physics* 42(1-2):146–165.

Biere, A.; Heule, M.; and van Maaren, H. 2009. *Handbook of satisfiability*, volume 185. ios press.

Coja-Oghlan, A., and Panagiotou, K. 2013. Going after the k-SAT threshold. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, 705–714. ACM.

Coja-Oghlan, A., and Panagiotou, K. 2016. The asymptotic k-SAT threshold. *Advances in Mathematics* 288:985–1068.

Coja-Oglan, A., and Panagiotou, K. 2012. Catching the k-NAESAT threshold. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, 899–908. ACM.

Cook, S. A. 1971. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, 151–158. ACM.

Creignou, N., and Daudé, H. 2003. Generalized satisfiability problems: minimal elements and phase transitions. *Theoretical Computer Science* 302(1):417–430.

Creignou, N.; Daudé, H.; and Dubois, O. 2007. Expected number of locally maximal solutions for random boolean CSPs. In *2007 Conference on Analysis of Algorithms, AofA 07*, 109–122. Discrete Mathematics and Theoretical Computer Science.

Ding, J.; Sly, A.; and Sun, N. 2015. Proof of the satisfiability conjecture for large k. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, 59–68. ACM.

Dubois, O., and Mandler, J. 2002. The 3-XORSAT threshold. *Comptes Rendus Mathematique* 335(11):963–966.

Dubois, O. 2001. Upper bounds on the satisfiability threshold. *Theoretical Computer Science* 265(1):187–197.

Dudek, J. M.; Meel, K. S.; and Vardi, M. Y. 2016. Combining the k-cnf and xor phase-transitions. In *IJCAI*.

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.

Mitchell, D.; Selman, B.; and Levesque, H. 1992. Hard and easy distributions of SAT problems. In *AAAI*, volume 92, 459–465.

Molloy, M. 2003. Models for random constraint satisfaction problems. *SIAM Journal on Computing* 32(4):935–949.

Montanari, A.; Restrepo, R.; and Tetali, P. 2011. Reconstruction and clustering in random constraint satisfaction problems. *SIAM Journal on Discrete Mathematics* 25(2):771–808.

O'Donnell, R. 2003. *Computational applications of noise sensitivity*. Ph.D. Dissertation, Massachusetts Institute of Technology.

O'Donnell, R. 2014. *Analysis of boolean functions*. Cambridge University Press.

Pittel, B., and Sorkin, G. B. 2016. The satisfiability threshold for k-XORSAT. *Combinatorics, Probability and Computing* 25(02):236–268.

Russell, S. J.; Norvig, P.; Canny, J. F.; Malik, J. M.; and Edwards, D. D. 2003. *Artificial intelligence: a modern approach*, volume 2. Prentice hall Upper Saddle River.

Vardi, M. Y. 2014. Boolean satisfiability: theory and engineering. *Commun. ACM* 57(3):5.