

Bounds on the Satisfiability Threshold for Power Law Distributed Random SAT

Tobias Friedrich¹, Anton Krohmer¹, Ralf Rothenberger¹, Thomas Sauerwald² and Andrew M. Sutton¹

- 1 Hasso Plattner Institute, Potsdam, Germany
firstname.lastname@hpi.de
- 2 University of Cambridge, United Kingdom
firstname.lastname@cl.cam.ac.uk

Abstract

Propositional satisfiability (SAT) is one of the most fundamental problems in computer science. The worst-case hardness of SAT lies at the core of computational complexity theory. The average-case analysis of SAT has triggered the development of sophisticated rigorous and non-rigorous techniques for analyzing random structures.

Despite a long line of research and substantial progress, nearly all theoretical work on random SAT assumes a *uniform* distribution on the variables. In contrast, real-world instances often exhibit large fluctuations in variable occurrence. This can be modeled by a *scale-free* distribution of the variables, which results in distributions closer to industrial SAT instances.

We study random k -SAT on n variables, $m = \Theta(n)$ clauses, and a power law distribution on the variable occurrences with exponent β . We observe a satisfiability threshold at $\beta = (2k - 1)/(k - 1)$. This threshold is tight in the sense that instances with $\beta \leq (2k - 1)/(k - 1) - \varepsilon$ for any constant $\varepsilon > 0$ are *unsatisfiable* with high probability (w. h. p.). For $\beta \geq (2k - 1)/(k - 1) + \varepsilon$, the picture is reminiscent of the uniform case: instances are *satisfiable* w. h. p. for sufficiently small constant clause-variable ratios m/n ; they are *unsatisfiable* above a ratio m/n that depends on β .

1998 ACM Subject Classification G.2.1 Combinatorics

Keywords and phrases satisfiability, random structures, random SAT, power law distribution, scale-freeness, phase transitions

Digital Object Identifier 10.4230/LIPIcs.ESA.2017.1

1 Introduction

Satisfiability of propositional formulas (SAT) is one of the most researched problems in theoretical computer science. SAT is widely used to model practical problems such as bounded model checking, hardware and software verification, automated planning and scheduling, and circuit design. Even *large industrial instances* with millions of variables can often be solved very efficiently by modern SAT solvers. The structure of these industrial SAT instances appears to allow a much faster processing than the theoretical worst-case of this NP-complete problem. It is an open and widely discussed question which structural properties make a SAT instance easy to solve for modern SAT solvers.

Random SAT: For modeling typical inputs, we study random propositional formulas. In random satisfiability, we have a distribution over Boolean formulas in conjunctive normal form (CNF). The degree of a variable in a CNF formula is the number of disjunctive clauses in which that variable appears either positively or negatively. Two interesting properties



of random models are its *degree distribution* and its *satisfiability threshold*. The degree distribution $F(x)$ of a formula Φ is the fraction of variables that occur more than x times (negated or unnegated). A satisfiability threshold is a critical value around which the probability that a formula is satisfiable changes from 0 to 1.

Uniform random SAT: In the classical uniform random model, the degree distribution is binomial. On uniform random k -SAT, the *satisfiability threshold conjecture* [1] asserts if Φ is a formula drawn uniformly at random from the set of all k -CNF formulas with n variables and m clauses, there exists a real number r_k such that

$$\lim_{n \rightarrow \infty} \Pr\{\Phi \text{ is satisfiable}\} = \begin{cases} 1 & m/n < r_k; \\ 0 & m/n > r_k. \end{cases}$$

A well-known result of Friedgut [22] establishes that the transition is sharp, even though its location is not known exactly for all values of k (and may also depend on n). For $k = 2$, the critical threshold is $r_2 = 1$ [14, 18, 23]. Recently, Coja-Oghlan and Panagiotou [16, 17] gave a sharp bound (up to lower order terms) with $r_k = 2^k \log 2 - \frac{1}{2}(1 + \log 2) \pm o_k(1)$. Ding, Sly, and Sun [20] derive an exact representation of the threshold for all $k \geq k_0$, where k_0 is a large enough constant. Explicit bounds also exist for low values of k , e.g., $3.52 \leq r_3 \leq 4.4898$ [19, 24, 25], and numerical estimates using the cavity method from statistical mechanics [29] suggest that $r_3 \approx 4.26$.

Other random SAT models: In the regular random model [11], formulas are constructed at random, but the degree distribution is fixed: each literal appears exactly $\lfloor \frac{km}{2n} \rfloor$ or $\lfloor \frac{km}{2n} \rfloor + 1$ times in the formula. Similarly, Bradonjic and Perkins [12] considered a random geometric k -SAT model in which $2n$ points are placed at random in $[0, 1]^d$. Each point corresponds to a unique literal, and clauses are formed by all k -sets of literals that lie together within a ball of diameter $\Theta(n^{-1/d})$. Again, this model has a binomial variable distribution.

Power law random SAT: Recently, there has been a paradigm shift when modeling real-world data. In many applications, it has been found that certain quantities do not cluster around a specific scale as suggested by a uniform distribution, but are rather inhomogeneous [15, 30]. In particular, the degree distribution in complex networks often follows a power law [31]. This means that the fraction of vertices of degree k is proportional to $k^{-\beta}$, where the constant β depends on the network. To mathematically study the behavior of such networks, random graph models that generate a power law degree distribution have been proposed [2, 10, 27, 32].

While there has been a large amount of research on power law random graphs in the past few years [33], there is little previous work on power law SAT formulas. Nevertheless, the observation that quantities follow a power law in real-world data has also emerged in the context of SAT [11]. As all aforementioned random SAT models assume strongly concentrated degree distributions, it was conjectured that this property might be modeled well by random formulas with a power law degree distribution.

To address this conjecture, and to help close the gap between the structure of uniform random and industrial instances, Ansótegui, Bonet, and Levy [4] recently proposed a power-law random SAT model. This model has been studied experimentally [4–7], and empirical investigations found that (1) indeed the constraint graphs of many families of industrial instances obey a power-law and (2) SAT solvers that are constructed to specialize on industrial instances perform better on power-law formulas than on uniform random formulas. To complement these experimental findings, we contribute with this paper the first theoretical results on this model.

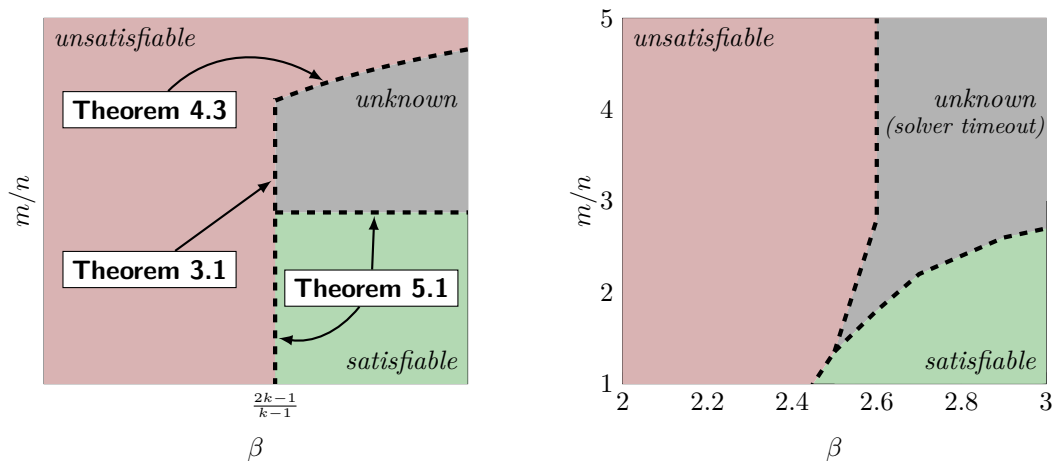


Figure 1 Illustration of our asymptotic results for the power law satisfiability threshold location when $n \rightarrow \infty$ (left) compared with empirical results for randomly generated power law 3-SAT formulas on $n = 10^6$ variables checked with the SAT solver MiniSAT (right). The timeout was set to one hour.

Our results: We study random k -SAT on n variables and $m = \Theta(n)$ clauses. Each clause contains $k = \Theta(1)$ different, independently sampled variables. Each variable x_i is chosen with non-uniform probability p_i and negated with probability $1/2$. A formal definition can be found in Section 2. We first study sufficient conditions under which the resulting k -SAT instances are *unsatisfiable*. Assume a probability distribution \vec{p} on the variables where p_i is non-decreasing in $i \in \{1, \dots, n\}$. If the k most frequent variables are sufficiently common, we prove in Section 3 the following statement:

► **Theorem 3.1.** *Let Φ be a random k -SAT formula with probability distribution \vec{p} on the variables (c.f. Definition 2.1), with $k \geq 2$ and $\frac{m}{n} = \Omega(1)$. If $p_{n-k+1} = \Omega((\frac{\log n}{n})^{1/k})$, then Φ is w. h. p. unsatisfiable.*

Our focus are power law distributions with some exponent β . Theorem 3.1 implies that power law random k -SAT formulas with $\beta = \frac{2k-1}{k-1} - \varepsilon$ for an arbitrary constant $\varepsilon > 0$ are unsatisfiable with high probability¹, cf. Corollary 3.2.

In Section 4 we show that something similar holds for the clause-variable ratio $\frac{m}{n}$, i.e. power law random k -SAT formulas with $\frac{m}{n}$ bigger than some constant are unsatisfiable with high probability. Although this already follows from basic observations, we derive a better bound on the value of the constant.

► **Theorem 4.1.** *Let Φ be a random k -SAT formula with probability distribution \vec{p} on the variables (c.f. Definition 2.1), with $k \geq 2$ and $r = \frac{m}{n}$. With high probability, Φ is unsatisfiable if*

$$\left(1 - \frac{1}{2^k}\right)^r \left[\prod_{i=1}^n \left[2 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right)^m \right] \right]^{\frac{1}{n}} < 1.$$

In Section 5 we prove the following positive result, which complements our picture of the satisfiability landscape:

¹ We say that an event E holds w. h. p., if there exists a $\delta > 0$ such that $\Pr[E] \geq 1 - \mathcal{O}(n^{-\delta})$.

► **Theorem 5.1.** *Let Φ be a random k -SAT formula whose variable probabilities follow a power law distribution (c.f. Definition 2.2). If the power law exponent is $\beta \geq \frac{2k-1}{k-1} + \varepsilon$ for an arbitrary $\varepsilon > 0$, Φ is satisfiable with high probability if $\frac{m}{n}$ is a small enough constant.*

Together our main theorems prove that random k -SAT instances whose variables follow power law distributions do not only exhibit a phase transition for some clause-variable ratio $r = \frac{m}{n}$, but also around the power law exponent $\beta = \frac{2k-1}{k-1}$. Figure 1 contains an overview of our results. To prove these statements, we borrow tools developed for the uniform random SAT model. Note, however, that many of their common techniques like the differential equation method seem difficult to apply to non-uniform distributions; as removing a variable results in a more complex rescaling of the rest of the distribution. It is therefore crucial to perform careful operations on the formulas that leave the distribution of variables intact. To this end, we use techniques known from the analysis of power law random graphs.

Clause length: We focus on power law variable distributions but fix the length of every clause to $k \geq 2$. Power law models have also been proposed in which clause length is distributed by a power law as well [4, 5]. As long as there is a constant *minimum clause length* $k_{\min} \geq 2$, our results can be extended to this case in the following way.

If the clause lengths are distributed as a power law, there will appear $\Theta(n)$ clauses of length k_{\min} , and all other clauses are of larger size. In that case, Theorems 3.1 and 5.1 are directly applicable to the linear number of clauses with size k_{\min} (obtaining different hidden constants); and we have that the formula is satisfiable with high probability if $\beta \geq \frac{2k_{\min}-1}{k_{\min}-1} + \varepsilon$ and m/n is a small enough constant. On the other hand, the formula is unsatisfiable with high probability, if $\beta \leq \frac{2k_{\min}-1}{k_{\min}-1} - \varepsilon$. Consequently, the satisfiability of the formula does (asymptotically) not depend on the second power law.

2 Definition of the Model and Preliminaries

We analyze random k -SAT on n variables and $m = \Theta(n)$ clauses, where $k \geq 2$. The constant $r := \frac{m}{n}$ is called *clause-variable ratio* or *constraint density*. We denote by x_1, \dots, x_n the Boolean variables. A clause is a disjunction of k literals $\ell_1 \vee \dots \vee \ell_k$, where each literal assumes a (possibly negated) variable. Finally, a formula Φ in conjunctive normal form is a conjunction of clauses $c_1 \wedge \dots \wedge c_m$. We conveniently interpret a clause c both as a Boolean formula and as a set of literals. Following standard notation, we write $|\ell|$ to refer to the indicator of the variable corresponding to literal ℓ . We say that Φ is satisfiable if there exists an assignment of variables x_1, \dots, x_n such that the formula evaluates to 1.

► **Definition 2.1** (Random k -SAT). Let m, n be given, and consider any probability distribution \vec{p} on n variables with $\sum_{i=1}^n p_i = 1$. To construct a random SAT formula Φ , we sample m clauses independently at random. Each clause is sampled as follows:

1. Select k variables independently at random from the distribution \vec{p} . Repeat until no variables coincide.
2. Negate each of the k variables independently at random with probability $1/2$.

Observe that by setting $p_i = \frac{1}{n}$ for all i , we obtain again the uniform random SAT model. One can show (see Appendix A) that for power law distributions, the probability to sample a specific clause c is

$$(1 + o(1)) \frac{k!}{2^k} \prod_{\ell \in c} p_{|\ell|}. \quad (2.1)$$

Power law Distributions. In this paper, we are mostly concerned with distributions p_i that follow a power law. To this end, we define two models: A *general* model to capture most power law distributions (which is harder to analyze), and a *concrete* model that gives us one instance of \vec{p} depending only on n that can be used to compute precise leading constants. We use the general model to derive some asymptotic results; and the concrete model to compare with the uniform random SAT model and for the experiments.

Before we define these two models, let us establish the concept of a *weight* w_i of a variable x_i . The weight gives us (roughly) the expected number of times the variable appears in the formula. That is,

$$p_i := \frac{w_i}{\sum_j w_j}.$$

Thus, fixing the weights $\vec{w} = (w_1, \dots, w_n)$ also fixes the probability distribution \vec{p} . It is important to distinguish between the initial distribution of variables \vec{p} and modified distributions that may arise as a result of stochastic considerations. For instance, the smallest-weight variable in a clause is clearly not distributed according to \vec{p} (except in 1-SAT). To avoid confusion, we identify a variable with its weight, as the weights stay fixed throughout the analysis. For convenience, we further assume w.l.o.g. that the variables are ordered increasingly by weight, i. e. for $i \leq j$ we have $w_i \leq w_j$. Note that our definition of power law ensures that for $\beta > 2$, we have $\sum_j w_j = \Theta(n)$.

We are now ready to define the two models.

► **Definition 2.2** (General Power Law). Let the weights $\vec{w} := w_1, \dots, w_n$ be given, and let W be a weight selected uniformly at random. We say that \vec{w} follows a power law with exponent β , if $w_1 = \Theta(1)$, $w_n = \Theta(n^{\frac{1}{\beta-1}})$, and for all $w \in [w_1, w_n]$ it holds

$$F(w) := \Pr[W \geq w] = \Theta(w^{1-\beta}) \quad (2.2)$$

Whenever we need the explicit constants bounding the distribution function, we refer to them by α_1, α_2 as in

$$\alpha_1 w^{1-\beta} \leq F(w) \leq \alpha_2 w^{1-\beta}. \quad (2.3)$$

We point out that Definition 2.2 assumes a deterministic weight sequence; but it can be easily generalized to also support randomly generated weights.

For the concrete model, we define the weights as follows.

► **Definition 2.3** (Concrete Power Law). Given a power law exponent β , we call \vec{w} the concrete power law sequence, if

$$w_{n-i+1} := \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}}. \quad (2.4)$$

One can check that for these concrete weights, it holds $n \cdot F(w) = \lfloor nw^{1-\beta} \rfloor$, so in a sense, they are a canonical choice for producing a power law weight distribution.

It remains to show that using a power law distribution in Definition 2.1 indeed results in a power law distribution of variable occurrences. Ansótegui et al. [5] provide a proof sketch for this fact, we prove it rigorously in Appendix B.

► **Theorem 2.4.** Let Φ be a random k -SAT formula that follows an arbitrary power law distribution with exponent β (c.f. Definition 2.2) and $m = \Theta(n)$. Then, there are $d_{\min} = \Theta(w_{\min})$ and $d_{\max} = \Theta(w_{\max})$, such that for all $d_{\min} \leq d \leq d_{\max}$ w. h. p. it holds that

$$N_{\geq d} = \Theta(n \cdot d^{1-\beta}),$$

where $N_{\geq d}$ is the number of variables that appear at least d times in Φ .

To analyze power law distributions, we often make use of the following result of Bringmann, Keusch, and Lengler [13, Lemma B.1], which allows replacing sums by integrals.

► **Theorem 2.5** ([13]). *Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function, and let $F^>(w) := \Pr[W > w]$. Then, for any $0 \leq \underline{w} \leq \bar{w}$,*

$$\sum_{i \in [n], \underline{w} \leq w_i \leq \bar{w}} \frac{1}{n} f(w_i) = f(\underline{w}) \cdot F(\underline{w}) - f(\bar{w}) \cdot F^>(\bar{w}) + \int_{\underline{w}}^{\bar{w}} f'(w) \cdot F(w) dw.$$

Using this theorem, the following corollary can be shown (see Appendix B):

► **Corollary 2.6.** *Let the variables w_i be power law distributed with exponent $\beta > 2$, and define $W_{\geq w} := \sum_{i \in [n]: w_i \geq w} w_i$. Then, $W_{\geq w} = \Theta(nw^{2-\beta})$.*

Hence, $\sum_j w_j = W_{\geq w_1} = \Theta(n)$ and therefore $p_i = \Theta(\frac{w_i}{n})$. Finally, we denote by V the random variable describing the weight of a SAT variable chosen according to a power law distribution p_i , that is, $\Pr[V = w] = \sum_i p_i \cdot \mathbf{1}[w_i = w]$, where $\mathbf{1}$ denotes the indicator variable of the event. Note that this is not equivalent to W , since there is a subtle difference in the two random processes: W is a random variable drawn uniformly at random from w_1, \dots, w_n , whereas V is a random variable drawn from the same set, but with the non-uniform distribution p_1, \dots, p_n . Hence, by Corollary 2.6,

$$\Pr[V \geq w] = \Theta(w^{2-\beta}). \quad (2.5)$$

3 Small Power Law Exponents are Unsatisfiable

For small power law exponents, one can show that they result in formulas that are unsatisfiable (for large n) for all constant clause-variable ratios. The rationale behind this is that large variables with weight $\Theta(w_n)$ appear polynomially often together in a clause. For constant k , they thus appear in all 2^k configurations (negated and non-negated), making the formula trivially unsatisfiable. Theorem 3.1, already stated in the introduction, gives a sufficient condition on the variable distribution to make a random k -SAT formula unsatisfiable.

► **Theorem 3.1.** *Let Φ be a random k -SAT formula with probability distribution \vec{p} on the variables (c.f. Definition 2.1), with $k \geq 2$ and $\frac{m}{n} = \Omega(1)$. If $p_{n-k+1} = \Omega((\frac{\log n}{n})^{1/k})$, then Φ is w. h. p. unsatisfiable.*

Proof. Recall that p_i is without loss of generality increasing in i . Consider the k largest variables $n-k+1, \dots, n$. We call \mathcal{E}_i the event that clause i consists of these variables. Then,

$$\Pr[\mathcal{E}_i] = \Omega(p_{n-k+1}^k) = \Omega(\frac{\log n}{n}).$$

Since each clause is drawn independently at random, we obtain by a Chernoff bound (see for example Theorem 1.1 in [21]) that with high probability, the total number of clauses consisting of these variables is

$$|\mathcal{E}| := \sum_{i=1}^m \mathbf{1}[\mathcal{E}_i] = \Omega(\log n).$$

In other words, the number of clauses in which the k largest variables appear together increases as a logarithm in n . Since in each of these clauses, the literals appear negated or non-negated with constant probability $1/2$, we have that all 2^k possible combinations of negated and non-negated literals appear in the formula with probability at least

$$1 - 2^k \cdot (\frac{2^k - 1}{2^k})^{|\mathcal{E}|} = 1 - n^{-\Omega(1)}$$

by the union bound. Since all 2^k combinations cannot be satisfied at once, the resulting formula is unsatisfiable. \blacktriangleleft

By applying Theorem 3.1 to a power law distribution on the variables, we obtain the following power law threshold for unsatisfiability.

► **Corollary 3.2.** *Let Φ be a random k -SAT formula that follows an arbitrary power law distribution fulfilling Definition 2.2. If the power law exponent is $\beta \leq \frac{2k-1}{k-1} - \varepsilon$ for an arbitrary $\varepsilon > 0$, Φ is unsatisfiable with high probability.*

Proof. Observe that from $\beta = \frac{2k-1}{k-1} - \varepsilon$ it follows $k = \frac{\beta-1}{\beta-2} - \varepsilon'$ for some constant ε' . By setting $nF(w) \leq k$ we obtain that the largest k variables all have weight $\Theta(w_n) = \Theta(n^{\frac{1}{\beta-1}})$. Consequently, when $\beta > 2$,

$$(p_{n-k})^k = \Theta(n^{-k \frac{\beta-2}{\beta-1}}) = \Theta(n^{-1+\varepsilon' \frac{\beta-2}{\beta-1}}) = \omega\left(\frac{\log n}{n}\right),$$

and the statement follows from Theorem 3.1. For the case where $\beta \leq 2$, one can show using Theorem 2.5 that $\sum_i w_i = \Theta(n^{\frac{1}{\beta-1}})$, and therefore $p_{n-k} = \Omega(1)$. Again, the statement follows from Theorem 3.1. \blacktriangleleft

4 Large Clause-Variable Ratios are Unsatisfiable

It is a well-known result that random SAT on any probability distribution will result in unsatisfiable formulas if the clause-variable ratio is high. This follows from the probabilistic method: The expected number of assignments that satisfy a formula is $2^n(1-2^{-k})^m$. This is independent from the variable distribution as long as each variable is negated with probability $1/2$. Hence, if the clause-variable ratio exceeds $\ln(2)/\ln(\frac{2^k}{2^k-1})$, the resulting formula will be unsatisfiable with high probability. This constant is rather large, however: In the case of $k=3$ this yields an upper bound on the clause-variable ratio of ≈ 5.191 . For the concrete power law distribution in Definition 2.3, the true threshold is much smaller. In fact, it appears to be below the satisfiability threshold for uniform random SAT.

Let us restate the main result, which will be proven with the Single Flip Method [26].

► **Theorem 4.1.** *Let Φ be a random k -SAT formula with probability distribution \vec{p} on the variables (c.f. Definition 2.1), with $k \geq 2$ and $r = \frac{m}{n}$. With high probability, Φ is unsatisfiable if*

$$\left(1 - \frac{1}{2^k}\right)^r \left[\prod_{i=1}^n \left[2 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2)}\right)^m \right] \right]^{\frac{1}{n}} < 1.$$

The following is a corollary from this theorem:

► **Corollary 4.2.** *Let Φ be a random k -SAT formula that follows Definition 2.1 with $k \geq 2$, $r = \frac{m}{n}$ and $\|\vec{p}\|_2^2 = o(1)$. With high probability, Φ is unsatisfiable if*

$$\left(1 - \frac{1}{2^k}\right)^r \left(2 - \exp\left(-\left(\frac{k}{2^k - 1} r\right)(1 + o(1))\right) \right) < 1.$$

The proof can be found in Appendix B. Interestingly, the above corollary gives the same inequality as the Single-Flip Method for uniform random SAT [26]. This shows that the uniform distribution resembles a worst-case for this method; and all other distributions can only improve this bound.

If \vec{p} follows a power law distribution as in Definition 2.3, we can derive the following theorem, which gives an upper bound independent of n .

k	power law distribution with exponent β								uniform dist.
	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	
3				3.48	3.71	3.87	3.99	4.08	4.67
4			7.87	8.42	8.78	9.04	9.23	9.37	10.23
5		16.27	17.75	18.64	19.21	19.61	19.90	20.11	21.33
7	67.21	75.74	79.81	82.09	83.49	84.42	85.07	85.54	87.88
10	619.28	662.48	680.93	690.36	695.77	699.12	701.34	702.88	708.94

■ **Table 1** Numerical upper bounds on the density threshold obtained from the Single-Flip Method (cf. Theorems 4.1 and 4.3). Empty fields indicate unsatisfiability for *all* constant densities by Theorem 3.1. To the best of our knowledge, the bounds for uniform random SAT with $k \geq 4$ are the currently best known numerical upper bounds. For $k = 3$ the best known unconditional numerical upper bound is 4.4898 [19].

► **Theorem 4.3.** *Let Φ be a random k -SAT formula with $k \geq 2$ and $r = \frac{m}{n}$ that follows a power law distribution fulfilling Definition 2.3. Let further $N \in \mathbb{N}^+$ be any constant. If the power law exponent is $\beta > 2$, then Φ is w. h. p. unsatisfiable if*

$$\left(\left(1 - \frac{1}{2^k}\right)^r 2^{\frac{1}{N}} \prod_{l=1}^{N-1} \left[2 - \exp \left(- (1 + o(1)) r \frac{k}{2^k - 1} \frac{\beta - 2}{\beta - 1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}} \right) \right]^{\frac{1}{N}} \right) < 1.$$

The bound from this Theorem improves as $N \rightarrow \infty$. As this expression is rather terse, we also numerically determine in Table 1 the smallest constant r such that the formula is unsatisfiable. We compare these values to the upper bounds for uniform random SAT obtained from the Single-Flip Method. In the remainder of this section, we show Theorem 4.1 and defer the proof of Theorem 4.3 to the Appendix.

► **Definition 4.4 (Single-Flip Property).** For a random formula Φ a truth assignment A has the *single-flip property* iff A satisfies Φ and every assignment A' obtained from A by flipping exactly one zero to one does *not* satisfy Φ .

Let N_{SF} be the number of truth assignments with the single-flip property for Φ . As argued in [26], such an assignment exists if Φ is satisfiable. From Markov's Inequality, we thus know $\Pr[\Phi \text{ satisfiable}] \leq \mathbb{E}[N_{SF}]$.

In the following, we derive a bound on $\mathbb{E}[N_{SF}]$. By Lemma A.1, the probability of choosing a clause c is at most

$$\frac{k!}{2^k} \cdot \frac{\prod_{\ell \in c} p_{|\ell|}}{1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2}.$$

To bound the number of assignments with the single-flip property, we use the following result.

► **Lemma 4.5 ([26]).** *The expected number of assignments with the single-flip property is*

$$\mathbb{E}[N_{SF}] = \left(1 - \frac{1}{2^k}\right)^m \sum_{\text{assignment } A} \Pr[A \text{ single-flip} \mid A \text{ satisfying}].$$

Proof. Note that for a certain truth assignment A , the probability of choosing a clause which is not satisfied by A is $1/2^k$. Therefore, the probability that A is a satisfying assignment for Φ is exactly $\left(1 - \frac{1}{2^k}\right)^m$. ◀

We next bound the probability that a satisfying assignment A has the single-flip property.

► **Lemma 4.6.** For a satisfying assignment $A = (a_1, a_2, \dots, a_n) \in \{0, 1\}^n$ it holds that

$$\Pr[A \text{ single-flip} \mid A \text{ satisfying}] \leq \prod_{i: a_i=0} 1 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}\right)^m.$$

Proof. For a satisfying assignment A to have the single-flip property, all assignments A^i obtained by flipping a bit $a_i = 0$ of A must not satisfy Φ . To fulfill this property for A^i , we have to choose at least one clause which contains $\overline{X_i}$ and $k - 1$ other variables with appropriate signs so that A^i does not satisfy the clause. Let $S^i(c)$ denote the event that a clause c is satisfied by A , but not by A^i . Then,

$$\Pr[S^i(c)] = \frac{k! \cdot p_i \sum_{J \in \mathcal{P}_{k-1}([n] \setminus \{i\})} \prod_{j \in J} p_j}{2^k (1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)} \leq \frac{k \cdot p_i}{2^k (1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}$$

since $\sum_{J \in \mathcal{P}_{k-1}([n] \setminus \{i\})} \prod_{j \in J} p_j \leq \frac{\|\vec{p}\|_1^{k-1}}{(k-1)!}$. The probability of choosing a clause not satisfied by A^i under the condition that A is satisfying is then

$$\Pr[S^i(c) \mid A \text{ sat}] = \Pr[S^i(c) \mid A \text{ satisfies } c] \leq \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}$$

as the probability of choosing a clause which is satisfied by any assignment is exactly $\frac{2^k - 1}{2^k}$. For a fixed assignment A^i we conclude

$$\begin{aligned} \Pr[A^i \text{ unsat} \mid A \text{ sat}] &= 1 - (1 - \Pr[S^i(c) \mid A \text{ sat}])^m \\ &\leq 1 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}\right)^m. \end{aligned} \quad (4.1)$$

It remains to find the joint probability that all single-flipped assignments A^i for $1 \leq i \leq n$ with $a_i = 0$ are not satisfying. We show this using a correlation inequality by Farr [28]. The sets of clauses which are not satisfied by the A^i 's are pairwise disjoint as each clause in the set for A^i has to contain $\overline{X_i}$, whereas each clause in the set for A^j ($j \neq i$) can not contain $\overline{X_i}$. In the context of the correlation inequality from [28] we set $V = \{1, 2, \dots, m\}$, $I = \{i \in \{1, 2, \dots, n\} \mid a_i = 0\}$, $X_v = i$ iff the v -th clause is satisfied by A , but not by A^i , and \mathcal{F}_i the ‘‘increasing’’ collection of non-empty subsets of V . The application of the Theorem then directly yields

$$\begin{aligned} \Pr[A \text{ single-flip} \mid A \text{ sat}] &= \Pr\left[\bigcap_{i: a_i=0} A^i \text{ unsat} \mid A \text{ sat}\right] \\ &\leq \prod_{i: a_i=0} \left[1 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}\right)^m\right]. \end{aligned} \quad \blacktriangleleft$$

Combining Lemmas 4.5 and 4.6 we get that the expected number of assignments with single-flip property is at most

$$\begin{aligned} \mathbb{E}[N_{SF}] &\leq \left(1 - \frac{1}{2^k}\right)^m \sum_{I \subseteq \{1, 2, \dots, n\}} \prod_{i \in I} \left[1 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}\right)^m\right] \\ &= \left(1 - \frac{1}{2^k}\right)^m \prod_{i=1}^n \left[2 - \left(1 - \frac{k \cdot p_i}{2^k - 1} \frac{1}{(1 - \frac{1}{2}k^2 \|\vec{p}\|_2^2)}\right)^m\right]. \end{aligned}$$

This establishes Theorem 4.1.

Algorithm 1 Clause Shrinking Algorithm**Input:** k -SAT formula Φ ; weight distribution \vec{w}

- 1: **for all** $c \in \Phi$ **do**
- 2: $\ell_1 \leftarrow \operatorname{argmin}_{\ell \in c} \{w_{|\ell|}\}$
- 3: $\ell_2 \leftarrow \operatorname{argmin}_{\ell \in c \setminus \{\ell_1\}} \{w_{|\ell|}\}$
- 4: $c \leftarrow (\ell_1 \vee \ell_2)$
- 5: Solve Φ using any polynomial time 2-SAT algorithm

5 Conditions for Satisfiability

In this section, we provide a complementary result to Theorems 3.1 and 4.3 proving that if $\beta \geq \frac{2k-1}{k-1} + \varepsilon$ and the clause-variable ratio $r = \frac{m}{n}$ does not exceed some small constant, then a random k -SAT formula with exponent β is satisfiable with high probability. Let us first restate the main result:

► **Theorem 5.1.** *Let Φ be a random k -SAT formula whose variable probabilities follow a power law distribution (c.f. Definition 2.2). If the power law exponent is $\beta \geq \frac{2k-1}{k-1} + \varepsilon$ for an arbitrary $\varepsilon > 0$, Φ is satisfiable with high probability if $\frac{m}{n}$ is a small enough constant.*

We show this statement by constructing an algorithm that satisfies Φ w. h. p. if the clause-variable ratio is small. Algorithm 1 contains a formal description. The main idea is to shrink all clauses to size 2 by selecting the literals with smallest weight in each clause; and then running any well-known (polynomial time) 2-SAT algorithm (e. g. [8]).

In the following, we seek to establish that Algorithm 1 will find a satisfying assignment (for small constraint densities) with high probability. To this end, we first analyze the probability distribution of a clause c after it has been shrunk.

► **Lemma 5.2.** *Let ℓ_1, ℓ_2 be the selected literals of an arbitrary clause $c \in \Phi$ in Algorithm 1. Then,*

$$\Pr[|\ell_1| = i, |\ell_2| = j] + \Pr[|\ell_1| = j, |\ell_2| = i] \leq \mathcal{O}\left(\frac{1}{n^2} (w_i w_j)^{1 - \frac{1}{2}(k-2)(\beta-2)}\right).$$

Proof. W.l.o.g., we assume that $w_i \leq w_j$. Then, $\Pr[|\ell_1| = j, |\ell_2| = i] = 0$ by the definition of Algorithm 1. For the event $|\ell_1| = i, |\ell_2| = j$ to happen, all other $k-2$ literals in the clause must be of larger weight. By Equations (2.1) and (2.5),

$$\begin{aligned} \Pr[|\ell_1| = i, |\ell_2| = j] &= \frac{1}{2} \cdot \binom{k}{2} \cdot (1 + o(1)) \cdot p_i \cdot p_j \cdot \Pr[V \geq w_j]^{k-2} \\ &= \Theta\left(\frac{1}{n^2}\right) \cdot w_i w_j^{1 - (k-2)(\beta-2)} \\ &\leq \mathcal{O}\left(\frac{1}{n^2}\right) \cdot (w_i w_j)^{1 - \frac{1}{2}(k-2)(\beta-2)}. \end{aligned}$$

The last statement holds since $w_i \leq w_j$. ◀

Having derived a bound on the probability distribution of a shrunk clause, it is possible to compute the probability that the resulting 2-SAT formula is satisfiable. We use that the clauses are sampled independently. To avoid confusion, we write Φ' and c' , whenever we talk about the shrunk formula and clauses. To upper bound the probability of Φ not being satisfiable, we look at so-called *bi-cycles* in Φ' .

► **Definition 5.3.** A *bi-cycle* of length l is a sequence of $l + 1$ clauses of the form

$$(u, \ell_1), (\bar{\ell}_1, \ell_2), \dots, (\bar{\ell}_{l-1}, \ell_l), (\bar{\ell}_l, v),$$

where ℓ_1, \dots, ℓ_l are literals of distinct variables and $u, v \in \{\ell_1, \dots, \ell_l, \bar{\ell}_1, \dots, \bar{\ell}_l\}$.

Chvatal and Reed [14, Theorem 3] show that if the formula Φ' is unsatisfiable, it must contain a bi-cycle. Consequently, by upper bounding the probability that a bi-cycle appears, we immediately obtain an upper bound on the probability that Φ' and henceforth Φ is unsatisfiable.

► **Theorem 5.4** ([14]). *Let Φ' be any 2-SAT formula. If Φ' contains no bi-cycle, it is satisfiable.*

Before we are able to prove the main Theorem, we need the following auxiliary Lemma, the proof of which can be found in the Appendix.

► **Lemma 5.5.** *Let $\beta = \delta + 1 + \varepsilon$ for some $\varepsilon > 0$. For all $1 \leq l \leq n$, there is a constant c with*

$$\sum_{\substack{S \subseteq [n]: \\ |S|=l}} \prod_{i \in S} w_i^\delta \leq n^l \cdot c^{\frac{1}{l}}.$$

We are now able to show Theorem 5.1. As discussed above, we do this by upper bounding the probability that a bi-cycle appears in Φ' . To this end, we calculate the expected number of bi-cycles in Φ' , observe that it is $\text{poly}(n)^{-1}$, and apply Markov's inequality. This yields that w. h. p., Φ' and thus Φ are satisfiable.

Proof of Theorem 5.1. We calculate the expected number of bi-cycles in Φ' . First, we fix a set $S \subseteq [n]$ of $l \geq 2$ variables to appear in a bi-cycle. Let X_B denote the random variable counting how many times a *specific* bi-cycle B with the variables from S appears in F . Then

$$\mathbb{E}[X_B] \leq \binom{m}{l+1} (l+1)! \cdot \Pr[u \vee x_1] \Pr[\bar{x}_l \vee v] \cdot \prod_{i=1}^{l-1} \Pr[\bar{x}_i \vee x_{i+1}].$$

The factor $\binom{m}{l+1} (l+1)!$ counts the possible positions of B in F . By Lemma 5.2,

$$\mathbb{E}[X_B] \leq m^{l+1} \cdot \left(\frac{c_1}{n^{\frac{1}{2}}}\right)^{l+1} \cdot \left(w_{|u|} w_{|v|} \prod_{i \in S} w_i^2\right)^{1 - \frac{1}{2}(k-2)(\beta-2)}$$

for some suitable constant c_1 . Now let X_S denote the random variable counting how many times *any* bi-cycle with the variables from S appears in F . There are $l!$ permutations of the l variables; and 2^l combinations of literals on l variables. Similarly, literals u and v have 4 possible sign combinations. Thus,

$$\mathbb{E}[X_S] \leq m^{l+1} \cdot l! \cdot 2^l \cdot \left(\frac{c_1}{n^{\frac{1}{2}}}\right)^{l+1} \cdot 4 \left(\sum_{i \in S} w_i^{1 - \frac{1}{2}(k-2)(\beta-2)}\right)^2 \prod_{i \in S} w_i^{2 - (k-2)(\beta-2)}.$$

To estimate the sum, we upper bound $w_i \leq w_n$ for all sets up to a certain size l_0 , which we will determine later. We set $\delta := 2 - (k-2)(\beta-2)$ and define $\alpha(l)$ as

$$\left(\sum_{i \in S} w_i^{\delta/2}\right)^2 \leq \alpha(l) := \begin{cases} \mathcal{O}(l^2), & \text{if } \delta \leq 0, \\ l_0^2 \cdot w_n^\delta, & \text{if } \delta > 0 \text{ and } l \leq l_0, \\ \mathcal{O}(n^2), & \text{otherwise.} \end{cases}$$

Now let X denote the random variable counting the number of bi-cycles that appear in F .

$$\mathbb{E}[X] \leq \sum_{l=2}^n 2^{l+2} \cdot m^{l+1} \cdot l! \cdot \left(\frac{c_1}{n^2}\right)^{l+1} \cdot \alpha(l) \sum_{\substack{S \subseteq [n] \\ |S|=l}} \prod_{i \in S} w_i^\delta.$$

Since $\delta + 1 = 2 - (k - 2)(\beta - 2) + 1 < \beta$ by our assumption $\beta \geq \frac{2k-1}{k-1} + \varepsilon$, we can apply Lemma 5.5. Using $r := m/n$, we obtain that the right-hand side is at most

$$\mathbb{E}[X] \leq \sum_{l=2}^n 2^{l+2} \cdot m^{l+1} \cdot l! \cdot \left(\frac{c_1}{n^2}\right)^{l+1} \cdot \alpha(l) \cdot n^l \cdot c^{\frac{l}{l}} \leq \frac{1}{n} \sum_{l=2}^n c_2^l \cdot r^l \cdot \alpha(l), \quad (5.1)$$

for some suitable constant c_2 . Since r is a small enough constant we thus have $c_2 \cdot r < 1$. If $\delta \leq 0$, we are finished, since then

$$\frac{1}{n} \sum_{l=2}^n c_2^l \cdot r^l \cdot \alpha(l) \leq \frac{1}{n} \sum_{l=2}^n (c_2 \cdot r)^l \cdot l^2 \leq \mathcal{O}\left(\frac{1}{n}\right).$$

Otherwise, if $\delta > 0$, we choose $l_0 := -4 \cdot \ln^{-1}(c_2 r) \ln(n)$, which ensures $(r \cdot c_2)^l = \mathcal{O}(n^{-4})$ for all $l > l_0$. For $l = 2, \dots, l_0$, equation (5.1) sums up to at most

$$\frac{1}{n} \sum_{l=2}^{l_0} (c_2 r)^l \cdot l^2 \cdot w_n^\delta = \mathcal{O}(\log^3(n) \cdot n^{1-k \frac{\beta-2}{\beta-1}}),$$

where we substituted $w_n = \Theta(n^{\frac{1}{\beta-1}})$ and $\delta = 2 - (k - 2)(\beta - 2)$. Since $\beta \geq \frac{2k-1}{k-1} + \varepsilon$, the exponent $1 - k \frac{\beta-2}{\beta-1} < -\varepsilon'$ is negative, and we thus have

$$\mathbb{E}[X] \leq \frac{1}{n} \sum_{l=2}^n c_2^l r^l \alpha(l) \leq \mathcal{O}(\log^3(n) \cdot n^{-\varepsilon'}) + \mathcal{O}\left(\frac{1}{n}\right),$$

which proves the Theorem by Markov's inequality. ◀

6 Discussion of the Results

In this work, we have shown that with high probability, a power law random k -SAT formula is satisfiable, if $\beta \geq \frac{2k-1}{k-1} + \varepsilon$ and the clause-variable ratio is not too large; and that it is unsatisfiable if $\beta \leq \frac{2k-1}{k-1} - \varepsilon$, or if the clause-variable ratio is too large. Here, we give a few observations following these results.

First, as explained in Section 1 our results translate directly to the model where clause lengths are power law distributed. This observation might help to explain a phenomenon that arose in [5]: The authors experimentally observed that a random-sat formula with double power law distribution (both variables and clause lengths are drawn from a power law) can be solved extremely fast by MiniSAT. Although the formula was of length $5 \cdot 10^5$, MiniSAT already gave an answer after 4 seconds! Using our results, we are now able to provide a potential explanation for this phenomenon: Disregarding the double power law distribution, the smallest clause length k_{\min} occurring in their generated formulas is one. Thus, there will be $\Theta(n)$ clauses of length one and by Theorem 3.1 the formula is likely unsatisfiable.

Second, we observe a sharp threshold in the sense of Friedgut [22] (for small constraint densities r) for β at the point $\frac{2k-1}{k-1}$. In contrast, it is unclear whether such a sharp threshold exists (and can be analytically derived) for fixed β but variable r . Considering however, that decades of research were dedicated to the same question in the uniform case—an arguably simpler model—it is unlikely that we obtain a satisfying answer any time soon; at least for all k . As in the uniform model, however, it might be more tractable to get sharp thresholds for $k \rightarrow \infty$.

References

- 1 D. Achlioptas, A. Coja-Oghlan, and F. Ricci-Tersenghi. On the solution-space geometry of random constraint satisfaction problems. *Random Structures & Algorithms*, 38:251–268, 2011.
- 2 W. Aiello, F. Chung, and L. Lu. A random graph model for power law graphs. *Experimental Mathematics*, 10:53–66, 2001.
- 3 D. Alistarh, T. Sauerwald, and M. Vojnović. Lock-free algorithms under stochastic schedulers. In *34th Symp. Principles of Distributed Computing (PODC)*, pp. 251–260, 2015.
- 4 C. Ansótegui, M. L. Bonet, and J. Levy. On the structure of industrial SAT instances. In *15th Intl. Conf. Principles and Practice of Constraint Programming (CP)*, pp. 127–141, 2009.
- 5 C. Ansótegui, M. L. Bonet, and J. Levy. Towards industrial-like random SAT instances. In *21st Intl. Joint Conf. Artificial Intelligence (IJCAI)*, pp. 387–392, 2009.
- 6 C. Ansótegui, M. L. Bonet, J. Giráldez-Cru, and J. Levy. The fractal dimension of SAT formulas. In S. Demri, D. Kapur, and C. Weidenbach, editors, *Automated Reasoning - 7th International Joint Conference, IJCAR 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 19-22, 2014. Proceedings*, Vol. 8562 of *Lecture Notes in Computer Science*, pp. 107–121. Springer, 2014.
- 7 C. Ansótegui, M. L. Bonet, J. Giráldez-Cru, and J. Levy. On the classification of industrial SAT families. In E. Armengol, D. Boixader, and F. Grimaldo, editors, *Artificial Intelligence Research and Development - Proceedings of the 18th International Conference of the Catalan Association for Artificial Intelligence, Valencia, Catalonia, Spain, October 21-23, 2015.*, Vol. 277 of *Frontiers in Artificial Intelligence and Applications*, pp. 163–172. IOS Press, 2015.
- 8 B. Aspvall, M. F. Plass, and R. E. Tarjan. A linear-time algorithm for testing the truth of certain quantified boolean formulas. *Information Processing Letters*, 8:121–123, 1979.
- 9 A. Auger and B. Doerr. *Theory of randomized search heuristics: Foundations and recent developments*, Vol. 1. World Scientific, 2011.
- 10 A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- 11 Y. Boufkhad, O. Dubois, Y. Interian, and B. Selman. Regular random k -sat: Properties of balanced formulas. *J. Automated Reasoning*, 35:181–200, 2005.
- 12 M. Bradonjic and W. Perkins. On sharp thresholds in random geometric graphs. In *18th Intl. Workshop on Randomization and Computation (RANDOM)*, pp. 500–514, 2014.
- 13 K. Bringmann, R. Keusch, and J. Lengler. Geometric inhomogeneous random graphs. *arXiv preprint arXiv:1511.00576*, 2015.
- 14 V. Chvatal and B. Reed. Mick gets some (the odds are on his side). In *33rd Symp. Foundations of Computer Science (FOCS)*, pp. 620–627, 1992.
- 15 A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM Review*, 51:661–703, 2009.
- 16 A. Coja-Oghlan. The asymptotic k -SAT threshold. In *46th Symp. Theory of Computing (STOC)*, pp. 804–813, 2014.
- 17 A. Coja-Oghlan and K. Panagiotou. The asymptotic k -sat threshold. *Advances in Mathematics*, 288:985–1068, 2016.
- 18 W. F. de la Vega. Random 2-SAT: results and problems. *Theoretical Computer Science*, 265:131–146, 2001.

- 19 J. Díaz, L. M. Kirousis, D. Mitsche, and X. Pérez-Giménez. On the satisfiability threshold of formulas with three literals per clause. *Theoretical Computer Science*, 410:2920–2934, 2009.
- 20 J. Ding, A. Sly, and N. Sun. Proof of the satisfiability conjecture for large k . In *47th Symp. Theory of Computing (STOC)*, pp. 59–68, 2015.
- 21 D. Dubhashi and A. Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009.
- 22 E. Friedgut. Sharp thresholds of graph properties, and the k -SAT problem. *J. ACM*, 12: 1017–1054, 1999.
- 23 A. Goerdt. A threshold for unsatisfiability. *J. Computer & System Sciences*, 53:469–486, 1996.
- 24 M. T. Hajiaghayi and G. B. Sorkin. The satisfiability threshold of random 3-SAT is at least 3.52. Technical Report RC22942, IBM, 2003.
- 25 A. C. Kaporis, L. M. Kirousis, and E. G. Lalas. The probabilistic analysis of a greedy satisfiability algorithm. *Random Structures & Algorithms*, 28:444–480, 2006.
- 26 L. M. Kirousis, E. Kranakis, D. Krizanc, Y. C. Stamatiou, et al. Approximating the unsatisfiability threshold of random formulas. *Random Structures & Algorithms*, 12: 253–269, 1998.
- 27 D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82:036106, 2010.
- 28 C. McDiarmid. On a correlation inequality of Farr. *Combinatorics, Probability & Computing*, 1:157–160, 1992.
- 29 M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812–815, 2002.
- 30 M. E. Newman. Power laws, pareto distributions and zipf’s law. *Contemporary physics*, 46:323–351, 2005.
- 31 M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45: pp. 167–256, 2003.
- 32 B. Söderberg. General formalism for inhomogeneous random graphs. *Phys. Rev. E*, 66: 066121, 2002.
- 33 R. van der Hofstad. Random graphs and complex networks. Available at www.win.tue.nl/~rhofstad/NotesRGCN.pdf, 2011.

A Exact Clause Sampling Probability

The probability to draw a specific clause c is

$$\frac{\prod_{\ell \in c} p_{|\ell|}}{2^k \sum_{J \in \mathcal{P}_k(\{1,2,\dots,n\})} \prod_{j \in J} p_j}, \quad (\text{A.1})$$

where $\mathcal{P}_k(\cdot)$ denotes the set of cardinality- k elements of the power set. The factor 2^k in the denominator comes from the different possibilities to negate variables. Note that $k! \sum_{J \in \mathcal{P}_k(\{1,2,\dots,n\})} \prod_{j \in J} p_j$ is the probability of choosing a k -clause that contains no variable more than once. To see that this probability is almost 1 for most distributions, we apply the following result from [3].

► **Lemma A.1** (Non-Uniform Birthday Paradox). *Let $\vec{p} = (p_1, \dots, p_n)$ be any probability distribution on n items. Assume you sample t items from \vec{p} . Let $\mathcal{E}(t)$ be the event that there is a collision, i.e. that at least 2 of t items are equal. Then,*

$$\Pr[\mathcal{E}(t)] \leq \frac{1}{2} t^2 \|\vec{p}\|_2^2 = \frac{1}{2} t^2 \sum_{i=1}^n p_i^2.$$

The probability that a sampled k -clause thereby has collisions is at most $\frac{1}{2} k^2 \|\vec{p}\|_2^2$; so for $\|\vec{p}\|_2^2 = o(1)$ and constant k we obtain that the probability to draw a specific clause c is

$$(1 + o(1)) \frac{k!}{2^k} \prod_{\ell \in c} p_{|\ell|}. \quad (\text{A.2})$$

Using Theorem 2.5, we can show that $\|\vec{p}\|_2^2 = o(1)$ indeed holds for Definition 2.2 with exponent $\beta > 2$ as required in Equation (A.2):

$$\|\vec{p}\|_2^2 = \sum_{i=1}^n p_i^2 = \Theta(n^{-2}) \sum_{i=1}^n w_i^2 = \Theta(n^{-1}) \cdot n^{\frac{3-\beta}{\beta-1}} = o(1).$$

B Omitted Proofs

► **Theorem 2.4.** *Let Φ be a random k -SAT formula that follows an arbitrary power law distribution with exponent β (c.f. Definition 2.2) and $m = \Theta(n)$. Then, there are $d_{\min} = \Theta(w_{\min})$ and $d_{\max} = \Theta(w_{\max})$, such that for all $d_{\min} \leq d \leq d_{\max}$ w. h. p. it holds that*

$$N_{\geq d} = \Theta(n \cdot d^{1-\beta}),$$

where $N_{\geq d}$ is the number of variables that appear at least d times in Φ .

Proof. Let f_x be the number of appearances of x or \bar{x} in Φ . Observe that $\mathbb{E}[f_x] \leq k \cdot m \cdot p_x$, since each variable can appear at most once in a clause. On the other hand, it holds $\mathbb{E}[f_x] \geq m \cdot p_x$, since this is the expected number of appearances of x in a 1-SAT formula. Thus, since $m = \Theta(n)$ and $k = \Theta(1)$ by assumption, it holds that

$$\mathbb{E}[f_x] = \Theta(w_x). \quad (\text{B.1})$$

We first prove the statement for $d \geq 2c \ln n$, where $c > 0$ is some suitable constant. By applying Chernoff bounds, we can derive that w. h. p. all variables x with $\mathbb{E}[f_x] < \frac{d}{2}$ appear fewer than d times; and all variables x with $\mathbb{E}[f_x] \geq 2d$ appear at least d times. The

requirement $d_{\min} \leq d \leq d_{\max}$ is needed so that the Chernoff bounds work, which might not be the case if d is too close to w_{\min} or w_{\max} . Due to Definition 2.2 and Equation (B.1) this implies

$$N_{\geq d} = \Theta(n \cdot d^{1-\beta}).$$

Now let us consider the case $d < 2c \ln n$ and let Y_i be random variables indicating if $f_{x_i} \geq d$ for $i = 1, 2, \dots, n$. To show a lower bound on $N_{\geq d}$, we again look at variables x with $\mathbb{E}[f_x] \geq 2d$. For those it holds that

$$\Pr(Y_i = 0) \leq \Pr\left(f_{x_i} < \frac{1}{2}\mathbb{E}[f_{x_i}]\right) \leq e^{-\frac{\mathbb{E}[f_{x_i}]}{8}} \leq e^{-\frac{d}{4}}, \quad (\text{B.2})$$

again due to Chernoff Bounds. Also, by Equation (B.1), it holds for variables x with $\mathbb{E}[f_x] \geq 2d$ that $w_x = \Omega(2d)$. By the requirements on the weight distribution from Definition 2.2 there are $\Theta(n \cdot d^{1-\beta})$ such variables. Therefore, it holds that

$$\mathbb{E}[N_{\geq d}] \geq \mathbb{E}\left[\sum_{\substack{i \in [n]: \\ \mathbb{E}[f_{x_i}] \geq 2d}} Y_i\right] \geq c' \left(1 - e^{-\frac{d}{4}}\right) \cdot n \cdot d^{1-\beta}$$

for a suitable constant $c' > 0$. Observe that if we condition on $Y_i = 1$, i. e. that x_i appears at least d times, this lowers the probability of all other variables to appear d times, and vice versa. Thus, the random variables Y_1, \dots, Y_n are negatively correlated and we may apply a Chernoff bound [9, Theorem 1.16]. Since $1 - e^{-d/4} = \Omega(1)$ and $d = \mathcal{O}(\log n)$, we obtain that w. h. p.

$$N_{\geq d} \geq \frac{1}{2}c' \cdot \left(1 - e^{-\frac{d}{4}}\right) \cdot n \cdot d^{1-\beta} = \Omega(n \cdot d^{1-\beta}).$$

To show an upper bound on $N_{\geq d}$ we consider variables x with $\mathbb{E}[f_x] \leq \frac{d}{2e}$ of which there are $n \cdot (1 - \Theta(d^{1-\beta}))$ due to Definition 2.2. For these variables, by a Chernoff bound [21] it holds that

$$\Pr(Y_i = 1) \leq \Pr(f_{x_i} > 2e \cdot \mathbb{E}[f_{x_i}]) \leq 2^{-d}.$$

Now let $N'_{\geq d}$ be the number of variables with $\mathbb{E}[f_x] \leq \frac{d}{2e}$ and $f_x \geq d$. Thus, there exists a constant $c'' > 0$ such that

$$\mathbb{E}[N'_{\geq d}] \leq n \cdot 2^{-d} (1 - c'' \cdot d^{1-\beta}).$$

Due to negative association of the Y_i 's we can again use a Chernoff bound, yielding that w. h. p.,

$$N'_{\geq d} \leq n \cdot 2^{1-d} (1 - c'' \cdot d^{1-\beta}).$$

Observe that $1 - c''d^{1-\beta} = \mathcal{O}(1)$. Furthermore, for variables x with $\mathbb{E}[f_x] > \frac{d}{2e}$, we pessimistically assume $f_x \geq d$. This gives us

$$N_{\geq d} = \mathcal{O}(n \cdot (d^{1-\beta} + 2^{-d})) = \mathcal{O}(n \cdot d^{1-\beta}),$$

since $2^{-d} = \mathcal{O}(d^{1-\beta})$ for constant β . ◀

► **Corollary B.1.** *Let the variables w_i be power law distributed with exponent $\beta > 2$, and define $W_{\geq w} := \sum_{i \in [n]: w_i \geq w} w_i$. Then, $W_{\geq w} = \Theta(nw^{2-\beta})$.*

Proof. Observe that $\sum_{w' \geq w} w' = \sum_{i \in [n], w_i \geq w} w_i$. We apply Theorem 2.5 to obtain

$$\begin{aligned} \frac{1}{n} \sum_{i \in [n], w_i \geq w} w_i &= w \cdot F(w) + \int_w^{w_n} F(v) \, dv \\ &\leq \alpha_2 w^{2-\beta} + \left[\frac{\alpha_2}{2-\beta} v^{2-\beta} \right]_w^{w_n} \\ &\leq \alpha_2 \frac{\beta-1}{\beta-2} w^{2-\beta}. \end{aligned}$$

In a similar fashion, one may show that $\frac{1}{n} \sum_{i \in [n], w_i \geq w} w_i \geq \alpha_1 \frac{\beta-1}{\beta-2} (1 - o(1)) w^{2-\beta}$. \blacktriangleleft

► Corollary B.2. *Let Φ be a random k -SAT formula that follows Definition 2.1 with $k \geq 2$, $r = \frac{m}{n}$ and $\|\vec{p}\|_2^2 = o(1)$. With high probability, Φ is unsatisfiable if*

$$\left(1 - \frac{1}{2^k}\right)^r \left(2 - \exp\left(-\left(\frac{k}{2^k-1}\right)r(1+o(1))\right)\right) < 1.$$

Proof. We can upper-bound the left-hand side of the inequality as follows

$$\begin{aligned} &\left(1 - \frac{1}{2^k}\right)^r \left[\prod_{i=1}^n \left[2 - \left(1 - \frac{k \cdot p_i}{2^k-1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right)^m \right] \right]^{\frac{1}{n}} \\ &\leq \left(1 - \frac{1}{2^k}\right)^r \left[\frac{1}{n} \sum_{i=1}^n \left[2 - \left(1 - \frac{k \cdot p_i}{2^k-1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right)^m \right] \right] \\ &= \left(1 - \frac{1}{2^k}\right)^r \left[2 - \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{k \cdot p_i}{2^k-1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right)^m \right] \end{aligned}$$

by applying the inequality of arithmetic and geometric means. Since $\frac{k \cdot p_i}{2^k-1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}$ is upper-bounding a probability, we can assume it to be at most 1. It now holds that

$$\begin{aligned} \left(1 - \frac{k \cdot p_i}{2^k-1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right)^m &> \exp\left(-\frac{m}{\frac{(2^k-1)\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}{k \cdot p_i} - 1}\right) \\ &= \exp\left(-\frac{k \cdot p_i}{2^k-1} \frac{m}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)} (1 + o(1))\right), \end{aligned}$$

since $\|\vec{p}\|_2^2 = o(1)$ implies $\max_i (p_i) = o(1)$. By plugging this into the inequality from before and applying the inequality of arithmetic and geometric means again, we get

$$\begin{aligned} &\left(1 - \frac{1}{2^k}\right)^r \left[\prod_{i=1}^n \left[2 - \left(1 - \frac{k \cdot p_i}{2^k-1} \frac{1}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right)^m \right] \right]^{\frac{1}{n}} \\ &\leq \left(1 - \frac{1}{2^k}\right)^r \left[2 - \frac{1}{n} \sum_{i=1}^n \exp\left(-\left(\frac{k \cdot p_i}{2^k-1} \frac{m}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right) (1 + o(1))\right) \right] \\ &\leq \left(1 - \frac{1}{2^k}\right)^r \left[2 - \left(\prod_{i=1}^n \exp\left(-\left(\frac{k \cdot p_i}{2^k-1} \frac{m}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right) (1 + o(1))\right)\right)^{\frac{1}{n}} \right] \\ &= \left(1 - \frac{1}{2^k}\right)^r \left[2 - \exp\left(-\left(\frac{k}{2^k-1} \frac{r}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right)}\right) (1 + o(1))\right) \right]. \end{aligned}$$

For $\|\vec{p}\|_2^2 = o(1)$ this is roughly

$$\left(1 - \frac{1}{2^k}\right)^r \left(2 - \exp\left(-\left(\frac{k}{2^k-1}\right)r(1+o(1))\right)\right). \quad \blacktriangleleft$$

► **Theorem 4.3.** *Let Φ be a random k -SAT formula with $k \geq 2$ and $r = \frac{m}{n}$ that follows a power law distribution fulfilling Definition 2.3. Let further $N \in \mathbb{N}^+$ be any constant. If the power law exponent is $\beta > 2$, then Φ is w. h. p. unsatisfiable if*

$$\left(\left(1 - \frac{1}{2^k}\right)^r 2^{\frac{N-1}{N}} \prod_{l=1}^{N-1} \left[2 - \exp\left(-\left(1+o(1)\right)r \frac{k}{2^k-1} \frac{\beta-2}{\beta-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}\right)\right]^{\frac{1}{N}}\right) < 1.$$

Proof. We apply Theorem 4.1. If \vec{p} follows a power law distribution as in Definition 2.3, we can further simplify $\mathbb{E}[N_{SF}]$ to

$$\mathbb{E}[N_{SF}] \leq \left(1 - \frac{1}{2^k}\right)^m \prod_{i=1}^n \left[2 - \exp\left(-\frac{\frac{k \cdot p_i}{2^k-1} m}{\left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right) - \frac{k \cdot p_i}{2^k-1}}\right)\right]$$

using the inequality $1-x \geq e^{-\frac{x}{1-x}}$ which holds for all $x < 1$. We upper bound the probabilities p_i by choosing an integer $N \geq 2$ and dividing the set of variables into N buckets of equal size. For $i \in \left[\left\lceil \frac{l-1}{N}n \right\rceil + 1, \left\lceil \frac{l}{N}n \right\rceil\right]$ and $1 \leq l \leq N-1$ we can estimate

$$p_i \leq \frac{\left(\frac{n}{n - \left\lceil \frac{l}{N}n \right\rceil + 1}\right)^{\frac{1}{\beta-1}}}{\sum_{i=1}^n \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}}} = \frac{\left(\frac{N}{N-l}\right)^{\frac{1}{\beta-1}}}{\sum_{i=1}^n \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}}}.$$

The last bucket $i \in \left[\left\lceil \frac{N-1}{N}n \right\rceil + 1, n\right]$ is simply bounded by $2^{\frac{n}{N}}$ in the overall product. W.l.o.g. we assume that there are exactly $\frac{n}{N}$ variables in each bucket, as we could split the factor for $\left\lceil \frac{l}{N}n \right\rceil$ into appropriate parts which both obey the upper bound on p_i for bucket l and $l+1$ respectively. We can now upper-bound $\mathbb{E}[N_{SF}]$ by

$$\left(1 - \frac{1}{2^k}\right)^m 2^{\frac{n}{N}} \prod_{l=1}^{N-1} \left[2 - \exp\left(-\frac{\frac{k}{2^k-1} \cdot m \cdot \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}}{\left(\sum_{i=1}^n \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}}\right) \left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right) - \frac{k}{2^k-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}}\right)\right]^{\frac{n}{N}}$$

We are now interested in what happens to the expression in the exponent when n tends to infinity. First, $\sum_{i=1}^n \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}} \rightarrow \frac{\beta-1}{\beta-2}n$ for $\beta > 2$. Second, by Theorem 2.5 we have that $\|\vec{p}\|_j^j = \Theta\left(n^{-j \frac{\beta-2}{\beta-1}}\right) \rightarrow 0$ whenever $j > \beta-1$ and $\|\vec{p}\|_j^j = \Theta\left(n^{-j+1}\right) \rightarrow 0$ whenever $j \leq \beta-1$. Finally, for every constant N we have that $\frac{k}{2^k-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}$ is also constant. Using $m = r \cdot n$ we can thus simplify

$$\frac{\frac{k}{2^k-1} \cdot m \cdot \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}}{\left(\sum_{i=1}^n \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}}\right) \left(1 - \frac{1}{2} k^2 \|\vec{p}\|_2^2\right) - \frac{k}{2^k-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}} = (1+o(1)) r \frac{k}{2^k-1} \frac{\beta-2}{\beta-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}.$$

Plugging this into our inequality we get

$$\begin{aligned} \mathbb{E}[N_{SF}] &\leq \left(1 - \frac{1}{2^k}\right)^m 2^{\frac{n}{N}} \prod_{l=1}^{N-1} \left[2 - \exp\left(-\left(1+o(1)\right)r \frac{k}{2^k-1} \frac{\beta-2}{\beta-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}\right)\right]^{\frac{n}{N}} \\ &= \left(\left(1 - \frac{1}{2^k}\right)^r 2^{\frac{1}{N}} \prod_{l=1}^{N-1} \left[2 - \exp\left(-\left(1+o(1)\right)r \frac{k}{2^k-1} \frac{\beta-2}{\beta-1} \left(\frac{N}{l}\right)^{\frac{1}{\beta-1}}\right)\right]^{\frac{1}{N}}\right)^n \end{aligned}$$

This establishes Theorem 4.3. ◀

► **Lemma B.3.** *Let $\beta = \delta + 1 + \varepsilon$ for some $\varepsilon > 0$. For all $1 \leq l \leq n$, there is a constant c with*

$$\sum_{\substack{S \subseteq [n]: \\ |S|=l}} \prod_{i \in S} w_i^\delta \leq n^l \cdot c^{\frac{1}{l}}.$$

Proof. We begin by observing that the term on the left side of the equation is obviously monotone in w_i : If $\delta \geq 0$ ($\delta < 0$), then increasing (decreasing) w_i increases the sum. Thus, instead of considering the true distribution function $F(w)$, we may consider the upper (lower) bound on $F(w)$, see Equation (2.3). For the sake of brevity, we consider the distribution $\widehat{F}(w) = \alpha w^{1-\beta}$, where α is chosen to be either α_1 if $\delta \geq 0$, or α_2 otherwise.

To estimate this sum, we arrange the elements of S increasingly by weight, such that $w_{s_1} < w_{s_2} < \dots < w_{s_l}$. This gives us

$$\sum_{\substack{S \subseteq [n] \\ |S|=l}} \prod_{i \in S} w_i^\delta = \sum_{s_1=1}^{n-l+1} \left(w_{s_1}^\delta \sum_{s_2=s_1+1}^{n-l+2} \left(w_{s_2}^\delta \dots \sum_{s_l=s_{l-1}+1}^n w_{s_l}^\delta \right) \right).$$

We are now inductively estimating these sums, beginning with the innermost. Recall that $\widehat{F}(w) = \alpha w^{1-\beta}$. Let d be a large enough constant. We establish the following induction hypothesis:

$$\sum_{s_{l-i}=s_{l-i-1}+1}^{n-i} \left(w_{s_{l-i}}^\delta \sum_{s_{l-i+1}=s_{l-i}+1}^{n-i+1} \left(w_{s_{l-i+1}}^\delta \dots \sum_{s_l=s_{l-1}+1}^n w_{s_l}^\delta \right) \right) \leq n^{i+1} \cdot w_{s_{l-i-1}}^{(i+1) \cdot (\delta+1-\beta)} \cdot d^{i+1} \frac{1}{(i+1)!}$$

Now we apply Theorem 2.5 to prove the induction basis. For $i = 0$, we have

$$\begin{aligned} \sum_{s_l=s_{l-1}+1}^n w_{s_l}^\delta &\leq n \cdot \alpha w_{s_{l-1}}^{\delta+1-\beta} + n \int_{w_{s_{l-1}}}^{w_n} \alpha \delta w^{\delta-\beta} dw \\ &\leq n \cdot \alpha w_{s_{l-1}}^{\delta+1-\beta} + n \cdot \frac{\alpha \delta}{\beta-\delta-1} w_{s_{l-1}}^{\delta+1-\beta} \\ &= n \cdot \frac{\alpha(\beta-1)}{\beta-\delta-1} w_{s_{l-1}}^{\delta+1-\beta} \end{aligned}$$

as desired, since $\beta > \delta + 1$.

Now suppose the induction hypothesis holds for $i - 1$. For i we get

$$\sum_{s_{l-i}=s_{l-i-1}+1}^{n-i} \left(w_{s_{l-i}}^\delta \sum_{s_{l-i+1}=s_{l-i}+1}^{n-i+1} \left(w_{s_{l-i+1}}^\delta \dots \sum_{s_l=s_{l-1}+1}^n w_{s_l}^\delta \right) \right) \leq n^i \cdot d^i \frac{1}{i!} \sum_{s_{l-i}=s_{l-i-1}+1}^{n-i} w_{s_{l-i}}^{i \cdot (\delta+1-\beta) + \delta} \quad (\text{B.3})$$

To bound the sum, we distinguish two cases. If $i(\delta + 1 - \beta) + \delta \geq 0$, then

$$\begin{aligned}
 & \frac{1}{n} \sum_{s_{l-i}=s_{l-i-1}+1}^{n-i} w_{s_{l-i}}^{i \cdot (\delta+1-\beta) + \delta} \\
 & \leq \alpha w_{s_{l-i-1}}^{(i+1) \cdot (\delta+1-\beta)} + \int_{w_{s_{l-i-1}}}^{w_n} \alpha (i \cdot (\delta + 1 - \beta) + \delta) w^{i \cdot (\delta+1-\beta) + \delta - \beta} dw \\
 & = \alpha w_{s_{l-i-1}}^{(i+1) \cdot (\delta+1-\beta)} + \left[\frac{\alpha (i \cdot (\delta+1-\beta) + \delta)}{(i+1) \cdot (\delta+1-\beta)} w^{(i+1) \cdot (\delta+1-\beta)} \right]_{w_{s_{l-i-1}}}^{w_n} \\
 & = \left(\alpha + \frac{\alpha (i \cdot (\delta+1-\beta) + \delta)}{(i+1) \cdot (\delta+1-\beta)} \right) w_{s_{l-i-1}}^{(i+1) \cdot (\delta+1-\beta)} - \frac{\alpha (i \cdot (\delta+1-\beta) + \delta)}{(i+1) \cdot (\delta+1-\beta)} w_n^{(i+1) \cdot (\delta+1-\beta)} \\
 & = \frac{\alpha(\beta-1)}{(i+1)(\beta-\delta-1)} w_{s_{l-i-1}}^{(i+1) \cdot (\delta+1-\beta)} \cdot \left(1 - \frac{i \cdot (\delta+1-\beta) + \delta}{\beta-1} \left(\frac{w_n}{w_{s_{l-i-1}}} \right)^{(i+1) \cdot (\delta+1-\beta)} \right).
 \end{aligned}$$

For the integration, we need to make sure that the special case $-1 = i \cdot (\delta + 1 - \beta) + \delta - \beta$ does not occur. By rearranging, one can see that this is only true for $i = -1$, however, our weights begin at $i = 1$.

We now bound the error term that appears from the integration limit w_n . Note that we only need to consider the case where $i(\delta + 1 - \beta) < -\delta$, otherwise the error term is smaller than 1 and may simply be omitted.

Observe from Equation (B.3) that $w_{s_{l-i-1}} \leq w_{n-i}$. Further, by Equation (2.3) we have that $w_n = \Theta(n^{\frac{1}{\beta-1}})$. Similarly,

$$\frac{i}{n} = \widehat{F}(w_{n-i}) = \alpha w_{n-i}^{1-\beta}, \tag{B.4}$$

therefore $w_{n-i} = \Theta(1) \cdot \left(\frac{n}{i}\right)^{\frac{1}{\beta-1}}$. Therefore, we have

$$\frac{w_n}{w_{s_{l-i-1}}} \leq \frac{w_n}{w_{n-i}} = \Theta\left(i^{\frac{1}{\beta-1}}\right). \tag{B.5}$$

Recall that we are in the case where the error term is positive; and that the exponent $(\delta + 1 - \beta)$ is negative. Substituting the above inequality, we obtain

$$\begin{aligned}
 1 - \frac{i \cdot (\delta+1-\beta) + \delta}{\beta-1} \left(\frac{w_n}{w_{s_{l-i-1}}} \right)^{(i+1) \cdot (\delta+1-\beta)} & \leq 1 - \frac{i \cdot (\delta+1-\beta)}{\beta-1} i^{\frac{i+1}{\beta-1} \cdot (\delta+1-\beta)} \\
 & = 1 - \frac{\delta+1-\beta}{\beta-1} i^{1 + \frac{i+1}{\beta-1} \cdot (\delta+1-\beta)}
 \end{aligned}$$

By inspecting the exponent $1 + \frac{i+1}{\beta-1} \cdot (\delta + 1 - \beta)$, we observe that it is of order $\mathcal{O}(1)$. In particular, once i is a large enough constant, the exponent becomes negative. Therefore, we may conclude that

$$1 - \frac{\delta+1-\beta}{\beta-1} i^{1 + \frac{i+1}{\beta-1} \cdot (\delta+1-\beta)} = \mathcal{O}(1), \tag{B.6}$$

where the constant is not dependent on the iteration i . Thus, as d was chosen large enough,

$$\frac{1}{n} \sum_{s_{l-i}=s_{l-i-1}+1}^{n-i} w_{s_{l-i}}^{i \cdot (\delta+1-\beta) + \delta} \leq \frac{d}{i+1} w_{s_{l-i-1}}^{(i+1) \cdot (\delta+1-\beta)}, \tag{B.7}$$

Plugging this into inequality (B.3) proves the induction step.

Choosing $i = l - 1$ and setting $s_0 = 0$ yields

$$\sum_{\substack{S \subseteq [n]: \\ |S|=l}} \prod_{i \in S} w_i^\delta \leq n^l \cdot w_1^{l \cdot (\delta+1-\beta)} \cdot d^l \frac{1}{l!}.$$

Since $w_1^{\delta+1-\beta} = \Theta(1)$, we can choose an appropriate constant c such that the statement holds. ◀