Going after the k-SAT Threshold

Amin Coja-Oghlan* Konstantinos Panagiotou†

December 7, 2012

Abstract

Random k-SAT is the single most intensely studied example of a random constraint satisfaction problem. But despite substantial progress over the past decade, the threshold for the existence of satisfying assignments is not known precisely for any $k \geq 3$. The best current results, based on the second moment method, yield upper and lower bounds that differ by an additive $k \cdot \frac{\ln 2}{2}$, a term that is unbounded in k (Achlioptas, Peres: STOC 2003). The basic reason for this gap is the inherent asymmetry of the Boolean value 'true' and 'false' in contrast to the perfect symmetry, e.g., among the various colors in a graph coloring problem. Here we develop a new *asymmetric second moment method* that allows us to tackle this issue head on for the first time in the theory of random CSPs. This technique enables us to compute the k-SAT threshold up to an additive $\ln 2 - \frac{1}{2} + O(1/k) \approx 0.19$. Independently of the rigorous work, physicists have developed a sophisticated but non-rigorous technique called the "cavity method" for the study of random CSPs (Mézard, Parisi, Zecchina: Science 2002). Our result matches the best bound that can be obtained from the so-called "replica symmetric" version of the cavity method, and indeed our proof directly harnesses parts of the physics calculations.

1 Introduction

Since the early 2000s physicists have developed a sophisticated but highly non-rigorous technique called the "cavity method" for the study of random constraint satisfaction problems. This method allowed them to put forward a very detailed *conjectured* picture according to which various phase transitions affect both computational and structural properties of random CSPs. In addition, the cavity method has inspired new message passing algorithms called *Belief/Survey Propagation guided decimation*. Over the past few years there has been significant progress in turning bits and pieces of the physics picture into rigorous theorems. Examples include results on the interpolation method [2, 7] or the geometry of the solution space [1, 28, 29] and their algorithmic implications [3, 9].

In spite of this progress, substantial gaps remain. Perhaps most importantly, in most random CSPs the threshold for the existence of solutions is not known precisely. In the relatively simple case of the random k-NAESAT ("Not-All-Equal-Satisfiability") problem the difference between the best current lower and upper bounds is as tiny as $2^{-\Omega(k)}$ [11]. By contrast, in random graph k-coloring, a problem already studied by Erdős and Rényi in the 1960s, the best current bounds differ by $\Theta(\ln k)$ [5]. Hence, the difference is *unbounded* in terms of the number of colors. Even worse, in random k-SAT the gap is as big as $\Theta(k)$ [6]. Yet random k-SAT is probably the single most important example of a random CSP, not least due to the great amount of experimental and algorithmic work conducted on it (e.g., [22, 24]).

The reason for the large gap in random k-SAT is that the satisfiability problem lacks a certain *symmetry property*. This property is vital to the current rigorous proof methods, particularly the *second moment method*, on which most of the previous work is based (e.g., [4, 5, 6]). More precisely, in random graph coloring the different colors all play the exact same role: for any proper coloring of a graph, another proper coloring can be obtained by simply permuting the color classes (e.g., color all red vertices blue and vice versa). Similarly, in k-NAESAT, where the requirement is that in each clause at least one literal must be true and at least one false, the binary inverse of any NAE-solution is a

*Goethe University, Mathematics Institute, Frankfurt 60054, Germany, acoghlan@math.uni-frankfurt.de. Supported by ERC Starting Grant 278857–PTCC (FP7).

†University of Munich, Mathematics Institute, Theresienstr. 39, 80333 München, Germany, kpanagio@math.lmu.de, Supported by DFG grant PA 2080/2-1.

NAE-solution as well. By contrast, in k-SAT there is an inherent asymmetry between the Boolean values 'true' and 'false'.

As has been noticed in prior work [4, 6], the second moment method is fundamentally ill-posed to deal with such asymmetries. Roughly speaking, the second moment method is based on the assumption that in a random CSP instance, two randomly chosen solutions are perfectly uncorrelated. But in random k-SAT, this is simply not the case. Indeed, suppose that a variable x appears much more often positively than negatively throughout the formula. Then it seems reasonable to expect that most satisfying assignments set x to 'true', thereby satisfying all clauses where x appears positively. More generally, define the *majority vote* σ_{maj} to be the assignment that sets variable x to true if it appears more often positively than negatively, and to false otherwise. Then we expect that the satisfying assignments of a random formula "gravitate toward" σ_{maj}. Unfortunately, the correlations among satisfying assignments induced by this drift toward σ_{maj} doom the second moment method. Previously this issue was sidestepped by symmetrizing the problem artificially [4, 6]. But this inevitably leaves a $\Theta(k)$ gap.

The main contribution of the present work is a new *asymmetric second moment method* that enables us to tackle this problem head on. A key feature of this method is that we harness the Belief Propagation calculation from physics, called the "replica symmetric case" of the cavity method in physics jargon. We are going to employ Belief Propagation directly as an "educated guess" in the design the random variable upon which our proof is based in order to quantify how much a typical satisfying assignment leans toward σ_{maj}.

This is in contrast to most prior work on the subject, where individual statements hypothesized on the basis of physics arguments were proved via completely different methods (with the notable exception of the interpolation technique [2, 7, 17]). Hence, we view the present work as a pivotal step in the long-term effort of providing a rigorous foundation for the physicists' cavity method. In fact, the general approach developed here does not hinge on particular properties of the k-SAT problem, and thus we expect that the technique will extend to other asymmetric problems as well. Examples include not only other random CSPs that are asymmetric per se, but also instances of random problems that arise at intermediate steps of message passing algorithms such as *Belief/Survey Propagation guided decimation*, even if the initial problem is symmetric. In particular, we believe that getting a handle on asymmetric problems is a necessary step to analyze such message passing algorithms accurately.

To state our results precisely, we let $k \geq 3, n > 0$ be integers and we let $V = \{x_1, \ldots, x_n\}$ be a set of n Boolean variables. Further, let $\mathbf{\Phi} = \mathbf{\Phi}_k(n, m)$ denote a Boolean formula with m clauses of length k over the variables V chosen uniformly at random among all $(2n)^{km}$ such formulas. Let $r = m/n$ denote the *density*. We say that an event occurs *with high probability* ('w.h.p.') if its probability tends to 1 as $n \to \infty$.

Friedgut [18] showed that for any $k \geq 3$ there exists a *threshold sequence*[1] $r_{k-\mathrm{SAT}}(n)$ such that for any (fixed) $\varepsilon > 0$ w.h.p. $\mathbf{\Phi}$ is satisfiable if $m/n < (1-\varepsilon)r_{k-\mathrm{SAT}}(n)$, while for $m/n > (1+\varepsilon)r_{k-\mathrm{SAT}}(n)$ $\mathbf{\Phi}$ is unsatisfiable w.h.p.

Upper bounds on $r_{k-\mathrm{SAT}}$ can be obtained via the *first moment method*. The best current ones [17, 23] are

$$r_{k-\mathrm{SAT}} \leq r_{\mathrm{upper}} = 2^k \ln 2 - (1+\ln 2)/2 + o_k(1), \tag{1}$$

where $o_k(1)$ hides a term that tends to 0 for large k. The best prior lower bound is due to Achlioptas and Peres [6], who used a "symmetric" second moment argument to show

$$r_{k-\mathrm{SAT}} \geq r_{\mathrm{bal}} = 2^k \ln 2 - k \cdot \frac{\ln 2}{2} - \left(1 + \frac{\ln 2}{2}\right) + o_k(1). \tag{2}$$

The bounds (1) and (2) leave an additive gap of $k \cdot \frac{\ln 2}{2} + \frac{1}{2} + o_k(1)$, i.e., the gap is unbounded in terms of k.

Theorem 1.1 *There is $\varepsilon_k = o_k(1)$ such that*

$$r_{k-\mathrm{SAT}} \geq r_{\mathrm{BP}} = 2^k \ln 2 - \frac{3\ln 2}{2} - \varepsilon_k. \tag{3}$$

Achlioptas and Peres asked whether the gap $r_{\mathrm{upper}} - r_{k-\mathrm{SAT}}$ is bounded by an absolute constant (independent of k). Theorem 1.1 answers this question, reducing the gap to $\ln 2 - \frac{1}{2} \approx 0.19$. No attempt at optimizing the error term ε_k has been made, but our proofs yield rather directly that $\varepsilon_k = O(1/k)$.

[1] It is widely conjecture but as yet unproved that $r_{k-\mathrm{SAT}}(n)$ converges for any $k \geq 3$.

Apart from the quantitative improvement, the main point of this paper is that we manage to solve the problem of asymmetry in random CSPs for the first time. To explain this point, we start by discussing what we mean by asymmetry and how it derails the second moment method. That this is so was already intuited in [4, 6]. In the next section, we are going to verify and elaborate on those discussions.

2 Asymmetry and the second moment method

The second moment method. In general, the second moment method works as follows. Suppose that $Z = Z(\mathbf{\Phi})$ is a non-negative random variable such that $Z > 0$ only if $\mathbf{\Phi}$ is satisfiable. Moreover, suppose that for some density $r > 0$ there is a number $C = C(k) > 0$ that may depend on k but not on n such that

$$0 < \mathrm{E}\left[Z^2\right] \leq C \cdot \mathrm{E}\left[Z\right]^2. \tag{4}$$

We claim that then $r_{k-\mathrm{SAT}} \geq r$. Indeed, the *Paley-Zygmund inequality*

$$\mathrm{P}\left[Z > 0\right] \geq \frac{\mathrm{E}\left[Z\right]^2}{\mathrm{E}\left[Z^2\right]} \tag{5}$$

implies that $\mathrm{P}\left[\mathbf{\Phi} \text{ is satisfiable}\right] \geq \mathrm{P}\left[Z > 0\right] \geq 1/C$. Because the right hand side remains bounded away from 0 as $n \to \infty$, the following simple consequence of Friedgut's sharp threshold result implies $r_{k-\mathrm{SAT}} \geq r$.

Lemma 2.1 ([18]) *Let $k \geq 3$. If for some r we have*

$$\liminf_{n\to\infty} \mathrm{P}\left[\mathbf{\Phi} \textit{ is satisfiable}\right] > 0,$$

then $r_{k-\mathrm{SAT}} \geq r - o(1)$.

Hence, we "just" need to find a random variable that satisfies (5). Let $\mathcal{S}(\mathbf{\Phi})$ denote the set of satisfying assignments; then certainly $Z = |\mathcal{S}(\mathbf{\Phi})|$ is the most obvious choice. However, this "vanilla" second moment argument turns out to fail spectacularly. We need to understand why.

Asymmetry and the majority vote. The origin of the problem is that k-SAT is asymmetric in the following sense. Suppose that all we know about the random formula $\mathbf{\Phi}$ is for each variable x the number d_x of times that x appears as a positive literal in the formula, and the number $d_{\neg x}$ of negative occurrences. Then our best stab at constructing a satisfying assignment seems to be the "majority vote" assigment σ_{maj} where we set x to true if $d_x > d_{\neg x}$ and to false otherwise. Indeed, by maximizing the total number of true literal occurrences, of which a satisfying assignment must put one in every clause, σ_{maj} also maximizes the probability of being satisfiable.

Our proof of Theorem 1.1 allows us to formalize this observation, thereby verifying a conjecture from [6]. Let $\mathrm{dist}(\cdot,\cdot)$ denote the Hamming distance.

Corollary 2.2 *There is a number $\delta = \delta(k) > 0$ such that for $2^k/k < r < r_{\mathrm{BP}}$ w.h.p. we have*

$$\sum_{\sigma\in\mathcal{S}(\mathbf{\Phi})} \frac{\mathit{dist}(\sigma,\sigma_{maj})}{|\mathcal{S}(\mathbf{\Phi})|} \leq \left(\frac{1}{2} - \delta\right)\cdot n. \tag{6}$$

Hence, the average Hamming distance of $\sigma \in \mathcal{S}(\mathbf{\Phi})$ from σ_{maj} is strictly smaller than $n/2$, i.e., the set $\mathcal{S}(\mathbf{\Phi})$ is "skewed toward" σ_{maj} w.h.p.

This asymmetry dooms the second moment method. To see why, let

$$w_{\mathrm{maj}} = w_{\mathrm{maj}}(\mathbf{\Phi}) = \sum_{x\in V} \frac{\max\left\{d_x, d_{\neg x}\right\}}{km}$$

denote the ***majority weight*** of $\mathbf{\Phi}$. Then the larger w_{maj}, the more likely σ_{maj} and assignments close to it are to be satisfying. In effect, the bigger w_{maj}, the more satisfying assignments we expect to have. The consequence of this

is that the number $|\mathcal{S}(\mathbf{\Phi})|$ of satisfying assignments behaves like a "lottery": its expectation is driven up by a tiny fraction of "lucky" formulas with w_{maj} much bigger than expected.

Let us highlight this tradeoff, as it is characteristic of the kind of trouble that asymmetry causes. For $\xi > 0$ independent of n but sufficiently small it turns out that for a certain constant $c > 0$,

$$\mathrm{P}\left[w_{\mathrm{maj}} \sim \mathrm{E}\left[w_{\mathrm{maj}}\right] + \xi\right] = \exp\left[-(c\xi^2 + O(\xi^3))n\right]. \tag{7}$$

That is, the probability is exponentially small but, like in the Chernoff bound, the exponent is a *quadratic* function of ξ. By comparison, increasing the majority weight by ξ boosts the expected number of satisfying assignments by a *linear* exponential factor: there is $c' > 0$ such that

$$\mathrm{E}\left[\,|\mathcal{S}(\mathbf{\Phi})| \,\middle|\, w_{\mathrm{maj}} \sim \mathrm{E}\left[w_{\mathrm{maj}}\right] + \xi\right] \quad = \quad \exp\left[(c'\xi + O(\xi^2))n\right] \cdot \mathrm{E}\left[|\mathcal{S}(\mathbf{\Phi})| \,\middle|\, w_{\mathrm{maj}} \sim \mathrm{E}\left[w_{\mathrm{maj}}\right]\right]. \tag{8}$$

The exponent in (8) is linear because for a typical assignment τ at distance $(\frac{1}{2} - \delta)n$ from σ_{maj} increasing w_{maj} by ξ boosts the number of literals that are true under τ by $2\delta\xi \cdot km$, a term that is linear in ξ.

Since the exponent is linear in (8) but quadratic in (7), there is a (small but) strictly positive $\xi > 0$ such that the "gain" $\exp\left[(c'\xi + O(\xi^2))n\right]$ in the expected number of satisfying assignments exceeds the "penalty" $\exp\left[-(c\xi^2 + O(\xi^3))n\right]$ for deviating from $\mathrm{E}\left[w_{\mathrm{maj}}\right]$. With little extra work, this observation leads to

Lemma 2.3 *For any $k \geq 3$ and $r > 2^k/k$ we have*

$$|\mathcal{S}(\mathbf{\Phi})| \leq \exp\left(-\Omega(4^{-k}) \cdot n\right) \cdot \mathrm{E}\left[|\mathcal{S}(\mathbf{\Phi})|\right] \quad \textit{w.h.p.}$$

Lemma 2.3 entails rather easily that the "vanilla" second moment argument fails dramatically. Indeed, as already noticed in [4, 6], we have $\mathrm{E}\left[|\mathcal{S}(\mathbf{\Phi})|^2\right] \geq \exp(\Omega(n)) \cdot \mathrm{E}\left[|\mathcal{S}(\mathbf{\Phi})|\right]^2$. Hence, we miss our mark (4) by an exponential factor. But Lemma 2.3 is witness to an even worse failure: not only does (4) fail to hold, but even the normally much more dependable *first* moment overshoots the "actual" number of satisfying assignments by an exponential factor! (Lemma 2.3 is an improvement of an observation from [1], showing that $|\mathcal{S}(\mathbf{\Phi})| \leq \exp(-\xi n)\mathrm{E}\left[|\mathcal{S}(\mathbf{\Phi})|\right]$ w.h.p. for some tiny $\xi = \xi(k) > 0$; we conjecture that the 4^{-k} term in Lemma 2.3 is tight.)

In summary, the drift toward σ_{maj} and the resulting fluctuations of the majority weight induce a tremendous source of variance, derailing the "vanilla" second moment argument.

Balanced assignments. A natural way to sidestep this issue is to work with a 'symmetric' subset of $\mathcal{S}(\mathbf{\Phi})$. Perhaps the most obvious choice is the set $\mathcal{S}_{\mathrm{NAE}}(\mathbf{\Phi})$ of NAE-solutions. In a landmark paper, Achlioptas and Moore [4] proved that indeed there is $C = C(k) > 0$ such that for $Z_{\mathrm{NAE}} = |\mathcal{S}_{\mathrm{NAE}}(\mathbf{\Phi})|$ we have

$$\mathrm{E}\left[Z_{\mathrm{NAE}}^2\right] \leq C \cdot \mathrm{E}\left[Z_{\mathrm{NAE}}\right]^2 \quad \text{for } r \leq 2^{k-1}\ln 2 - O_k(1). \tag{9}$$

As we saw above (cf. Lemma 2.1), this implies that $r_{k-\mathrm{SAT}} \geq 2^{k-1}\ln 2 - O(1)$. However, a simple (first moment) calculation shows that for $r > 2^{k-1}\ln 2$, the set $\mathcal{S}_{\mathrm{NAE}}(\mathbf{\Phi})$ is empty w.h.p. Thus, the idea of working with NAE-solutions stops working at $r \sim 2^{k-1}\ln 2$, about a factor of two below the satisfiability threshold.

Achlioptas and Peres [6] obtained a better bound by precipitating symmetry in a more subtle manner. Let us call $\sigma \in \{0,1\}^n$ ***balanced*** if under σ out of the km literal occurrences in $\mathbf{\Phi}$ *exactly half* are true (i.e., $\frac{km}{2} \pm 1$). Thus, balanced assignments are expressly forbidden from pandering to σ_{maj}. Now, let $\mathcal{S}_{\mathrm{bal}}(\mathbf{\Phi})$ be the set of all balanced satisfying assignments, and set $Z_{\mathrm{bal}} = |\mathcal{S}_{\mathrm{bal}}(\mathbf{\Phi})|$. Achlioptas and Peres used a clever weighting scheme to prove that

$$\mathrm{E}\left[Z_{\mathrm{bal}}^2\right] \leq C \cdot \mathrm{E}\left[Z_{\mathrm{bal}}\right]^2 \qquad \text{for } r \leq r_{\mathrm{bal}} \quad \text{(cf. (2))}. \tag{10}$$

As before, this implies that $r_{k-\mathrm{SAT}} \geq r_{\mathrm{bal}}$ (Lemma 2.1).

Yet as in the case of NAE-solutions, balanced satisfying assignments cease to exist way before the satisfiability threshold. Indeed, Achlioptas and Peres observed that $\mathcal{S}_{\mathrm{bal}}(\mathbf{\Phi}) = \emptyset$ for $r > 2^k \ln 2 - k\frac{\ln 2}{2}$ w.h.p. In effect, to close in further on $r_{k-\mathrm{SAT}}$ we will have to accommodate assignments that lean toward σ_{maj}. How can this be accomplished?

A quick fix? We saw that to make an asymmetric second moment argument work, we need to rule out fluctuations of the majority weight. A sensible way of implementing this is by actually fixing the entire vector $\boldsymbol{d} = (d_x, d_{\neg x})_{x \in V}$

that counts the positively/negatively occurrences of each variable. More precisely, given a non-negative integer vector $\boldsymbol{d} = (d_x, d_{\neg x})_{x \in V}$ with $\sum_{x \in V} d_x + d_{\neg x} = km$ let $\boldsymbol{\Phi_d}$ denote a uniformly random k-CNF in which each variable x appears d_x times positively and $d_{\neg x}$ times negatively. Then we can split the generation of a random formula $\boldsymbol{\Phi}$ into two steps:

First, choose the occurrence vector $\boldsymbol{d}$ randomly from the "correct" distribution $\boldsymbol{D}$.

Then, choose a random formula $\boldsymbol{\Phi_d}$.

The "correct" $\boldsymbol{D}$ is as follows. Let $\boldsymbol{e} = (e_x, e_{\neg x})_{x \in V}$ be a family of independent Poisson variables with mean $kr/2$ each. Moreover, let $\mathcal{E}$ be the event that $\sum_{x \in V} e_x + e_{\neg x} = km$. Let $\boldsymbol{D}$ be the conditional distribution of $\boldsymbol{e}$ given $\mathcal{E}$. Then standard arguments show that the outcome of first choosing $\boldsymbol{d}$ and then $\boldsymbol{\Phi_d}$ is exactly the uniformly random $\boldsymbol{\Phi}$.

The point of generating $\boldsymbol{\Phi}$ in two steps as above is that given the outcome $\boldsymbol{d}$ of the first step, the majority weight is *fixed*. Hence, if we could show that *given* a "typical" $\boldsymbol{d}$, the second moment succeeds for $|\mathcal{S}(\boldsymbol{\Phi_d})|$ we would obtain a lower bound on r_{k-SAT}. Unfortunately, matters are not so simple.

Lemma 2.4 *W.h.p. for a vector $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ we have* $\mathrm{E}[|\mathcal{S}(\boldsymbol{\Phi_d})|^2] \geq \exp(\Omega(n)) \cdot \mathrm{E}[|\mathcal{S}(\boldsymbol{\Phi_d})|]^2$.

Let us stress the two levels of randomness in Lemma 2.4. First, there is the choice of $\boldsymbol{d}$. Then, for a *given* $\boldsymbol{d}$, we compare $\mathrm{E}[|\mathcal{S}(\boldsymbol{\Phi_d})|^2]$ and $\mathrm{E}[|\mathcal{S}(\boldsymbol{\Phi_d})|]^2$. Of course, both of these quantities depend on $\boldsymbol{d}$, and we find that w.h.p. $\boldsymbol{d}$ is such that the first exceeds the second by an exponential factor.

The explanation for this is that even if we fix $\boldsymbol{d}$, various other types of fluctuations remain, turning $|\mathcal{S}(\boldsymbol{\Phi_d})|$ into a "lottery". For instance, even given $\boldsymbol{d}$ the number of clauses that are unsatisfied under σ_{maj} fluctuates. Hence, the inherent asymmetry of k-SAT puts not only the majority weight but also various other parameters on a slippery slope. What we need is a way of controlling all these fluctuations simultaneously. We will present our solution in Section 5.

Catching the k-SAT threshold? Before we come to that, let us discuss what it would take to eliminate the (small but non-zero) gap left by Theorem 1.1, i.e., how far we are from "catching" the k-SAT threshold. The physicists' cavity method comes in two installments. The (relatively speaking) simpler "replica symmetric" version is based on Belief Propagation. Theorem 1.1 provides a rigorous proof of the best possible bound on the k-SAT threshold that can be obtained from this version of the cavity method (up to possibly the precise error term ε_k) [25].

Unfortunately, for $r > r_{\mathrm{BP}}$ the replica symmetric version (and in particular the Belief Propagation predictions that we depend upon) are conjectured to break down. According to the more sophisticated "1-step replica symmetry breaking" (1RSB) version of the cavity method, the reason for this is that at $r \sim r_{\mathrm{BP}}$ a new type of correlation amongst satisfying assignments arises. To deal with these correlations, the physics methods replace Belief Propagation by the *much* more intricate Survey Propagation technique.

In [11] we managed to prove rigorously that the 1RSB prediction for the random k-NAESAT threshold is correct (up to an additive $2^{-\Omega(k)}$). However, [11] depends *heavily* on the fact that k-NAESAT is symmetric. While it would be very interesting to combine the merits of the present paper with those of [11], this appears to be quite challenging. Thus, putting the 1RSB calculation for random k-SAT on a rigorous foundation remains an important open problem. That said, we believe that any such attempt would need to build upon the techniques developed in this paper.

3 Related work

The interest in random k-SAT originated largely from the *experimental* observation that there seems to be a sharp threshold for satisfiability and, moreover, that for certain densities $r < r_{k-SAT}$ no polynomial time algorithm is known to find a satisfying assignment w.h.p. [22, 24]. Currently, the precise k-SAT threshold is known (rigorously) only in two cases. Chvatal and Reed [8] and Goerdt [21] proved independently that $r_{2-\mathrm{SAT}} = 1$. Of course, 2-SAT is special because there is a simple criterion for (un)satisfiability, which enables the proofs of [8, 21]. Unsurprisingly, these methods do not extend to $k > 2$. Additionally, the threshold is known precisely when $k > \log_2 n$, i.e., the clause length *diverges* as a function of n [20]. In this case, the problem of asymmetry evaporates because the majority weight is sufficiently concentrated for the "vanilla" second moment method to succeed. (Note that Proposition 2.3 holds for

any *fixed* k, but not for $k = k(n) \to \infty$.) The issue of asymmetry also disappears in the case of *strongly regular* formulas [31] where for some fixed d we have $d_x = d_{\neg x} = d$ for all $x \in V$.

Also in random k-XORSAT (random linear equations mod 2) the threshold for the existence of solutions is known precisely [14]. The proof relies on computing the second moment of the number of solutions (after the instance has been stripped down to a suitable core). In contrast to random k-SAT, the random k-XORSAT problem is symmetric (cf. Remark 5.5 below), albeit in a more subtle way than k-NAESAT.

Other problems where the second moment method succeeds are symmetric as well. Pioneering the use of the second moment method in random CSPs, Achlioptas and Moore [4] computed the random k-NAESAT threshold within an additive $1/2$. By enhancing this argument with insights from physics this gap can be narrowed to a mere $2^{-\Omega(k)}$ [11, 12]. Moreover, the best current bounds on the random (hyper)graph k-colorability thresholds are based on "vanilla" second moment arguments as well [5, 15]. In summary, in all the previous second moment arguments, the issue of asymmetry either did not appear at all by the nature of the problem [4, 5, 11, 12, 14, 15, 20], or it was sidestepped [6].

The best current algorithms for random k-SAT find satisfying assignments w.h.p. for densities up to $1.817 \cdot 2^k/k$ (better for small k) resp. $2^k \ln(k)/k$ (better for large k) [9, 19], a factor of $\Theta(k/\ln k)$ below the satisfiability threshold. By comparison, the Lovász Local Lemma and its algorithmic version succeed up to $r = \Theta(2^k/k^2)$ [30].

Apart from experimental work [24], very little is known about the physics-inspired message passing algorithms ("Belief/Survey Propagation guided decimation") [27]. The most basic variant of Belief Propagation guided decimation is known to fail w.h.p. on random formulas if $r > c \cdot 2^k/k$ for some constant $c > 0$ [10]. However, it is conceivable that Survey Propagation and/or other variants of Belief Propagation perform better.

4 Preliminaries

We shall make repeated use of the following local limit theorem for the sums of independent random variables, see [16] and [11].

Lemma 4.1 *Let $X_1, \ldots, X_n$ be independent random variables with support on $\mathbf{N}_0$ with probability generating function $P(z)$. Let $\mu = \mathrm{E}[X_1]$ and $\sigma^2 = \mathrm{Var}[X_1]$. Assume that $P(z)$ is an entire and aperiodic function. Then, uniformly for all $T_0 < \alpha < T_\infty$, where $T_x = \lim_{z \to x} \frac{zP'(z)}{P(z)}$, as $n \to \infty$*

$$\Pr[X_1 + \cdots + X_n = \alpha n] = (1 + o(1)) \frac{1}{\zeta \sqrt{2\pi n \xi}} \left(\frac{P(\zeta)}{\zeta^\alpha} \right)^n, \tag{11}$$

where ζ and ξ are the solutions to the equations

$$\frac{\zeta P'(\zeta)}{P(\zeta)} = \alpha \qquad \text{and} \qquad \xi = \frac{d^2}{dz^2} \left(\ln P(z) - \alpha \ln z \right) \Big|_{z=\zeta}. \tag{12}$$

Moreover, there is a $\delta_0 > 0$ such that for all $0 \leq |\delta| \leq \delta_0$ the following holds. If $\alpha = \mathrm{E}[X_1] + \delta\sigma$, then

$$\Pr[X_1 + \cdots + X_n = \alpha n] = (1 + O(\delta)) \frac{1}{\sqrt{2\pi n \sigma}} e^{(-\delta^2/2 + O(\delta^3))n}. \tag{13}$$

From this we can rather easily derive the following well-known statement about the rate function of the binomial distribution.

Lemma 4.2 *Let $0 < p, q < 1$. Let*

$$\psi(p, q) \quad = \quad -q \ln \left(\frac{q}{p} \right) - (1 - q) \ln \left(\frac{1 - q}{1 - p} \right),$$

If p, q remain fixed as $n \to \infty$, then

$$\mathrm{P}\left[\mathrm{Bin}(n, p) = qn \right] = \Theta(n^{-1/2}) \exp\left[\psi(p, q) n \right].$$

The following form of the chain rule will prove useful.

Lemma 4.3 *Let $g : \mathbf{R}^a \to \mathbf{R}^b$ and $f : \mathbf{R}^b \to \mathbf{R}$ be of class C^2, i.e, with continuous second derivatives. Then for any $x_0 \in \mathbf{R}^a$ and with $y_0 = g(x_0)$ we have for any $i, j \in [a]$*

$$\left.\frac{\partial^2 f \circ g}{\partial x_i \partial x_j}\right|_{x_0} = \sum_{k=1}^{b} \left.\frac{\partial f}{\partial y_k}\right|_{y_0} \left.\frac{\partial^2 g_k}{\partial x_i \partial x_j}\right|_{x_0} + \sum_{k,l=1}^{b} \left.\frac{\partial^2 f}{\partial y_k \partial y_l}\right|_{y_0} \left.\frac{\partial g_k}{\partial x_i}\right|_{x_0} \left.\frac{\partial g_l}{\partial x_j}\right|_{x_0}.$$

Finally, we need the following version of the inverse function theorem that states under which conditions a given system of equations can be solved around a specific point u. A detailed exposition can be found in [32].

Lemma 4.4 *Let $U \subset \mathbf{R}^h$ be open and let $f \in C^1(U)$. Assume that $u \in U$ and $\lambda > 0$ are such that*

$$\left\{x \in \mathbf{R}^h : \|x - u\|_2 \leq \lambda\right\} \subset U.$$

Let $Df(x)$ be the Jacobian matrix of f at x, id *the identity matrix, and $\|\cdot\|$ denote the operator norm over $L^2(\mathbf{R}^h)$. Assume that $Df(u) = \mathrm{id}$ and*

$$\|Df(x) - \mathrm{id}\| \leq \frac{1}{3} \qquad \textit{for all } x \in \mathbf{R}^h \textit{ such that } \|x - u\|_2 \leq \lambda,$$

Then for each $y \in \mathbf{R}^h$ such that $\|y - f(u)\| \leq \lambda/2$ there is precisely one $x \in \mathbf{R}^h$ such that $\|x - u\| \leq r$ and $f(x) = y$. Furthermore, the inverse map f^{-1} is C^1 on $\left\{x \in \mathbf{R}^h : \|x - u\|_2 < \lambda\right\}$, and $Df^{-1}(x) = (Df(x))^{-1}$ on this set.

Notation. We will generally assume that $n > n_0, k > k_0$ for certain large enough constants n_0, k_0. We are going to use the asymptotic symbols $O(f(x))$, $\Omega(f(x))$, etc. It is understood that the asymptotic is with respect to the parameter x of the function $f(x)$. Thus, if f is a function of n, then the asymptotic notation refers to the limit $n \to \infty$, and if f is a function of k, then the notation refers to k being large. We use the following convention for the O-notation in the case that f is a constant: we let $O(1)$ be a term that remains bounded in the limit of large n, but that may by unbounded in terms of k. By constrast, $O_k(1)$ refers to a term that remains bounded both in the limit of large k and large n. Expressions such as $o_k(1)$ are to be interpreted analogously. Generally, all asymptotics are *uniform* in the various other parameters (such as the degree sequence $\boldsymbol{d}$ or r). For a function $f(k) > 0$ use the symbol $\tilde{O}(f(k))$ to denote a function $g(k)$ such that for some constant $c > 0$ we have $g(k) = O(f(k) \cdot \ln^c f(k))$. For vectors ξ, η we use the symbol

$$\eta \doteq \xi$$

to denote the fact that $\|\xi - \eta\|_\infty \leq O(1/n)$.

Let $V = \{x_1, \ldots, x_n\}$ and let $L = \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$. For a literal $l \in L$ we let $|l|$ denote the underlying variable. Moreover, $\mathrm{sign}(l) = 1$ if l is a positive literal, and $\mathrm{sign}(l) = -1$ otherwise. For a k-CNF Φ we let Φ_i denote the ith clause of Φ and Φ_{ij} the jth literal of Φ_i.

From here on out, we let

$$r = 2^{-k} \ln 2 - \rho \quad \text{with} \quad \rho = \frac{3}{2} \ln 2 - \varepsilon_k \tag{14}$$

for some sequence $\varepsilon_k = o_k(1)$ that tends to 0 sufficiently slowly.

5 The random variable

5.1 The construction

Our goal is to make the second moment method work for a random variable that counts "asymmetric" satisfying assignments. In this section, we develop this random variable. The starting point, and the key ingredient, is simply a

map $p : \mathbf{Z} \to [0,1]$. For the sake of clarity, we start by setting up the framework for generic maps p; below we will use the Belief Propagation formalism to pick the "optimal" p.

The idea is that p prescribes how strongly the assignments that we work with lean toward the majority vote. Informally speaking, we are going to work with assignments such that a variable x that occurs d_x times positively and $d_{\neg x}$ times negatively has a $p(d_x - d_{\neg x})$ chance of being set to 'true'. Before we give a formal definition, we need to fix the number of times that each variable appears positively or negatively.

Fixing the majority weight. As we saw in Section 2, in order to make the second moment argument work, we need to rule out fluctuations of the majority weight. To achieve this, we follow the strategy outlined in Section 2. That is, we are going to work with formulas $\boldsymbol{\Phi_d}$ with a given vector $\boldsymbol{d} = (d_x, d_{\neg x})_{x \in V}$ of occurrence counts, where each variable x appears precisely d_x times positively and $d_{\neg x}$ times negatively. As in Section 2, we let $\boldsymbol{D}$ denote the (conditional Poisson) distribution over sequences $\boldsymbol{d}$ such that first choosing $\boldsymbol{d}$ from $\boldsymbol{D}$ and then generating $\boldsymbol{\Phi_d}$ is equivalent to choosing a k-CNF $\boldsymbol{\Phi}$ uniformly at random.

Fixing the marginals. Now, fix one such vector $\boldsymbol{d}$. Then the map $p : \mathbf{Z} \to [0,1]$ induces a map $p_{\boldsymbol{d}}$ from the set $L = \{x, \neg x : x \in V\}$ of literals to $[0,1]$ in the natural way: we let

$$p_{\boldsymbol{d}}(x) = p(d_x - d_{\neg x}) \text{ and } p_{\boldsymbol{d}}(\neg x) = 1 - p(x). \tag{15}$$

The idea is that, given $\boldsymbol{d}$, we should set variable x to 'true' with probability $p_{\boldsymbol{d}}(x)$.

To formalize this, we call $p_{\boldsymbol{d}}(l)$ the $p_{\boldsymbol{d}}$***-type*** of the literal l. Let $\mathcal{T} = \mathcal{T}_{\boldsymbol{d}} = \{p_{\boldsymbol{d}}(l) : l \in L\}$ be the set of all possible $p_{\boldsymbol{d}}$-types. We say that $\sigma : V \to \{0,1\}$ has $p_{\boldsymbol{d}}$***-marginals*** if for any type $t \in \mathcal{T}_{\boldsymbol{d}}$ we have

$$\sum_{l \in L : p_{\boldsymbol{d}}(l) = t} (\sigma(l) - t) \cdot d_l = O(1).$$

i.e., among all occurrences of literals of type t, a t fraction is true under σ. This definition captures the above idea that variable x has a $p_{\boldsymbol{d}}(x)$ chance of being 'true'.

Fixing the clause types. We define the $p_{\boldsymbol{d}}$***-type*** of a clause $l_1 \vee \cdots \vee l_k$ as the k-tuple $(p_{\boldsymbol{d}}(l_1), \ldots, p_{\boldsymbol{d}}(l_k)) \in [0,1]^k$ comprising of the individual literal types. Let $\mathcal{L} = \mathcal{L}_{\boldsymbol{d}} = \mathcal{T}_{\boldsymbol{d}}^k$ be the set of all possible clause types. For each $\ell \in \mathcal{L}_{\boldsymbol{d}}$ let $M_{\boldsymbol{\Phi_d}}(\ell)$ be the set of indices $i \in [m]$ such that the ith clause of $\boldsymbol{\Phi_d}$ has type ℓ, and let $m_{\boldsymbol{\Phi_d}}(\ell) = |M_{\boldsymbol{\Phi_d}}(\ell)|$.

In addition to fluctuations of the majority weight, we also need to suppress fluctuations of the numbers $m_{\boldsymbol{\Phi_d}}(\ell)$. We are going to use the same trick as in the case of the majority weight. Namely, we split the generation of a random formula $\boldsymbol{\Phi_d}$ into two steps:

> First, choose a vector $\boldsymbol{m} = (m(\ell))_{\ell \in \mathcal{L}}$ from the "correct" distribution $\boldsymbol{M_d}$.
>
> Then, generate a formula $\boldsymbol{\Phi_{d,m}}$ uniformly at random in which each variable x appears exactly d_x times positively and exactly $d_{\neg x}$ times negatively and that has exactly $m(\ell)$ clauses of type ℓ for all $\ell \in \mathcal{L}$.

Formally, the "correct" $\boldsymbol{M_d}$ is just the distribution of the random vector $\boldsymbol{m}_{\boldsymbol{\Phi_d}} = (m_{\boldsymbol{\Phi_d}}(\ell))_{\ell \in \mathcal{L}}$ that counts the clauses by types in the "unrestricted" formula $\boldsymbol{\Phi_d}$. It is easily verified that the overall outcome of the above experiment is identical to $\boldsymbol{\Phi_d}$. From now on, we fix both $\boldsymbol{d}$ and $\boldsymbol{m}$.

Given $\boldsymbol{d}, \boldsymbol{m}$ there is a simple way of generating the random formula $\boldsymbol{\Phi_{d,m}}$. Namely, create d_l clones of each literal l, and put all the clones of a given $p_{\boldsymbol{d}}$-type on a pile. Then the formula $\boldsymbol{\Phi_{d,m}}$ is simply the result of matching the clones on the type t pile randomly to all the clauses where a literal of type t is required.

An assignment σ with $p_{\boldsymbol{d}}$-marginals splits each pile into two subsets, namely the clones that are true under σ and those that are false. For each type t, among the clones in the type t pile, a t-fraction are true, since σ has $p_{\boldsymbol{d}}$-marginals. Therefore, we expect that under the random matching, for each clause type $\ell = (\ell_1, \ldots, \ell_k)$ and each index j, in an ℓ_j-fraction of clauses the jth literal is matched to a 'true' clone.

Judicious assignment. This observation motivates the following definition. We say that an assignment σ is $p_{\boldsymbol{d}}$***-judicious*** in $\boldsymbol{\Phi_{d,m}}$ if for all clause types $\ell = (\ell_1, \ldots, \ell_k) \in \mathcal{L}$ and all $j \in [k]$ we have

$$\sum_{i \in M_{\boldsymbol{\Phi_{d,m}}}(\ell)} \sigma(\boldsymbol{\Phi_{d,m}}, i, j) = m(\ell) \cdot \ell_j + O(1), \tag{16}$$

where $\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m},i,j}$ denotes the jth literal of the ith clause of $\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$, and the sum is over all i such that the ith clause has type ℓ. Let $\mathcal{S}_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ be the set of p-judicious satisfying assignments, and set $Z_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}) = |\mathcal{S}_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})|$.

Given that σ is p-judicious, in order for σ to be satisfying we just need that for each type ℓ the 'true' clones are distributed so that each clause receives at least one. Thus, the event of being satisfying is merely a matter of how exactly the 'true' clones are "shuffled" amongst the clauses of type ℓ, while for each j the total number of 'true' clones of type ℓ_j is fixed. In particular, this shuffling occurs independently for each clause type. Such random shuffling problems tend to be amenable to the second moment method. Therefore, it seems reasonable to expect that a second moment argument succeeds for $Z_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$. This is indeed the case for $r < r_{\mathrm{BP}} - 1 + \ln 2 \approx r_{\mathrm{BP}} - 0.3$. However, to actually reach r_{BP} we need to control one further parameter.

Fixing the cluster size. According to the physics predictions [25, 27], for $r_{\mathrm{bal}} < r < r_{\mathrm{BP}}$ the set of satisfying assignments decomposes into an exponential number of well-separated 'clusters'. More precisely, we expect that w.h.p. for any two satisfying σ, τ either $\mathrm{dist}(\sigma,\tau) < 0.01n$ (if σ, τ belong to the same cluster), or $\mathrm{dist}(\sigma,\tau) > 0.49n$ (different clusters). Formally, we simply define the ***cluster of*** σ as

$$\mathcal{C}_\sigma(\Phi) = \left\{\tau \in \mathcal{S}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}) : \frac{\mathrm{dist}(\sigma,\tau)}{n} \notin \left[\frac{1}{2} - k^2 2^{-k/2}, \frac{1}{2} + k^2 2^{-k/2}\right]\right\}.$$

The intuitive reason why the second moment argument for $Z_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ breaks down for r close to r_{BP} is that the cluster sizes $|\mathcal{C}_\sigma(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})|$ fluctuate. A similar problem occurred in prior work on random k-NAESAT [11, 12].

As in those papers, the problem admits a remarkably simple solution: let us call an assignment σ ***good*** in $\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$ if

$$|\mathcal{C}_\sigma(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})| \leq \mathrm{E}\left[Z_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]. \tag{17}$$

Let $\mathcal{S}_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ be the set of all good $\sigma \in \mathcal{S}_p(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$. To avoid fluctuations of the cluster size, we are just going to work with $Z_{p,\mathrm{good}} = |\mathcal{S}_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})|$.

The second moment bound. We now face the task of estimating the first and the second moment of $Z_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$. The result can be summarized as follows.

Theorem 5.1 *Suppose $r_{\mathrm{bal}} < r < r_{\mathrm{BP}}$. There exists $C = C(k)$ and a map $p = p_{\mathrm{BP}} : \mathbf{Z} \to [0,1]$ such that for $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ and for $\boldsymbol{m}$ chosen from $\boldsymbol{M}_{\boldsymbol{d}}$ w.h.p.*

$$0 < \mathrm{E}\left[Z_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})^2\right] \leq C \cdot \mathrm{E}\left[Z_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2.$$

Together with Paley-Zygmund (5), Theorem 5.1 shows that with $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ and $\boldsymbol{m}$ chosen from $\boldsymbol{M}_{\boldsymbol{d}}$ w.h.p.

$$\mathrm{P}\left[\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}} \text{ is satisfiable}\right] \quad \geq \quad \mathrm{P}\left[Z_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}) > 0\right] \geq \frac{\mathrm{E}\left[Z_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2}{\mathrm{E}\left[Z_{p,\mathrm{good}}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})^2\right]} \geq \frac{1}{C}. \tag{18}$$

The construction of $\boldsymbol{D}$, $\boldsymbol{M}_{\boldsymbol{d}}$ ensures that choosing $\mathbf{\Phi}$ at random is the same as first picking $\boldsymbol{d}$ from $\boldsymbol{D}$ and $\boldsymbol{m}$ from $\boldsymbol{M}_{\boldsymbol{d}}$ and then generating $\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$. Therefore, (18) implies $\liminf_{n\to\infty} \mathrm{P}\left[\mathbf{\Phi} \text{ is satisfiable}\right] > 0$, so that Lemma 2.1 yields $r_{k-\mathrm{SAT}} \geq r_{\mathrm{BP}}$. Hence, we are left to prove Theorem 5.1. We begin by constructing the map p_{BP}.

Guessing the marginals. For a set $\emptyset \neq S \subset \{0,1\}^V$ and a variable x we define the ***S-marginal of*** x as

$$\mu_S(x) = \sum_{\sigma \in S} \frac{\sigma(x)}{|S|}. \tag{19}$$

The definition of '$p_{\boldsymbol{d}}$-judicious' is guided by the idea that $p_{\boldsymbol{d}}(x)$ should prescribe the marginal of x in the set of all $p_{\boldsymbol{d}}$-judicious satisfying assignments. Hence, in order to make the set of $p_{\boldsymbol{d}}$-judicious assignments as good an approximation of the *entire* set of satisfying assignments as possible, we better pick p so that $p_{\boldsymbol{d}}(x)$ is a good approximation to the actual marginal $\mu_{\mathcal{S}(\mathbf{\Phi}_{\boldsymbol{d}})}(x)$ of x in the set of *all* satisfying assignments. The problem is that, because of the asymmetry of the k-SAT problem, these marginals are highly non-trivial quantities. Indeed, on general formulas Φ the marginals $\mu_{\mathcal{S}(\Phi)}(x)$ are $\#P$-hard to compute.

However, according to the physicists' cavity method, on random formulas with density $r < r_{\mathrm{BP}}$ the marginals can be computed by means of an efficient message passing algorithm called *Belief Propagation* [25]. While the mechanics of this are not important in our context, the result is.

Conjecture 5.2 *Suppose that $r_{\mathrm{bal}} < r < r_{\mathrm{BP}}$. Let $\boldsymbol{d}$ be chosen from $\boldsymbol{D}$ and let x be a variable. Then w.h.p.*

$$\mu_{\mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})}(x) = \frac{1}{2} + \frac{d_x - d_{\neg x}}{2^{k+1}} + O\left(\frac{d_x - d_{\neg x}}{2^k}\right)^2. \tag{20}$$

We observe that (20) is in line with the notion that $\mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})$ is "skewed toward" σ_{maj}. Indeed, the conjecture quantifies how much so. Motivated by Conjecture 5.2, we define

$$p_{\mathrm{BP}}(z) = \begin{cases} \frac{1}{2} + \frac{z}{2^{k+1}} & \text{if } |z| \leq 10\sqrt{k2^k \ln k}, \\ \frac{1}{2} & \text{otherwise.} \end{cases} \tag{21}$$

Under the distribution $\boldsymbol{D}$, the random variables $d_x, d_{\neg x}$ are asymptotically independent Poisson with mean $kr/2$ (cf. Section 2). Therefore,

$$\mathrm{E}_{\boldsymbol{d}}\left[(d_x - d_{\neg x})^2\right] = kr \leq k2^k \ln 2,$$

and standard concentration inequalities show that w.h.p. there are no more than n/k^{30} variables x with $(d_x - d_{\neg x})^2 > 100k2^k \ln k$. Hence, $p_{\boldsymbol{d}} = p_{\mathrm{BP},\boldsymbol{d}}$ is (asymptotically) equal to the conjectured value on the bulk of variables w.h.p.

In summary, the problem with the "vanilla" second moment argument is that the drift toward σ_{maj} induces correlations amongst the satisfying assignments. Indeed, they are correlated with the majority assignment and thus with each other. We circumvent this problem by explicitly prescribing the marginal probability that each variable is set to 'true'. One could think of this as working with the intersection of $\mathcal{S}(\boldsymbol{\Phi})$ with a particular "surface" within the Hamming cube $\{0,1\}^n$, namely the assignments with $p_{\boldsymbol{d}}$-marginals. Within this surface, all assignments are slanted equally toward σ_{maj}. The Belief Propagation-informed definition of p_{BP} is meant to ensure that the surface that we consider with is (about) the most populous one, i.e., the one with the largest number of satisfying assignments in it. The core of our argument will be to show that *with respect to the marginal distribution* p_{BP}, i.e., within the surface that p_{BP} defines, two random elements of $\mathcal{S}_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ are typically uncorrelated. But before we come to that, we need to compute the "first moment", i.e., the expected number of good p_{BP}-judicious satisfying assignments.

Remark 5.3 *Belief Propagation actually leads to a stronger prediction than Conjecture 5.2. Namely, it yields a conjecture for $\mu_{\mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})}(x)$ up to an additive error then tends to 0 as $n \to \infty$. However, (a) this stronger conjecture is not in explicit form, and (b) it does not only depend on $d_x, d_{\neg x}$, but also on various other parameters. In any case, even a more accurate prediction would not yield a better constant than $\frac{3}{2}\ln 2$ in Theorem 1.1.*

Remark 5.4 *In the present framework, the notion of balanced satisfying assignments from [6] simply corresponds to working with the constant map $p_{bal} : \mathbf{Z} \to [0,1],\ z \mapsto \frac{1}{2}$. This hightlights that the improvement that we obtain here stems from choosing the non-constant map p_{BP} inspired by Belief Propagation.*

Remark 5.5 *The definition (19) of the marginal of a set gives rise to a formal notion of 'symmetric problem'. Namely, we could call a (binary) random CSP **symmetric** if its set $\mathcal{S}_{\mathrm{CSP}}(\boldsymbol{\Phi})$ of solutions is such that for each variable x w.h.p. we have $\mu_x(\mathcal{S}_{\mathrm{CSP}}(\boldsymbol{\Phi})) = \frac{1}{2} + o(1)$. Clearly, k-NAESAT passes this test as $\mu_x(\mathcal{S}_{\mathrm{NAE}}(\boldsymbol{\Phi})) = \frac{1}{2}$ for all x with certainty. Similarly, the problem of having a balanced satisfying assignment is symmetric [6], as is random k-XORSAT.*

From here on out we keep the assumptions of Theorem 5.1. In particular, we assume $r_{\mathrm{bal}} < r < r_{\mathrm{BP}}$. Let $\boldsymbol{d}$ be chosen from $\boldsymbol{D}$, and let $\boldsymbol{m}$ be chosen from $M_{\boldsymbol{d}}$. Let $p = p_{\mathrm{BP}}$ be as in (21) and $p_{\boldsymbol{d}}$ as in (15).

5.2 Typical degree sequences

We need to collect a few basic properties of the sequence $\boldsymbol{d}$ chosen from $\boldsymbol{D}$. Let us call a sequence $\boldsymbol{d} = (d_l)_{l \in L}$ of non-negative integers such that $\sum_{l \in L} d_l = km$ a ***signed degree sequence***. For a k-CNF Φ let $\boldsymbol{d}(\Phi) = (d_l(\Phi))_{l \in L}$ denote the vector whose entry $d_l(\Phi)$ is equal to the number of times that literal l occurrs in Φ. Then $\boldsymbol{D} = \boldsymbol{D}_k(n,m)$ is just the distribution of the signed degree sequence $\boldsymbol{d}(\boldsymbol{\Phi})$.

The *signature* of a literal $l \in L$ with respect to a signed degree sequence $\boldsymbol{d}$ is the triple $(\mathrm{sign}(l), d_{|l|}, d_{\neg|l|})$. We omit the reference to $\boldsymbol{d}$ if it is clear from the context. Let $T = T(\boldsymbol{d})$ be the set of all possible signatures. For

each literal l we let $T(l)$ denote its signature. Furthermore, for a signature $\theta = (\mathrm{sign}(l), d_{|l|}, d_{\neg|l|}) \in T$ we let $\neg\theta = (-\mathrm{sign}(l), d_{|l|}, d_{\neg|l|})$.

Let $\boldsymbol{d}$ be a signed degree sequence. A k-CNF Φ over V is ***d*-compatible** if $\boldsymbol{d}(\Phi) = \boldsymbol{d}$. Thus,

$$\boldsymbol{\Phi}_{\boldsymbol{d}} = \boldsymbol{\Phi}_{\boldsymbol{d},1} \wedge \cdots \wedge \boldsymbol{\Phi}_{\boldsymbol{d},m}$$

is a uniformly random $\boldsymbol{d}$-compatible k-CNF.

In the sequel we are going to prove statements about the random formula $\boldsymbol{\Phi}_{\boldsymbol{d}}$ for a "typical" signed degree sequence $\boldsymbol{d}$. Formally, this means that we first choose $\boldsymbol{d}$ from the distribution $\boldsymbol{D}$ at random. Then, conditioning on $\boldsymbol{d}$, we will study the random formula $\boldsymbol{\Phi}_{\boldsymbol{d}}$. Thus, there are *two levels* of randomness: the distribution of $\boldsymbol{d}$ and then, given $\boldsymbol{d}$, the choice of the random formula $\boldsymbol{\Phi}_{\boldsymbol{d}}$. When referring to the random choice of $\boldsymbol{d}$ we use the notation $\mathrm{P}_{\boldsymbol{d}}[\cdot]$, $\mathrm{E}_{\boldsymbol{d}}[\cdot]$. By contrast, if we choose $\boldsymbol{\Phi}_{\boldsymbol{d}}$ randomly for $\boldsymbol{d}$ fixed, then we use $\mathrm{P}[\cdot]$, $\mathrm{E}[\cdot]$.

Lemma 5.6 *1. Let $\mathcal{E}$ be an event such that $\mathrm{P}[\boldsymbol{\Phi} \in \mathcal{E}] = o(1)$. Then w.h.p. a signed degree sequence $\boldsymbol{d}$ chosen from the distribution $\boldsymbol{D}$ is such that $\mathrm{P}[\boldsymbol{\Phi}_{\boldsymbol{d}} \in \mathcal{E}] = o(1)$. Conversely, if w.h.p. for a random $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ we have $\mathrm{P}[\boldsymbol{\Phi}_{\boldsymbol{d}} \in \mathcal{E}] = o(1)$, then $\mathrm{P}[\boldsymbol{\Phi} \in \mathcal{E}] = o(1)$.*

2. For any random variable $X \geq 0$ and any $\varepsilon > 0$ we have $\mathrm{P}_{\boldsymbol{d}}[\mathrm{E}[X(\boldsymbol{\Phi}_{\boldsymbol{d}})] > \mathrm{E}[X(\boldsymbol{\Phi})]/\varepsilon] \leq \varepsilon$.

Proof. The first claim follows from Markov's inequality as $\mathrm{P}[\boldsymbol{\Phi} \in \mathcal{E}] = \mathrm{E}_{\boldsymbol{d}}[\mathrm{P}[\boldsymbol{\Phi}_{\boldsymbol{d}} \in \mathcal{E}]]$. The second claim follows from from Markov's inequality as well because $\mathrm{E}[X(\boldsymbol{\Phi})] = \mathrm{E}_{\boldsymbol{d}}[\mathrm{E}[X(\boldsymbol{\Phi}_{\boldsymbol{d}})]]$. □

Lemma 5.7 *For $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ the following statements hold w.h.p.*

1. $\sum_{x \in V}(d_x - d_{\neg x})^2 \sim km$.
2. $\frac{1}{n}\sum_{x \in V}|d_x - d_{\neg x}| = \tilde{O}(2^{k/2})$.
3. *Let $\mathcal{M}$ contain the n literals of largest degree. Then* $\frac{1}{km}\sum_{l \in \mathcal{M}} d_l = \frac{1}{2} + \tilde{O}(2^{-k/2})$.

Proof. We use the following description of the distribution $\boldsymbol{D}$. Let $\boldsymbol{e} = (e_l)_{l \in L}$ be a family of indepedent $\mathrm{Po}(kr/2)$ variables. Moreover, let $\mathcal{E}$ be the event that $\sum_{l \in L} e_l = km$. It is well known that $\boldsymbol{e}$ given $\mathcal{E}$ has distribution $\boldsymbol{D}$. Furthermore, a simple calculation based on Stirling's formula yields

$$\mathrm{P}[\mathcal{E}] = \Theta(n^{-1/2}). \tag{22}$$

Let $\hat{e}_l = \min\{e_l, \ln^2 n\}$. Employing Stirling's formula once more, we find that $\mathrm{P}[\hat{e}_l \neq e_l] \leq n^{-10}$ for all $l \in L$. Hence, by the union bound,

$$\mathrm{P}[\forall l \in L : \hat{e}_l = e_l] \geq 1 - n^{-9}. \tag{23}$$

Furthermore, as $e_x, e_{\neg x}$ are independent for any $x \in V$, we have

$$\mathrm{E}\left[(\hat{e}_x - \hat{e}_{\neg x})^2\right] = 2\mathrm{Var}(\hat{e}_x) = 2\mathrm{Var}(e_x) + O(n^{-1}) = kr + O(n^{-1}). \tag{24}$$

Because $\hat{e}_l \leq \ln^2 n$ and the random variables $\{(\hat{e}_x - \hat{e}_{\neg x})^2\}_{x \in V}$ are mutually independent, Azuma's inequality yields

$$\mathrm{P}\left[\left|\sum_{x \in V}(\hat{e}_x - \hat{e}_{\neg x})^2 - \mathrm{E}\sum_{x \in V}(\hat{e}_x - \hat{e}_{\neg x})^2\right| > n^{2/3}\right] \leq 2\exp\left[-\frac{n^{1/3}}{8\ln^8 n}\right] \leq n^{-10}. \tag{25}$$

Combining (22)–(25), we find

$$\begin{aligned}
\mathrm{P}_{\boldsymbol{d}}\left[\left|\sum_{x \in V}(d_x - d_{\neg x})^2 - km\right| > n^{3/4}\right] &= \mathrm{P}\left[\left|\sum_{x \in V}(e_x - e_{\neg x})^2 - km\right| > n^{2/3} \,\middle|\, \mathcal{E}\right] \\
&\leq \Theta(n^{1/2})\mathrm{P}\left[\left|\sum_{x \in V}(e_x - e_{\neg x})^2 - km\right| > n^{3/4}\right] \\
&\leq o(1) + \Theta(n^{1/2})\mathrm{P}\left[\left|\sum_{x \in V}(\hat{e}_x - \hat{e}_{\neg x})^2 - \mathrm{E}\sum_{x \in V}(\hat{e}_x - \hat{e}_{\neg x})^2\right| > n^{2/3}\right] \\
&= o(1),
\end{aligned}$$

thereby proving the first claim. The second claim follows from the first by means of the Cauchy-Schwarz inequality: w.h.p.

$$\left[\frac{1}{n}\sum_{x\in V}|d_x - d_{\neg x}|\right]^2 \leq \frac{1}{n}\sum_{x\in V}(d_x - d_{\neg x})^2 \sim kr.$$

Finally, the third assertion is immediate from the second. □

For a set $S \subset L$ we let $\mathrm{Vol}(S) = \mathrm{Vol}_{\boldsymbol{d}}(S) = \sum_{l\in S} d_l$.

Lemma 5.8 *Let $\mathbf{d}$ be chosen from $\boldsymbol{D}$. Then w.h.p. the following is true.*

$$\text{For any set } S \subset L \text{ of literals we have } \mathrm{Vol}(S) \leq 10|S|\max\{kr, \ln(n/|S|)\}. \text{ Furthermore, if } |S| \geq n2^{-0.8k}, \text{ then } \mathrm{Vol}(S) \geq \tfrac{1}{3}|S|kr. \tag{26}$$

Proof. We use the alternative description of $\boldsymbol{D}$ from the proof of Lemma 5.7. That is, $\boldsymbol{e} = (e_l)_{l\in L}$ is a family of indepedent $\mathrm{Po}(kr/2)$ variables, and $\mathcal{E}$ is the event that $\sum_{l\in L} e_l = km$. Let $\lambda = kr/2$. For any fixed set $S \subset L$ the random variable $X_S = \sum_{l\in S} e_l$ has distribution $\mathrm{Po}(|S|\lambda)$ (because the sum of two independent Poisson variables is Poisson). Therefore, letting $\mu = 10|S|\max\{kr, \ln(n/|S|)\}$, we obtain from Stirling's formula

$$\mathrm{P}\left[X_S > \mu\right] \quad\leq\quad O(\sqrt{n})\mathrm{P}\left[X_S = \lceil\mu\rceil\right] \leq O(\sqrt{n})\cdot\frac{\lambda^\mu}{\mu!\exp(\lambda)} \leq O(\sqrt{n})\cdot\left(\frac{\mathrm{e}\lambda}{\mu}\right)^\mu\exp(-\lambda). \tag{27}$$

For $1 \leq s \leq 2n$ let $X_s = \sum_{S:|S|=s}\mathbf{1}_{X_S>\mu}$. Then (27) yields

$$\mathrm{E}X_s \quad\leq\quad O(\sqrt{n})\binom{2n}{s}\cdot\exp(-\lambda-\mu) \leq O(\sqrt{n})\left(\frac{2\mathrm{e}n}{s}\right)^s\cdot\exp(-\lambda-\mu) = o(1/n^2),$$

because $\mu \geq 10s\ln(n/s)$. Thus, the first claim follows from (22) and the union bound.

To prove the second claim, we use Lemma 5.6. For $S \subset L$ we let Y_S be the total number of occurrences of literals from S in $\boldsymbol{\Phi}$. Then Y_S has distribution $\mathrm{Bin}(km, |S|/2n)$ with mean $|S|kr/2$. By the Chernoff bound,

$$\mathrm{P}\left[Y_S < kr|S|/3\right] \quad\leq\quad \exp\left[-\frac{kr|S|}{100}\right]. \tag{28}$$

Hence, letting $Y_s = \sum_{S:|S|=s}\mathbf{1}_{Y_S<kr|S|/3}$, we get from (28) for $s \geq n2^{-0.8k}$

$$\mathrm{E}\left[Y_s\right] \quad\leq\quad \binom{2n}{s}\exp\left[-\frac{krs}{100}\right] \leq \exp\left[s(2+k) - \frac{krs}{100}\right] = o(n^{-2}).$$

Thus, by the union bound $\mathrm{P}\left[\forall s \geq n2^{-0.8k} : Y_s = 0\right] = 1 - o(1/n)$. Applying Lemma 5.6 completes the proof. □

For any $t \in \mathcal{T}$ we let $n(t)$ be the number of variables $x \in V$ such that $p_{\boldsymbol{d}}(x) = t$.

Lemma 5.9 *Let $\mathbf{d}$ be chosen from $\boldsymbol{D}$. Then w.h.p. for any type $t \in \mathcal{T}$ we have*

$$n(t) \geq 2^{-3k/4}n.$$

Proof. We use the alternative description of the distribution $\boldsymbol{D}$ from the proof of Lemma 5.7. That is, let $\boldsymbol{e} = (e_l)_{l\in L}$ be a family of indepedent $\mathrm{Po}(kr/2)$ variables, and $\mathcal{E}$ be the event that $\sum_{l\in L} e_l = km$. For any s, Δ let $X(s,\Delta)$ denote the number of literals l such that $\mathrm{sign}(l) = s$ and $e_{|l|} - e_{\neg|l|} = \Delta$. Since $\mathrm{Var}(e_l) = kr/2 = \Omega_k(k2^k)$, for any $s \in \{\pm 1\}$ and any Δ such that $\Delta^2 \leq 100k2^k\ln k$ we have $\mathrm{E}\left[X(s,\Delta)\right] \geq nk^{-c}$ for some absolute constant $c > 0$. Furthermore, because the random variables $(e_l)_{l\in L}$ are mutually independent, the Chernoff bound implies that

$$\mathrm{P}\left[X(s,\Delta) \leq \frac{1}{2}nk^{-c}\right] \quad\leq\quad \exp(-\Omega(n)) \qquad \text{provided that } \Delta^2 \leq 100k2^k\ln k. \tag{29}$$

Similarly, if we let X'_s denote the number of literals l such that $\mathrm{sign}(l) = s$ and $|e_{|l|} - e_{\neg|l|}| > 100k2^k \ln k$, then $\mathrm{E}[X'(s)] \geq nk^{-c'}$ for some absolute constant c' and

$$\mathrm{P}\left[X'(s) \leq \frac{1}{2}nk^{-c'}\right] \quad \leq \quad \exp(-\Omega(n)). \tag{30}$$

Thus, the assertion follows by combining (22), (29) and (30). □

For each $t \in \mathcal{T}$ we let $\pi(t)$ denote the fraction of literal occurrences of p-type t, i.e.,

$$\pi(t) = \sum_{l \in L: p_{\boldsymbol{d}}(l) = t} \frac{d_l}{km}.$$

For each $\ell \in \mathcal{L}$ let

$$\gamma_\ell = \frac{1}{n}\mathrm{E}\left[m_{\boldsymbol{\Phi_d}}(\ell)\right].$$

Lemma 5.10 *Let $\boldsymbol{d}$ be chosen from $\boldsymbol{D}$. Then w.h.p. $\gamma_\ell \sim \prod_{j=1}^k \pi(\ell_j)$ for all $\ell = (\ell_1, \ldots, \ell_k) \in \mathcal{L}$.*

Proof. By the linearity of expectation, we just need to compute the probability that the first clause $\boldsymbol{\Phi}_{\boldsymbol{d},1}$ has type ℓ. Since $\left|\mathcal{T}^{-1}(\boldsymbol{t})\right| = \Omega(n)$ for all $\boldsymbol{t} \in \mathcal{T}$, the types of the k literals of $\boldsymbol{\Phi}_{\boldsymbol{d},1}$ are asymptotically independent. Thus, the assertion follows from the fact that $\pi(\ell_j)$ equals the marginal probability that a random lityal has type ℓ_j. □

Lemma 5.11 *W.h.p. for $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ we have* $\mathrm{P}\left[\forall \ell \in \mathcal{L} : |m_{\boldsymbol{\Phi_d}}(\ell) - \gamma_\ell n| \leq n^{2/3}\right] = 1 - o(1)$.

Proof. Fix a type $\ell = (\ell_1, \ldots, \ell_j)$. Because p is a feasible marginal, for any $j \in [k]$ there are $\Omega(n)$ literals l with $p(l) = p(\ell_j)$. Therefore, a straightforward calculation shows that

$$\mathrm{P}\left[\boldsymbol{\Phi}_{d,i} \text{ has type } \ell | \boldsymbol{\Phi}_{d,h} \text{ has type } \ell\right] = \mathrm{P}\left[\boldsymbol{\Phi}_{d,i} \text{ has type } \ell\right] \cdot (1 + O(1/n)) \qquad \text{for any } i \neq h.$$

Consequently, $\mathrm{Var}(m_{\boldsymbol{\Phi_d}}(\ell)) \sim \mathrm{E}\left[m_{\boldsymbol{\Phi_d}}(\ell)\right] = O(n)$. Hence, by Chebyshev's inequality

$$\mathrm{P}\left[\left|m_{\boldsymbol{\Phi_d}}(\ell) - \mathrm{E}\left[m_{\boldsymbol{\Phi_d}}(\ell)\right]\right| > n^{2/3}\right] = O(n^{-1/3}) = o(1). \tag{31}$$

Since $|\mathcal{L}| = O(1)$ as $n \to \infty$ by the construction of p, the assertion follows from (31) and the union bound. □

6 The first moment

6.1 Outline

Let $\rho > \frac{3}{2}\ln 2$ be such that $r = 2^k \ln 2 - \rho$.

Proposition 6.1 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that*

$$\mathrm{E}\left[Z_{p,\mathrm{good}}(\boldsymbol{\Phi_{d,m}})\right] \quad = \quad \exp\left[\frac{n}{2^k}\left(\rho - \frac{\ln 2}{2} + o_k(1)\right)\right].$$

We begin by computing $\mathrm{E}\left[Z_p(\boldsymbol{\Phi_{d,m}})\right]$. By definition, any assignment that is $p_{\boldsymbol{d}}$-judicious has $p_{\boldsymbol{d}}$-marginals. Thus, let $\mathcal{H}_p(\boldsymbol{d}) \subset \{0,1\}^V$ denote the set of all assignments that have $p_{\boldsymbol{d}}$-marginals. Then by the linearity of expectation,

$$\mathrm{E}\left[Z_p(\boldsymbol{\Phi_{d,m}})\right] \quad = \quad \sum_{\sigma \in \mathcal{H}_p(\boldsymbol{d})} \mathrm{P}\left[\sigma \in \mathcal{S}_p(\boldsymbol{\Phi_{d,m}})\right]. \tag{32}$$

Hence, we need to compute $|\mathcal{H}_p(\boldsymbol{d})|$ and the probability $\mathrm{P}\left[\sigma \in \mathcal{S}_p(\boldsymbol{\Phi_{d,m}})\right]$ for any $\sigma \in \mathcal{H}_p(\boldsymbol{d})$. Using basic properties of the entropy, we obtain

Lemma 6.2 *Let $\chi(z) = -z\ln z - (1-z)\ln(1-z)$ denote the entropy function. Then w.h.p. $\boldsymbol{d}$ is such that*

$$\ln|\mathcal{H}_p(\boldsymbol{d})| \sim n \cdot \sum_{x\in V}\chi(p(x)).$$

Taylor expanding $\chi(z)$ around $z = 1/2$ and plugging in the definition (21) of p, we obtain that w.h.p. $\boldsymbol{d}$ is such that

$$\frac{1}{n}\ln|\mathcal{H}_p(\boldsymbol{d})| = \ln 2 - \frac{k\ln 2}{2^{k+1}} + o_k(2^{-k}). \tag{33}$$

As a next step, we compute the probability of $\sigma \in \mathcal{S}_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ for $\sigma \in \mathcal{H}_p(\boldsymbol{d})$.

Lemma 6.3 *W.h.p. $\boldsymbol{d}$, $\boldsymbol{m}$ are such that for any $\sigma \in \mathcal{H}_p(\boldsymbol{d})$,*

$$\frac{1}{n}\ln\mathrm{P}\left[\sigma \in \mathcal{S}_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \quad = \quad -\ln 2 + \frac{k\ln 2}{2^{k+1}} + 2^{-k}\left[\rho - \frac{\ln 2}{2} + o_k(1)\right]. \tag{34}$$

Let us defer the proof of Lemma 6.3, which is the core of the first moment computation, for a little while. Combining (32)–(34), we see that w.h.p. over the choice of $\boldsymbol{d}, \boldsymbol{m}$ we have

$$\begin{aligned}\ln\mathrm{E}\left[Z_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] &= \ln|\mathcal{H}_p(\boldsymbol{d})| + \ln\mathrm{P}\left[\sigma \in \mathcal{S}_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \\ &\sim 2^{-k}\left[\rho - \frac{\ln 2}{2} + o_k(1)\right]\cdot n \end{aligned} \tag{35}$$

To obtain the expectation of $Z_{p,\mathrm{good}}$, we show the following.

Lemma 6.4 *W.h.p. over the choice of $\boldsymbol{d}, \boldsymbol{m}$ we have*

$$\mathrm{E}\left[Z_{p,\mathrm{good}}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \sim \mathrm{E}\left[Z_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right].$$

The proof of Lemma 6.4 is based on arguments developed in [1] for analyzing the geometry of the set of satisfying assignments. Combining (35) and Lemma 6.4 yields Proposition 6.1.

6.2 Proof of Lemma 6.3

For a sequence $\boldsymbol{m} = (m(\ell))_{\ell\in\mathcal{L}}$ of non-negative integers we let $\Gamma_{\boldsymbol{m}}$ denote the event that $m_{\boldsymbol{\Phi}_{\boldsymbol{d}}}(\ell) = m(\ell)$ for all $\ell \in \mathcal{L}$. Let us call $\boldsymbol{m}$ *feasible* if $\mathrm{P}_{\boldsymbol{m}}[\Gamma_{\boldsymbol{m}}] > 0$ and $|m(\ell) - \gamma_\ell n| \le n^{2/3}$ for all $\ell \in \mathcal{L}$. Let Z be the number of $p_{\boldsymbol{d}}$-judicious satisfying assignments.

Proposition 6.5 *Let $\mathbf{d}$ be chosen from $\boldsymbol{D}$. Then w.h.p. for any feasible $\boldsymbol{m} = (m(\ell))_{\ell\in\mathcal{L}}$ the following statements hold.*

1. *We have*

$$2^{-k}\left[\rho - \frac{\ln 2}{2} - k^{-9}\right] \quad \le \quad \frac{1}{n}\ln\mathrm{E}\left[Z(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \le 2^{-k}\left[\rho - \frac{\ln 2}{2} + k^{-9}\right].$$

2. *For any $t \in \mathcal{T}$ we have*

$$\sum_{l\in L: p_{\boldsymbol{d}}(l)=t} d_l^2 \le \frac{2km\pi(t)}{n(t)}.$$

3. *For any $\sigma \in \{0,1\}^V$ with p-marginals we have*

$$\frac{1}{km}\sum_{l\in L} d_l \mathbf{1}_{\sigma(l)=1} = \frac{1}{2} + O(2^{-k}).$$

The proof of Proposition 6.5 consists of two steps. We defer the proof of the following lemma to Section 6.3.

Lemma 6.6 *With the assumptions of Proposition 6.5 and with δ, δ' defined by*

$$\begin{aligned}\frac{1}{km}\sum_{x\in V}\left(p(x)-\frac{1}{2}\right)^2 &= (1+\delta)\,2^{-2k-2} \quad \textit{and}\\ \Sigma &= \frac{1}{km}\sum_{x\in V}(1-2p(x))(d_x-d_{\neg x}) = -(1+\delta')2^{-k}\end{aligned}$$

we have w.h.p.

$$\frac{1}{n}\ln \mathrm{E}\left[Z(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] = 2^{-k}\left[\rho-\frac{\ln 2}{2}\right]+O\left(\frac{k(\delta+\delta')}{2^k}\right)+\tilde{O}(2^{-3k/2}).$$

Proof of Proposition 6.5. Let $\Delta = 100k2^k\ln k$ and let δ,δ' be as in Lemma 6.6. Using the alternative description of the distribution $\boldsymbol{D}$ from the proof of Lemma 5.7 and applying Azuma's inequality, one can easily verify that w.h.p.

$$\sum_{x\in V}\mathbf{1}_{(d_x-d_{\neg x})^2\le\Delta}\cdot(d_x-d_{\neg x})^2 \ge (1-k^{-12})\sum_{x\in V}(d_x-d_{\neg x})^2. \tag{36}$$

Therefore, Lemma 5.7 entails that w.h.p.

$$\frac{1}{km}\sum_{x\in V}\left(p(x)-\frac{1}{2}\right)^2 = \frac{1+O_k(k^{-12})}{km}\sum_{x\in V}\frac{(d_x-d_{\neg x})^2}{4^{k+1}} = \frac{1+O_k(k^{-12})}{4^{k+1}}.$$

Consequently, w.h.p. we have

$$\delta = O_k(k^{-12}). \tag{37}$$

Similarly, invoking (36) once more, we see that w.h.p.

$$-\Sigma = \frac{1}{km}\sum_{x\in V}(2p(x)-1)(d_x-d_{\neg x}) = \frac{1}{2^kkm}\sum_{x\in V}\mathbf{1}_{(d_x-d_{\neg x})^2\le\Delta}\cdot(d_x-d_{\neg x})^2 = \frac{1+O_k(k^{-12})}{2^k},$$

whence

$$\delta' = O_k(k^{-12}) \tag{38}$$

w.h.p. Thus, Proposition 6.5 is a direct consequence of Lemmas 5.9 and 6.6 and (37), (38). □

6.3 Proof of Lemma 6.6

We begin by determining the number $\sigma \in \{0,1\}^V$ with p-marginals. The following is an easy consequence of Lemma 6.2.

Corollary 6.7 *W.h.p. for $\mathbf{d}$ chosen from $\boldsymbol{D}$ we have*

$$\frac{1}{n}\ln|\mathcal{H}(p)| = \ln 2-\frac{2}{n}\sum_{x\in V}\left(p(x)-\frac{1}{2}\right)^2+\tilde{O}(2^{-3k/2}).$$

Proof. This follows from Lemma 6.2 by Taylor expanding $\chi(\cdot)$ around $\frac{1}{2}$. □

We need to compute the probability that an assignment $\sigma \in \{0,1\}^V$ with p-marginals is a p-judicious satisfying assignment. To this end, we introduce a new probability space $(\hat{\Omega}, \hat{\mathrm{P}})$. Let $\boldsymbol{q} = (q_{\ell,j})_{\ell\in\mathcal{L}, j\in[k]}$ be a matrix with entries in $[0,1]$. The elements of our new probability space $\hat{\Omega}$ are all $0/1$ vectors

$$(\hat{\boldsymbol{\sigma}}_{ij}(\ell))_{\ell\in\mathcal{L}, i\in[m(\ell)], j\in[k]}.$$

The distribution $\hat{\mathrm{P}}$ is such that the entries $\hat{\boldsymbol{\sigma}}_{ij}(\ell)$ are mutually independent, and for each $\ell = (\ell_1, \ldots, \ell_k) \in \mathcal{L}$, $i \in [m(\ell)]$, $j \in [k]$ we let $\hat{\boldsymbol{\sigma}}_{ij}(\ell) = \mathrm{Be}(q_{\ell,j})$ be a Bernoulli random variable. (It may be helpful to think of $\hat{\boldsymbol{\sigma}}_{ij}(\ell)$ as the truth value of the jth literal of the ith clause of type ℓ in a random formula $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$.)

For $\ell = (\ell_1, \ldots, \ell_k) \in \mathcal{L}$ let $S_i(\ell)$ be the event that

$$\max_{j\in[k]} \hat{\boldsymbol{\sigma}}_{ij}(\ell) = 1$$

(the intuition is that this corresponds to the event that the clause i of type ℓ is satisfied). Let $S(\ell) = \bigcap_{i\in[m(\ell)]} S_i(\ell)$ and $S = \bigcap_{\ell\in\mathcal{L}} S(\ell)$. Moreover, for $j \in [k]$ let $B(\ell, j)$ be the event that

$$\frac{1}{m(\ell)} \sum_{i\in[m(\ell)]} \hat{\boldsymbol{\sigma}}_{ij}(\ell) \doteq p(t).$$

Let $B(\ell) = \bigcap_{j=1}^{k} B(\ell, j)$ and $B = \bigcap_{\ell\in\mathcal{L}} B(\ell)$. The connection between the probability space $\hat{\Omega}$ and Lemma 6.6 is as follows.

Lemma 6.8 *Suppose that $\sigma \in \{0,1\}^V$ has p-marginals. Let $\mathcal{S}(\sigma)$ be the event that σ is a satisfying assignment of $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$ and let $\mathcal{B}(\sigma)$ be the event that σ is $p_{\boldsymbol{d}}$-judicious. Then* $\mathrm{P}\left[\mathcal{S}(\sigma)|\mathcal{B}(\sigma)\right] = \hat{\mathrm{P}}\left[S|B\right]$.

Proof. Note that in $\mathrm{P}\left[\mathcal{S}(\sigma)|\mathcal{B}(\sigma)\right]$ probability is taken over the choice of the random formula $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$, while in $\hat{\mathrm{P}}\left[S|B\right]$ probability is taken over $\hat{\boldsymbol{\sigma}}$ chosen from the above distribution. Thus, we need to relate the two probability spaces.

For any $\boldsymbol{d}$-compatible formula $\Phi \in \Gamma_{\boldsymbol{m}}$ we can define a map

$$\sigma \in \{0,1\}^V \mapsto \hat{\sigma}\big|_{\Phi} = \left(\sigma_{ij}(\ell)\big|_{\Phi}\right)_{\ell\in\mathcal{L}, i\in[m(\ell)], j\in[k]},$$

by letting $\hat{\sigma}_{ij}(\ell)|_{\Phi}$ be the truth value of the jth literal of the ith clause of type ℓ in Φ under σ. In other words, $\hat{\sigma}|_{\Phi}$ is the string of truth values that we get by "plugging the assignment σ into Φ". Then σ is judicious iff $\hat{\sigma}_{\Phi} \in B$. Furthermore, σ is satisfying iff $\hat{\sigma}_{\Phi} \in S$. Finally, if σ has p-marginals, then $\hat{\sigma}|_{\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}}$ becomes a random vector. Given $\mathcal{B}(\sigma)$ its distribution is identical to the conditional distribution of $\hat{\boldsymbol{\sigma}}$ given B. □

Corollary 6.9 *With the notation of Lemma 6.8 we have* $\mathrm{P}\left[\mathcal{S}(\sigma) \cap \mathcal{B}(\sigma)\right] = \hat{\mathrm{P}}\left[S|B\right]\exp(o(n))$. *Moreover, for any σ with p-marginals we have* $\mathrm{P}\left[\mathcal{B}(\sigma)\right] = \Theta\left(n^{(|\mathcal{T}|-k|\mathcal{L}|)/2}\right)$.

Proof. Since the total number $|\mathcal{L}|$ of clause types is bounded, the assertion follows from a repeated application of Lemma 4.1 (the local limit theorem). □

Thus, we have reduced the proof of Lemma 6.6 to the computation of $\hat{\mathrm{P}}\left[S|B\right]$. The benefit of the probability space $\hat{\Omega}$ is that S, B can be decomposed easily into independent events. Indeed, for any $\ell \in \mathcal{L}$ and any $i \in [m(\ell)]$ we have

$$\hat{\mathrm{P}}\left[S_i(\ell)\right] = 1 - \prod_{j=1}^{k} 1 - q_{\ell,j},$$

because the $\hat{\boldsymbol{\sigma}}_{ij}(\ell)$ are independent. Moreover, due to independence and because $\boldsymbol{m}$ is feasible,

$$\frac{1}{n}\ln\hat{\mathrm{P}}\left[S(\ell)\right] = \frac{1}{n}\sum_{i\in[m(\ell)]}\ln\hat{\mathrm{P}}\left[S_i(\ell)\right] \sim \gamma_\ell \ln\left[1 - \prod_{j=1}^{k} 1 - q_{\ell,j}\right]$$

and thus

$$\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right]\sim\sum_{\ell\in\mathcal{L}}\gamma_{\ell}\ln\left[1-\prod_{j=1}^{k}1-q_{\ell,j}\right]. \tag{39}$$

Similarly,

$$\frac{1}{n}\ln\hat{\mathrm{P}}\left[B\right]=\frac{1}{n}\sum_{\ell\in\mathcal{L}}\ln\hat{\mathrm{P}}\left[B(\ell)\right]=\frac{1}{n}\sum_{\ell\in\mathcal{L}}\sum_{j=1}^{k}\ln\hat{\mathrm{P}}\left[B(\ell,j)\right]. \tag{40}$$

A further benefit of the space $\hat{\Omega}$ is that we are free to choose the vector $\boldsymbol{q}$ as we please (subject only to the condition that $\hat{\mathrm{P}}\left[B\right]>0$). To facilitate the computation of $\hat{\mathrm{P}}\left[S|B\right]$, we are going to choose $\boldsymbol{q}$ such that

$$\hat{\mathrm{P}}\left[B|S\right]=\exp(o(n)). \tag{41}$$

For if (41) holds, then

$$\hat{\mathrm{P}}\left[S|B\right]=\frac{\hat{\mathrm{P}}\left[S\right]}{\hat{\mathrm{P}}\left[B\right]}\cdot\exp(o(n)),$$

where $\hat{\mathrm{P}}\left[S\right]$, $\hat{\mathrm{P}}\left[B\right]$ can be calculated rather easily via (39) and (40). Thus, as a next step we need to find $\boldsymbol{q}$ such that (41) is true. To this end, we define

$$\hat{q}_{\ell,j}=\hat{\mathrm{E}}\left[\hat{\boldsymbol{\sigma}}_{ij}(\ell)|S_{i}(\ell)\right]=\frac{q_{\ell,j}}{1-\prod_{l=1}^{k}1-q_{\ell,l}}\qquad\qquad(\ell\in\mathcal{L},j\in[k]). \tag{42}$$

Lemma 6.10 *There exists $\boldsymbol{q}$ such that $\hat{q}_{\ell,j}=\ell_{j}$ for all $\ell=(\ell_{1},\ldots,\ell_{k})\in\mathcal{L},j\in[k]$. Furthermore, this $\boldsymbol{q}$ satisfies*

$$q_{\ell,j}=\ell_{j}-2^{-k-1}+\tilde{O}(2^{-3k/2}). \tag{43}$$

Proof. For any ℓ,j we have

$$\begin{aligned}\frac{\partial\hat{q}_{\ell,j}}{\partial q_{\ell,j}}&=\frac{1-(1-2q_{\ell,j})\prod_{l\neq j}1-q_{\ell,l}}{\left(1-\prod_{l}1-q_{\ell,l}\right)^{2}},\\ \frac{\partial\hat{q}_{\ell,j}}{\partial q_{\ell,h}}&=-\frac{q_{\ell,j}\prod_{l\neq h}1-q_{\ell,l}}{\left(1-\prod_{l}1-q_{\ell,l}\right)^{2}}\qquad\qquad(h\neq j).\end{aligned}$$

Hence, for k large enough and $0.01<q_{j}<0.99$ for all j, the $k\times k$ matrix $D\hat{q}$ is close to id. In particular, this is true for q_{j} close to $1/2$. Therefore, the assertion follows from the inverse function theorem (Lemma 4.4). □

Corollary 6.11 *With $\boldsymbol{q}$ from Lemma 6.10 we have $\hat{\mathrm{P}}\left[B|S\right]=\Theta(n^{-k|\mathcal{L}|/2})=\exp(o(n))$ and thus (41).*

Proof. Equation (42) shows that for the vector $\boldsymbol{q}$ from Lemma 6.10 we have $\hat{\mathrm{E}}\left[\hat{\boldsymbol{\sigma}}_{ij}(\ell)|S_{i}(\ell)\right]=\ell_{j}$ for all ℓ,j,i. Therefore, a repeated application of Lemma 4.1 yields $\hat{\mathrm{P}}\left[B|S\right]=\Theta(n^{-k|\mathcal{L}|/2})=\exp(o(n))$. □

From this point on we fix $\boldsymbol{q}$ as in Lemma 6.12.

Lemma 6.12 *Letting*

$$\Sigma=\frac{1}{km}\sum_{x\in V}(1-2p_{\boldsymbol{d}}(x))(d_{x}-d_{\neg x}), \tag{44}$$

we have $\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right]=-\ln 2+2^{-k}\left[\rho-\frac{\ln 2}{2}-k\ln 2\right]-k\Sigma\ln 2+\tilde{O}(2^{-3k/2})$.

Proof. Starting from (39), we obtain

$$\begin{aligned}\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right] &\sim \sum_{\ell\in\mathcal{L}}\gamma_\ell \ln\left[1-\prod_{j=1}^{k}1-q_{\ell,j}\right] \\ &= -\sum_{\ell\in\mathcal{L}}\gamma_\ell\left[\left(\prod_{j=1}^{k}1-q_{\ell,j}\right)+\frac{1}{2}\left(\prod_{j=1}^{k}1-q_{\ell,j}\right)^2+\tilde{O}(8^{-k})\right],\end{aligned} \tag{45}$$

where we used the approximation $\ln(1+x)=x-\frac{1}{2}x^2+O(x^3)$. Thus, Lemma 5.10 yields

$$\begin{aligned}\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right] &= -\sum_{\ell\in\mathcal{L}}\gamma_\ell\left[\left(\prod_{j=1}^{k}1-q_{\ell,j}\right)+\frac{1}{2}\cdot 4^{-k}+\tilde{O}(2^{-5k/2})\right] \\ &= -\frac{r}{2}\cdot 4^{-k}+\tilde{O}(2^{-3k/2})-r\sum_{\ell\in\mathcal{L}}\prod_{j=1}^{k}\pi(\ell_j)(1-q_{\ell,j}).\end{aligned}$$

Further, by (43)

$$\begin{aligned}\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right] &= -\frac{r}{2}\cdot 4^{-k}+\tilde{O}(2^{-3k/2})-r\sum_{\ell\in\mathcal{L}}\prod_{j=1}^{k}\pi(\ell_j)(1-\ell_j+2^{-k-1}) \\ &= -\frac{r}{2}\cdot 4^{-k}+\tilde{O}(2^{-3k/2})-r\left[\sum_{t\in\mathcal{T}}\pi(t)(1-t+2^{-k-1})\right]^k \\ &= -\frac{r}{2}\cdot 4^{-k}+\tilde{O}(2^{-3k/2})-r\left[2^{-k-1}+\sum_{t\in\mathcal{T}}\pi(t)(1-t)\right]^k \\ &= -\frac{r}{2}\cdot 4^{-k}-kr\cdot 4^{-k}+\tilde{O}(2^{-3k/2})-r\left[\sum_{t\in\mathcal{T}}\pi(t)(1-t)\right]^k \\ &= -\frac{r}{2}\cdot 4^{-k}-kr\cdot 4^{-k}+\tilde{O}(2^{-3k/2})-r\left[\frac{1}{2}-\sum_{t\in\mathcal{T}}\pi(t)\left(t-\frac{1}{2}\right)\right]^k.\end{aligned}$$

Now,

$$\begin{aligned}\sum_{t\in\mathcal{T}}\pi(t)\left(t-\frac{1}{2}\right) &= \sum_{x\in V}\frac{d_x}{km}\left(p(x)-\frac{1}{2}\right)+\frac{d_{\neg x}}{km}\left(\frac{1}{2}-p(x)\right) \\ &= \frac{1}{km}\sum_{x\in V}(d_x-d_{\neg x})\left(p(x)-\frac{1}{2}\right)=-\Sigma/2.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right] &= -\frac{r}{2}\cdot 4^{-k}-kr\cdot 4^{-k}-r\left(\frac{1+\Sigma}{2}\right)^k+\tilde{O}(2^{-3k/2}) \\ &= -r\cdot 2^{-k}-\frac{r}{2}\cdot 4^{-k}-kr\cdot 4^{-k}-kr\Sigma 2^{-k}+\tilde{O}(2^{-3k/2}).\end{aligned}$$

Plugging in $r=2^k\ln 2-\rho$, we get

$$\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right] = -\ln 2+2^{-k}\left[\rho-\frac{\ln 2}{2}-k\ln 2\right]-k\Sigma\ln 2+\tilde{O}(2^{-3k/2}),$$

as claimed. □

Lemma 6.13 *We have* $\frac{1}{n}\ln\hat{\mathrm{P}}\left[B\right]=-\frac{k\ln 2}{2^{k+1}}+\tilde{O}(2^{-3k/2})$.

Proof. Due to (40) we just need to estimate $\ln\hat{\mathrm{P}}\left[B(\ell,j)\right]$ for any $\ell=(\ell_1,\ldots,\ell_k)\in\mathcal{L}$ and $j\in[k]$. By construction,

$$\hat{\mathrm{P}}\left[B(t,\ell)\right]=\mathrm{P}\left[\mathrm{Bin}(m(\ell),q_{\ell,j})=\ell_j m(\ell)\right].$$

By Lemma 6.10 we have $q_{\ell,j}=\ell_j-2^{-k-1}+\tilde{O}(2^{-3k/2})=\frac{1}{2}+\tilde{O}(2^{-k/2})$. Hence, using Lemma 4.2, we find

$$\begin{aligned}\frac{1}{m(\ell)}\ln\hat{\mathrm{P}}\left[B(\ell,j)\right] &\sim \psi(q_{\ell,j},\ell_j) \qquad (46)\\ &= -\frac{(q_{\ell,j}-\ell_j)^2}{2q_{\ell,j}}-\frac{(\ell_j-q_{\ell,j})^2}{2(1-q_{\ell,j})}+O_k(8^{-k})\\ &= -\frac{(q_{\ell,j}-\ell_j)^2}{2}\left(\frac{1}{q_{\ell,j}}+\frac{1}{1-q_{\ell,j}}\right)+O_k(8^{-k})\\ &= -\left(2+\tilde{O}(2^{-k/2})\right)(q_{\ell,j}-\ell_j)^2\\ &= -\left(2+\tilde{O}(2^{-k/2})\right)(2^{-k-1}+\tilde{O}(2^{-3k/2}))^2=-\left(\frac{1}{2}+\tilde{O}(2^{-k/2})\right)2^{-2k}.\end{aligned}$$

Hence, (40) yields

$$\begin{aligned}\frac{1}{n}\ln\hat{\mathrm{P}}\left[B\right] &= -\sum_{\ell\in\mathcal{L}}\sum_{j=1}^{k}\frac{m(\ell)}{n}\cdot\left[\frac{1}{2}\cdot 2^{-2k}+\tilde{O}(2^{-5k/2})\right]\\ &= -kr\cdot\left[\frac{1}{2}\cdot 2^{-2k}+\tilde{O}(2^{-5k/2})\right]=-\frac{k\ln 2}{2^{k+1}}+\tilde{O}(2^{-3k/2}),\end{aligned}$$

as claimed. □

Remark 6.14 *In the second moment calculation we will need to know that*

$$\ln\hat{\mathrm{P}}\left[S|B\right]=\sum_{\ell\in\mathcal{L}}m(\ell)\left[\ln\left(1-\prod_{j=1}^{k}1-q_{\ell,j}\right)-\sum_{j=1}^{k}\psi(q_{\ell,j},\ell_j)\right]$$

which follows from (45) and (46).

Corollary 6.15 *Let* $\delta,\delta'>0$ *be such that*

$$\begin{aligned}\sum_{x\in V}\left(p_{\boldsymbol{d}}(x)-\frac{1}{2}\right)^2 &= \frac{(1+\delta)km}{2^{2k+2}},\\ \Sigma &= (1+\delta')2^{-k} \qquad \textit{with } \Sigma \textit{ from (44).}\end{aligned}$$

Then with $r=2^{-k}\ln 2-c$ *we have*

$$\frac{\ln|\mathcal{H}(p)|+\ln\hat{\mathrm{P}}\left[S|B\right]}{n}=2^{-k}\left[\rho-\frac{\ln 2}{2}\right]+O\left(\frac{k(\delta+\delta')}{2^k}\right)+\tilde{O}(2^{-3k/2}).$$

Proof. By the above,

$$\begin{aligned}\frac{1}{n}\ln\hat{\mathrm{P}}\left[S\right] &= -\ln 2+2^{-k}\left(\rho-\frac{\ln 2}{2}-k\ln 2\right)-k\Sigma\ln 2+\tilde{O}(2^{-3k/2})\\ &= -\ln 2+2^{-k}\left(\rho-\frac{\ln 2}{2}\right)+\frac{k\delta'\ln 2}{2^k}+\tilde{O}(2^{-3k/2}),\\ \frac{1}{n}\ln\hat{\mathrm{P}}\left[B\right] &= -\frac{k\ln 2}{2^{k+1}}+\tilde{O}(2^{-3k/2}),\\ \frac{1}{n}\ln|\mathcal{H}(p)| &= \ln 2-\frac{2}{n}\sum_{x\in V}\left(p_{\boldsymbol{d}}(x)-\frac{1}{2}\right)^2=\ln 2-\frac{k\ln 2}{2^{k+1}}-\frac{\delta k\ln 2}{2^{k+1}}.\end{aligned}$$

Summing up yields the result. □

Proof of Lemma 6.6. Lemma 6.6 is a direct consequence of Corollaries 6.7, 6.9, 6.11 and 6.15. □

6.4 Proof of Lemma 6.4

Assume that $\boldsymbol{m}$ is feasible. Let $\mathcal{Z}$ denote the number of good p-satisfying assignments.

Proposition 6.16 *Let $\mathbf{d}$ be chosen from $\boldsymbol{D}$ and let $\boldsymbol{m}$ be chosen from $\boldsymbol{M}_{\mathbf{d}}$. Then $\mathrm{E}\left[\mathcal{Z}(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}})\right] \sim \mathrm{E}\left[Z(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}})\right]$ w.h.p.*

The proof of Proposition 6.16 is based on three lemmas.

Lemma 6.17 *Let $\mathbf{d}$ be chosen from $\boldsymbol{D}$ and let $\boldsymbol{m}$ be chosen from $\boldsymbol{M}_{\mathbf{d}}$.*

1. *Let $\mathcal{E}$ be an event such that $\mathrm{P}\left[\boldsymbol{\Phi} \in \mathcal{E}\right] = o(1)$. Then $\mathrm{P}\left[\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}} \in \mathcal{E}\right] = o(1)$.*
2. *For any random variable $X \geq 0$ and any $\varepsilon > 0$ we have $\mathrm{P}_{\mathbf{d},\boldsymbol{m}}\left[\mathrm{E}\left[X(\boldsymbol{\Phi}_{\mathbf{d}})\right] > \mathrm{E}\left[X(\boldsymbol{\Phi})\right]/\varepsilon\right] \leq \varepsilon$.*

Proof. This follows from a similar application of Markov's inequality as in the proof of Lemma 5.6. □

Lemma 6.18 *With the assumptions of Proposition 6.16 the random variable*

$$Z' = \left|\left\{\sigma \in \mathcal{S}(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}}) : \left|\left\{\tau \in \mathcal{S}(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}}) : \mathit{dist}(\sigma,\tau) < 2^{-0.99k}n\right\}\right| \leq \mathrm{E}\left[Z(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}})\right]\right\}\right|$$

satisfies $\mathrm{E}\left[Z'(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}})\right] \sim \mathrm{E}\left[Z(\boldsymbol{\Phi}_{\mathbf{d},\boldsymbol{m}})\right]$ w.h.p.

The proof of Lemma 6.18 can be found in Section 6.5. Moreover, in Section 6.6 we prove the following.

Lemma 6.19 *Suppose that $r \leq 2^k \ln 2$. Let $\xi = k2^{-k/2}$. Let Z'' be the number of pairs $(\sigma,\tau) \in \mathcal{S}(\boldsymbol{\Phi})^2$ such that*

$$\mathit{dist}(\sigma,\tau) \in \left[k2^{-k},1\right] \setminus \left[\frac{1}{2}-\xi, \frac{1}{2}+\xi\right].$$

Then $\mathrm{E}\left[Z''\right] = o(1)$.

Finally, Proposition 6.16 follows immediately from Lemmas 6.17, 6.18 and 6.19.

6.5 Proof of Lemma 6.18

Let Φ be a k-CNF and $\sigma \in \mathcal{S}(\Phi)$. We say that a variable x is ξ-*rigid* in (Φ,σ) if for any $\tau \in \mathcal{S}(\Phi)$ with $\tau(x) \neq \sigma(x)$ we have $\mathrm{dist}(\sigma,\tau) \geq \xi n$. Let $\lambda = kr/(2^k-1)$.

Lemma 6.20 *1. The expected number of $\sigma \in \mathcal{S}(\boldsymbol{\Phi})$ in which more than $k^{12}2^{-k}n$ variables support at most 12 clauses is $\leq \exp(-nk^9/2^k)\mathrm{E}\left|\mathcal{S}(\boldsymbol{\Phi})\right|$.*

2. The expected number of $\sigma \in \mathcal{S}(\boldsymbol{\Phi})$ in which more than $(1+1/k^2)2^{-k}n$ variables support no clause at all is $\leq \exp(-n/(k^6 2^k))\mathrm{E}\left|\mathcal{S}(\boldsymbol{\Phi})\right|$.

Proof. Fix an assignment $\sigma \in \{0,1\}^V$, say $\sigma = \mathbf{1}$. Then the number of clauses supported by each $x \in V$ is asymptotically Poisson with mean λ. Let $\mathcal{E}_x$ be the event that x supports no more than 12 clauses. Then

$$\mathrm{P}\left[\mathcal{E}_x\right] \leq \lambda^{12}\exp(-\lambda) \leq \frac{1}{2}k^{12}2^{-k}.$$

The events $(\mathcal{E}_x)_{x\in V}$ are negatively correlated. Therefore, the total number X of variables $x \in V$ for which $\mathcal{E}_x$ occurs is stochastically dominated by a binomial variable $\mathrm{Bin}(n, \frac{1}{2}k^{12}2^{-k})$. Hence, the first assertion follows from Chernoff bounds.

With respect to the second assertion, let $\mathcal{E}'_x$ be the event that x supports no clause at all. Then $\mathrm{P}\left[\mathcal{E}_x\right] \leq \exp(-\lambda)$. Using negative correlation and Chernoff bounds once more completes the proof. □

Let us call a set $S \subset V$ *self-contained* if each variable in S supports at least ten clauses that consist of variables in S only. There is a simple process that yields a (possibly empty) self-contained set S.

- For each variable x that supports at least one clause, choose such a clause C_x randomly.
- Let R be the set of all variables that support at least 12 clauses.
- While there is a variable $x \in R$ that supports fewer than ten clauses $\mathbf{\Phi}_i \neq C_x$ that consist of variables of R only, remove x from R.

The clauses C_x will play a special role later.

Lemma 6.21 *The expected number of solutions $\sigma \in \mathcal{S}(\mathbf{\Phi})$ for which the above process yields a set R of size $|R| \leq (1-k^{15}/2^k)n$ is bounded by $\exp(-nk^3/2^k)\mathrm{E}\,|\mathcal{S}(\mathbf{\Phi})|$.*

Proof. Let $\sigma \in \{0,1\}^V$ be an assignment, say $\sigma = \mathbf{1}$. Let Q be the set of all variables that support fewer than 12 clauses. By Lemma 6.20 we may condition on $|Q| \leq k^{12}2^{-k}n$. Assume that $|R| \leq (1-k^{15}/2^k)n$. Then there exists a set $S \subset V \setminus (R \cup Q)$ of size $\frac{1}{2}k^{15}n/2^k \leq S \leq k^{15}n/2^k$ such that each variable in S supports ten clauses that contain another variable from $S \cup Q$. With $s = |S|/n$ the probability of this event is bounded by

$$\binom{m}{10sn}\left[\frac{2^{1-k}}{1-2^{1-k}} \cdot \frac{k^2|S \cup Q|^2}{n^2}\right]^{10sn} \quad \leq \quad \left[4\mathrm{e}k^2 s\right]^{10sn}.$$

Hence, the expected number of set S for which the aforementioned event occurs is bounded by

$$\binom{n}{s}\left[4\mathrm{e}k^2 s\right]^{10sn} \quad \leq \quad \left[\frac{\mathrm{e}}{s} \cdot (4\mathrm{e}k^2 s)^2\right]^{sn} \leq \exp(-sn),$$

which implies the assertion. □

Let us call a variable x is *attached* if x supports a clause whose other $k-1$ variables belong to R.

Corollary 6.22 *The expected number of $\sigma \in \mathcal{S}(\mathbf{\Phi})$ in which more than $n/(k^2 2^k)$ variables $x \notin R$ that support at least one clause are not attached is bounded by $\mathrm{E}\,|\mathcal{S}(\mathbf{\Phi})| \cdot \exp(-n/(k^6 2^k))$.*

Proof. Let $F = V \setminus R$. By Proposition 6.21 we may assume that $|F| \leq nk^{15}/2^k$. Therefore, for each of the "special" clause C_x that we reserved for each x that supports at least one clause the probability of containing a variable from $F \setminus \{x\}$ is bounded by

$$(1+o_k(1))k \cdot \frac{|F|}{n} \leq \frac{3k^{16}}{2^k}.$$

Furthermore, these events are independent (because the clauses C_x were disregarded in the construction of R). Hence, the number of variables $x \notin R$ that support at least one clause but that are not attached is dominated by $\mathrm{Bin}(|F|, \frac{3k^{16}}{2^k})$. The assertion thus follows from Chernoff bounds. □

Let us call $S \subset V$ *dense* if each variable in S supports at least ten clauses and at most $2k$ clauses such that at least ten of them feature another variable from S.

Lemma 6.23 *For $\boldsymbol{d}$ chosen from $\boldsymbol{D}$, $\boldsymbol{m}$ chosen from $\boldsymbol{M}_{\boldsymbol{d}}$ and any $\sigma \in \{0,1\}^V$ the following holds w.h.p. Let $\mathcal{A}$ be the event that σ is a p-satisfying assignment of $\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$. Then*

$$\mathrm{P}\left[\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}} \text{ has a dense } S \subset V, |S| \leq n2^{-0.99k} \mid \mathcal{A}\right] = o(1).$$

Proof. We may assume that $\boldsymbol{d}$ satisfies (26); we emphasize that this is a property of $\boldsymbol{d}$ only, regardless of $\boldsymbol{m}$ or the event $\mathcal{A}$. Let $\mathcal{D}(S)$ be the event that $S \subset V$ is dense. We may fix (i.e., condition on) the specific clauses supported by each variable $x \in S$. Let $x \in S$ and let $i \in [m]$ be the index of a clause supported by x. Let ℓ be the type of clause i. For each $t \in \mathcal{T}$ let V_t be the set of literals l of type t. Then the probability that clause i contains another variable from S is bounded by

$$\sum_{j \in [k]} \frac{\mathrm{Vol}(V_{\ell_j} \cap \sigma^{-1}(0) \cap S)}{\mathrm{Vol}(V_{\ell_j} \cap \sigma^{-1}(0))} \leq k \max_{t \in \mathcal{T}} \left(\frac{\mathrm{Vol}(V_t \cap \sigma^{-1}(0) \cap S)}{\mathrm{Vol}(V_t \cap \sigma^{-1}(0))}\right).$$

Since $|V_t| \geq n2^{-0.8k}$ for all t w.h.p. by Lemma 5.9, we have $\mathrm{Vol}(V_t) \geq \frac{1}{3}kr|V_t| \geq 30 \cdot 2^{0.2k}n$. Furthermore, $\mathrm{Vol}(V_t \cap \sigma^{-1}(0)) \geq \frac{1}{3}\mathrm{Vol}(V_t)$ by the choice of $p(t)$. Hence, (26) yields

$$\frac{\mathrm{Vol}(V_t \cap S \cap \sigma^{-1}(0))}{\mathrm{Vol}(V_t \cap \sigma^{-1}(0))} \leq \frac{\mathrm{Vol}(S)}{\frac{1}{3}\mathrm{Vol}(V_t)} \leq \frac{\max\{kr, \ln(n/|S|)\}|S|}{2^{0.2k}n}.$$

Due to negative correlation, in total we obtain

$$\mathrm{P}\left[\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}} \in \mathcal{D}(S) | \mathcal{A}\right] \quad \leq \quad \binom{2k}{10}^{|S|} \cdot \left(\frac{k\max\{kr, \ln(n/|S|)\}|S|}{2^{0.2k}n}\right)^{10|S|}.$$

(The factor $\binom{2k}{10}^{|S|}$ accounts for the number of ways to choose 10 out of the at most $2k$ clauses that each variable in S supports.)

For $0 < s \leq 1/k^5$ let X_s be the number of sets S of size $|S| = sn$ for which $\mathcal{D}(S)$ occurs. Then

$$\begin{aligned}\mathrm{E}\left[X_s|\mathcal{A}\right] \quad &\leq \quad \binom{n}{sn}\binom{2k}{10}^{|S|}\left(\frac{k\max\{kr, \ln(n/|S|)\}|S|}{2^{0.2k}n}\right)^{10|S|} \leq \left[\frac{\mathrm{e}\left(k^2\max\{kr, -\ln(s)\}s\right)^{10}}{s4^k}\right]^{sn} \\ &= \quad \left[\frac{\mathrm{e}k^{20}\max\left\{s^9(kr)^{10}, s^9\ln^{10}(s)\right\}}{4^k}\right]^{sn}.\end{aligned}$$

There are two cases to consider. First, if $s \leq \ln(n)/n$, then the term in the brackets is clearly $o(1)$. Second, if $s \geq \ln(n)/n$, then we have the following bound. Since $s \leq s_{\max} = 2^{-0.99k}$ and as $x \mapsto x^9\ln^{10}x$ is monotonically increasing for $x < 0.1$, we have

$$s^9\ln^{10}(s) \leq s_{\max}^9\ln^{10}s_{\max} \leq s_{\max}^9(kr)^{10} \leq 2^{10k-8.91k} = 2^{1.09k}.$$

Hence, the entire bracket is bounded by $2^{-k/2}$. Summing over all possible s and using Markov's inequality completes the proof. □

Let us call a variable $x \in V$ *ξ-rigid* in $\sigma \in \mathcal{S}(\Phi)$ if for any $\tau \in \mathcal{S}(\Phi)$ with $\tau(x) \neq \sigma(x)$ we have $\mathrm{dist}(\sigma, \tau) \geq \xi n$.

Corollary 6.24 *W.h.p. for $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ and for $\boldsymbol{m}$ chosen from $\boldsymbol{M}_{\boldsymbol{d}}$ the following is true. Let $\sigma \in \{0,1\}^V$ and let $\mathcal{A}$ be the event that σ is a p-satisfying assignment of $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$. Moreover, let Y be the number of variables that are not $2^{-0.99k}$-rigid. Then*

$$\mathrm{P}\left[Y(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}) \leq (1 + 2k^{-2})2^{-k}n \mid \mathcal{A}\right] = 1 - o(1).$$

Proof. Let $\xi = 2^{-0.99k}$. We condition on the event $\mathcal{A}$. Consider a variable z that is either attached or in R. Let $\tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ be such that $\tau(z) \neq \sigma(x)$ and $\mathrm{dist}(\sigma, \tau) < n/2^{0.99k}$. Because z is attached or in R, the set

$$\Delta = \{x \in R : \tau(x) \neq \sigma(x)\}$$

is non-empty. Moreover, Δ is dense by the construction of R. Thus, Lemma 6.23 shows that $\mathrm{dist}(\sigma, \tau) \geq |\Delta| \geq n/2^{0.99k}$ w.h.p. Hence, w.h.p. all z that are either attached or in R are ξ-rigid.

Further, let $\mathcal{R}$ be the event that

- no more than $(1 + 1/k^2)2^{-k}n$ variables support no clause at all and
- at most $n/(k^2 2^k)$ variables $x \notin R$ that support at least one clause are not attached

Then Lemma 6.20 and Corollary 6.22 imply together with Proposition 6.5 that

$$\mathrm{P}\left[\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}} \in \mathcal{R} \mid \mathcal{A}\right] = 1 - o(1).$$

Hence, the total number of vertices that either do not support a clause or that are not attached is bounded by $(1 + 2/k^2)2^{-k}n$ □

Proof of Lemma 6.18. Suppose that $r = 2^k \ln 2 - c$. By Proposition 6.5 we have

$$\frac{1}{n} \ln \mathrm{E}\left[Z(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \geq 2^{-k}\left(c - \frac{\ln 2}{2} + o_k(1)\right)$$

w.h.p. Now, assume that in $\sigma \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ all but at most $(1+2k^{-2})2^{-k}n$ variables are ξ-rigid with $\xi = 2^{-0.99k}$. If $\tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$ is such that $\mathrm{dist}(\sigma,\tau) \leq \xi n$, then σ, τ agree on all ξ-rigid variables of σ. Hence,

$$\frac{1}{n} \ln \left\{\tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}) : \mathrm{dist}(\sigma,\tau) \leq \xi n\right\} \leq (1+2k^{-2})2^{-k} = (1+o_k(1))2^{-k}\ln 2.$$

As $c - \frac{\ln 2}{2} + o_k(1) > (1+o_k(1))\ln 2$ for $c > \frac{3}{2}\ln 2 + \varepsilon$ and k large enough, the assertion follows. □

6.6 Proof of Lemma 6.19

By Markov's inequality, it suffices to bound the expected number of paris $(\sigma,\tau) \in \mathcal{S}(\boldsymbol{\Phi})$ at the given Hamming distances. More precisely, let $\mathcal{Z}_x$ be the number of pairs $(\sigma,\tau) \in \mathcal{S}(\boldsymbol{\Phi})$ such that $\mathrm{dist}(\sigma,\tau)/n = x$. Let $h(x) = -x\ln x - (1-x)\ln(1-x)$ and set

$$q(x) = r \cdot \ln\left(1 - 2^{1-k} + 2^{-k}(1-x)^k\right).$$

Then

$$\frac{1}{n}\ln \mathrm{E}\left[\mathcal{Z}_x\right] \leq \ln 2 + h(x) + q(x). \tag{47}$$

We consider several cases.

Case 1: $k2^{-k} \leq x \leq (2k)^{-1}$**.** We have

$$\begin{aligned}
h(x)+q(x)+\ln 2 &\leq \ln 2 + x(1-\ln x) + r\left(-2^{1-k} + 2^{-k}(1-x)^k\right)\\
&\leq \ln 2 + x(1-\ln x) + 2^k \ln 2\left(-2^{1-k} + 2^{-k}(1-x)^k\right) + c2^{1-k} \quad [\text{as } r = 2^k\ln 2 - c]\\
&\leq x(1-\ln x) - \ln 2 + (1-x)^k \ln 2\\
&\leq x(1-\ln x) - \ln 2 + (1-kx+k^2x^2)\ln 2\\
&\leq x(1-\ln x) - kx + k^2x^2 = x\left[1 - \ln x - k + k^2 x\right].
\end{aligned}$$

If $k2^{-k} \leq x \leq k^{-2}$, then $1 - \ln x - k + k^2 x \leq 1 - \ln k + 1 < 0$. Moreover, if $k^{-2} \leq x \leq (2k)^{-1}$, then $1 - \ln x - k + k^2 x \leq 1 + 2\ln k - \frac{3}{4}k < 0$.

Case 2: $(2k)^{-1} < x < 0.01$**.** We have

$$\begin{aligned}
h(x)+q(x)+\ln 2 &\leq \ln 2 + x(1-\ln x) + r\left(-2^{1-k} + 2^{-k}(1-x)^k\right)\\
&\leq \ln 2 + x(1-\ln x) - \frac{r}{2^{k-1}} + \frac{r}{2^k}\exp(-kx)\\
&\leq x(1-\ln x) - \ln 2 + \frac{c}{2^{k-1}} + \exp(-kx)\ln 2\\
&\leq x(1-\ln x) + \frac{c}{2^{k-1}} + (\exp(-1/2) - 1)\ln 2
\end{aligned}$$

The last expression is negative for $x < 0.05$ (and k not too small).

Case 3: $0.01 < x < \frac{1}{2} - k2^{-k/2}$**.** We have

$$\begin{aligned}
h'(x) &= -\ln x + \ln(1-x),\\
q'(x) &= -\frac{kr(1-x)^{k-1}}{2^k - 2 + (1-x)^k} \geq -\frac{kr(1-x)^{k-1}}{2^k - 2} = \exp(-\Omega(k)).
\end{aligned}$$

Hence, for $0.01 \leq x < \frac{1}{2} - k^{-2}$ we have $h'(x) + q'(x) > 0$. Thus, $h(x) + q(x) + \ln 2$ is monotonically increasing in this interval. Now, let $x = \frac{1}{2} - \varepsilon$ for $k^{-2} \leq \varepsilon \leq k2^{-k/2}$. Then

$$\begin{aligned} h(x) &= \ln 2 - 2\varepsilon^2 + O(\varepsilon^3), \\ q(x) &= (2^k \ln 2 - c)\left(-2^{1-k} + 2^{1-2k} + 2^{-k}\left(\frac{1}{2} - \varepsilon\right)^k + \tilde{O}(8^{-k}\right) \\ &= -2\ln 2 + 2^{1-k}(c + \ln 2) + \tilde{O}(4^{-k}). \end{aligned}$$

Consequently,

$$h(x) + q(x) + \ln 2 \quad = \quad -2\varepsilon^2 + O(\varepsilon^3) + O(2^{-k}) < 0.$$

Case 4: $\frac{1}{2} + k2^{-k/2} \leq x < 1$**.** The function $h(x)$ satisfies $h(1-y) = h(y)$ for $0 < y < 1/2$. Furthermore, $q(x)$ is monotonically decreasing. Therefore, for any $x \geq \frac{1}{2} + k2^{-k/2}$ we have

$$\ln 2 + h(x) + q(x) \leq \ln 2 + h\left(\frac{1}{2} - k2^{-k}\right) + q\left(\frac{1}{2} - k2^{-k}\right) < 0.$$

In each case we have $\ln 2 + h(x) + q(x) < 0$. Thus, the assertion follows from (47) and Markov's inequality.

7 The second moment

Throughout this section we assume that $r = 2^{-k}\ln 2 - \rho$ with $\rho = \frac{3}{2}\ln 2 - \varepsilon_k$ for some sequence $\varepsilon_k = o_k(1)$ that tends to 0 sufficiently slowly. We also assume that $k \geq k_0$ for a large enough constant $k_0 > 3$. We let $\boldsymbol{d}$ denote a signed degree sequence $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ and we let $\boldsymbol{m}$ denote a vector chosen from $\boldsymbol{M}_{\boldsymbol{d}}$. By Lemma 5.11 we may assume that $|m(\ell) - \gamma_\ell n| \leq n^{2/3}$ for all ℓ. Let $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \{0,1\}^V$ denote a pair of assignments chosen uniformly and independently from the set of all assignments with p-marginals. Finally, let $\xi = k2^{-k/8}$.

7.1 Outline

The ***overlap*** of two assignments $\sigma, \tau \in \{0,1\}^V$ is the vector

$$\mathcal{O}(\sigma,\tau) = \left[\frac{1}{km\pi(t)} \sum_{l \in L : p_{\boldsymbol{d}}(l) = t} d_l \cdot \mathbf{1}_{\sigma(l)=1} \cdot \mathbf{1}_{\tau(l)=1}\right]_{t \in \mathcal{T}}.$$

In words, $\mathcal{O}(\sigma,\tau)$ captures the fraction of occurrences of literals of each type t that are true under both σ, τ. Since $\boldsymbol{\sigma}, \boldsymbol{\tau}$ are independent and have p-marginals, we have

$$\mathrm{E}\left[\mathcal{O}(\boldsymbol{\sigma}, \boldsymbol{\tau})\right] = \left[t^2\right]_{t \in \mathcal{T}}.$$

Set $\mathcal{O}^* = \left[t^2\right]_{t \in \mathcal{T}}$.

Let Z'' be the number of pairs (σ, τ) of $p_{\boldsymbol{d}}$-judicious satisfying assignments of $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$ such that

$$\mathrm{dist}(\sigma,\tau) \in \left[\frac{1}{2} - k^2 2^{-k/2}, \frac{1}{2} + k^2 2^{-k/2}\right]. \tag{48}$$

Moreover, let Z' be the number of pairs (σ, τ) of $p_{\boldsymbol{d}}$-judicious satisfying assignments of $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$ such that

$$\|\mathcal{O}(\sigma,\tau) - \mathcal{O}^*\|_\infty \leq \xi.$$

Proposition 7.1 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that* $\mathrm{E}\left[Z''(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \leq \mathrm{E}\left[Z'(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] + o(1)$.

The proof of Proposition 7.1 can be found in Section 7.2. Let Z denote the number of p-satisfying assignments of $\mathbf{\Phi}_{\boldsymbol{d}}$. Furthermore, let $\mathcal{Z}$ signify the number of good p-satisfying assignments of $\mathbf{\Phi}_{\boldsymbol{d}}$. In Section 8 we are going to establish the following.

Proposition 7.2 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that* $\mathrm{E}\left[Z'(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \leq C \cdot \mathrm{E}\left[Z(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2$.

Corollary 7.3 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that* $\mathrm{E}\left[\mathcal{Z}^2(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \leq C' \cdot \mathrm{E}\left[\mathcal{Z}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2$.

Proof. Let Y be the number of pairs (σ, τ) of good p-satisfying assignments of $\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$ such that

$$\mathrm{dist}(\sigma, \tau) \notin \left[\frac{1}{2} - k^2 2^{-k}, \frac{1}{2} + k^2 2^{-k}\right]. \tag{49}$$

By definition, for any good σ there are at most $\mathrm{E}\left[Z(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]$ p-satisfying τ such that (49) holds. Therefore,

$$\mathrm{E}\left[Y(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \leq \mathrm{E}\left[Z(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2. \tag{50}$$

Combining (50) with Proposition 7.1 and 7.2, we obtain for $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ w.h.p.

$$\begin{aligned}\mathrm{E}\left[\mathcal{Z}^2(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] &\leq \mathrm{E}\left[(Y + Z'')(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \\ &\leq \mathrm{E}\left[(Y + Z')(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] + o(1) \leq (C+1)\mathrm{E}\left[Z(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2 + o(1). \end{aligned} \tag{51}$$

By Proposition 6.5 we have $\mathrm{E}\left[Z(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] = \exp(\Omega(n))$. Furthermore, Proposition 6.16 yields $\mathrm{E}\left[Z(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \sim \mathrm{E}\left[\mathcal{Z}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]$. Consequently, (51) implies $\mathrm{E}\left[\mathcal{Z}^2(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right] \leq (C+2)\mathrm{E}\left[\mathcal{Z}(\mathbf{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2$, as desired. □

The second part of Theorem 5.1 follows directly from Corollary 7.3.

7.2 Proof of Proposition 7.1

We begin by relating the overlap to the Hamming distance.

Lemma 7.4 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that for all pairs $\sigma, \tau \in \{0,1\}^V$ satisfying (48) we have*

$$\bar{\mathcal{O}}(\sigma, \tau) = \frac{1}{km}\sum_{l \in L} d_l \mathbf{1}_{\sigma(l)=1}\mathbf{1}_{\tau(l)=1} = \frac{1}{4} + \tilde{O}(2^{-k/2}).$$

Proof. By Lemma 5.7 w.h.p.

$$\begin{aligned}\mathcal{O} &= \frac{1}{km}\sum_{x \in V}\frac{d_x}{2}\mathbf{1}_{\sigma(x)=\tau(x)} + O(|d_x^+ - d_x^-|) \\ &= \tilde{O}(2^{-k/2}) + \frac{1}{km}\sum_{x \in V}\frac{d_x}{2}\mathbf{1}_{\sigma(x)=\tau(x)} \\ &= \tilde{O}(2^{-k/2}) + \frac{1}{km}\sum_{x \in \mathcal{M}}\frac{d_x}{2} = \frac{1}{4} + \tilde{O}(2^{-k/2}),\end{aligned}$$

as claimed. □

Lemma 7.5 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that for any $\sigma, \tau \in \{0,1\}^V$ that satisfy (48) and that have p-marginals we have*

$$\frac{1}{n}\ln \mathrm{P}\left[\sigma, \tau \in \mathcal{S}(\mathbf{\Phi}_{\boldsymbol{d}})\right] \leq -2\ln 2 + O(k2^{-k}).$$

Proof. Much as in the first moment calculation in Section 6.3, here it is convenient to work with a different probability space. Namely, we let $\hat{\Omega}$ be the set of all vectors $(\hat{\sigma}_{ij}, \hat{\tau}_{ij})_{i\in[m],j\in[k]}$ of $0/1$ pairs. We define a probability distribution on $\hat{\Omega}$ in which the *pairs* $(\hat{\sigma}_{ij}, \hat{\tau}_{ij})_{i\in[m],j\in[k]}$ are mutually independent random variables. For any $i \in [m], j \in [k]$ we let $\hat{\mathrm{P}}\left[(\hat{\sigma}_{ij}, \hat{\tau}_{ij}) = (a, b)\right] = q^{ab}$, where the parameters q^{ab} are chosen so that the following equations hold:

$$\begin{aligned} q^{11} &= \bar{\mathcal{O}}(\sigma, \tau), \\ q^{10} &= q^{01}, \\ q^{11} + q^{10} &= \frac{1}{km} \sum_{l \in L} d_l \mathbf{1}_{\sigma(l)=1}, \\ \sum_{a,b=0}^{1} q^{ab} &= 1. \end{aligned}$$

Let $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\tau}})$ denote a random pair chosen from this distribution.

Proposition 6.5 and Lemma 7.4 ensure that w.h.p. $\boldsymbol{d}$ is such that

$$q^{11} = \frac{1}{4} + \tilde{O}(2^{-k/2}), \qquad q^{11} + q^{10} = \frac{1}{2} + O(2^{-k}). \tag{52}$$

Thus, we may assume that (52) holds.

Let B be the event that

$$\sum_{i,j} \hat{\boldsymbol{\sigma}}_{ij} = \sum_{l \in L} d_l \mathbf{1}_{\sigma(l)=1}, \quad \sum_{i,j} \hat{\boldsymbol{\tau}}_{ij} = \sum_{l \in L} d_l \mathbf{1}_{\tau(l)=1} \quad \text{and} \quad \sum_{i,j} \hat{\boldsymbol{\sigma}}_{ij} \hat{\boldsymbol{\tau}}_{ij} = km\bar{\mathcal{O}}(\sigma, \tau).$$

In addition, let S be the event that $\max_{j\in[k]} \hat{\boldsymbol{\sigma}}_{ij} = \max_{j\in[k]} \hat{\boldsymbol{\tau}}_{ij}$ for all $i \in [m]$. We claim that

$$\mathrm{P}\left[\sigma, \tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})\right] = \mathrm{P}\left[S | B\right]. \tag{53}$$

Indeed, any d-compatible formula Φ induces a pair $(\hat{\sigma}|_\Phi, \hat{\tau}|_\Phi) \in \hat{\Omega}$ defined by $\hat{\sigma}_{ij}|_\Phi = \sigma(\Phi_{ij})$, $\hat{\tau}_{ij}|_\Phi = \tau(\Phi_{ij})$. Clearly, the distribution of the random pair $(\hat{\sigma}|_{\boldsymbol{\Phi}_{\boldsymbol{d}}}, \hat{\tau}|_{\boldsymbol{\Phi}_{\boldsymbol{d}}})$ is identical to the distribution of $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\tau}})$ given B.

Due to independence, the probability of the event S is easy to compute. Indeed, with $q = q^{10} + q^{11}$ inclusion/exclusion yields

$$\hat{\mathrm{P}}\left[S\right] = \left[1 - 2q^k + (1 - 2q + q^{11})^k\right]^m$$

Furthermore, $\hat{\mathrm{P}}\left[B\right] = \exp(o(n))$ by the local limit theorem for the multinomial distribution. Hence, (53) yields

$$\begin{aligned} \frac{1}{n} \ln \mathrm{P}\left[\sigma, \tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})\right] &= \frac{1}{n} \ln \hat{\mathrm{P}}\left[S | B\right] \leq o(1) + \frac{1}{n} \ln \frac{\hat{\mathrm{P}}\left[S\right]}{\hat{\mathrm{P}}\left[B\right]} = o(1) + \frac{1}{n} \ln \hat{\mathrm{P}}\left[S\right] \\ &\sim r \ln \left[1 - 2q^k + (1 - 2q + q^{11})^k\right] \leq -r \left[2q^k - (1 - 2q + q^{11})^k\right]. \end{aligned}$$

Using (52) and simplifying completes the proof. □

Lemma 7.6 *Let $\lambda > 2^{-k}$ and $t \in \mathcal{T}$. For $\boldsymbol{d}$ chosen from $\boldsymbol{D}$ the following is true w.h.p. Let $\mathcal{H}''$ be the set of all pairs $\sigma, \tau \in \{0,1\}^V$ such that $|\mathcal{O}_t(\sigma, \tau) - 1/4| > \lambda$. Then*

$$|\mathcal{H}''| \leq 4^n \exp\left[-\frac{\lambda^2 n(t)}{18}\right].$$

Proof. Let $\boldsymbol{\sigma}'', \boldsymbol{\tau}'' \in \{0,1\}^V$ be chosen uniformly and independently. Then $\mathrm{E}\left[\mathcal{O}_t(\boldsymbol{\sigma}'', \boldsymbol{\tau}'')\right] = \frac{1}{4}$. Furthermore, $\mathcal{O}_t$ satisfies the following Lipschitz condition.

If $\sigma', \tau'', \sigma'', \tau'' \in \{0,1\}^V$ are such that there is a literal l_0 with $\mathcal{T}(l_0) = t$ such that $\sigma''(l) = \sigma'(l), \tau''(l) = \tau'(l)$ for all $l \notin \{l_0, \neg l_0\}$, then

$$|\mathcal{O}_t(\sigma'', \tau'') - \mathcal{O}_t(\sigma', \tau')| \leq \frac{2d_{l_0}}{km\pi(t)}.$$

Therefore, by Azuma's inequality for any $\lambda > 0$ we have

$$\mathrm{P}\left[\left|\mathcal{O}_t(\boldsymbol{\sigma}'', \boldsymbol{\tau}'') - \frac{1}{4}\right| > \lambda\right] \leq \exp\left[-\frac{\lambda^2 (km\pi(t))^2}{9\sum_{l \in \mathcal{L}: \mathcal{T}(l)=t} d_l^2}\right] \leq \exp\left[-\frac{\lambda^2 n(t)}{18}\right],$$

where the last step follows from part 2 of Proposition 6.5. □

Proof of Proposition 7.1. Let H'' be the set of pairs (σ, τ) such that

- σ, τ satisfy (48) and have p-marginals, and
- $\|\mathcal{O}(\sigma, \tau) - \mathcal{O}^*\|_\infty > \xi$.

Then by Lemma 7.6 and the second part of Proposition 6.5 w.h.p. (over the choice of $\boldsymbol{d}$) we have

$$|H''| \quad \leq \quad 4^n \exp\left[-\frac{\xi^2 n(t)}{36}\right] \leq 4^n \exp\left[-\frac{k^2 n}{36 \cdot 2^k}\right]. \tag{54}$$

Furthermore, by Lemma 7.5 w.h.p. (again over the choice of $\boldsymbol{d}$) we have

$$\mathrm{P}\left[\sigma, \tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})\right] \quad \leq \quad 4^{-n} \exp\left[\frac{O(k)}{2^k}\right] \qquad \text{for any } (\sigma, \tau) \in H''. \tag{55}$$

Combining (54) and (55), we obtain that w.h.p. $\boldsymbol{d}$ is such that

$$\mathrm{E}\left[(Z'' - Z')(\boldsymbol{\Phi}_{\boldsymbol{d}})\right] \quad \leq \quad \sum_{(\sigma,\tau) \in H''} \mathrm{P}\left[\sigma, \tau \in \mathcal{S}(\boldsymbol{\Phi}_{\boldsymbol{d}})\right] \leq |H''|\, 4^{-n} \exp\left[\frac{O(k)}{2^k}\right] = o(1).$$

Therefore, the definition of the distribution $\boldsymbol{M}_{\boldsymbol{d}}$ entails that w.h.p. $\boldsymbol{d}$ is such that

$$\mathrm{E}_{\boldsymbol{m}}\left[\mathrm{E}\left[(Z'' - Z')(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]\right] = \mathrm{E}\left[(Z'' - Z')(\boldsymbol{\Phi}_{\boldsymbol{d}})\right] = o(1).$$

Thus, the assertion follows from Markov's inequality. □

8 Proof of Proposition 7.2

We keep the notation and the assumptions of Section 7.

8.1 Overview

For two assignments σ, τ and a formula Φ with signed degree distribution $\boldsymbol{d}$ we define a matrix

$$\omega(\sigma, \tau, \Phi) = (\omega_{\ell,j}(\sigma, \tau, \Phi))_{\ell \in \mathcal{L}, j \in [k]}$$

by letting $\omega_{\ell,j}(\sigma, \tau, \Phi)$ be equal to the fraction of clauses of type ℓ whose jth literal is true under both σ, τ. We call $\omega_{\ell,j}(\sigma, \tau, \Phi)$ the ***overlap matrix*** of σ, τ in Φ. Recalling that $\boldsymbol{\sigma}, \boldsymbol{\tau}$ denote two independent uniformly distributed assignments with p-marginals, we define $\boldsymbol{\omega} = \omega(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})$; thus, $\boldsymbol{\omega}$ is a random matrix. We use the symbol ω to denote (fixed, non-random) matrices $\omega = (\omega_{\ell,j})_{\ell \in \mathcal{L}, j \in [k]}$ with entries in $[0,1]$. Furthermore, we let $\omega_\ell = (\omega_{\ell,j})_{j \in [k]}$ denote the ℓ-row of such a matrix ω. Finally, let $\omega^* = (\omega^*_{\ell,j})$ be the matrix with entries $\omega^*_{\ell,j} = \ell_j^2$ for all ℓ, j.

In addition, let $\mathcal{S}(\ell)$ be the event that both $\boldsymbol{\sigma}, \boldsymbol{\tau}$ satisfy all clauses of type ℓ of $\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}$. Let $\mathcal{S} = \bigcap_{\ell \in \mathcal{L}} \mathcal{S}(\ell)$. Further, let $\mathcal{B}(\ell, j)$ be the event that under both

$$\frac{1}{m(\ell)} \sum_{i \in M_{\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}}(\ell)} \boldsymbol{\sigma}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m},i,j}) \boldsymbol{\tau}(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m},i,j}) \doteq \ell_j,$$

i.e., the fraction of clauses of type ℓ whose jth literal is true equals $\ell_j + O(1/n)$. Let

$$\mathcal{B} = \bigcap_{\ell \in \mathcal{L},\, j \in [k]} \mathcal{B}(\ell, j).$$

In Section 9 we are going to prove the following.

Proposition 8.1 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that the following holds. Let $\mathcal{L}' \subset \mathcal{L}$ be a set of clause types and let $\mathcal{S}' = \bigcap_{\ell \in \mathcal{L}'} \mathcal{S}(\ell)$.*

1. *For all $\omega = (\omega_{\ell,j})$ such that $|\omega_{\ell,j} - \omega^*_{\ell,j}| \leq k^{-12}$ for all $\ell \in \mathcal{L}', j \in [k]$ we have the bound*

$$\mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega,\, \mathcal{B}\right] \quad \leq \quad \mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega^*,\, \mathcal{B}\right] \exp\left[\tilde{O}(4^{-k}) \sum_{\ell \in \mathcal{L}'} m(\ell) \left\|\omega_\ell - \omega^*_\ell\right\|_2^2\right].$$

2. *We have*

$$\mathrm{P}\left[\mathcal{S} | \boldsymbol{\omega} \doteq \omega^*,\, \mathcal{B}\right] \quad \leq \quad \mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega^*,\, \mathcal{B}\right] \exp\left[-\Theta(2^{-k}) \sum_{\ell \notin \mathcal{L}'} m(\ell)\right].$$

3. *For any assignment σ with p-marginals we have*

$$\mathrm{P}\left[\mathcal{S} | \boldsymbol{\omega} \doteq \omega^*,\, \mathcal{B}\right] \quad \leq \quad O(1) \cdot \mathrm{P}\left[\sigma \in \mathcal{S}_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}) | \sigma \text{ is } p\text{-judicious}\right]^2.$$

For $\omega = (\omega_{\ell,j})$ define $\mathcal{O}(\omega) \in [0,1]^{\mathcal{T}}$ by letting

$$\mathcal{O}_t(\omega) = \sum_{\ell \in \mathcal{L}} \sum_{j \in [k]: \ell_j = t} \frac{m(\ell)\omega_{\ell,j}}{km\pi(t)}.$$

We also let $\bar{\omega}$ denote the matrix with entries $\bar{\omega}_{\ell,j} = \mathcal{O}_{\ell_j}(\omega)$ for all ℓ, j. We say that ω is ***compatible*** with $\mathcal{O} \in [0,1]^{\mathcal{T}}$ if $\mathcal{O} = \mathcal{O}(\omega)$. In Section 8.2 we are going to prove the following.

Proposition 8.2 *W.h.p. $\boldsymbol{d}$ has the following property. For any $\omega = (\omega_{\ell,j})$ such that $\left\|\mathcal{O}(\omega) - \frac{1}{4}\mathbf{1}\right\|_\infty \leq 2\xi$ we have*

$$\mathrm{P}\left[\boldsymbol{\omega} \doteq \omega | \mathcal{O}(\boldsymbol{\omega}) \doteq \mathcal{O}(\omega),\, \mathcal{B}\right] \quad \leq \quad \mathrm{P}\left[\boldsymbol{\omega} \doteq \bar{\omega} | \mathcal{O}(\boldsymbol{\omega}) \doteq \mathcal{O}(\omega),\, \mathcal{B}\right] \exp\left[-\Omega_k(1) \cdot \sum_{\ell \in \mathcal{L}} m(\ell) \left\|\omega_\ell - \bar{\omega}_\ell\right\|_2^2\right].$$

Recall that $\mathcal{O}^* = (t^2)_{t \in \mathcal{T}}$. In Section 8.3 we will prove the following.

Corollary 8.3 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that the following holds. For any $\mathcal{O} = (\mathcal{O}_t)_{t \in \mathcal{T}}$ such that $\left\|\mathcal{O} - \frac{1}{4}\mathbf{1}\right\|_\infty \leq 2\xi$ we have*

$$\mathrm{P}\left[\mathcal{S} | \mathcal{O}(\boldsymbol{\sigma}, \boldsymbol{\tau}) \doteq \mathcal{O},\, \mathcal{B}\right] \quad \leq \quad O(1) \cdot \mathrm{P}\left[\mathcal{S} | \boldsymbol{\omega} \doteq \omega^*,\, \mathcal{B}\right] \exp\left[n \cdot \tilde{O}(2^{-k}) \sum_{t \in \mathcal{T}} \pi(t)(\mathcal{O}_t - \mathcal{O}^*_t)^2\right].$$

In Section 8.4 we will show the following.

Proposition 8.4 *There exists a constant $\eta > 0$ such that w.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that the following holds. For all $\mathcal{O}$ with $\left\|\mathcal{O} - \frac{1}{4}\mathbf{1}\right\|_\infty \leq 2\xi$ we have*

$$\mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right] \quad \leq \quad \eta \cdot \mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}^*\right].$$

Furthermore, $\mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \mathcal{O}^*\right] = \Theta(n^{|\mathcal{T}|-k|\mathcal{L}|})$.

Recall that $n(t)$ is the number of variables of type $t \in \mathcal{T}$. In Section 10 we are going to prove the following.

Proposition 8.5 *W.h.p. $\boldsymbol{d}, \boldsymbol{m}$ are such that the following holds. For all vectors $\lambda = (\lambda_t)_{t\in\mathcal{T}}$ with $\|\lambda\|_\infty \leq 1/8$ we have*

$$\mathrm{P}\left[\forall t \in \mathcal{T} : |\mathcal{O}_t(\boldsymbol{\sigma},\boldsymbol{\tau}) - \mathcal{O}^*_t| \geq \lambda_t\right] \quad \leq \quad \exp\left[-n \cdot \Omega_k(1)\sum_{t\in\mathcal{T}}\pi(t)\lambda_t^2\right].$$

Proof of Proposition 7.2. Suppose that $\mathcal{O} \in [0,1]^{\mathcal{T}}$ satisfies $\|\mathcal{O} - \mathcal{O}^*\|_\infty \leq \xi$. By Proposition 8.4 w.h.p.

$$\begin{aligned}\mathrm{P}\left[\mathcal{S},\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right] &= \mathrm{P}\left[\mathcal{S}|\mathcal{B},\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right]\mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right] \\ &\leq \eta \cdot \mathrm{P}\left[\mathcal{S}|\mathcal{B},\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right]\mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}^*\right]. \end{aligned} \tag{56}$$

Furthermore, by Corollary 8.3 w.h.p.

$$\mathrm{P}\left[\mathcal{S}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O},\mathcal{B}\right] \quad \leq \quad O(1)\cdot\mathrm{P}\left[\mathcal{S}|\boldsymbol{\omega} \doteq \omega^*,\mathcal{B}\right]\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t - \mathcal{O}^*_t)^2\right]. \tag{57}$$

Combining (56) and (57), we see that

$$\begin{aligned}\mathrm{P}\left[\mathcal{S},\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right] \quad \leq \quad & O(1)\cdot\mathrm{P}\left[\mathcal{S}|\mathcal{B},\boldsymbol{\omega} \doteq \omega^*\right]\cdot\mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}^*\right] \\ & \cdot\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t - \mathcal{O}^*_t)^2\right]. \end{aligned} \tag{58}$$

For an assignment σ with p-marginals let

$$b = \mathrm{P}\left[\sigma \text{ is } p\text{-judicious in } \boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}\right], \qquad s = \mathrm{P}\left[\sigma \in \mathcal{S}_p(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})|\sigma \text{ is } p\text{-judicious in } \boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}}\right].$$

Then by part 3 of Proposition 8.1, Corollary 8.4 and Corollary 6.9 we have

$$\mathrm{P}\left[\mathcal{S}|\mathcal{B},\boldsymbol{\omega} \doteq \omega^*\right]\cdot\mathrm{P}\left[\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}^*\right] \leq O(1)\cdot(bs)^2.$$

Therefore, (58) yields

$$\mathrm{P}\left[\mathcal{S},\mathcal{B}|\mathcal{O}(\boldsymbol{\sigma},\boldsymbol{\tau}) \doteq \mathcal{O}\right] \quad \leq \quad O(1)\cdot(bs)^2\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t - \mathcal{O}^*_t)^2\right]. \tag{59}$$

For a vector $\lambda = (\lambda_t)_{t\in\mathcal{T}}$ let

$$h(\lambda) = \mathrm{P}\left[\forall t \in \mathcal{T} : |\mathcal{O}_t(\boldsymbol{\sigma},\boldsymbol{\tau}) - \mathcal{O}^*_t| \geq \lambda_t\right].$$

Moreover, for $c = c(k) > 0$ a sufficiently large number let $\Lambda = \frac{c}{\sqrt{n}}\mathbf{Z}^{\mathcal{T}}_{\geq 0}$ be the positive $\mathcal{T}$-dimensional grid scaled by a factor of $c/\sqrt{n}$. In addition, let h be the number of assignments σ with p-marginals. Then by Proposition 8.5

and (59) there is a number $\zeta = \zeta(k) > 0$ such that

$$\begin{aligned}
\frac{\mathrm{E}\left[Z'(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]}{\mathrm{E}\left[Z(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]^2} &\leq O(1)\cdot\frac{\mathrm{E}\left[Z'(\boldsymbol{\Phi}_{\boldsymbol{d},\boldsymbol{m}})\right]}{(bhs)^2} \leq O(1)\cdot\sum_{\lambda\in\Lambda}h(\lambda)\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\lambda_t+c/\sqrt{n})^2\right]\\
&\leq O(1)\cdot\sum_{\lambda\in\Lambda}h(\lambda)\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)\lambda_t^2\right]\\
&\leq O(1)\cdot\sum_{\lambda\in\Lambda}\exp\left[n\sum_{t\in\mathcal{T}}\pi(t)\lambda_t^2\left[\tilde{O}(2^{-k})-\Omega_k(1)\right]\right]\\
&\leq O(1)\cdot\sum_{\lambda\in\Lambda}\exp\left[-n\cdot\Omega_k(1)\sum_{t\in\mathcal{T}}\pi(t)\lambda_t^2\right]\leq O(1)\cdot\sum_{\lambda\in\Lambda}\exp\left[-\zeta n\left\|\lambda\right\|_2^2\right]\\
&\leq O(1)\cdot\sum_{\boldsymbol{z}\in\mathbf{Z}^{\mathcal{T}}_{\geq 0}}\exp\left[-\zeta c^2\left\|\boldsymbol{z}\right\|_2^2\right]=O(1)\left[\sum_{z=0}^{\infty}\exp\left[-\zeta c^2 z^2\right]\right]^{|\mathcal{T}|}=O(1),
\end{aligned}$$

as desired. □

Notation for the proofs of Propositions 8.1–8.4. It will be convenient to work with a different probability space. Namely, let $\hat{\Omega}$ be the set of all pairs $(\hat{\sigma},\hat{\tau})$ of $0/1$ vectors

$$(\hat{\sigma},\hat{\tau}) = (\hat{\sigma}_{ij}(\ell),\hat{\tau}_{ij}(\ell))_{\ell\in\mathcal{L},i\in[m(\ell)],j\in[k]}.$$

Let $B_{\ell,j}\subset\hat{\Omega}$ be the event that

$$\frac{1}{m(\ell)}\sum_{i\in[m(\ell)]}\hat{\sigma}_{ij}(\ell)\doteq\ell_j \quad\text{and}\quad \frac{1}{m(\ell)}\sum_{i\in[m(\ell)]}\hat{\tau}_{ij}(\ell)\doteq\ell_j$$

for all $\ell\in\mathcal{L}, j\in[k]$. Let $B_\ell=\bigcap_{j\in[k]}B_{\ell,j}$ and let $B=\bigcap_{\ell\in\mathcal{L}}B_\ell$.

To define a measure $\hat{\mathrm{P}}$ on $\hat{\Omega}$, let $\boldsymbol{q}=(q^{ab}_{\ell,j})_{a,b\in\{0,1\},\ell\in\mathcal{L},j\in[k]}$ be a vector with entries in $[0,1]$ such that

$$\sum_{a,b=0}^{1}q^{ab}_{\ell,j}=1,\qquad q^{01}_{\ell,j}=q^{10}_{\ell,j} \tag{60}$$

for all ℓ, j. Define

$$q_{\ell,j}=q^{11}_{\ell,j}+q^{10}_{\ell,j} \tag{61}$$

so that

$$q^{00}_{\ell,j}=1-2q_{\ell,j}+q^{11}_{\ell,j}. \tag{62}$$

We define a measure $\hat{\mathrm{P}}=\hat{\mathrm{P}}_{\boldsymbol{q}}$ on $\hat{\Omega}$ as follows.

> For any $\ell=(\ell_1,\ldots,\ell_k)\in\mathcal{L}$, $i\in[m(\ell)]$ and $j\in[k]$ independently we choose a pair of values $(\hat{\boldsymbol{\sigma}}_{ij}(\ell),\hat{\boldsymbol{\tau}}_{ij}(\ell))\in\{0,1\}^2$ such that
>
> $$\hat{\mathrm{P}}\left[(\hat{\boldsymbol{\sigma}}_{ij}(\ell),\hat{\boldsymbol{\tau}}_{ij}(\ell))=(a,b)\right]=q^{ab}_{\ell,j}$$
>
> for any $a,b\in\{0,1\}$.

This probability space induces a random matrix $\hat{\boldsymbol{\omega}}=(\hat{\boldsymbol{\omega}}_{\ell,j})_{\ell,j}$ with entries

$$\hat{\boldsymbol{\omega}}_{\ell,j}=\frac{1}{m(\ell)}\sum_{i\in[m(\ell)]}\hat{\boldsymbol{\sigma}}_{ij}(\ell)\hat{\boldsymbol{\tau}}_{ij}(\ell).$$

We will use the probability space $(\hat{\Omega},\hat{\mathrm{P}})$ several times in the proof of the various propositions below, with various choices of $\boldsymbol{q}$.

8.2 Proof of Proposition 8.2

Consider any $\omega = (\omega_{\ell,j})$ such that $\left\|\mathcal{O}(\omega) - \frac{1}{4}\mathbf{1}\right\|_\infty \leq 2\xi$. We use the probability space $(\hat{\Omega}, \hat{\mathrm{P}})$ with the vector $\boldsymbol{q}$ defined by

$$q^{11}_{\ell,j} = \bar{\omega}_{\ell,j}, \qquad q_{\ell,j} = \ell_j \qquad \text{for all } \ell, j;$$

the remaining entries of $\boldsymbol{q}$ are determined by (60)–(62). Then the following is immediate from the construction.

Fact 8.6 *We have* $\mathrm{P}\left[\boldsymbol{\omega} \doteq \omega | \mathcal{O}(\boldsymbol{\omega}) \doteq \mathcal{O}(\omega),\, \mathcal{B}\right] = \mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega | \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right].$

Now,

$$\begin{aligned}
\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega | \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right] &= \frac{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega,\, \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right]}{\mathrm{P}\left[\mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right]} \doteq \frac{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega,\, B\right]}{\mathrm{P}\left[\mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right]} \\
&= \frac{\mathrm{P}\left[B | \hat{\boldsymbol{\omega}} \doteq \omega\right]}{\mathrm{P}\left[\mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right]} \cdot \mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega\right] \\
&\leq O(1) \cdot \frac{\mathrm{P}\left[B | \hat{\boldsymbol{\omega}} \doteq \bar{\omega}\right]}{\mathrm{P}\left[\mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\bar{\omega}),\, B\right]} \cdot \mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega\right],
\end{aligned}$$

The last step follows from the local limit theorem for the multinomial distribution because

$$\mathrm{E}\left[\sum_{i \in [m(\ell)]} \hat{\boldsymbol{\sigma}}_{ij}(\ell) \middle| \hat{\boldsymbol{\omega}} \doteq \bar{\omega}\right] = \mathrm{E}\left[\sum_{i \in [m(\ell)]} \hat{\boldsymbol{\tau}}_{ij}(\ell) \middle| \hat{\boldsymbol{\omega}} \doteq \bar{\omega}\right] = m(\ell) \cdot \ell_j$$

for all ℓ, j. Hence,

$$\frac{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega | \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right]}{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \bar{\omega} | \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}(\omega),\, B\right]} \leq \frac{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega\right]}{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \bar{\omega}\right]}.$$

For each $\ell \in \mathcal{L}$, $j \in [k]$ the sum $\sum_{i \in [m(\ell)]} \hat{\boldsymbol{\sigma}}_{ij}(\ell)\hat{\boldsymbol{\tau}}_{ij}(\ell)$ has a binomial distribution $\mathrm{Bin}(m(\ell), \bar{\omega}_{\ell,j})$. Furthermore, these random variables are mutually independent. Therefore, Chernoff bounds yield

$$\frac{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega\right]}{\mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \bar{\omega}\right]} \leq \exp\left[-\Omega_k(1) \sum_{\ell \in \mathcal{L}} \sum_{j \in [k]} m(\ell)(\omega_{\ell,j} - \bar{\omega}_{\ell,j})^2\right],$$

whence the assertion follows.

8.3 Proof of Corollary 8.3

Let ω be an overlap matrix such that $\mathcal{O} \doteq \mathcal{O}(\omega)$. Let $\mathcal{L}' = \mathcal{L}'(\boldsymbol{\omega})$ be the set of all $\ell \in \mathcal{L}$ such that $|\omega_{\ell,j} - 1/4| \leq \xi$ for all $j \in [k]$. Let $\mathcal{S}' = \bigcap_{\ell \in \mathcal{L}'} \mathcal{S}(\ell)$. Then

$$\begin{aligned}
P(\omega) &= \mathrm{P}\left[\mathcal{S}',\, \boldsymbol{\omega} \doteq \omega | \mathcal{O}(\boldsymbol{\omega}) \doteq \mathcal{O},\, \mathcal{B}\right] \\
&= \mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega,\, \mathcal{B}\right] \cdot \mathrm{P}\left[\boldsymbol{\omega} \doteq \omega | \mathcal{O}(\boldsymbol{\omega}) \doteq \mathcal{O},\, \mathcal{B}\right].
\end{aligned}$$

Let

$$\bar{P} = \mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega^*,\, \mathcal{B}\right] \cdot \mathrm{P}\left[\boldsymbol{\omega} \doteq \bar{\omega} | \mathcal{O}(\boldsymbol{\omega}) \doteq \mathcal{O},\, \mathcal{B}\right];$$

observe that $\bar{P}$ depends on $\mathcal{O}$ but not on the specific choice of ω. Then by Propositions 8.1 and 8.2

$$\begin{aligned}
P(\omega) &\leq \bar{P} \cdot \exp\left[\sum_{\ell \in \mathcal{L}} m(\ell) \left[\mathbf{1}_{\ell \in \mathcal{L}'} \cdot \tilde{O}(4^{-k}) \left\|\omega_\ell - \omega^*_\ell\right\|_2^2 - \Omega_k(1) \left\|\bar{\omega}_\ell - \omega_\ell\right\|_2^2\right]\right] \\
&\leq \bar{P} \cdot \exp\left[\sum_{\ell \in \mathcal{L}} m(\ell) \left[\mathbf{1}_{\ell \in \mathcal{L}'} \cdot \tilde{O}(4^{-k}) \left(\left\|\bar{\omega}_\ell - \omega_\ell\right\|_2^2 + \left\|\bar{\omega}_\ell - \omega^*_\ell\right\|_2^2\right) - \Omega_k(1) \left\|\bar{\omega}_\ell - \omega_\ell\right\|_2^2\right]\right] \\
&\leq \bar{P} \cdot \exp\left[\sum_{\ell \in \mathcal{L}} m(\ell) \left[\tilde{O}(4^{-k}) \left\|\bar{\omega}_\ell - \omega^*_\ell\right\|_2^2 - \Omega_k(1) \left\|\bar{\omega}_\ell - \omega_\ell\right\|_2^2\right]\right].
\end{aligned}$$

By the second part of Proposition 8.1,

$$\frac{1}{n}\ln\frac{\mathrm{P}\left[\mathcal{S}'|\boldsymbol{\omega}\doteq\omega^*,\,\mathcal{B}\right]}{\mathrm{P}\left[\mathcal{S}|\boldsymbol{\omega}\doteq\omega^*,\,\mathcal{B}\right]} \quad = \quad \Theta(2^{-k})\sum_{\ell\notin\mathcal{L}'}\frac{m(\ell)}{n}\leq\sum_{\ell\notin\mathcal{L}'}\frac{m(\ell)}{kn}\xi^2\leq\frac{1}{k}\sum_{\ell\notin\mathcal{L}'}\frac{m(\ell)}{n}\left\|\bar{\omega}_\ell-\omega_\ell\right\|_2^2.$$

Hence, letting $\tilde{P}=\mathrm{P}\left[\mathcal{S}|\boldsymbol{\omega}\doteq\omega^*,\,\mathcal{B}\right]\cdot\mathrm{P}\left[\boldsymbol{\omega}\doteq\bar{\omega}|\mathcal{O}(\boldsymbol{\omega})\doteq\mathcal{O},\,\mathcal{B}\right]$, we obtain

$$P(\omega) \quad \leq \quad \tilde{P}\exp\left[\tilde{O}(4^{-k})\sum_{\ell\in\mathcal{L}}m(\ell)\left\|\bar{\omega}_\ell-\omega_\ell^*\right\|_2^2-\sum_{\ell\in\mathcal{L}}\Omega_k(1)m(\ell)\left\|\bar{\omega}_\ell-\omega_\ell\right\|_2^2\right].$$

To proceed, we note that

$$\begin{aligned}\sum_{\ell\in\mathcal{L}}m(\ell)\left\|\bar{\omega}_\ell-\omega_\ell^*\right\|^2 &= \sum_{j\in[k]}\sum_{\ell\in\mathcal{L}}m(\ell)(\mathcal{O}^*_{\ell_j}-\mathcal{O}_{\ell_j})^2\\ &= \sum_{t\in\mathcal{T}}\sum_{\ell\in\mathcal{L}}\sum_{j\in[k]}m(\ell)(\mathcal{O}_t^*-\mathcal{O}_t)^2\cdot\mathbf{1}_{\ell_j=t}=km\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t^*-\mathcal{O}_t)^2.\end{aligned}$$

Thus,

$$P(\omega) \quad \leq \quad \tilde{P}\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t^*-\mathcal{O}_t)^2-\sum_{\ell\in\mathcal{L}}\Omega_k(1)m(\ell)\left\|\bar{\omega}_\ell-\omega_\ell\right\|^2\right].$$

Summing over all possible overlap matrices ω of assignments with p-marginals, we get

$$P \quad = \quad \sum_{\omega:\mathcal{O}(\omega)\doteq\mathcal{O}}P(\omega)=\mathrm{P}\left[\mathcal{S}'|\mathcal{O}(\boldsymbol{\omega})\doteq\mathcal{O},\,\mathcal{B}\right]\geq\mathrm{P}\left[\mathcal{S}|\mathcal{O}(\boldsymbol{\omega})\doteq\mathcal{O},\,\mathcal{B}\right],$$

which we can bound by

$$\begin{aligned}P \quad &\leq \quad \tilde{P}\cdot\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t^*-\mathcal{O}_t)^2\right]\sum_{\omega:\mathcal{O}(\omega)\doteq\mathcal{O}}\exp\left[-\sum_{\ell\in\mathcal{L}}\Omega_k(1)m(\ell)\left\|\bar{\omega}_\ell-\omega_\ell\right\|_2^2\right]\\ &= \quad \mathrm{P}\left[\mathcal{S}|\boldsymbol{\omega}\doteq\omega^*,\,\mathcal{B}\right]\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t^*-\mathcal{O}_t)^2\right]\\ &\qquad\cdot\sum_{\omega:\mathcal{O}(\omega)\doteq\mathcal{O}}\exp\left[-\sum_{\ell\in\mathcal{L}}\Omega(1)m(\ell)\left\|\bar{\omega}_\ell-\omega_\ell\right\|_2^2\right]\mathrm{P}\left[\boldsymbol{\omega}\doteq\bar{\omega}|\mathcal{O}(\boldsymbol{\omega})\doteq\mathcal{O},\,\mathcal{B}\right]\\ &\leq \quad O(1)\cdot\mathrm{P}\left[\mathcal{S}|\boldsymbol{\omega}\doteq\omega^*,\,\mathcal{B}\right]\exp\left[n\tilde{O}(2^{-k})\sum_{t\in\mathcal{T}}\pi(t)(\mathcal{O}_t^*-\mathcal{O}_t)^2\right],\end{aligned}$$

as desired.

8.4 Proof of Proposition 8.4

Let ω be such that $\mathcal{O}\doteq\mathcal{O}(\omega)$ and $\left\|\omega-\bar{\omega}\right\|_\infty\leq n^{-1/3}$. We are going to work with the probability space $(\hat{\Omega},\hat{\mathrm{P}})$ defined by letting

$$q^{11}_{\ell,j}=\omega_{\ell,j},\qquad q_{\ell,j}=\ell_j.$$

We claim that there exist numbers $0<c_k<c_k'$ (independent of ω) such that w.h.p. $\boldsymbol{d}$ is such that

$$c_k\leq n\mathrm{P}\left[B_{\ell,j}|\hat{\boldsymbol{\omega}}\doteq\omega\right]\leq c_k'\qquad\text{for all }\ell,j. \tag{63}$$

Indeed, given $\hat{\boldsymbol{\omega}}_{\ell,j} \doteq \omega_{\ell,j}$ the total number of indices $i \in [m(\ell)]$ such that $(\hat{\boldsymbol{\sigma}}_{ij}(\ell), \hat{\boldsymbol{\tau}}_{ij}(\ell)) = (1,0)$ has distribution

$$\mathrm{Bin}\left((1-\omega_{\ell,j})m(\ell), \frac{\ell_j - \omega_{\ell,j}}{1-\omega_{\ell,j}}\right).$$

Therefore, the probability that the total number of such i equals its expectation is in the interval $\left[c_{k,1}n^{-1/2}, c_{k,2}n^{-1/2}\right]$ for certain $c_{k,2} > c_{k,1} > 0$. Furthermore, given this event, the number of $i \in [m(\ell)]$ such that $(\hat{\boldsymbol{\sigma}}_{ij}(\ell), \hat{\boldsymbol{\tau}}_{ij}(\ell)) = (0,1)$ has distribution

$$\mathrm{Bin}\left((1-\ell_j)m(\ell), \frac{\ell_j - \omega_{\ell,j}}{1-\ell_j}\right).$$

Once more, the conditional probability that this random variable equals its expectation lies in $\left[c_{k,3}n^{-1/2}, c_{k,4}n^{-1/2}\right]$ for certain $c_{k,4} > c_{k,3} > 0$. Hence, setting $c_k = c_{k,1}c_{k,3}$ and $c'_k = c_{k,2}c_{k,4}$, we obtain (63).

Summing (63) over all (finitely many) possible ω with $\mathrm{P}\left[\omega \doteq \hat{\boldsymbol{\omega}}\right] > 0$ and $\mathcal{O}(\omega) \doteq \mathcal{O}$ and invoking Proposition 8.2, we find that w.h.p. over the choice of $\boldsymbol{d}$,

$$\begin{aligned}
\mathrm{P}\left[B_{\ell,j} | \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}\right] &= \sum_{\omega} \mathrm{P}\left[B_{\ell,j} | \hat{\boldsymbol{\omega}} \doteq \omega\right] \mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega\right] \\
&\leq o(1/n) + \sum_{\omega : \|\omega - \bar{\omega}\|_\infty \leq n^{-1/3}} \mathrm{P}\left[B_{\ell,j} | \hat{\boldsymbol{\omega}} \doteq \omega\right] \mathrm{P}\left[\hat{\boldsymbol{\omega}} \doteq \omega\right] \\
&\leq o(1/n) + c'_k/n \leq 2c'_k/n.
\end{aligned}$$

A similar calculation shows $\mathrm{P}\left[B_{\ell,j} | \mathcal{O}(\hat{\boldsymbol{\omega}}) \doteq \mathcal{O}\right] \geq \frac{1}{2}c_k/n$. As c_k, c'_k are independent of the specific vector $\mathcal{O}$, the assertion follows.

9 Proof of Proposition 8.1

We keep the notation and the assumptions of Section 7.

9.1 Outline

In Section 9.2 we will establish the following.

Proposition 9.1 *There exist C^2-functions $\mathcal{P}_\ell(\cdot)$ that range over matrices $\omega = (\omega_{\ell,j})_{\ell \in \mathcal{L}, j \in [k]}$ such that*

$$\|\omega_\ell - \omega^*_\ell\|_\infty < k^{-12} \quad \textit{for all} \quad \ell \in \mathcal{L}'$$

with the following properties.

1. *For all such ω we have*

$$\mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega, \mathcal{B}\right] = \exp\left[O(1) + \sum_{\ell \in \mathcal{L}'} m(\ell) \cdot \mathcal{P}_\ell(\omega_\ell)\right].$$

2. *For each ℓ, $\mathcal{P}_\ell$ is a function of the row ω_ℓ only.*

We need to analyse the functions $\mathcal{P}_\ell$ from Proposition 9.1. Crucially, ω^* turns out to be a stationary point.

Proposition 9.2 *The differentials of the functions $\mathcal{P}_\ell$ from Proposition 9.1 satisfy $D\mathcal{P}_\ell\left(\omega^*_\ell\right) = 0$ for all ℓ.*

The proof of Proposition 9.2 can be found in Section 9.3. Furthermore, in Section 9.4 we derive the following bound on the second derivatives of $\mathcal{P}_\ell$.

Proposition 9.3 *The functions $\mathcal{P}_\ell$ from Proposition 9.1 have the following property. For any j, j', ℓ we have*

$$\frac{\partial^2 \mathcal{P}_\ell}{\partial \omega_{\ell,j} \partial \omega_{\ell,j'}} \leq \tilde{O}(4^{-k})$$

on the entire domain of $\mathcal{P}_\ell$.

Corollary 9.4 *For any ω in the domain of $\mathcal{P}$ we have*

$$\mathcal{P}_\ell(\omega_\ell) \leq \mathcal{P}(\omega_\ell^*) + \tilde{O}(4^{-k}) \left\|\omega_\ell - \omega_\ell^*\right\|_2^2.$$

Proof. This follows directly from Propositions 9.2 and 9.3 and Taylor's formula. □

Finally, in Section 9.6 we will show of Proposition 8.1 follows from Proposition 9.1 and Corollary 9.4.

9.2 Proof of Proposition 9.1

To construct the functions $\mathcal{P}_\ell$, we are going to work with the probability space $(\hat{\Omega}, \hat{\mathrm{P}})$ from Section 8 once more; we are going to define the vector $\boldsymbol{q}$ that determines the measure $\hat{\mathrm{P}}$ so as to facilitate the definition of $\mathcal{P}_\ell$ in due course. Fix $\omega = (\omega_{\ell,j})_{\ell \in \mathcal{L}, j \in [k]}$ such that $\left\|\omega_\ell - \omega_\ell^*\right\|_\infty < k^{-12}$ for all $\ell \in \mathcal{L}'$. Let $B' = \bigcap_{\ell \in \mathcal{L}'} B_\ell$. Further, for $\ell \in \mathcal{L}$ and $j \in [k]$ let $C_{\ell,j}$ be the event that $\hat{\boldsymbol{\omega}}_{\ell,j} \doteq \omega_{\ell,j}$. Let $C_\ell = \bigcap_{j \in [k]} C'_{\ell,j}$ and let $C' = \bigcap_{\ell \in \mathcal{L}'} C'_\ell$. Finally, let $S' = \bigcap_{\ell \in \mathcal{L}'} S(\ell)$. The following two facts are direct consequences of the definition of $\hat{\mathrm{P}}$.

Fact 9.5 *If $\boldsymbol{q}$ is such that $\hat{\mathrm{P}}\left[B' \cap C'\right] > 0$, then $\hat{\mathrm{P}}\left[\cdot | B' \cap C'\right]$ is the uniform distribution over the set $B' \cap C'$.*

Fact 9.6 *Suppose that $\boldsymbol{q}$ is such that the conditional distribution $\hat{\mathrm{P}}\left[\cdot | B' \cap C'\right]$ is uniform. Then $\hat{\mathrm{P}}\left[S' | B',\, C'\right] = \mathrm{P}\left[\mathcal{S}' | \boldsymbol{\omega} \doteq \omega,\, \mathcal{B}\right]$.*

Thus, our goal is pick $\boldsymbol{q}$ such that $\hat{\mathrm{P}}\left[S' | B',\, C'\right]$ is easy to compute. Roughly speaking, we are going to accomplish this by choosing $\boldsymbol{q}$ so that $\hat{\mathrm{P}}\left[B',\, C' | S'\right]$ is as big as possible. To implement this, we first need to determine the unconditional probabilities $\hat{\mathrm{P}}\left[S'\right]$, $\hat{\mathrm{P}}\left[B', C'\right]$ as functions of $\boldsymbol{q}$.

Lemma 9.7 *Suppose that $\boldsymbol{q}$ is such that $q_{\ell,j} \in (0,1)$ for all $\ell \in \mathcal{L}'$, $j \in [k]$. Then*

$$\hat{\mathrm{P}}\left[S_i(\ell)\right] = 1 - 2\prod_{j=1}^{k}(1 - q_{\ell,j}) + \prod_{j=1}^{k}(1 - 2q_{\ell,j} + q_{\ell,j}^{11}) \tag{64}$$

for all $\ell \in \mathcal{L}'$, $i \in [m(\ell)]$, and

$$\frac{1}{n}\ln \hat{\mathrm{P}}\left[S'\right] \quad = \quad \sum_{\ell \in \mathcal{L}} \frac{m(\ell)}{n} \ln\left[1 - 2\prod_{j=1}^{k}(1 - q_{\ell,j}) + \prod_{j=1}^{k}(1 - 2q_{\ell,j} + q_{\ell,j}^{11})\right].$$

Proof. The first statement follows by inclusion/exclusion. The probability that $\max_{j \in [k]} \hat{\boldsymbol{\sigma}}_{ij}(\ell) = 0$ equals $\prod_{j=1}^{k}(1 - q_{\ell,j})$ as the components $\hat{\boldsymbol{\sigma}}_{ij}(\ell)$ are the results of independent $\mathrm{Be}(q_{\ell,j})$ experiments. For the event $\max_{j \in [k]} \hat{\boldsymbol{\tau}}_{ij}(\ell) = 0$ we get the exact same expression. Furthermore, the probability of $\max_{j \in [k]} \hat{\boldsymbol{\sigma}}_{ij}(\ell) = \max_{j \in [k]} \hat{\boldsymbol{\tau}}_{ij}(\ell) = 0$ equals $\prod_{j=1}^{k}(1 - 2q_{\ell,j} + q_{\ell,j}^{11})$. To see this, note that for each individual j we have

$$\mathrm{P}\left[\hat{\boldsymbol{\sigma}}_{ij}(\ell) = \hat{\boldsymbol{\tau}}_{ij}(\ell) = 0\right] = 1 - 2q_{\ell,j} + q_{\ell,j}^{11}$$

by inclusion/exclusion, and these events are independent for $j \in [k]$. The second one is due to independence over ℓ and i. □

Lemma 9.8 *For any $\mathbf{q}$ and any ℓ, j we have*

$$\hat{\mathrm{P}}\left[C_{\ell,j}\right] = \hat{\mathrm{P}}\left[\mathrm{Bin}(m(\ell), q^{11}_{\ell,j}) = \omega_{\ell,j} m(\ell) + O(1)\right]. \tag{65}$$

Furthermore, if $q^{11}(\ell, j) < 1$ then

$$\hat{\mathrm{P}}\left[B_{\ell,j} | C_{\ell,j}\right] \quad = \quad \Theta(n^{-1/2}) \cdot \hat{\mathrm{P}}\left[\mathrm{Bin}\left((1-\omega_{\ell,j})m(\ell), \frac{1 - 2q_{\ell,j} + q^{11}_{\ell,j}}{1 - q^{11}_{\ell,j}}\right) = m(\ell)(1 - 2\ell_j + \omega_{\ell,j})\right].$$

Proof. Recall that $C_{\ell,j}$ is the event that

$$\sum_{i \in [m(\ell)]} \hat{\boldsymbol{\sigma}}_{ij}(\ell) \cdot \hat{\boldsymbol{\tau}}_{ij}(\ell) = \omega_{\ell,j} m(\ell) + O(1).$$

By construction, the random variables $\hat{\boldsymbol{\sigma}}_{ij}(\ell) \cdot \hat{\boldsymbol{\tau}}_{ij}(\ell)$ are independent $\mathrm{Be}(q^{11}_{\ell,j})$ variables, and thus their sum has distribution $\mathrm{Bin}(m(\ell), q^{11}_{\ell,j})$. Hence we get (65).

Furthermore, once we condition on the event $C_{\ell,j}$, the remaining $(1-\omega_{\ell,j})m(\ell)$ pairs $(\hat{\boldsymbol{\sigma}}_{ij}(\ell), \hat{\boldsymbol{\tau}}_{ij}(\ell))$ are chosen conditional on the outcome being different from $(1,1)$. Hence, by construction each such pair takes the value $(0,0)$ with probability $\frac{1-2q_{\ell,j}+q^{11}_{\ell,j}}{1-q^{11}_{\ell,j}}$ independently (with the numerator resulting from (62)). In effect, the probability that the total number of $(0,0)$s equals $m(\ell)(1-2\ell_j+\omega_{\ell,j})$ is just

$$\mathrm{P}\left[\mathrm{Bin}\left((1-\omega_{\ell,j})m(\ell), \frac{1 - 2q_{\ell,j} + q^{11}_{\ell,j}}{1 - q^{11}_{\ell,j}}\right) = m(\ell)(1 - 2\ell_j + \omega_{\ell,j}) + O(1)\right].$$

Now, given that both this event and $C_{\ell,j}$ occur, the remaining $2(\ell_j - \omega)m(\ell)$ pairs $(\hat{\boldsymbol{\sigma}}_{ij}(\ell), \hat{\boldsymbol{\tau}}_{ij}(\ell))$ come up either $(1,0)$ or $(0,1)$ with probability $1/2$. By Stirling's formula, the probability that both outcomes occur an equal number of times is $\Theta(n^{-1/2})$. □

Note that

$$\hat{\mathrm{P}}\left[B', C'\right] \quad = \quad \prod_{\ell \in \mathcal{L}'} \hat{\mathrm{P}}\left[B(\ell) \cap C(\ell)\right] = \prod_{\ell \in \mathcal{L}'} \prod_{j=1}^{k} \hat{\mathrm{P}}\left[B(t_j, \ell) \cap C(t_j, \ell)\right] \tag{66}$$

because under $\hat{\mathrm{P}}$ the components of the vector $(\hat{\boldsymbol{\sigma}}_{ij}(\ell), \hat{\boldsymbol{\tau}}_{ij}(\ell))_{\ell,i,j}$ are independent.

Lemma 9.9 *There exists a vector $\mathbf{q}$ such that*

$$\ell_j \quad = \quad \frac{q_{\ell,j} - (q_{\ell,j} - q^{11}_{\ell,j}) \prod_{h \neq j}(1 - q_{\ell,h})}{1 - 2\prod_{h=1}^{k}(1 - q_{\ell,h}) + \prod_{h=1}^{k}(1 - 2q_{\ell,h} + q^{11}_{\ell,h})}, \tag{67}$$

$$\omega_{\ell,j} \quad = \quad \frac{q^{11}_{\ell,j}}{1 - 2\prod_{h=1}^{k}(1 - q_{\ell,h}) + \prod_{h=1}^{k}(1 - 2q_{\ell,h} + q^{11}_{\ell,h})}. \tag{68}$$

for all $\ell \in \mathcal{L}', j \in [k]$. This vector $\mathbf{q}$ satisfies

$$q_{\ell,j} = \ell_j - 2^{-k-1} + \tilde{O}(2^{-3k/2}) \text{ and } q^{11}_{\ell,j} = \omega_{\ell,j} + O(2^{-k}).$$

Proof. This follows from applying the inverse function theorem in a similar way as in the proof of Lemma 6.10. □

In the rest of this section, we fix $\mathbf{q}$ as in Lemma 9.9.

Lemma 9.10 *Let*

$$\begin{aligned}\mathcal{P}_\ell(\omega) &= \ln\left[1-2\prod_{j=1}^{k}(1-q_{\ell,j})+\prod_{j=1}^{k}(1-2q_{\ell,j}+q_{\ell,j}^{11})\right]\\ &\quad-\sum_{j\in[k]}\left[\psi(q_{\ell,j}^{11},\omega_{\ell,j})+(1-\omega_{\ell,j})\psi\left(\frac{1-2q_{\ell,j}+q_{\ell,j}^{11}}{1-q_{\ell,j}^{11}},\frac{1-2\ell_j+\omega_{\ell,j}}{1-\omega_{\ell,j}}\right)\right].\end{aligned}$$

Furthermore, let

$$\mathcal{P}(\omega)=\sum_{\ell\in\mathcal{L}'}\frac{m(\ell)}{n}\mathcal{P}_\ell(\omega). \tag{69}$$

Then

$$\hat{\mathrm{P}}\left[S'|B',C'\right]=\exp\left[n\mathcal{P}(\omega)+O(1)\right].$$

Proof. The choice of $\boldsymbol{q}$ ensures that for any ℓ and j,

$$\hat{\mathrm{E}}\left[\sum_{i\in[m(\ell)]}\hat{\boldsymbol{\sigma}}_{ij}(\ell)\cdot\hat{\boldsymbol{\tau}}_{ij}(\ell)\middle|S'\right] = \frac{m(\ell)q_{\ell,j}^{11}}{1-2\prod_{h=1}^{k}(1-q_{\ell,h})+\prod_{h=1}^{k}(1-2q_{\ell,h}+q_{\ell,h}^{11})}=\omega_{\ell,j}m(\ell); \tag{70}$$

indeed, by (64) the denominator in the middle term equals the probability of the event $S_i(\ell)$. Furthermore, by construction for any i,j,ℓ we have

$$\hat{\mathrm{P}}\left[\hat{\boldsymbol{\sigma}}_{ij}(\ell)=1,\hat{\boldsymbol{\tau}}_{ij}(\ell)=0,S_i(\ell)\right]=q_{\ell,j}^{10}\left(1-\prod_{h\neq j}(1-q_{\ell,j})\right)=(q_{\ell,j}-q_{\ell,j}^{11})\left(1-\prod_{h\neq j}(1-q_{\ell,j})\right).$$

As a consequence, (67) ensures that

$$\hat{\mathrm{E}}\left[\sum_{i\in[m(\ell)]}\hat{\boldsymbol{\sigma}}_{ij}(\ell)|S'\right] = \frac{q_{\ell,j}-(q_{\ell,j}-q_{\ell,j}^{11})\prod_{h\neq j}(1-q_h)}{1-2\prod_{h=1}^{k}(1-q_{\ell,h})+\prod_{h=1}^{k}(1-2q_{\ell,h}+q_{\ell,h}^{11})}=\ell_j m(\ell). \tag{71}$$

By inclusion/exclusion, we obtain from (70) and (71) that

$$\hat{\mathrm{E}}\left[\sum_{i\in[m(\ell)]}(1-\hat{\boldsymbol{\sigma}}_{ij}(\ell))\cdot(1-\hat{\boldsymbol{\tau}}_{ij}(\ell))\middle|S'\right] = (1-2\ell_j+\omega_{\ell,j})m(\ell). \tag{72}$$

Due to (70) and (72), a repeated application of Lemma 4.1 (the local limit theorem) yields

$$\hat{\mathrm{P}}\left[B',C'|S'\right] = \Theta(n^{-3k|\mathcal{L}'|/2}). \tag{73}$$

Invoking Lemma 9.8 and using the large deviations principle for the binomial distribution (Lemma 4.2), we can easily determine the *unconditional* probability of $B'\cap C'$: we have

$$\begin{aligned}\hat{\mathrm{P}}\left[B',C'\right] &= \prod_{\ell,j}\hat{\mathrm{P}}\left[C'_{\ell,j}\right]\hat{\mathrm{P}}\left[B'_{\ell,j}|C'_{\ell,j}\right]\\ &= \Theta(n^{-k|\mathcal{L}'|/2})\prod_{\ell,j}\hat{\mathrm{P}}\left[\mathrm{Bin}(m(\ell),q_{\ell,j}^{11})=\omega_{\ell,j}m(\ell)\right]\\ &\qquad\cdot\hat{\mathrm{P}}\left[\mathrm{Bin}\left((1-\omega_{\ell,j})m(\ell),\frac{1-2q_{\ell,j}+q_{\ell,j}^{11}}{1-q_{\ell,j}^{11}}\right)=m(\ell)(1-2\ell_j+\omega_{\ell,j})\right]\\ &= \Theta(n^{-3k|\mathcal{L}'|/2})\exp\left[\sum_{\ell,j}m(\ell)\left[\psi(q_{\ell,j}^{11},\omega_{\ell,j})+(1-\omega_{\ell,j})\psi\left(\frac{1-2q_{\ell,j}+q_{\ell,j}^{11}}{1-q_{\ell,j}^{11}},\frac{1-2\ell_j+\omega_{\ell,j}}{1-\omega_{\ell,j}}\right)\right]\right].\end{aligned}$$

Thus,

$$\begin{aligned}\ln \hat{\mathrm{P}}\left[S'|B',C'\right] &= \ln\left(\frac{\hat{\mathrm{P}}\left[S'\right]\hat{\mathrm{P}}\left[B',C'|S'\right]}{\hat{\mathrm{P}}\left[B',C'\right]}\right)\\ &= O(1)+\ln\hat{\mathrm{P}}\left[S'\right]\\ &\quad -\sum_{\ell,j}m(\ell)\left[\psi(q^{11}_{\ell,j},\omega_{\ell,j})+(1-\omega_{\ell,j})\psi\left(\frac{1-2q_{\ell,j}+q^{11}_{\ell,j}}{1-q^{11}_{\ell,j}},\frac{1-2\ell_j+\omega_{\ell,j}}{1-\omega_{\ell,j}}\right)\right].\end{aligned}$$

The assertion follows by plugging in the expression for $\mathrm{P}\left[S'\right]$ from Lemma 9.7. □

Finally, Proposition 9.1 follows from Fact 9.6 and Lemma 9.10.

9.3 Proof of Proposition 9.2

We start with the following observation.

Lemma 9.11 *Let $\boldsymbol{q}$ be the solution to (67) and (68) for $\omega=\omega^*$. There is $\gamma=\gamma(k)>0$ such that for any $\varepsilon>0$ and any $\ell\in\mathcal{L}'$ we have*

$$\hat{\mathrm{P}}\left[\left\|\hat{\boldsymbol{\omega}}_\ell-\omega^*_\ell\right\|_\infty>\varepsilon\,|\,S_\ell,\,B_\ell\right] \leq \exp(-\gamma\varepsilon^2 n+o(n)) \qquad \textit{and} \tag{74}$$

$$\hat{\mathrm{P}}\left[\left\|\hat{\boldsymbol{\omega}}_\ell-\omega^*_\ell\right\|_\infty>\varepsilon\,|\,B_\ell\right] \leq \exp(-\gamma\varepsilon^2 n+o(n)). \tag{75}$$

Proof. Equation (73) from the proof of Lemma 9.10 shows that

$$\hat{\mathrm{P}}\left[B_\ell|S_\ell\right]=\exp(o(n)). \tag{76}$$

Therefore, it is going to be sufficient to estimate $\hat{\mathrm{P}}\left[\left\|\hat{\boldsymbol{\omega}}_\ell-\omega^*_\ell\right\|_\infty>\varepsilon\,|\,S_\ell\right]$. If we just condition on the event S_ℓ, then the k-tuples $(\hat{\boldsymbol{\sigma}}_{ij}(\ell),\hat{\boldsymbol{\tau}}_{ij}(\ell))_{j\in[k]}$ of $0/1$ pairs are mutually independent for all $i\in[m(\ell)]$. Furthermore, given S_ℓ modifying just one such k-tuple can alter any entry $\hat{\boldsymbol{\omega}}_{\ell,j}$ by at most c/n, for some number $c=c(k)>0$. Therefore, Azuma's inequality yields

$$\hat{\mathrm{P}}\left[\left|\hat{\boldsymbol{\omega}}_{\ell,j}-\mathrm{E}\left[\hat{\boldsymbol{\omega}}_{\ell,j}\right]\right|>\varepsilon|S_\ell\right]\leq 2\exp(-\gamma\varepsilon^2 n), \tag{77}$$

for some $\gamma=\gamma(k)>0$. Since (68) ensures that $\hat{\mathrm{E}}\left[\hat{\boldsymbol{\omega}}_\ell|S_\ell\right]=\omega^*_\ell$, (74) follows from (76), (77) and the union bound.

To obtain (75), let $\boldsymbol{q}'$ be the vector with entries $\boldsymbol{q}'_{\ell,j}=p(\ell_j)$ for all ℓ,j. Then

$$\hat{\mathrm{P}}_{\boldsymbol{q}'}\left[B_\ell\right]=\exp(o(n)). \tag{78}$$

Furthermore, applying Azuma's inequality just as in the previous paragraph, we find that

$$\hat{\mathrm{P}}_{\boldsymbol{q}'}\left[\left|\hat{\boldsymbol{\omega}}_{\ell,j}-\mathrm{E}\left[\hat{\boldsymbol{\omega}}_{\ell,j}\right]\right|>\varepsilon\right]\leq 2\exp(-\gamma\varepsilon^2 n) \tag{79}$$

for some $\gamma=\gamma(k)>0$. Moreover, $\hat{\mathrm{E}}_{\boldsymbol{q}'}\left[\hat{\boldsymbol{\omega}}_\ell\right]=\omega^*_\ell$ by the choice of $\boldsymbol{q}'$. Thus, (75) follows from (78), (79) and the union bound. □

Proof of Proposition 9.2. Let $\ell\in\mathcal{L}'$. Let $\boldsymbol{q}$ be the solution to (67) and (68) for $\omega=\omega^*$. Then $\hat{\mathrm{P}}\left[\,\cdot\,|B'\right]$ is the uniform distribution over pairs $(\hat{\sigma},\hat{\tau})\in\Omega$ such that $(\hat{\sigma},\hat{\tau})\in B'$. Indeed, for $\omega=\omega^*$ the solution $\boldsymbol{q}$ to (67) and (68) satisfies $q^{11}_{\ell,j}=q^2_{\ell,j}$ for all ℓ,j. Therefore, for any $(\hat{\sigma},\hat{\tau})\in\Omega$ we have

$$\hat{\mathrm{P}}\left[\hat{\boldsymbol{\sigma}}=\hat{\sigma},\,\hat{\boldsymbol{\tau}}=\hat{\tau}\right]=q^{\sum_{\ell,i,j}\hat{\sigma}_{i,j}(\ell)+\hat{\tau}_{i,j}(\ell)}_{\ell,j}(1-q_{\ell,j})^{km-\sum_{\ell,i,j}\hat{\sigma}_{i,j}(\ell)+\hat{\tau}_{i,j}(\ell)} \tag{80}$$

Since the sums $\sum_{\ell,i,j}\hat{\sigma}_{i,j}(\ell)+\hat{\tau}_{i,j}(\ell)$ coincide for all $\hat{\sigma},\hat{\tau}\in B'$, (80) shows that $\hat{\mathrm{P}}\left[\,\cdot\,|B'\right]$ is uniform.

Let $H(\omega)$ be the number of pairs $(\hat{\sigma},\hat{\tau}) \in \hat{\Omega}$ such $(\hat{\sigma},\hat{\tau}) \in B'$ and $\hat{\boldsymbol{\omega}}(\hat{\sigma},\hat{\tau}) = \omega$. We claim that

$$\frac{1}{n} D \ln H(\omega^*) = o(1). \tag{81}$$

This can be verified either by representing $H(\omega)$ as a product of binomial coefficients and applying Stirlings formula or, alternatively, by using (75). Indeed, assume that (81) is false. Then for small enough $\varepsilon > 0$ there is $\delta > 0$ such that for some ω' with $\|\omega' - \omega^*\|_\infty \sim \varepsilon$ we have

$$\ln H(\omega') \geq \delta n + \max_{\omega:\|\omega-\omega^*\|_\infty<\varepsilon/2} \ln H(\omega) \tag{82}$$

(with both ε, δ possibly dependent on k but not on n). Letting

$$\bar{H} = \sum_{(\hat{\sigma},\hat{\tau})\in B'} H(\hat{\omega}(\hat{\sigma},\hat{\tau})),$$

we obtain from (75) that

$$\begin{aligned} 1 &\sim \hat{\mathrm{P}}\left[\|\hat{\boldsymbol{\omega}} - \omega^*\|_\infty < \varepsilon/2 | B'\right] = \frac{1}{\bar{H}} \sum_{(\hat{\sigma},\hat{\tau})\in B'} \mathbf{1}_{\|\hat{\omega}(\hat{\sigma},\hat{\tau})-\omega^*\|_\infty<\varepsilon/2} \cdot H_{\hat{\omega}(\hat{\sigma},\hat{\tau})} \\ &= \exp(o(n)) \cdot \max_{\omega:\|\omega-\omega^*\|_\infty<\varepsilon/2} H(\omega)/\bar{H}. \end{aligned} \tag{83}$$

However, combining (82) and (83) we get

$$\begin{aligned} \hat{\mathrm{P}}\left[\|\hat{\boldsymbol{\omega}} - \omega^*\|_\infty > \varepsilon/2 | B'\right] &\geq H(\omega')/\bar{H} \geq \exp(\delta n) \max_{\omega:\|\omega-\omega^*\|_\infty<\varepsilon/2} H(\omega)/\bar{H} \\ &\geq \exp(\delta n - o(n))\hat{\mathrm{P}}\left[\|\hat{\boldsymbol{\omega}} - \omega^*\|_\infty < \varepsilon/2 | B'\right] > 1, \end{aligned}$$

which is a contradiction. Hence, (81) follows.

Now, assume for contradiction that $D\mathcal{P}_\ell(\omega^*) \neq 0$. Because the function $\mathcal{P}_\ell(\,\cdot\,)$ remains fixed as $n \to \infty$, there exists a fixed $\varepsilon' > 0$ such that $\|D\mathcal{P}_\ell(\omega^*)\|_\infty > \varepsilon'$. Therefore, (81) entails that for any $\varepsilon > 0$ small enough exist ω', $\delta > 0$ such that $\|\omega' - \omega^*\|_\infty \sim \varepsilon$ and

$$\ln H(\omega') + n \cdot \mathcal{P}_\ell(\omega') \geq \delta n + \max_{\omega:\|\omega-\omega^*\|_\infty<\varepsilon/2} \ln H(\omega) + n \cdot \mathcal{P}_\ell(\omega), \tag{84}$$

with ε, δ independent of n. Let

$$\bar{H}_\ell = \sum_{(\hat{\sigma},\hat{\tau})\in B'} H(\hat{\omega}(\hat{\sigma},\hat{\tau})) \exp\left[n\mathcal{P}_\ell(\hat{\omega}(\hat{\sigma},\hat{\tau}))\right].$$

Then by (74),

$$\begin{aligned} 1 &\sim \hat{\mathrm{P}}\left[\|\hat{\boldsymbol{\omega}}_\ell - \omega^*_\ell\|_\infty < \varepsilon/2 | S_\ell, B'\right] \\ &= \frac{1}{\bar{H}} \sum_{(\hat{\sigma},\hat{\tau})\in B'} \mathbf{1}_{\|\hat{\omega}(\hat{\sigma},\hat{\tau})-\omega^*\|_\infty<\varepsilon/2} \cdot H_{\hat{\omega}(\hat{\sigma},\hat{\tau})} \exp\left[n\mathcal{P}_\ell(\hat{\omega}(\hat{\sigma},\hat{\tau})) + O(1)\right] \\ &= \exp(o(n)) \cdot \max_{\omega:\|\omega-\omega^*\|_\infty<\varepsilon/2} H(\omega) \exp(n\mathcal{P}_\ell(\omega))/\bar{H}_\ell. \end{aligned} \tag{85}$$

However, combining (84) and (85) we get

$$\begin{aligned} \hat{\mathrm{P}}\left[\|\hat{\boldsymbol{\omega}}_\ell - \omega^*_\ell\|_\infty > \varepsilon/2 | S_\ell, B'\right] &\geq \frac{H(\omega') \exp\left[n\mathcal{P}_\ell(\omega') + O(1)\right]}{\bar{H}_\ell} \\ &\geq \exp(\delta n) \max_{\omega:\|\omega-\omega^*\|_\infty<\varepsilon/2} H(\omega) \exp(n\mathcal{P}_\ell(\omega))/\bar{H}_\ell \\ &\geq \exp(\delta n - o(n))\hat{\mathrm{P}}\left[\|\hat{\boldsymbol{\omega}} - \omega^*\|_\infty < \varepsilon/2 | S_\ell, B'\right] > 1. \end{aligned}$$

This contradiction shows that $D\mathcal{P}_\ell(\omega^*) = 0$ for all ℓ. □

9.4 Proof of Proposition 9.3

We need to compute the second derivative of $\mathcal{P}_\ell$. In particular, we also need to differentiate $\boldsymbol{q} = \boldsymbol{q}(\omega)$ the solution to (67)–(68). Furthermore, we fix some type $\ell \in \mathcal{L}$ for the rest of this section. Let $\mathcal{W}_\ell$ denote the set of all vectors ω_ℓ such that $\left|\omega_{\ell,j} - \frac{1}{4}\right| \le k^{-4}$ for all $j \in [k]$. In Section 9.5 we are going to establish the following.

Lemma 9.12 *On $\mathcal{W}_\ell$ we have*

$$\frac{\partial q^{11}_{\ell,h}}{\partial \omega_{\ell,i}} = \mathbf{1}_{h=i} + \tilde{O}(2^{-k}), \qquad \frac{\partial^2 q^{11}_{\ell,h}}{\partial \omega_{\ell,i}\partial\omega_{\ell,j}} = \tilde{O}(2^{-k}),$$
$$\frac{\partial q_{\ell,h}}{\partial \omega_{\ell,i}} = \tilde{O}(2^{-k}), \qquad \frac{\partial^2 q_{\ell,h}}{\partial \omega_{\ell,i}\partial\omega_{\ell,j}} = \tilde{O}(2^{-k}).$$

for any $h, i, j \in [k]$.

We split the function $\mathcal{P}_\ell$ into a sum of various contributions: let

$$\begin{aligned}
\phi_\ell(q) &= \ln\left[1 - 2\prod_{j=1}^{k}(1-q_{\ell,j}) + \prod_{j=1}^{k}(1-2q_{\ell,j}+q^{11}_{\ell,j})\right] \quad \text{and}\\
\psi_\ell(\omega, q) &= \sum_{j\in[k]} \psi_{\ell,j}(\omega,q) + \tilde{\psi}_{\ell,j}(\omega,q) \qquad \text{with}\\
\psi_{\ell,j}(\omega,q) &= \psi(q^{11}_{\ell,j}, \omega_{\ell,j}),\\
\tilde{\psi}_{\ell,j}(\omega,q) &= (1-\omega_{\ell,j})\psi\left(\frac{1-2q_{\ell,j}+q^{11}_{\ell,j}}{1-q^{11}_{\ell,j}}, \frac{1-2\ell_j+\omega_{\ell,j}}{1-\omega_{\ell,j}}\right).
\end{aligned}$$

Lemma 9.13 *On $\mathcal{W}_\ell$ we have*

$$\frac{\partial^2 \phi_\ell(\boldsymbol{q})}{\partial \omega_{\ell,h}\,\partial\omega_{\ell,j}} \le \tilde{O}(4^{-k}) \qquad \textit{for all } h, j \in [k]\,.$$

Proof. By Lemma 9.9 for all $\omega \in \mathcal{W}_\ell$ we have $|q_{\ell,j} - 1/2| \le 1/k^2$ and $|q^{11}_{\ell,j} - 1/4| \le 1/k^2$ for all $j \in [k]$. For such vectors $\boldsymbol{q}_\ell$ we obtain the bounds

$$\begin{aligned}
\frac{\partial \phi_\ell}{\partial q_{\ell,j}}, \frac{\partial^2 \phi_\ell}{\partial q_{\ell,j}\partial q_{\ell,h}}, \frac{\partial^2 \phi_\ell}{\partial q^{11}_{\ell,j}\partial q_{\ell,h}} &= \tilde{O}(2^{-k}),\\
\frac{\partial \phi_\ell}{\partial q^{11}_{\ell,j}}, \frac{\partial^2 \phi_\ell}{\partial q^{11}_{\ell,j}\partial q^{11}_{\ell,h}} &= \tilde{O}(4^{-k})
\end{aligned}$$

for all $i, j, h \in [k]$. Therefore, the assertion follows from Lemma 4.3 (the chain rule) and Lemma 9.12. □

Let $\varepsilon > 0$. We say that $\Psi \in C^2((0,1)^2, \mathbf{R})$ is *ε-tame* on $\mathcal{Y} \subset (0,1)^2$ if the following conditions hold:

T1. For all $y \in (0,1)$ we have $\Psi(y,y) = 0$.

T2. On $\mathcal{Y}$ we have $\left|\sum_{i=1}^{2} \frac{\partial^2 \Psi}{\partial z_i \partial z_j}\right| \le \varepsilon$ for any $j = 1, 2$.

T3. On $\mathcal{Y}$ we have $\left|\sum_{i,j=1}^{2} \frac{\partial^2 \Psi}{\partial z_i \partial z_j}\right| \le \varepsilon^2$.

T4. On $\mathcal{Y}$ we have $|\frac{\partial^2 \Psi}{\partial z_i \partial z_j}| \le 100$ for any $i, j = 1, 2$.

Let $f : (0,1)^k \to \mathbf{R}^2$, $(z_1, \ldots, z_k) \mapsto (f_1(z_1, \ldots, z_k), f_2(z_1, \ldots, z_k))$ be a C^2-function. We say that f is *ε-benign* on $\mathcal{W}$ if the following statements are true on $\mathcal{W}$:

B1. $\left|\frac{\partial f_1}{\partial z_1} - \frac{\partial f_2}{\partial z_1}\right| < \varepsilon$.

B2. $\left|\frac{\partial f_i}{\partial z_j}\right| < \varepsilon$ for any $1 < j \le k$ and $i = 1, 2$ and $\left|\frac{\partial f_i}{\partial z_1}\right| \le 100$.

B3. $\left|\frac{\partial^2 f_i}{\partial z_h \partial z_j}\right| < \varepsilon$ for any i and $(h, j) \neq (1, 1)$.

B4. $\left|\frac{\partial^2 f_1}{\partial z_1^2} - \frac{\partial^2 f_2}{\partial z_1^2}\right| < \varepsilon$ and $|\frac{\partial^2 f_1}{\partial z_1^2}| \le 100$.

Lemma 9.14 *There is an absolute constant $C > 0$ such that the following is true. Assume that f is ε-benign on $\mathcal{W}$ and that Ψ is ε-tame on $f(\mathcal{W})$. Then on $\mathcal{W}$ we have*

$$\frac{\partial^2 \Psi \circ f}{\partial z_i \partial z_j} \le C\varepsilon^2 \qquad \text{for any } i, j \in [k].$$

Proof. By Lemma 4.3 (the chain rule), we have

$$\frac{\partial^2 \Psi \circ f}{\partial z_i \partial z_j} = \sum_{h=1}^{2} \frac{\partial \Psi}{\partial y_h} \frac{\partial^2 f_h}{\partial z_i \partial z_j} + \sum_{a,b=1}^{2} \frac{\partial^2 \Psi}{\partial y_a \partial y_b} \frac{\partial f_a}{\partial z_i} \frac{\partial f_b}{\partial z_j}.$$

Since by **T4** and Taylor's formula we have $\frac{\partial \Psi}{\partial y_h} = O_k(\varepsilon)$, **B3** implies that for $(i, j) \neq (1, 1)$

$$\sum_{h=1}^{2} \frac{\partial \Psi}{\partial y_h} \frac{\partial^2 f_h}{\partial z_i \partial z_j} = O_k(\varepsilon^2).$$

Furthermore, as $\frac{\partial \Psi}{\partial y_h} = O_k(\varepsilon)$, **B4** yields

$$\sum_{h=1}^{2} \frac{\partial \Psi}{\partial y_h} \frac{\partial^2 f_h}{\partial z_1^2} = \frac{\partial^2 f_1}{\partial z_1^2} \sum_{h=1}^{2} \frac{\partial \Psi}{\partial y_h} + \sum_{h=1}^{2} \frac{\partial \Psi}{\partial y_h} \left[\frac{\partial^2 f_h}{\partial z_1^2} - \frac{\partial^2 f_1}{\partial z_1^2}\right] = O_k(1) \sum_{h=1}^{2} \frac{\partial \Psi}{\partial y_h} + O_k(\varepsilon^2) = O_k(\varepsilon^2);$$

the last step follows from **T2** and Taylor's formula.

To deal with the second sum, we consider four cases.

Case 1: $i \neq 1, j \neq 1$**.** By **B2** we have $\frac{\partial f_a}{\partial z_i} \frac{\partial f_b}{\partial z_j} \le O_k(\varepsilon^2)$, and thus

$$\frac{\partial^2 \Psi}{\partial y_a \partial y_b} \frac{\partial f_a}{\partial z_i} \frac{\partial f_b}{\partial z_j} = O_k(\varepsilon^2)$$

by **T4**.

Case 2: $i = 1, j \neq 1$**.** We have

$$\begin{aligned}
\sum_{a,b=1}^{2} \frac{\partial^2 \Psi}{\partial y_a \partial y_b} \frac{\partial f_a}{\partial z_1} \frac{\partial f_b}{\partial z_j} &= \sum_{b=1}^{2} \frac{\partial f_b}{\partial z_j} \sum_{a=1}^{2} \frac{\partial^2 \Psi}{\partial y_a \partial y_b} \frac{\partial f_a}{\partial z_1} \overset{\textbf{B2}}{=} \sum_{b=1}^{2} O_k(\varepsilon) \sum_{a=1}^{2} \frac{\partial^2 \Psi}{\partial y_a \partial y_b} \frac{\partial f_a}{\partial z_1} \\
&\overset{\textbf{B1, T4}}{=} O_k(\varepsilon^2) + \frac{\partial f_1}{\partial z_1} \sum_{b=1}^{2} O_k(\varepsilon) \sum_{a=1}^{2} \frac{\partial^2 \Psi}{\partial y_a \partial y_b} \\
&\overset{\textbf{B2}}{=} O_k(\varepsilon^2) + \sum_{b=1}^{2} O_k(\varepsilon) \sum_{a=1}^{2} \frac{\partial^2 \Psi}{\partial y_a \partial y_b} \overset{\textbf{T2}}{=} O_k(\varepsilon^2).
\end{aligned}$$

Case 3: $i \neq 1, j = 1$**.** The same argument as in case 2 applies.

Case 4: $i = j = 1$**.** We have

$$\begin{aligned}
\sum_{a,b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\frac{\partial f_a}{\partial z_1}\frac{\partial f_b}{\partial z_1} &= \left(\frac{\partial f_1}{\partial z_1}\right)^2 \sum_{a,b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b} + \sum_{a,b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\left[\frac{\partial f_a}{\partial z_1}\frac{\partial f_b}{\partial z_1} - \left(\frac{\partial f_1}{\partial z_1}\right)^2\right] \\
&\overset{\text{B2, T3}}{=} O_k(\varepsilon^2) + \sum_{a,b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\frac{\partial f_a}{\partial z_1}\left[\frac{\partial f_b}{\partial z_1} - \frac{\partial f_1}{\partial z_1}\right] + \sum_{a,b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\frac{\partial f_1}{\partial z_1}\left[\frac{\partial f_a}{\partial z_1} - \frac{\partial f_1}{\partial z_1}\right] \\
&\overset{\text{B1}}{=} O_k(\varepsilon^2) + \sum_{b=1}^{2} O_k(\varepsilon) \sum_{a=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\frac{\partial f_a}{\partial z_1} + \sum_{a=1}^{2} O_k(\varepsilon) \sum_{b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\frac{\partial f_1}{\partial z_1} \\
&\overset{\text{B1}}{=} O_k(\varepsilon^2) + \sum_{a=1}^{2} O_k(\varepsilon) \sum_{b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b}\frac{\partial f_1}{\partial z_1} \overset{\text{B1}}{=} O_k(\varepsilon^2) + \sum_{a=1}^{2} O_k(\varepsilon) \sum_{b=1}^{2} \frac{\partial^2\Psi}{\partial y_a \partial y_b} \overset{\text{T2}}{=} O_k(\varepsilon^2).
\end{aligned}$$

Hence, in all cases we obtain a bound of $O_k(\varepsilon^2)$. □

Lemma 9.15 *The functions* $(y_1, y_2) \mapsto \psi(y_1, y_2)$ *and* $(y_1, y_2) \mapsto (1 - y_1)\psi(y_1, y_2)$ *are* $\tilde{O}(2^{-k})$*-tame on*

$$\mathcal{Y} = \left\{(y_1, y_2) \in (0, 1)^2 : |y_1 - y_2| \le k^3 2^{-k},\ \max_{i=1,2} |y_i - 1/4| \le 1/k^2\right\}.$$

Proof. It is straightforward to work out the differentials of ψ: we have

$$\begin{aligned}
\frac{\partial\psi}{\partial y_1} &= \frac{y_2}{y_1} - \frac{1 - y_2}{1 - y_1}, \qquad \frac{\partial\psi}{\partial y_2} = -\ln\left(\frac{y_2}{y_1}\right) + \ln\left(\frac{1 - y_2}{1 - y_1}\right), \\
\frac{\partial^2\psi}{\partial {y_1}^2} &= -\frac{y_2}{{y_1}^2} - \frac{1 - y_2}{(1 - y_1)^2}, \qquad \frac{\partial^2\psi}{\partial y_1 \partial y_2}\psi = \frac{1}{y_1} + \frac{1}{1 - y_1}, \qquad \frac{\partial^2\psi}{\partial {y_2}^2} = -\frac{1}{y_2} - \frac{1}{1 - y_2}.
\end{aligned}$$

Differentiating once more with respect to y_1, we get

$$\frac{\partial^3\psi}{\partial y_1^3} = \frac{2y_2}{y_1^3} - \frac{2(1 - y_2)}{(1 - y_1)^3}, \quad \frac{\partial^3\psi}{\partial y_1^2 \partial y_2} = -\frac{1}{y_1^2} + \frac{1}{(1 - y_1)^2}, \quad \frac{\partial^3\psi}{\partial y_1 \partial y_2^2} = 0.$$

Therefore, at $y_1 = y_2 + \varepsilon$ the second derivatives work out to be

$$\begin{aligned}
\frac{\partial^2\psi}{\partial y_1^2}(y_2 + \varepsilon, y_2) &= -\frac{1}{y_2} - \frac{1}{1 - y_2} + 2\varepsilon\left(\frac{1}{{y_2}^2} - \frac{1}{(1 - y_2)^2}\right) + O(\varepsilon^2), \\
\frac{\partial^2\psi}{\partial y_1 \partial y_2}(y_2 + \varepsilon, y_2) &= \frac{1}{y_2} + \frac{1}{1 - y_2} + \varepsilon\left(-\frac{1}{{y_2}^2} + \frac{1}{(1 - y_2)^2}\right) + O(\varepsilon^2), \\
\frac{\partial^2\psi}{\partial {y_2}^2}(y_2 + \varepsilon, y_2) &= -\frac{1}{y_2} - \frac{1}{1 - y_2}.
\end{aligned}$$

Hence, ψ is tame. Furthermore, differentiating $(y_1, y_2) \mapsto (1 - y_2)\psi(y_1, y_2)$ yields

$$\begin{aligned}
\frac{\partial}{\partial y_1}(1 - y_2)\psi(y_1, y_2) &= (1 - y_2)\frac{\partial}{\partial y_1}\psi(y_1, y_2), \\
\frac{\partial}{\partial y_2}(1 - y_2)\psi(y_1, y_2) &= (1 - y_2)\frac{\partial}{\partial y_2}\psi(y_1, y_2) - \psi(y_1, y_2), \\
\frac{\partial^2}{\partial y_1^2}(1 - y_2)\psi(y_1, y_2) &= (1 - y_2)\frac{\partial^2}{\partial y_1^2}\psi(y_1, y_2), \\
\frac{\partial^2}{\partial y_2^2}(1 - y_2)\psi(y_1, y_2) &= (1 - y_2)\frac{\partial^2}{\partial y_2^2}\psi(y_1, y_2) - 2\frac{\partial}{\partial y_2}\psi(y_1, y_2), \\
\frac{\partial^2}{\partial y_1 \partial y_2}(1 - y_2)\psi(y_1, y_2) &= (1 - y_2)\frac{\partial^2}{\partial y_1 \partial y_2}\psi(y_1, y_2) - \frac{\partial}{\partial y_1}\psi(y_1, y_2).
\end{aligned}$$

Hence, the fact that $(1-y_2)\psi(y_1,y_2)$ is ε-tame follows from the fact that ψ is. □

Lemma 9.16 *With $\boldsymbol{q}=\boldsymbol{q}(\omega)$ the functions*

$$\begin{aligned}\xi_{\ell,j} &: \omega \mapsto \left(q^{11}_{\ell,j},\omega_{\ell,j}\right),\\ \zeta_{\ell,j} &: \omega \mapsto (\zeta_{1,\ell,j},\zeta_{2,\ell,j}) = \left(\frac{1-2q_{\ell,j}+q^{11}_{\ell,j}}{1-q^{11}_{\ell,j}},\frac{1-2\ell_j+\omega_{\ell,j}}{1-\omega_{\ell,j}}\right)\end{aligned}$$

are $\tilde{O}(2^{-k})$-benign on $\mathcal{W}=\left\{\omega:\left\|\omega-\frac{1}{4}\mathbf{1}\right\|_\infty\le k^{-4}\right\}$.

Proof. The fact that $\xi_{\ell,j}$ is benign follows directly from Lemma 9.12. With respect to $\zeta_{\ell,j}$ we have

$$\begin{aligned}\frac{\partial\zeta_{2,\ell,j}}{\partial\omega_{\ell,j}} &= \frac{2(1-\ell_j)}{(1-\omega_{\ell,j})^2}, \qquad \frac{\partial^2\zeta_{2,\ell,j}}{\partial\omega^2_{\ell,j}}=\frac{4(1-\ell_j)}{(1-\omega_{\ell,j})^3},\\ \frac{\partial\zeta_{2,\ell,j}}{\partial\omega_{\ell,h}} &= 0, \qquad \frac{\partial^2\zeta_{2,\ell,j}}{\partial\omega_{\ell,h}\partial\omega_{\ell,i}}=0 \qquad (h\ne j),\\ \frac{\partial\zeta_{1,\ell,j}}{\partial\omega_{\ell,j}} &= \frac{(1-q^{11}_{\ell,j})\left[-2\frac{\partial q_{\ell,j}}{\partial\omega_{\ell,j}}+\frac{\partial q^{11}_{\ell,j}}{\partial\omega_{\ell,j}}\right]+\frac{\partial q^{11}_{\ell,j}}{\partial\omega_{\ell,j}}(1-2q_{\ell,j}+q^{11}_{\ell,j})}{(1-q^{11}_{\ell,j})^2}=\frac{2(1-q_{\ell,j})}{(1-q^{11}_{\ell,j})^2}+\tilde{O}(2^{-k}),\\ \frac{\partial^2\zeta_{1,\ell,j}}{\partial\omega^2_{\ell,j}} &= \frac{4(1-q_{\ell,j})}{(1-q_{\ell,j})^4}+\tilde{O}(2^{-k}),\\ \frac{\partial\zeta_{1,\ell,j}}{\partial\omega_{\ell,h}} &= \tilde{O}(2^{-k}), \qquad \frac{\partial^2\zeta_{1,\ell,j}}{\partial\omega_{\ell,h}\partial\omega_{\ell,i}}=\tilde{O}(2^{-k}) \qquad (h\ne j).\end{aligned}$$

Since $|q_{\ell,j}-\ell_j|\le\tilde{O}(2^{-k})$ and $|q^{11}_{\ell,j}-\omega_{\ell,j}|\le\tilde{O}(2^{-k})$ by Lemma 9.9, the assertion follows. □

Finally, Proposition 9.3 follows directly from Lemmas 9.13, 9.14, 9.15 and 9.16.

9.5 Proof of Lemma 9.12

Let

$$\begin{aligned}P_{\ell,j} &: \boldsymbol{q}\mapsto\frac{q_{\ell,j}-(q_{\ell,j}-q^{11}_{\ell,j})\prod_{h\ne j}(1-q_{\ell,h})}{1-2\prod_{h=1}^k(1-q_{\ell,h})+\prod_{h=1}^k(1-2q_{\ell,h}+q^{11}_{\ell,h})},\\ \Omega_{\ell,j} &: \boldsymbol{q}\mapsto\frac{q^{11}_{\ell,j}}{1-2\prod_{h=1}^k(1-q_{\ell,h})+\prod_{h=1}^k(1-2q_{\ell,h}+q^{11}_{\ell,h})}.\end{aligned}$$

A straightforward calculation shows that for $\boldsymbol{q}$ such that $|q_{\ell,j}-1/2|\le 1/k^2$ and $|q^{11}_{\ell,j}-1/4|\le 1/k^2$ we have

$$\begin{aligned}\frac{\partial P_{\ell,j}}{\partial q_{\ell,h}}&=\mathbf{1}_{j=h}+\tilde{O}(2^{-k}), \qquad \frac{\partial P_{\ell,j}}{\partial q^{11}_{\ell,h}}=\tilde{O}(2^{-k}),\\ \frac{\partial\Omega_{\ell,j}}{\partial q_{\ell,h}}&=\tilde{O}(2^{-k}), \qquad \frac{\partial\Omega_{\ell,j}}{\partial q^{11}_{\ell,h}}=\mathbf{1}_{j=h}+\tilde{O}(2^{-k})\end{aligned}$$

for any $j,h\in[k]$. Let $F:\boldsymbol{q}\mapsto\binom{(P_{\ell,j}(\boldsymbol{q}))_{j\in[k]}}{(\Omega_{\ell,j}(\boldsymbol{q}))_{j\in[k]}}$. Then the differential of F satisfies

$$DF=\begin{bmatrix}\left(\left(\frac{\partial P_{\ell,j}}{\partial q_{\ell,h}}\right)_{h\in[k]},\left(\frac{\partial P_{\ell,j}}{\partial q^{11}_{\ell,h}}\right)_{h\in[k]}\right)_{j\in[k]}\\ \left(\left(\frac{\partial P_{\ell,j}}{\partial q_{\ell,h}}\right)_{h\in[k]},\left(\frac{\partial P_{\ell,j}}{\partial q^{11}_{\ell,h}}\right)_{h\in[k]}\right)_{j\in[k]}\end{bmatrix}=\mathrm{id}+\tilde{O}(2^{-k})\mathbf{1}, \tag{86}$$

where id is the matrix with ones on the diagonal and zeros elsewhere, and $\mathbf{1}$ signifies the matrix with all entries equal to one. By the inverse function theorem, we have $D(F^{-1}) = (DF)^{-1}$. Furthermore, by (86) and Cramer's rule,

$$(DF)^{-1} = \mathrm{id} + \tilde{O}(2^{-k})\mathbf{1}. \tag{87}$$

Since $\boldsymbol{q}(\omega)$ is the solution to $F(\boldsymbol{q}) = \binom{(p(\ell_j))_{j\in[k]}}{(\omega_{\ell,j})_{j\in[k]}}$, (87) yields the assertions on the first derivatives $\frac{\partial q^{11}_{\ell,h}}{\partial \omega_{\ell,i}}$, $\frac{\partial q_{\ell,h}}{\partial \omega_{\ell,i}}$ in Lemma 9.12.

Proceeding to the second derivative, we highlight the following (folklore) fact.

Lemma 9.17 *Let $\varepsilon, \delta = \exp(-\Omega(k))$. Let $\mathcal{A}$ be the set of all $k \times k$ matrices $A = (A_{ij})$ such that $|A_{ii} - 1| < \varepsilon$ for all i and $|A_{ij}| < \delta$ for all $i \neq j$. Then A is regular and the operator $\mathrm{inv} : A \in \mathcal{A} \mapsto A^{-1} = (\mathrm{inv}_{st}A)_{s,t=1,\ldots,k}$ satisfies*

$$\left.\frac{\partial \mathrm{inv}_{st}}{\partial a_{ij}}\right|_A \leq \tilde{O}(\delta) - \mathbf{1}_{i=j=s=t}(1 + \tilde{O}(\varepsilon)) \quad \textit{for any } i, j, s, t \in [k].$$

Proof. This is a simple consequence of Cramer's rule. Indeed, let A'_{ij} be the matrix obtained from A by omitting row i and column j. Then

$$\mathrm{inv}_{st}A = (-1)^{s+t}\frac{\det A'_{ts}}{\det A}.$$

Thus, we need to differentiate $\det A'_{ts}$ and $\det A$. For any $i \neq j$ we have

$$\frac{\partial}{\partial a_{ii}} \det A \quad = \quad \prod_{h\neq i} a_{hh} + \tilde{O}(\delta) = 1 + \tilde{O}(\varepsilon) + \tilde{O}(\delta), \quad \frac{\partial}{\partial a_{ij}} \det A = \tilde{O}(\delta).$$

Similarly, for $i \neq j$ and $s \neq t$ we have

$$\frac{\partial}{\partial a_{ii}} \det A'_{tt} \quad = \quad \mathbf{1}_{i\neq t} \cdot (1 + \tilde{O}(\varepsilon)), \quad \frac{\partial}{\partial a_{ii}} \det A'_{ts} = \tilde{O}(\delta), \quad \frac{\partial}{\partial a_{ij}} \det A'_{ts} = \tilde{O}(\delta).$$

Thus, the assertion follows from the quotient rule. □

A direct calculation shows that for $\boldsymbol{q}$ such that $|q_{\ell,j} - 1/2| \leq 1/k^2$ and $|q^{11}_{\ell,j} - 1/4| \leq 1/k^2$ we have

$$\begin{aligned}
\frac{\partial^2 P_{\ell,j}}{\partial q_{\ell,h}\partial q_{\ell,i}}, \frac{\partial^2 P_{\ell,j}}{\partial q_{\ell,h}\partial q^{11}_{\ell,i}}, \frac{\partial^2 P_{\ell,j}}{\partial q^{11}_{\ell,h}\partial q^{11}_{\ell,i}} &= \tilde{O}(2^{-k}),\\
\frac{\partial^2 \Omega_{\ell,j}}{\partial q_{\ell,h}\partial q_{\ell,i}}, \frac{\partial^2 \Omega_{\ell,j}}{\partial q_{\ell,h}\partial q^{11}_{\ell,i}}, \frac{\partial^2 \Omega_{\ell,j}}{\partial q^{11}_{\ell,h}\partial q^{11}_{\ell,i}} &= \tilde{O}(2^{-k})
\end{aligned}$$

for any $h, i, j \in [k]$. Thus,

$$\left\|D^2 F\right\|_\infty \leq \tilde{O}(2^{-k}). \tag{88}$$

Because by the chain rule $D(\mathrm{inv} \circ DF) = (D\mathrm{inv}) \circ (D^2 F)$, the assertion on the second derivatives follows from Lemma 9.17, (87) and (88).

9.6 Completing the proof of Proposition 8.1

The first assertion is a direct consequence of Proposition 9.1 and Corollary 9.4. Similarly, the second assertion follows from Proposition 9.1 because $\mathcal{P}_\ell(\omega) \leq -\Omega_k(2^{-k})$ for all ℓ.

Finally, let $\omega = \omega^*$. It is straightforward to verify that by letting $q_{\ell,j}$ be as in Lemma 6.10 and by setting $q^{11}_{\ell,j} = q^2_{\ell,j}$ we obtain the unique solution to (67)–(68). We need to plug this solution into $\mathcal{P}(\omega)$: we have

$$\begin{aligned}
\ln\left[1 - 2\prod_{j=1}^k (1 - q_{\ell,j}) + \prod_{j=1}^k (1 - 2q_{\ell,j} + q^{11}_{\ell,j})\right] &= \ln\left[1 - 2\prod_{j=1}^k (1 - q_{\ell,j}) + \prod_{j=1}^k (1 - q_{\ell,j})^2\right]\\
&= 2\ln\left(1 - \prod_{j=1}^k 1 - q_{\ell,j}\right).
\end{aligned} \tag{89}$$

Moreover,

$$\begin{aligned}\psi(q^{11}_{\ell,j},\omega_{\ell,j}) &= \psi(q^2_{\ell,j},\ell^2_j) = -2\ell^2_j\ln\left(\frac{\ell_j}{q_{\ell,j}}\right)-(1-\ell^2_j)\ln\left(\frac{1-\ell^2_j}{1-q^2_{\ell,j}}\right)\\ &= -2\ell^2_j\ln\left(\frac{\ell_j}{q_{\ell,j}}\right)-(1-\ell^2_j)\left[\ln\left(\frac{1-\ell_j}{1-q_{\ell,j}}\right)+\ln\left(\frac{1+\ell_j}{1+q_{\ell,j}}\right)\right]. \qquad (90)\end{aligned}$$

Further,

$$\begin{aligned}(1-\ell^2_j)\psi\left(\frac{1-2q_{\ell,j}+q^{11}_{\ell,j}}{1-q^{11}_{\ell,j}},\frac{1-2\ell_j+\omega_{\ell,j}}{1-\omega_{\ell,j}}\right) &= (1-\ell^2_j)\psi\left(\frac{(1-q_{\ell,j})^2}{1-q^2_{\ell,j}},\frac{(1-\ell_j)^2}{1-\ell^2_j}\right)\\ = (1-\ell^2_j)\psi\left(\frac{1-q_{\ell,j}}{1+q_{\ell,j}},\frac{1-\ell_j}{1+\ell_j}\right)\\ = -(1-\ell_j)^2\ln\left(\frac{1-\ell_j}{1-q_{\ell,j}}\right)-(1-\ell^2_j)\ln\left(\frac{1+q_{\ell_j}}{1+\ell_j}\right)-2\ell_j(1-\ell_j)\ln\left(\frac{\ell_j}{q_{\ell,j}}\right). \qquad (91)\end{aligned}$$

Summing up (89)–(91), we find

$$\frac{n\mathcal{P}(\omega)}{2} = \sum_{\ell\in\mathcal{L}}m(\ell)\left[\ln\left(1-\prod_{j=1}^{k}1-q_{\ell,j}\right)-\sum_{j\in[k]}\psi(q_{\ell,j},\ell_j)\right].$$

Therefore, the third assertion follows from Remark 6.14.

10 Enumeration of Assignments with p-Marginals

In this section we will prove Lemma 6.2 and Proposition 8.5. Before we present the actual details we will introduce an appropriate framework, which will enable us to perform the enumeration of assignments with p-marginals, and pairs of such assignments with a given overlap.

In Section 5 we said that an assignment $\sigma\in\{0,1\}^V$ has $p_{\boldsymbol{d}}$*-marginals* if for any type $t\in\mathcal{T}$ we have

$$\sum_{l\in L:\mathcal{T}(l)=t}\mathbf{1}_{\sigma(l)=1}\cdot\frac{d_l}{km}\doteq p(t)\pi(t).$$

In words, the fraction of literal occurrences of type t that are true under σ equals $p(t)$ up to an error of $O(1/n)$. However, due to technical reasons and because it simplifies some of our calculations significantly, we will actually work with a slightly refined definition. Let us say that a signature (s,d^+,d^-) is *good*, if $d^+,d^-<3kr/4$ and $0<(d^+-d)^2\leq 100k2^k\ln k$. Instead of requiring that the fraction of literal occurrences of type t equals $p(t)$, we require that this is true for *every good signature*. That is, we say that an assignment $\sigma\in\{0,1\}^V$ has $p_{\boldsymbol{d}}$*-marginals* if for any good $s\in T$

$$\sum_{l\in L:T(l)=s}\mathbf{1}_{\sigma(l)=1}\cdot\frac{d_l}{km}=p(s)\sum_{l\in L:T(l)=s}\frac{d_l}{km},$$

and moreover, that fraction of literal occurrences of all other variables is $1/2$, i.e.,

$$\sum_{l\in L:p(l)=1/2}\mathbf{1}_{\sigma(l)=1}\cdot\frac{d_l}{km}=\frac{1}{2}\sum_{l\in L:p(l)=1/2}\frac{d_l}{km}.$$

We are going to prove Lemma 6.2 and Proposition 8.5 with this modified definition. It is easily checked that this modification does not affect any of the arguments in the previous sections.

Let $s \in T$ be any signature and set $L_s = \{\ell \in L : T(\ell) = s\}$. Moreover, denote by $V_s = \{|\ell| : \ell \in L_t\}$ and observe that $V_s = V_{\neg s}$. For any $\sigma \in \{0,1\}^n$ let us denote by the s-weight $w_s(\sigma)$ the number of satisfied literal occurrences, where only literals of signature s are considered, i.e.,

$$w_s(\sigma) = \sum_{\ell \in L_s} \mathbf{1}_{[\sigma(\ell)=1]} d_\ell.$$

Let us also define similar quantities with respect to the types. Let $t \in \mathcal{T}$ and set, as previously, $L_t = \{\ell \in L : \mathcal{T}(\ell) = t\}$. Denote by $\neg t \in \mathcal{T}$ the type satisfying $p(\neg t) = 1 - p(t)$. Note that $\neg t$ exists, and we have $L_{\neg t} = \{\neg \ell : \ell \in L_t\}$. Moreover, note that if $p(t) \neq 1/2$ we have $L_t \cap L_{\neg t} = \emptyset$, and $L_t = L_{\neg t}$ otherwise. Finally, set $V_t = \{|\ell| : \ell \in L_t\} = \{|\ell| : \ell \in L_{\neg t}\}$. In accordance with the case of signatures, let us for any $\sigma \in \{0,1\}^n$ denote by the t-weight $w_t(\sigma)$ the number of satisfied literal occurrences, where only literals of type t are considered, i.e.,

$$w_t(\sigma) = \sum_{\ell \in L_t} \mathbf{1}_{[\sigma(\ell)=1]} d_\ell.$$

Let $t_{1/2}$ be the type such that $p(t_{1/2}) = 1/2$. Since $L_{t_{1/2}} = L_{\neg t_{1/2}}$ it follows that in this special case

$$w_{t_{1/2}}(\sigma) = \sum_{v \in V_{t_{1/2}}} \mathbf{1}_{[\sigma(v)=1]} d_v + \mathbf{1}_{[\sigma(v)=0]} d_{\neg v}. \tag{92}$$

With the above notation, an assignment σ has p-marginals if and only if

$$\forall s \in T \setminus t_{1/2} : \; w_s(\sigma) = p(s)\pi(s)km \quad \text{and} \quad w_{t_{1/2}}(\sigma) = \frac{1}{2}\pi(t_{1/2})km.$$

The next proposition is the first step towards the estimation of the total number of assignments with p-marginals, c.f. Lemma 6.2. We denote by $H(x) = -x\ln x - (1-x)\ln(1-x)$ the entropy of x, and with $[z^n]f(z)$ the n-th coefficient in the Taylor series expansion of an analytic function f around 0.

Proposition 10.1 *W.h.p.* $\mathbf{d}$ *chosen from* $\mathbf{D}$ *has the following property. There is a constant* $C > 0$ *such that if we denote by* $\mathcal{S}$ *the set of signatures* $s \in T$ *with the property* $p(s) > 1/2$*, then*

$$|\mathcal{H}| = (C + o(1))\, n^{-|\mathcal{S}|/2} \exp\left\{\sum_{s \in \mathcal{S}} |V_s| H(p(s))\right\} \cdot [z^{\pi(t_{1/2})km/2}] \prod_{v \in V_{t_{1/2}}} (z^{d_v} + z^{d_{\neg v}}). \tag{93}$$

Proof. First of all, note that if for an assignment σ and a signature $s \in T$ with $p(s) > 1/2$ we have $w_s(\sigma) = \pi(s)km$, then the fraction of variables in V_s that are set to true is $p(s)$. Thus, the fraction of variables set to false is $1 - p(s) = p(\neg s)$, and we infer that

$$w_{\neg s}(\sigma) = \sum_{\ell \in L_{\neg s}} \mathbf{1}_{[\sigma(\ell)=1]} d_\ell = \sum_{v \in V_s} \mathbf{1}_{[\sigma(v)=0]} d_{\neg v} = p(\neg s)\pi(\neg s)km.$$

Consequently, for any such s the number of partial assignments $\sigma_s : V_s \to \{0,1\}$, with the property that the fraction of satisfied variables is $p(s)$ is

$$\binom{|V_s|}{p(s)|V_s|} = \frac{1}{\sqrt{2\pi p(s)(1-p(s))|V_s|}} e^{|V_s| H(p(s))}.$$

Since w.h.p. $\mathbf{d}$ is such that $|V_s| = (1 + o(1))\alpha_s n$ for some $\alpha_s = \alpha_s(k)$, this provides the exponential terms in (93).

It remains to bound the number of partial assignments $\sigma' : V_{t_{1/2}} \to \{0,1\}$ such that $w_{t_{1/2}} = \frac{1}{2}\pi(t_{1/2})km$. Define the generating function

$$F(z) = \sum_{\sigma' : V_{t_{1/2}} \to \{0,1\}} z^{w_{t_{1/2}}(\sigma')}$$

By definition, the sought quantity is $[z^{\pi(t_{1/2})km/2}]F(z)$. Moreover, the definition of $F(z)$ and (92) imply that

$$F(z) = \sum_{\sigma' : V_{t_{1/2}} \to \{0,1\}} \prod_{v \in V_{t_{1/2}}} (\mathbf{1}_{[\sigma'(v)=1]} z^{d_v} + \mathbf{1}_{[\sigma'(v)=0]} z^{d_{\neg v}})$$

The assertion follows. □

Lemma 6.2 follows immediately from the next statement, which is shown in Section 10.1.

Proposition 10.2 *W.h.p.* $\mathbf{d}$ *chosen from* $\mathbf{D}$ *has the following property. There is a constant* $C = C(k) > 0$ *such that is we write* $N = |V_{t_{1/2}}|$*, then*

$$[z^{\pi(t_{1/2})km/2}] \prod_{v \in V_{t_{1/2}}} (z^{d_v} + z^{d_{\neg v}}) = (C + o(1))N^{-1/2}2^N.$$

We proceed with the proof of Proposition 8.5, i.e., we want to enumerate pairs of assignments with p-marginals that have a specific overlap. Let $s \in T$ be a signature. For any $\sigma, \tau \in \{0,1\}^n$ denote the by the s-overlap $o_s(\sigma, \tau)$ the number of literal occurrences that are satisfied in both σ and τ, where we consider only literals of signature s, i.e.,

$$o_s(\sigma, \tau) = \sum_{\ell \in L_s} \mathbf{1}_{[\sigma(\ell)=\tau(\ell)=1]} d_\ell.$$

Similarly, for any type $t \in \mathcal{T}$ we denote by $o_t(\sigma, \tau)$ the number of satisfied literal occurrences in both σ and τ, where only literals of type t are considered. Note that $o_t(\sigma, \tau) = \mathcal{O}(\sigma, \tau)_t \pi(t) km$, where $\mathcal{O}$ is defined in Section 7.1. For the special case $t = t_{1/2}$ it follows

$$o_{t_{1/2}}(\sigma, \tau) = \sum_{v \in V_{t_{1/2}}} \mathbf{1}_{[\sigma(v)=\tau(v)=1]} d_v + \mathbf{1}_{[\sigma(\neg v)=\tau(\neg v)=0]}. \tag{94}$$

Let us begin with a simple observation. Let $s \in t$ such that $p(s) > 1/2$, and let σ, τ be two assignments with p-marginals. Note that if $w_s(\sigma, \tau) = (1+\delta)p(s)^2\pi(s)km$, for some $\delta \geq -1$, then the fraction of variables in V_s that are set to true in σ and τ is $(1+\delta)p(s)^2$. Consequently, the number of variables that are set to false in both assignments is $(1-p(s))|V_s| - (p(s)|V_s| - (1+\delta)p(s)^2|V_s|)$, and therefore

$$\begin{aligned} w_{\neg s}(\sigma, \tau) &= (1-p(s))\pi(\neg s)km - \big(p(s)\pi(\neg s)km - (1+\delta)p(s)^2\pi(\neg s)km\big) \\ &= \left(1 - \delta\frac{(1-p(\neg s))^2}{p(\neg s)^2}\right) p(\neg s)^2\pi(\neg s)km. \end{aligned}$$

In words, the overlap in s determines the overlap in $\neg s$. However, note that the s'-overlap, for any $s' \neq s, \neg s$, is not affected by the quantities $w_s(\sigma, \tau)$ and $w_{\neg s}(\sigma, \tau)$.

Let $t \in \mathcal{T}$ be a type. With the previous observation at hand we are able to estimate the number of pairs of p-satisfying assignments with a given t- and $\neg t$-overlap. The proof can be found in Section 10.2.

Proposition 10.3 *There is a* $c > 0$ *such that the following is true. Let* $\varepsilon, \varepsilon' > 0$*. Let* $t \in \mathcal{T}$ *be a type such that* $p(t) \neq 1/2$*. Denote by* $\mathcal{H}^2_{t,\neg t}(\varepsilon, \varepsilon')$ *the set of pairs* σ*,* τ *of assignments with* p*-marginals, such that*

$$|w_t(\sigma, \tau) - p(t)^2\pi(t)km| \geq \varepsilon p(t)^2\pi(t)km \quad \text{and} \quad |w_{\neg t}(\sigma, \tau) - p(\neg t)^2\pi(\neg t)km| \geq \varepsilon' p(\neg t)^2\pi(\neg t)km.$$

Then,

$$|\mathcal{H}^2_t(\varepsilon, \varepsilon')| \leq |\mathcal{H}|^2 \cdot \exp\left\{-cn\left(\varepsilon^2\pi(t) + \varepsilon'^2\pi(\neg t)\right)\right\}.$$

What remains is to enumerate pairs of p-satisfying assignments with a given $t_{1/2}$-overlap. The next proposition provides this number as the coefficient of an appropriately defined generating function.

Proposition 10.4 *Let $\varepsilon \in (-1/4, 1/4)$. Let $\mathcal{H}^2_{1/2}(\varepsilon)$ denote the set of pairs σ', τ' of assignments to the variables in $V_{t_{1/2}}$ such that*

$$o_{t_{1/2}}(\sigma', \tau') = \left(\frac{1}{4} + \varepsilon\right) \pi(t_{1/2})km$$

and

$$w_{t_{1/2}}(\sigma') = w_{t_{1/2}}(\sigma') = \pi(t_{1/2}km)/2.$$

Then $\mathcal{H}^2_{1/2}(\varepsilon) = [(xy)^{\pi(t_{1/2})km/2}\, u^{(1/4+\varepsilon)\pi(t_{1/2})km}]F(x, y, u)$, where

$$F(x, y, u) = \prod_{v \in V_{t_{1/2}}} \left((xyu)^{d_v} + (xyu)^{d_{\neg v}} + x^{d_v} y^{d_{\neg v}} + x^{d_{\neg v}} y^{d_v}\right).$$

Proof. Assign to a pair of assignments σ', τ' to the variables in $V_{t_{1/2}}$ the weight $x^{w_{t_{1/2}}(\sigma')}\, y^{w_{t_{1/2}}(\tau')}\, u^{o_{t_{1/2}}(\sigma', \tau')}$. Then, by using (92) and (94)

$$\begin{aligned}
&\sum_{\sigma', \tau' : V_{t_{1/2}} \to \{0,1\}} x^{w_{t_{1/2}}(\sigma')}\, y^{w_{t_{1/2}}(\tau')}\, u^{o_{t_{1/2}}(\sigma', \tau')} \\
&= \sum_{\sigma', \tau' : V_{t_{1/2}} \to \{0,1\}} \prod_{v \in V_{t_{1/2}}} \mathbf{1}_{[\sigma(v)=\tau(v)=1]}(xyu)^{d_v} + \mathbf{1}_{[\sigma(v)=\tau(v)=0]}(xyu)^{d_{\neg v}} \\
&\qquad\qquad + \mathbf{1}_{[\sigma(v)=1, \tau(v)=0]} x^{d_v} y^{d_{\neg v}} + \mathbf{1}_{[\sigma(v)=0, \tau(v)=1]} x^{d_{\neg v}} y^{d_v}.
\end{aligned}$$

Summing this expression up yields the claimed statement. □

The next statement provides the asymptotic value of the sought coefficients of $F(x, y, u)$ from the previous proposition. The proof can be found in Section 10.3.

Proposition 10.5 *W.h.p. $\mathbf{d}$ chosen from $\mathbf{D}$ has the following property. There is a constant $C = C(k, \varepsilon) > 0$ such that if we write $N = |V_{t_{1/2}}|$ and $M = \pi(t_{1/2})km$, then*

$$[(xy)^{M/2}\, u^{(1/4+\varepsilon)M}]F(x, y, u) = (C + o(1)) \cdot E \cdot N^{-3/2},$$

where

$$E = \rho^{-(1-4\varepsilon)M/2} \prod_{v \in V_{t_{1/2}}} (2 + 2\rho^{d_v + d_{\neg v}}) \tag{95}$$

and ρ is the solution to the equation

$$(1/4 + \varepsilon)M = \sum_{v \in V_{t_{1/2}}} \frac{d_v + d_{\neg v}}{2 + 2\rho^{d_v + d_{\neg v}}}. \tag{96}$$

In order to complete the proof of Proposition 8.5 we will estimate the exponential term in the previous statement as a function of ε. Note that if $\varepsilon = 0$, then clearly $\rho = 1$ and $E = 4^N$. Let $|\varepsilon| < 1/100$. We begin with providing bounds for the value of ρ from Equation (96). Let $f_g(\rho) = g/(2 + 2\rho^g)$, where $g \geq 3$. Then $f_g(1) = g/4$, $f'_g(1) = -g^2/8$ and

$$f''_g(\rho) = \frac{g^2\left(g\rho^{2\,g-2} - g\rho^{g-2} + \rho^{g-2} + \rho^{2\,g-2}\right)}{2\left(1 + \rho^g\right)^3}.$$

Note that if $0 \leq \rho \leq 1$, then, with room to spare, $|f''_g(\rho)| \leq g^3$. Moreover, if $\rho > 1$, then we may estimate f''_g as follows:

$$|f''_g(\rho)| < \frac{g^2\left((g+1)\rho^{2\,g-2} + g\rho^{g-2} + \rho^{g-2}\right)}{2\rho^{3g}} \leq \frac{g^3}{\rho^g} \leq g^3.$$

Let us write $\rho = 1+\delta$. Taylor's theorem then implies that $|f_g(\rho) - (g/4 - g^2\delta/8)| \le g^2\delta^2$. By writing $g_v = d_v + d_{\neg v}$ and recalling that $M = \sum_{v\in V_{t_{1/2}}} g_v$ we infer from (96)

$$-\delta\frac{S_2}{8} - \delta^2 S_3 \le \varepsilon M \le -\delta\frac{S_2}{8} + \delta^2 S_3, \quad \text{where} \quad S_i = \sum_{v\in V_{t_{1/2}}} g_v^i, \quad \text{for } i\in\{2,3\}.$$

In view of these inequalities we might expect that whenever ε is not too large, then $\delta \approx -\varepsilon 8M/S_2$. This can be made precise as follows. By solving the quadratic equations explicitly we infer that δ satisfies

$$\frac{1}{16}\frac{-S_2+\sqrt{S_2^2-256S_3\varepsilon M}}{S_3} \le \delta \le -\frac{1}{16}\frac{-S_2+\sqrt{S_2^2+256S_3\varepsilon M}}{S_3}$$

Note that $\mathbf{d}$ is such that w.h.p. $S_2 = \Theta(krM)$ and $S_3 = \Theta((kr)^2 M)$. Thus, for sufficiently large k

$$\sqrt{S_2^2+256S_3\varepsilon M} = S_2\sqrt{1+\frac{256S_3\varepsilon M}{S_2^2}} = S_2 + \frac{128S_3\varepsilon M}{S_2} + O\left(\frac{S_3^2\varepsilon^2M^2}{S_2^3}\right).$$

The square-root with the minus sign can be estimated analogously. We infer that

$$\rho = 1+\delta, \quad \text{where} \quad \delta = -\varepsilon\frac{8M}{S_2} + O((kr)^{-1}\varepsilon^2). \tag{97}$$

With the approximate value of ρ at hand we can proceed with estimating the exponential term in (95). First of, we rearrange terms to obtain

$$E = \rho^{-(1-4\varepsilon)M/2}\prod_{v\in V_{t_{1/2}}}(2+2\rho^{d_v+d_{\neg v}}) = 4^N\cdot\rho^{2\varepsilon M}\cdot\prod_{v\in V_{t_{1/2}}}(\rho^{-g_v/2}+\rho^{g_v/2})/2. \tag{98}$$

The bounds on ρ imply that

$$\rho^{2\varepsilon M} = \left(1-\varepsilon\frac{8M}{S_2}+O((kr)^{-1}\varepsilon^2)\right)^{2\varepsilon M} \le \exp\left\{-16\varepsilon^2\frac{M^2}{S_2}+O(\varepsilon^3(kr)^{-1}M)\right\}. \tag{99}$$

Regarding the last term involving the product in (98), we bound it by the following probabilistic considerations. Note that

$$\prod_{v\in V_{t_{1/2}}}(\rho^{-g_v/2}+\rho^{g_v/2})/2 = \sum_{(s_v):v\in V_{t_{1/2}},s_v\in\{-1,+1\}} 2^{-N}\rho^{-1/2\sum_v s_v g_v}.$$

Let $(S_v)_{v\in V_{t_{1/2}}}$ be a family of independent random variables, which are uniformly distributed in $\{-1,+1\}$. Then the last expression in the previous display is equal to the expected value of $\rho\{-1/2\sum_v s_v g_v\}$. We obtain

$$\mu := \mathrm{E}\left[\rho^{-\frac{1}{2}\sum_{v\in V_{t_{1/2}}} S_v g_v}\right] \le 2\sum_{t\ge 0}\mathrm{P}\left[|\sum_{v\in V_{t_{1/2}}} S_v g_v| = t\right](\rho^{t/2}+\rho^{-t/2})$$

Note that since either $\rho^{t/2}\ge 1$ or $\rho^{-t/2}\ge 1$ we may assume without loss of generality that $\rho\ge 1$. The advantage of the above formulation is that we can estimate rather easily the probability for a large deviation of the sum $S = \sum_v S_v g_v$. Indeed, if we change the value of any S_v to obtain a new sum S', then $|S-S'| = 2g_v$. By applying Azuma-Hoeffding we obtain

$$\mathrm{P}\left[|\sum_{v\in V_{t_{1/2}}} S_v g_v| = t\right] \le \exp\left\{-2t^2/\sum_v (2g_v)^2\right\} = \exp\{-t^2/2S_2\}.$$

Thus, by using (97) and noting that $\varepsilon \leq 0$ due to our assumption $\rho \geq 1$ we obtain the bound

$$\mu \leq 4 \sum_{t \geq 0} e^{-t^2/2S_2} \cdot \rho^{t/2} \leq 4 \sum_{t \geq 0} e^{-t^2/2S_2} \cdot \left(1 - \varepsilon \frac{8M}{S_2}\right)^{t/2} \leq 4 \sum_{t \geq 0} \exp\left\{-\frac{t^2}{2S_2} + |\varepsilon| \frac{4Mt}{S_2}\right\}.$$

Since the exponent is convex in t, it can easily be seen that it is maximized at $t = 4M|\varepsilon|$, where its value equals

$$-\frac{(4M|\varepsilon|)^2}{2S_2} + |\varepsilon| \frac{4M(4M|\varepsilon|)}{S_2} = 8\varepsilon^2 \frac{M^2}{S_2}.$$

Thus, $\mu = O(\sqrt{N}) e^{8\varepsilon^2 \frac{M^2}{S_2}}$, and by combining (98) and (99) we infer that $E \leq \sqrt{N} e^{-8\varepsilon^2 \frac{M^2}{S_2}}$. But since $S_2 = \Theta(krM)$ and $M = \Theta(krN)$, this is at most $\sqrt{N} e^{-c\varepsilon^2 N}$, for some $c > 0$.

Proposition 8.5 then follows immediately from Propositions 10.3-10.5, and the (aforementioned) observation that the t- and t'-overlap of $\boldsymbol{\sigma}, \boldsymbol{\tau}$ are independent for $t \neq t', \neg t$.

10.1 Proof of Proposition 10.2

Set $M = \pi(t_{1/2})km$. By the virtue of Cauchy's integral formula we obtain

$$I := [z^{M/2}]F(z) = \frac{1}{2\pi i} \oint_C F(z) z^{-M/2-1} dz.$$

Since F is analytic in $\mathbf{C}$, C can be any curve enclosing the origin. To estimate the integral we will use the saddle point method, which is commonly used to determine the asymptotic behavior of integrals that involve a large parameter, and are simultaneously subject to huge variations. For an excellent overview and numerous applications we refer the reader to [16].

The main idea is to choose C such that the integrand 'peaks' at a unique point, so that the main contribution to the integral comes from a small neighborhood of this maximum. We choose C to be the unit circle centered at the origin, i.e., $C = \{e^{i\theta} : -\pi < \theta < \pi\}$. Moreover, let $\theta_0 = \theta_0(n) = N^{-2/5}$, and write $C_0 = \{e^{i\theta} : |\theta| \leq \theta_0(n)\}$ for the restriction of C to the segment with $|\theta| \leq \theta_0(n)$. Then we may write $I = I_0 + I_1$, where

$$I_0 = \frac{1}{2\pi i} \oint_{C_0} F(z) z^{-M/2-1} dz \text{ and } I_1 = \frac{1}{2\pi i} \oint_{C \setminus C_0} F(z) z^{-M/2-1} dz.$$

By changing variables, the first integral becomes

$$I_0 = \frac{1}{2\pi} \int_{-\theta_0}^{\theta_0} H(\theta) d\theta, \quad \text{where} \quad H(\theta) = e^{-i\theta M/2} \cdot \prod_{v \in V_{t_{1/2}}} (e^{i\theta d_v} + e^{i\theta d_{\neg v}}). \tag{100}$$

Moreover, by using the trivial bound for complex integrals and the fact $|z| = 1$ on C we obtain

$$I_1 \leq 2\pi \cdot \sup_{z \in C \setminus C_0} |F(z)| \tag{101}$$

Our subsequent proof strategy is as follows. We will first compute the asymptotic value of the integral over the 'central region'; in particular, we show that

$$I_0 = (c + o(1)) N^{-1/2} 2^N \tag{102}$$

for an appropriate $c > 0$. Then, by using (101) we show that $I_1 = o(I_0)$. The two statements combined yield then immediately the conclusion of the proposition.

We proceed with showing (102). Recall that $|\theta| \leq \theta_0 = N^{-2/5}$, and note that for any d, d', by applying Taylor's Theorem

$$e^{i\theta d} + e^{i\theta d'} = 2 + i(d + d')\theta - \frac{1}{2}(d^2 + d'^2)\theta^2 + O\Big((1+i)(d^3 + d'^3)\theta^3\Big) \quad \text{uniformly for all } d, d' \in \mathbf{N}, |\theta| \leq \theta_0.$$

Let us write

$$S_2 = \frac{1}{4}\sum_{v\in V_{t_{1/2}}} d_v^2 + d_{\neg v}^2 + (d_v + d_{\neg v})^2 \quad \text{and} \quad S_j = \sum_{v\in V_{t_{1/2}}} (d_v^j + d_{\neg v}^j). \quad \text{for } j \geq 3.$$

Observe that $\mathbf{d}$ is w.h.p. such that $S_j = (1+o(1))c_j N$ for some $c_j = c_j(k) > 0$, where $2 \leq j \leq 9$. Using (100) we infer that the integrand satisfies

$$\begin{aligned} H(\theta) &= e^{-i\theta M/2}\cdot \prod_{v\in V_{t_{1/2}}} \Big(2 + i(d_v + d_{\neg v})\theta - \frac{1}{2}(d_v^2 + d_{\neg v}^2)\theta^2 + O\Big((1+i)(d_v^3 + d_{\neg v}^3)\theta^3\Big)\Big) \\ &= 2^N \exp\left\{-S_2\theta^2 + O\left((1+i)(S_3\theta^3 + S_4\theta^4 + \cdots + S_9\theta^9)\right)\right\} \\ &= (1+o(1))2^N \exp\Big\{-S_2\theta^2\Big\}, \qquad \text{since } \theta \leq N^{-2/5}. \end{aligned}$$

Thus,

$$\begin{aligned} (2\pi)I_0 = \int_{-\theta_0}^{\theta_0} H(\theta)d\theta &= (1+o(1))\,2^N \int_{-\theta_0}^{\theta_0} e^{-S_2\theta^2} d\theta \\ &= (1+o(1))\,\frac{2^N}{\sqrt{c_2 N}} \int_{-\sqrt{c_2}N^{1/10}}^{\sqrt{c_2}N^{1/10}} e^{-x^2}dx = (1+o(1))\,\frac{2^N}{\sqrt{2\pi c_2 N}}. \end{aligned}$$

This proves (102). To complete the proof we will show that $\sup_{z\in C\setminus C_0}|F(z)|$ is asymptotically negligible compared to I_0. First, for any $v \in V_{t_{1/2}}$

$$f_v(\theta) := |e^{i\theta d_v} + e^{i\theta d_{\neg v}}| = \sqrt{2 + 2\,\cos\left(\theta\,(d_v - d_{\neg v})\right)}$$

Let us collect some basic properties of f_v. Note that if $d_v = d_{\neg v}$, then $f_v(\theta) = 2$ for any $-\pi < \theta < \pi$. Otherwise, f is maximized for any

$$\theta \in \mathcal{M}_{d_v - d_{\neg v}} = \left\{ j\frac{2\pi}{|d_v - d_{\neg v}|} : |j| < \frac{|d_v - d_{\neg v}|}{2}\right\},$$

where $f(\theta) = 2$.

For a pair $(d_+, d_-) \in \mathbf{N}^2$ let $V_{d_+,d_-} \subseteq V_{t_{1/2}}$ denote the set of variables v such that $d_v = d_+$ and $d_{\neg v} = d_-$, and write $N_{d_+,d_-} = |V_{d_+,d_-}|$. Then,

$$|F(e^{i\theta})| = \prod_{v\in V_{t_{1/2}}} f_v(\theta) = \prod_{s=(d_+,d_-)} \big(2 + 2\cos(\theta(d_+ - d_-))\big)^{N_s/2}.$$

Note that $\sum_{s=(d_+,d_-)} N_s = N$. Thus, $|F(e^{i\theta})| \leq 2^N$ for all θ. However, this bound is achieved only if all factors are maximized simultaneously. We will argue in the sequel that if $|\theta| \in (\theta_0, \pi)$, then a linear (in N) fraction of the factors is $\leq 2 - O(N^{-4/5})$. It follows for some $\alpha > 0$ that

$$|F(e^{i\theta})| \leq 2^{(1-\alpha)N}\cdot(2 - O(N^{-4/5}))^{\alpha N} = 2^N \cdot e^{-O(N^{1/5})} = o(N^{-1/2}2^N) = o(I_0).$$

To see the claim, consider the specific pair $(d'_+, d'_-) = (kr, kr-1)$, and note that if k is sufficiently large, then $kr - 1 > kr/2 + 10\sqrt{k2^k \ln k}$. So, indeed $V_{d_+,d_-} \subseteq V_{t_{1/2}}$. Furthermore, $\mathbf{d}$ is such that w.h.p. there is a constant $\alpha = \alpha(k) > 0$ such that $N_{d_+,d_-} \geq \alpha N$. It follows that for all variables $v \in V_{d'_+,d'_-}$

$$f_v(\theta) = \sqrt{2 + 2\cos(\theta)}.$$

It can easily be verified that f_v is monotone increasing for $-\pi < \theta < 0$ and decreasing for $0 < \theta < \pi$. Thus, for any $|\theta| \in (\theta_0, \pi)$ we have $f_v(\theta) \leq \max\{f_v(\theta_0), f_v(-\theta_0)\}$. By using the Taylor series expansion of the cosine and the square root we obtain that

$$f_v(\varepsilon) = 2 - \frac{\theta^2}{4} + O(\theta^4), \quad \text{uniformly for all } -\pi < \theta < \pi.$$

We conclude that $f_v(\theta) \leq 2 - O(n^{-4/5})$ for at least αN variables v, and the proof is completed.

10.2 Proof of Proposition 10.3

We will exploit a concentration inequality due to McDiarmid [26]. We present it here in a simplified form that is appropriate for our purpose. Given a finite non-empty set B, we denote by $Sym(B)$ the set of all $|B|!$ permutations of the elements of B. Let $B_1, \dots, B_N$ be a family of finite non-empty sets, and denote by $\Omega = Sym(B_1) \times \cdots \times Sym(B_N)$. Moreover, let $\boldsymbol{\pi} = (\pi_1, \dots, \pi_N)$ be a family of independent random permutations, where π_i is drawn uniformly from $Sym(B_i)$.

Theorem 10.6 *Let c and r be positive constants. Suppose that $h : \Omega \to \mathbf{R}_+$ is such that for any $\pi \in \Omega$ the following conditions are satisfied.*

- *If π' can be obtained from π by swapping two elements, then $|h(\pi) - h(\pi')| \leq c$.*
- *If $h(\pi) \geq s$, then there is a set of at most rs coordinates such that $h(\pi') \geq s$ for any $\pi' \in \Omega$ that agrees with π on these coordinates.*

Let $Z = h(\boldsymbol{\pi})$ and let m be the median of Z. Then, for any $t > 0$

$$\mathrm{P}\left[|Z - m| > t\right] \leq 4 \exp\left(-\frac{t^2}{16rc^2(m+t)}\right).$$

Let us proceed with the proof of Proposition 10.3. We will assume without loss of generality that t is such that $p(t) > 1/2$. We will abbreviate $p = p(t)$, $q = p(\neg t)$. Let σ be an arbitrary assignment with p-marginals. Moreover, denote by $\boldsymbol{\tau}$ an assignment that is obtained by selecting for any signature $s \in t$ uniformly at random $p|V_s|$ variables from V_s and setting them to true, and setting all other variables in $V \setminus V_t$ arbitrarily so that $\boldsymbol{\tau}$ has p-marginals. Equivalently, we may generate $\boldsymbol{\tau}$ by permuting the variables in V_s randomly, and setting the first $p|V_s|$ variables in that permutation to true, for all $s \in t$. With this notation we obtain

$$|\mathcal{H}^2_{t,\neg t}(\varepsilon, \varepsilon')| \leq |\mathcal{H}|^2 \cdot \mathrm{P}\left[|w_t(\sigma, \boldsymbol{\tau}) - p^2\pi(t)km| \geq \varepsilon\pi(t)km\right]$$

The latter probability can be estimated with Theorem 10.6. Indeed, note that

- if τ, τ' have p-marginals and can be obtained by swapping the truth assignment of two variables, then

$$|w_t(\sigma, \tau) - w_t(\sigma, \tau')| \leq 2 \max_{v \in V_t} d_v \leq 4kr.$$

- if $w_t(\sigma, \tau) \geq s$, then there is a set S of $\leq s / \min_{v \in V_t} d_v \leq 2s/kr$ variables that are set to true, and any τ' with p-marginals that sets all variables is S to true satisfies $w_t(\sigma, \tau') \geq s$.

We thus may apply Theorem 10.6 with $c = 4kr$ and $r = 2/kr$. Moreover, trivially $\mathrm{E}\left[w_t(\sigma, \boldsymbol{\tau})\right] \leq \pi(t)km$. We infer that

$$\frac{|\mathcal{H}^2_{t,\neg t}(\varepsilon, \varepsilon')|}{|\mathcal{H}|^2} \leq 4 \exp\left(-\Theta(1)\frac{(\varepsilon\pi(t)km)^2}{kr \cdot \pi(t)km}\right) = 4 \exp\left(-\Theta(1)\, \varepsilon^2\pi(t)n\right).$$

Exactly the same argument, where we interchange the roles of t and $\neg t$, shows that also

$$\frac{|\mathcal{H}^2_{t,\neg t}(\varepsilon, \varepsilon')|}{|\mathcal{H}|^2} \leq 4 \exp\left(-\Theta(1)\frac{(\varepsilon'\pi(\neg t)km)^2}{kr \cdot \pi(\neg t)km}\right) = 4 \exp\left(-\Theta(1)\, \varepsilon^2\pi(\neg t)n\right).$$

The claim follows.

10.3 Proof of Proposition 10.5

Set $M = \pi(t_{1/2})km$. By applying Cauchy's integral formula we obtain

$$I := [(xy)^{M/2}\, u^{(1/4+\varepsilon)M}]F(x,y,u) = \frac{1}{(2\pi i)^3}\oint_{C_1}\oint_{C_2}\oint_{C_o} F(x,y,u)(xy)^{-M/2-1}u^{-(1/4+\varepsilon)M-1}dudydx.$$

The function F is analytic in $\mathbf{C}^3$, implying that C_1, C_2, C_o can be any curves enclosing the origin. We choose

$$C_1 = \{\rho e^{i\theta} : |\theta| < \pi\}, \quad C_2 = \{\rho e^{i\varphi} : |\varphi| < \pi\}, \quad C_o = \{\rho^{-2} e^{i\psi} : |\psi| < \pi\},$$

where ρ is the solution to the Equation (96). Some remarks are in place here. The choice of the integration paths may seem arbitrary at this point. Note, however, that F is symmetric with respect to x and y, and thus it is natural to assume similar integration curves for them. Moreover, the choice of ρ is guided by the general principles of the saddle-point method and is such that the integrand has a unique maximum at $(\theta, \varphi, \psi) = (0,0,0)$. Indeed, as we will show subsequently, the integrand is around $(0,0,0)$ of elliptic type; this allows us to reduce the estimation of the main terms to the evaluation of a 3-dimensional Gaussian integral.

Denote by $\mathcal{C}$ the restriction of the circles C_1, C_2, C_o to a small region around the origin, i.e.,

$$\mathcal{C} = \{\rho e^{i\theta} : |\theta| < N^{-2/5}\} \times \{\rho e^{i\varphi} : |\varphi| < N^{-2/5}\} \times \{\rho^{-2} e^{i\psi} : |\psi| < N^{-2/5}\}.$$

Then we may write $I = I_0 + I_1$, where

$$I_0 = \frac{1}{(2\pi i)^3}\oint_{\mathcal{C}} F(x,y,u)\,(xy)^{-M/2-1}z^{-(1/4+\varepsilon)M-1}dzdydx,$$

and I_1 is the integral over $(C_1 \times C_2 \times C_o) \setminus \mathcal{C}$. By changing variables we obtain

$$I_0 = \frac{1}{(2\pi)^3}\int\limits_{[-N^{-2/5},N^{-2/5}]^3} H(\theta,\varphi,\psi)d\psi d\varphi d\theta, \text{ where } H = \rho^{-\frac{(1-4\varepsilon)M}{2}} e^{-i\frac{(\theta+\varphi)M}{2} - i\psi(1/4+\varepsilon)M} \prod_{v\in V_{t_{1/2}}} h_v(\theta,\varphi,\psi), \tag{103}$$

and

$$h_v(\theta,\varphi,\psi) = e^{i(\theta+\varphi+\psi)d_v} + e^{i(\theta+\varphi+\psi)d_{\neg v}} + \rho^{d_v+d_{\neg v}} e^{i\theta d_v + i\varphi d_{\neg v}} + \rho^{d_v+d_{\neg v}} e^{i\theta d_{\neg v} + i\varphi d_v}.$$

Regarding I_1, we will use the trivial bound

$$I_1 \le (2\pi)^3 \sup_{(x,y,u)\in(C_1\times C_2\times C_o)\setminus\mathcal{C}} |H(x,y,u)| \tag{104}$$

to show that $I_1 = o(I_0)$.

We begin with estimating I_0 by providing an appropriate asymptotic expansion of it for points around the origin. First of all, note that for any $v \in V_{t_{1/2}}$ we have $h_v(0,0,0) = 2 + 2\rho^{d_v+d_{\neg v}}$ and thus

$$H(0,0,0) = \rho^{-(1-4\varepsilon)M/2} \prod_{v\in V_{t_{1/2}}} (2 + 2\rho^{d_v+d_{\neg v}}) = E.$$

Moreover,

$$\frac{\partial}{\partial\theta}h_v(0,0,0) = \frac{\partial}{\partial\varphi}h_v(0,0,0) = (2+2\rho^{d_v+d_{\neg v}})\,\frac{i}{2}(d_v + d_{\neg v}), \quad \text{and} \quad \frac{\partial}{\partial\psi}h_v(0,0,0) = i(d_v+d_{\neg v}).$$

The second derivatives at $(0,0,0)$ are given by

$$\frac{\partial^2}{\partial\theta^2}h_v = \frac{\partial^2}{\partial\varphi^2}h_v = -(d_v^2 + d_{\neg v}^2)(1+\rho^{d_v+d_{\neg v}}), \quad \text{and} \quad \frac{\partial^2}{\partial\psi^2}h_v = -(d_v^2+d_{\neg v}^2).$$

Furthermore, the mixed second derivatives are

$$\frac{\partial^2}{\partial\theta\partial\varphi}h_v = -(d_v^2 + d_{\neg v}^2 + 2d_v d_{\neg v}\rho^{d_v+d_{\neg v}}) \quad \text{and} \quad \frac{\partial^2}{\partial\theta\partial\psi}h_v = \frac{\partial^2}{\partial\phi\partial\psi}h_v = -(d_v^2 + d_{\neg v}^2).$$

We will also need crude bounds for the third-order derivatives in order to establish an accurate approximation for H around the origin. Note that h_v linearly exponential in θ, φ, ψ and $d_v, d_{\neg v}$. Thus, every time we take a derivative with respect to some variable, the norm of each single term in the expression of h_v can increase by at most $m_v = \max\{d_v, d_{\neg v}\}$. Thus, uniformly for $(\theta, \varphi, \psi) \in [-N^{2/5}, N^{2/5}]$ we have that

$$\left|\frac{\partial^3}{\partial\xi_1\partial\xi_2\partial\xi_3}h_v\right| \leq 2(1+\rho^{d_v+d_{\neg v}})(d_v+d_{\neg v})^3, \quad \text{where} \quad \xi_1,\xi_2,\xi_3 \in \{\theta,\varphi,\psi\}.$$

By using the uniform estimate $1 + x = e^{x-x^2/2+\Theta(x^3)}$, where we set $1 + x = h_v(\theta, \varphi, \psi)/h_v(0, 0, 0)$ we infer that

$$\ln\frac{h_v(\theta,\varphi,\psi)}{h_v(0,0,0)} = \frac{i}{2}(d_v+d_{\neg v})(\theta+\phi) + i\frac{d_v+d_{\neg v}}{2+2\rho^{d_v+d_{\neg v}}}\psi + \text{ 2nd order } + \text{error}, \tag{105}$$

where the 2nd order terms are

$$-\frac{(d_v-d_{\neg v})^2}{8}(\theta^2+\phi^2) - \frac{(d_v-d_{\neg v})^2+2\rho^{d_v+d_{\neg v}}(d_v^2+d_{\neg v}^2)}{2(2+2\rho^{d_v+d_{\neg v}})^2}\psi^2 + \frac{(d_v-d_{\neg v})^2(\rho^{d_v+d_{\neg v}}-1)}{2(2+2\rho^{d_v+d_{\neg v}})}\theta\varphi - \frac{(d_v-d_{\neg v})^2}{4+4\rho^{d_v+d_{\neg v}}}(\theta+\varphi)\psi.$$

Finally, since $(\theta, \varphi, \psi) \in [-N^{2/5}, N^{2/5}]$ the error term is of order at most $(d_v + d_{\neg v})^3 N^{-6/5}$. In order to obtain an approximation for H we form the product over all $v \in V_{t_{1/2}}$. Observe that the (linear in the variables) exponential factor $e^{-i(\theta+\varphi)M/2 - i\psi(1/4+\varepsilon)M}$ cancels exactly with the first order terms in (105). By abbreviating

$$S_{\theta,\theta} = \sum_{v\in V_{t_{1/2}}}\frac{(d_v-d_{\neg v})^2}{8}, \quad S_{\psi,\psi} = \sum_{v\in V_{t_{1/2}}}\frac{(d_v-d_{\neg v})^2+2\rho^{d_v+d_{\neg v}}(d_v^2+d_{\neg v}^2)}{2(2+2\rho^{d_v+d_{\neg v}})^2},$$

and

$$S_{\theta,\phi} = \sum_{v\in V_{t_{1/2}}}\frac{(d_v-d_{\neg v})^2(\rho^{d_v+d_{\neg v}}-1)}{4+4\rho^{d_v+d_{\neg v}}}, \quad S_{\theta,\psi} = \sum_{v\in V_{t_{1/2}}}\frac{(d_v-d_{\neg v})^2}{4+4\rho^{d_v+d_{\neg v}}}, \quad S_3 = \sum_{v\in V_{t_{1/2}}}(d_v+d_{\neg v})^3$$

we obtain uniformly for any $(\theta, \varphi, \psi) \in [-N^{-2/5}, N^{-2/5}]^3$

$$\ln\left(\frac{H}{E}\right) = -S_{\theta,\theta}(\theta^2+\varphi^2) - S_{\psi,\psi}\psi^2 + S_{\theta,\phi}\theta\phi - S_{\theta,\psi}(\theta+\varphi)\psi + O(S_3N^{-6/5}).$$

Observe that $\mathbf{d}$ is such that w.h.p. all quantities $S_{.,.}$ and S_3 are linear in N. Thus, we are left with computing

$$I_0 = (1+o(1))E\cdot\int_{[-N^{-2/5},N^{-2/5}]^3} e^{-S_{\theta,\theta}(\theta^2+\varphi^2) - S_{\psi,\psi}\psi^2 + S_{\theta,\phi}\theta\varphi - S_{\theta,\psi}(\theta+\varphi)\psi}d\psi d\varphi d\theta.$$

In order to compute this integral we rescale each variable with $N^{-1/2}$. By writing $s_{.,.}$ for $S_{.,.}/N$ we obtain

$$I_0 = (1+o(1))E\cdot N^{-3/2}\cdot\int_{[-N^{1/10},N^{1/10}]^3} e^{-s_{\theta,\theta}(\theta^2+\varphi^2) - s_{\psi,\psi}\psi^2 + s_{\theta,\varphi}\theta\varphi - s_{\theta,\psi}(\theta+\varphi)\psi}d\psi d\varphi d\theta.$$

A termwise comparison and elementary algebraic manipulations yield that

$$4S_{\theta,\theta}^2 - S_{\theta,\varphi}^2 \geq 0 \quad \text{and} \quad 2S_{\psi,\psi}S_{\theta,\theta} - S_{\theta,\psi}^2 - S_{\psi,\psi}S_{\theta,\varphi} \geq 0$$

Thus, the squares can be completed and the integral in the above expression equals a constant depending on the family $s_{\cdot,\cdot}$; this shows that asymptotically I_1 is proportional to $N^{-3/2} \cdot E$.

In order to complete the proof we will use (104) to show that I_1 is asymptotically negligible compared to I_0. Recall the definition of H from (103). It follows that the absolute value of H is given by

$$\rho^{-(1-4\varepsilon)M/2} \cdot \prod_{v \in V_{t_{1/2}}} f_v(\theta, \varphi, \psi), \quad \text{where} \quad f_v(\theta, \varphi, \psi) = |h_v(\theta, \varphi, \psi)|.$$

Let us abbreviate $D_v = d_v - d_{\neg v}$. A lengthy calculation, which can be performed easily with the help of MAPLE, yields that

$$\begin{aligned} f_v(\theta, \varphi, \psi)^2 = 2 + 2\rho^{2(d_v + d_{\neg v})} + 2\cos\big(D_v(\theta + \varphi + \psi)\big) + 2\rho^{2(d_v+d_{\neg v})}\cos\big(D_v(\theta - \varphi)\big) \\ + 2\rho^{d_v + d_{\neg v}}\left(\cos\big(D_v\varphi + d_v\psi\big) + \cos\big(D_v\theta + d_v\psi\big) + \cos\big(D_v\theta - d_{\neg v}\psi\big) + \cos\big(D_v\varphi - d_{\neg v}\psi\big)\right). \end{aligned}$$

Note that we can get an upper bound for f_v if we replace all terms involving a cosine by one; this implies that $|H| \leq \rho^{-(1-4\varepsilon)M/2} \prod_v (2 + 2\rho^{d_v + d_{\neg v}}) = E$. Moreover, the bound is achieved only if all factors are maximized simultaneously, and this happens for example when we choose $(\theta, \varphi, \psi) = (0, 0, 0)$. We will argue in the sequel that if $(\theta, \varphi, \psi) \in (C_1 \times C_2 \times C_o) \setminus \mathcal{C}$, i.e., at least one of the variables θ, φ, ψ is assigned a value not lying in $[-N^{-2/5}, N^{-2/5}]$, then there is a subset of variables $V' \subset V_{t_{1/2}}$ such that $|V'| \geq \alpha N$ for some $\alpha > 0$ and for all $v \in V'$ it holds $f_v(0,0,0) \leq f_v(0,0,0) - O(N^{-4/5})$. Indeed, if this is true, then

$$|H| \leq \rho^{-(1-4\varepsilon)M/2} \prod_{v \in V_{t_{1/2}} \setminus V'} (2 + 2\rho^{d_v + d_{\neg v}}) \prod_{v \in V'} (2 + 2\rho^{d_v + d_{\neg v}} - O(N^{-4/5})).$$

Since ρ is bounded and $\mathbf{d}$ is such that w.h.p. $d_v + d_{\neg v} = o(\log n)$, it follows that $|H|$ smaller that E by an exponential factor, which shows with (104) that $I_1 = o(I_0)$.

To see that a set V' with the desired properties exists, let us assume that at least one of θ, φ, ψ is in absolute value at least $N^{-2/5}$. For a pair $(d_+, d_-) \in \mathbf{N}^2$ let $V_{d_+,d_-} \subseteq V_{t_{1/2}}$ denote the set of variables v such that $d_v = d_+$ and $d_{\neg v} = d_-$, and write $N_{d_+,d_-} = |V_{d_+,d_-}|$. Consider the specific pair $(d_+, d_-) = (kr, kr - 1)$, and note that for all such variables we have $D_v = 1$. Furthermore, $\mathbf{d}$ is such that w.h.p. there is a constant $\beta = \beta(k) > 0$ such that $N_{d_+,d_-} \geq \beta N$. Then we may assume that

$$\text{for all } v \in N_{d_+,d_-} : \ f_v(\theta, \varphi, \psi) \geq (2 + 2\rho^{2kr-1} - O(N^{-4/5})),$$

as otherwise there is nothing to show. This impliesthat the arguments of all cosines appearing in the expression of f_v are close to multiples of 2π, and in particular,

$$|\theta + \phi + \psi|, \quad |\theta - \varphi|, \quad |\varphi + d_+\psi| = O(N^{-2/5}) \pmod{2\pi}; \tag{106}$$

this follows directly from the series expansion of the cosine around integer multiples of 2π, which lack a linear term. Next, consider the pair $(d'_+, d'_-) = (kr, kr - 2)$; again $\mathbf{d}$ is such that w.h.p. there is a constant $\beta' = \beta'(k) > 0$ such that $N_{d'_+,d'_-} \geq \beta' N$. Note that for these variables we have $D_v = 2$. Then, as previously, we may also assume that

$$\text{for all } v \in N_{d'_+,d'_-} : \ f_v(\theta, \varphi, \psi) \geq (2 + 2\rho^{2kr-2} - O(N^{-4/5})),$$

But then, by the same argument as above, $|2\varphi + d'_+\psi| = O(N^{-2/5}) \pmod{2\pi}$. Since $d_+ = d'_+$ and, by assumption, $|\varphi| < \pi$, by combining this with the third term in (106), we infer that $|\varphi| = O(N^{-2/5})$. In turn, together with the second term in (106), this implies that also $|\theta| = O(N^{-2/5})$. Finally, the fact $|\theta + \varphi + \psi| = O(N^{-2/5}) \pmod{2\pi}$ from (106) then also implies that $|\delta| = O(N^{-2/5})$. Everything together yields that $(\theta, \varphi, \psi) \in \mathcal{C}$, a contradiction.

11 Proof of Corollary 2.2

As a direct consequence of our second moment argument, the Paley-Zygmund inequality, and a concentration result on the number of satisfying assignments from [1] we obtain the following.

Proposition 11.1 *For r as in* (14) *we have $|\mathcal{S}(\mathbf{\Phi})| \geq \mathrm{E}\,|\mathcal{S}(\mathbf{\Phi})| \cdot \exp\left[-\frac{nr}{k^9 4^k}\right]$ w.h.p.*

We consider the following "planted model": let $\Lambda = \Lambda_k(n,m)$ be the the of all pairs (Φ,σ) of k-CNFs Φ over V with m clauses and satisfying assignments $\sigma \in \mathcal{S}(\Phi)$. Let P_Λ signify the uniform distribution over Λ; P_Λ is sometimes called the *planted model*. Moreover, let P_G be the distribution on Λ obtained by first choosing a random formula $\mathbf{\Phi}$ and then a uniformly random $\sigma \in \mathcal{S}(\mathbf{\Phi})$ (provided that $\mathbf{\Phi}$ is satisfiable); P_G is sometimes called the *Gibbs distribution*. Combining Proposition 11.1 with an argument from [], we obtain the following "transfer result".

Corollary 11.2 *For any $\mathcal{B} \subset \Lambda$ the following is true. If $\mathrm{P}_\Lambda[\mathcal{B}] \leq \exp\left[-\frac{2nr}{k^9 4^k}\right]$, then $\mathrm{P}_G[\mathcal{B}] = o(1)$.*

Thus, in order to show that some 'bad' event $\mathcal{B}$ is unlikely under P_G, we "just" need to show that $\mathrm{P}_\Lambda[\mathcal{B}] \leq \exp\left[-\frac{2nr}{k^9 4^k}\right]$ is exponentially small.

Lemma 11.3 *There is a number $\delta = \delta(k) > 0$ such that*

$$\mathrm{P}_\Lambda\left[\mathit{dist}(\sigma,\sigma_{maj}) > \frac{1}{2} - \delta\right] \quad \leq \quad \exp\left[-\frac{2nr}{k^9 4^k}\right].$$

Proof. We can generate a pair (Φ,σ) from the planted model as follows: first, choose $\sigma \in \{0,1\}^V$ uniformly; then, generate m clauses that are satisfied under σ uniformly and independently. Without loss of generality, we may assume that $\sigma = \mathbf{1}$ is the all-true assignment. We need to study the distribution $\boldsymbol{d} = (d_l)_{l\in L}$ of literal degrees. To this end, let $(e_l)_{l\in L}$ be a family of independent Poisson variables such that $\mathrm{E}[e_l] = \mathrm{E}[d_l]$ for all l. It is easily verified that there is $\zeta = \Theta(2^{-k})$ such that

$$\mathrm{E}[d_x] = \frac{kr}{2}(1+\zeta), \quad \mathrm{E}[d_{\neg x}] = \frac{kr}{2}(1-\zeta) \tag{107}$$

for all $x \in V$. Furthermore, if we let $\mathcal{E}$ be the event that $\sum_{l\in L} e_l = km$, then $\boldsymbol{e} = (e_l)_{l\in L}$ given $\mathcal{E}$ has the same distribution as $\boldsymbol{d}$. Moreover,

$$\mathrm{P}[\mathcal{E}] = \Omega(n^{-1/2}). \tag{108}$$

Let

$$Y = \frac{1}{n}\sum_{x\in V} \mathbf{1}_{e_x > e_{\neg x}} + \frac{1}{2}\mathbf{1}_{e_x = e_{\neg x}}.$$

Viewing the difference $e_x - e_{\neg x}$ as a random walk of length $\mathrm{Po}(kr)$ and using limit theorems for resulting distribution (the Skellam distribution), we obtain from (107) that $\mathrm{E}[Y] \geq \frac{1}{2} + \Omega(\sqrt{kr}/2^k)$. Further, applying Chernoff bounds to Y (which is a sum of independent contributions), we find that for a certain $\delta = \Omega(\sqrt{kr}/2^k)$

$$\mathrm{P}\left[Y < \frac{1}{2} + \delta\right] \leq \exp\left[-\Omega(\sqrt{kr}/2^k)^2 n\right] \leq \exp\left[-\frac{3nr}{k^9 4^k}\right]. \tag{109}$$

Finally, the assertion follows from (108) and (109). □

12 Proof of Lemma 2.3

The expected majority weight in $\mathbf{\Phi}$ is easily computed. In $\mathbf{\Phi}$, for each x the numbers d_x, $d_{\neg x}$ of positive/negative occurrences are asymptotically independently Poisson with mean $kr/2$. Therefore, for any $d = \Theta(kr)$ we obtain

$$\mathrm{E}[|d_x - d_{\neg x}| \mid d_x + d_{\neg x} = d] = \sqrt{2d/\pi} + O_k(1).$$

In effect,

$$\mathrm{E}\left[w_{maj}(\mathbf{\Phi})\right] \sim \frac{1}{2} + \sqrt{\frac{2}{\pi kr}} + O_k(1/kr). \tag{110}$$

By comparison, given that, say, the all-true assignment is satisfying, the number d_x of positive occurrences has distribution $\mathrm{Po}((1+1/(2^k-1))kr/2)$, while $d_{\neg x}$ has distribution $\mathrm{Po}((1-1/(2^k-1))kr/2)$. The normal approximation to the Poisson distribution yields for $d = \Theta(kr)$,

$$\mathrm{E}\left[|d_x - d_{\neg x}| \mid \mathbf{1} \in \mathcal{S}(\mathbf{\Phi}), d_x + d_{\neg x} = d\right] = \sqrt{2d/\pi} + \Theta(4^{-k}d^{3/2}) + O_k(1).$$

for a certain constant $c > 0$. Consequently,

$$\mathrm{E}\left[w_{maj}(\mathbf{\Phi}) \mid \mathbf{1} \in \mathcal{S}(\mathbf{\Phi})\right] \sim \frac{1}{2} + \sqrt{\frac{2}{\pi kr}} + \Theta(4^{-k}(kr)^{1/2}). \tag{111}$$

Both with and without conditioning on $\mathbf{1} \in \mathcal{S}(\mathbf{\Phi})$, w_{maj} enjoys the following Lipschitz property: changing one single clause can alter the value of w_{maj} by at most $k/(km) = 1/(rn)$. Therefore, Azuma's inequality yields

$$\begin{aligned}
\mathrm{P}\left[|w_{maj} - \mathrm{E}\left[w_{maj}\right]| > \lambda\right] &\leq 2\exp\left[-\frac{(r\lambda n)^2}{2m}\right] = 2\exp\left[-\frac{r\lambda^2 n}{2}\right], \\
\mathrm{P}\left[|w_{maj} - \mathrm{E}\left[w_{maj}\right]| > \lambda | \mathbf{1} \in \mathcal{S}(\mathbf{\Phi})\right] &\leq 2\exp\left[-\frac{r\lambda^2 n}{2}\right].
\end{aligned}$$

In effect, for a certain constant $\zeta > 0$ we have

$$\mathrm{P}\left[w_{maj} \geq \frac{1}{2} + \sqrt{\frac{2}{\pi kr}} + \zeta 4^{-k}(kr)^{1/2}\right] \leq \exp\left[-\Omega\left(k/4^k\right)n\right], \tag{112}$$

$$\mathrm{P}\left[w_{maj} \leq \frac{1}{2} + \sqrt{\frac{2}{\pi kr}} + \zeta 4^{-k}(kr)^{1/2} | \mathbf{1} \in \mathcal{S}(\mathbf{\Phi})\right] \leq \exp\left[-\Omega\left(k/4^k\right)n\right]. \tag{113}$$

Combining (112) and (113) with a simple counting argument yields Lemma 2 from the extended abstract.

Acknowledgment. The first author thanks Dimitris Achlioptas for helpful discussions on the second moment method. We also thank Charilaos Efthymiou for helpful comments that have led to an improved presentation.

References

[1] D. Achlioptas, A. Coja-Oghlan: Algorithmic barriers from phase transitions. Proc. 49th FOCS (2008) 793–802.

[2] D. Achlioptas, R. Menchaca-Mendez: Unsatisfiability Bounds for Random CSPs from an Energetic Interpolation Method. Proc. 39th ICALP (2012) 1–12.

[3] D. Achlioptas, R. Menchaca-Mendez: Exponential lower bounds for DPLL algorithms on satisfiable random 3-CNF formulas. Proc. 15th SAT (2012) 327–340.

[4] D. Achlioptas, C. Moore: Random k-SAT: two moments suffice to cross a sharp threshold. SIAM Journal on Computing **36** (2006) 740–762.

[5] D. Achlioptas, A. Naor: The two possible values of the chromatic number of a random graph. Annals of Mathematics **162** (2005) 1333–1349.

[6] D. Achlioptas, Y. Peres: The threshold for random k-SAT is $2^k \ln 2 - O(k)$. Journal of the AMS **17** (2004) 947–973.

[7] M. Bayati, D. Gamarnik, P. Tetali: Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. Proc. 42nd STOC (2010) 105–114.

[8] V. Chvátal, B. Reed: Mick gets some (the odds are on his side). Proc. 33th FOCS (1992) 620–627.

[9] A. Coja-Oghlan: A better algorithm for random k-SAT. SIAM J. Computing **39** (2010) 2823–2864.

[10] A. Coja-Oghlan: On belief propagation guided decmation for random k-SAT. Proc. 22nd SODA (2011) 957–966.

[11] A. Coja-Oghlan, K. Panagiotou: Catching the k-NAESAT threshold. Proc. 43rd STOC (2012) 899–908.

[12] A. Coja-Oghlan, L. Zdeborová: The condensation transition in random hypergraph 2-coloring. Proc. 23rd SODA (2012) 241–250.

[13] O. Dubois, Y. Boufkhad: A general upper bound for the satisfiability threshold of random r-SAT formulae. J. Algorithms **24** (1997) 395–420.

[14] O. Dubois, J. Mandler: The 3-XORSAT threshold. Proc. 43rd FOCS (2002) 769–778.

[15] M. Dyer, A. Frieze, C. Greenhill: On the chromatic number of a random hypergraph. Preprint (2012).

[16] F. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

[17] S. Franz, M. Leone: Replica bounds for optimization problems and diluted spin systems. J. Statist. Phys. **111** (2003) 535–564.

[18] E. Friedgut: Sharp Thresholds of Graph Proprties, and the k-SAT Problem. J. AMS **12** (1999) 1017–1054.

[19] A. Frieze, S. Suen: Analysis of two simple heuristics on a random instance of k-SAT. Journal of Algorithms **20** (1996) 312–355.

[20] A. Frieze, N. Wormald: Random k-Sat: a tight threshold for moderately growing k. Combinatorica **25** (2005) 297–305.

[21] A. Goerdt: A threshold for unsatisfiability. Proc. 17th MFCS (1992) 264–274.

[22] S. Kirkpatrick, B. Selman: Critical behavior in the satisfiability of random boolean expressions. Science **264** (1994) 1297–1301.

[23] L. Kirousis, E. Kranakis, D. Krizanc, Y. Stamatiou: Approximating the unsatisfiability threshold of random formulas. Random Structures Algorithms **12** (1998) 253–269.

[24] L. Kroc, A. Sabharwal, B. Selman: Message-passing and local heuristics as decimation strategies for satisfiability. Proc 24th SAC (2009) 1408–1414.

[25] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborová: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences **104** (2007) 10318–10323.

[26] C. McDiarmid, Concentration for independent permutations, *Combinatorics, Probability and Computing* (2002) **11**, 163–178.

[27] M. Mézard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfiability problems. Science **297** (2002) 812–815.

[28] M. Molloy: The freezing threshold for k-colourings of a random graph. Proc. 43rd STOC (2012) 921–930.

[29] A. Montanari, R. Restrepo, P. Tetali: Reconstruction and clustering in random constraint satisfaction problems. SIAM J. Discrete Math. **25** (2011) 771–808.

[30] R. Moser, G. Tardos: A constructive proof of the general Lovász local lemma. J. ACM **57** (2010).

[31] V. Rathi, E. Aurell, L. K. Rasmussen, M. Skoglund: Bounds on threshold of regular random k-SAT. Proc. 12th SAT (2010) 264–277.

[32] M. Spivak, *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*, W. A. Benjamin, Inc., New York-Amsterdam, 1997.