Typical random 3-SAT formulae and the satisfiability threshold

O. Dubois, Y. Boufkhad and J. Mandler

†LIP6, Box 169, CNRS-Université Paris 6, 4 place Jussieu, 75252 Paris cedex 05, France.
emails : Olivier.Dubois@lip6.fr, Jacques.Mandler@lip6.fr
‡LIAFA, Université Paris 7 Denis Diderot, 175 rue du Chevaleret, 75013 Paris, France
email : boufkhad@liafa.jussieu.fr

Abstract : We present a new structural (or syntactic) approach for estimating the satisfiability threshold of random 3-SAT formulae. We show its efficiency in obtaining a jump from the previous upper bounds, lowering them to 4.506. The method combines well with other techniques, and also applies to other problems, such as the 3-colourability of random graphs.

1 Introduction

The last decade has seen a growth of interest in phase transition phenomena in hard combinatorial decision problems, due to resulting insights into their computational complexity and that of the associated optimization problems. There is a fast growing body of theoretical investigations as well as ones exploring algorithmic solver implications. Latterly, moreover, statistical physics studies have also shed new light on these phenomena, whence a further surge in interest. Among the various and extensive contributions, let us single out a few: [8, 37, 8, 22, 36, 20, 10, 33, 4, 13, 8, 39]. Several surveys can be found in [16]. One of the most challenging phase transitions, with a long history of results, concerns the problem of 3-Satisfiability (to satisfy sets of clauses of length 3, i.e. disjunctions of 3 literals). [16] contains a survey which we briefly summarize and update here. Experiments strongly suggest that satisfiability of random 3-SAT formulae (the 3-SAT problem) exhibits a sharp threshold or a phase transition as a function of a control parameter, the ratio $c$ of the number of clauses to the number of variables. More precisely, this would mean the existence of a critical value $c_0$ such that for any $c < c_0$ the probability of satisfiability of a random 3-SAT formula tends to 1 as $n \to \infty$, and for $c > c_0$ it tends to 0. Over the years, two series of bounds for $c_0$ have been established, the lower bounds being : 2.9 (positive probability only), 2/3, 1.63, 3.003, 3.145, 3.26, 3.42 (see [4, 8, 4, 21, 1, 4, 25]), and the upper bounds: 5.191, 5.081, 4.762, 4.643, 4.602, 4.596, 4.571, 4.506 (see [13, 8, 24, 11, 33, 4, 24, 28, 33]). The last bound, 4.506, was briefly presented in [12]. The present paper gives a detailed proof, emphasizing the potential of the main innovation, which we called the structural or syntactic approach, in contrast to the semantic approach hitherto used to establish upper bounds. A few general comments are in order. Thanks to this structural approach, a jump from 4.643 to 4.506 was obtained. Developments since then have confirmed the interest and versatility of this technique. Further refinements of the semantic approach, together with subtle and sophisticated probabilistic and analytical results, have failed to match the 4.506 bound,

*A preliminary short version of this paper apperared in the Proceedings of the Eleventh ACM-SIAM Symposium on Discrete Algorithms, pages 124-126, San Francisco, California, January 2000.
giving 4.571 as announced in [28]. And we recently applied our structural approach to the equally challenging 3-colouring problem. It turned out to combine well with the decimation technique we had used for the 3-XORSAT problem [14], lowering the best upper bound from 2.4945 ([18] and references therein) to 2.427 [15].

In the remainder of this section, we present the probabilistic model for 3-SAT we work with, then give an overview of our approach leading to the bound of 4.506. The subsequent sections contain the detailed calculations.

1.1 Probabilistic model.

Let \( V_n = \{ x_1, ..., x_n \} \) be a set of \( n \) boolean variables, \( L_n = \{ x_1, \bar{x}_1, ..., x_n, \bar{x}_n \} \) the corresponding set of positively and negatively signed literals. In this paper we use the ordered-clauses model. Here an \( n \)-formula \( F \) is simply a map to \( L_n \) from the formula template \( \Lambda_c,n \), an array of \( cn \) clause templates consisting each of 3 ordered places or cells. If the literal \( l \) is the image under \( F \) of cell \( \xi \), we also say that it fills \( \xi \). The set \( \Omega(n,c) \) of \( n \)-formulae is made into a probability space by assigning each formula the probability \( 1/|\Omega(n,c)| = (2n)^{3cn} \).

Each truth assignment \( A : V_n \to \{0,1\} \) is conventionally extended to \( L_n \) so that \( A(\bar{x}_i) = 1 - A(x_i) \), and is said to satisfy the clause \( C_k \) if \( A(l) = 1 \) for some \( l \in C_k \), and the formula \( F \) if it satisfies all its clauses; in which case \( A \) is a solution of \( F \), and \( F \) is satisfiable. The probability of satisfiability of a random formula \( F \) of \( \Omega(n,c) \) is denoted by \( \Pr_{n,c}(SAT) \).

A few words in comparison with the non-ordered-clauses model, also very usual. Here a clause is a set of 3 literals with distinct underlying variables, and a random formula is a sequence of \( m = cn \) clauses drawn independently and uniformly among the \( 2^3(n) \) possible clauses. Convergence to 0 (resp. 1), as \( n \to \infty \), of \( \Pr_{n,c}(SAT) \) is readily seen to imply the same for the probability in the non-ordered-clauses model. Thus our upper bound of 4.506, once proven in the ordered-clauses model, will hold in both.

1.2 Outline.

We give first a general idea of our approach stemming from concrete experiments. A computer-based generator of random formulae churns out \textit{mechanically}, as the case may be, \textit{only} satisfiable, or \textit{only} contradictory formulae. To say that certain formulae are \textit{never} produced (within a realistic timeframe) simply means that they form a set of vanishingly small probability; and, due to the very dumbness of the generator, the distinction between ‘likely’ and ‘unlikely’ formulae must be possible on a very basic level, considering only their \textit{form} or \textit{structure}. Ideally, we would like an exact criterion for ‘likely’ or ‘typical’ formulae; possibly, then, the first moment method, restricted not to particular types of solutions, but to formulae with this particular property, might give us the exact value of \( c_0 \). Such an exact characterization is elusive, though, and unlikely to emerge in a simple, usable form. Rather, in this paper we show the usefulness of an uncomplicated partial characterization in terms of the numbers of occurrences and signs of the variables. The pure effect on the expectation of restricting the formulae becomes only part of the story. Equally important is
the fact that, far from interfering with other approaches, the added structure actually helps in otherwise difficult or hopeless enumerations. Thus we do not need, e.g., sophisticated probability results. And, particularly, we are able to introduce at virtually no cost some structural manipulations on the balancing of the signs of occurrences per variable which would be impractical in the purely semantic approach. On the other hand, to attain full rigour the method does require fairly lengthy calculations, notably to bound errors arising from the finite size of formulae, and thoroughly to justify the optimization procedures. These remain relatively elementary, though, and, in the case of the error estimates, fairly routine.

Practically we first characterize the asymptotic distribution of the signed occurrences per variable, namely:

**Lemma 1.1** For any integers $0 \leq p \leq x$, define $\kappa_{x,p} = 2^{-x}p(x,\lambda)$, where $\lambda = 3c$ and $p(x,\lambda)$ is the Poisson probability mass function of mean $\lambda$, i.e. $p(x,\lambda) = e^{-\lambda}\lambda^x/x!$. Let the random variable $\omega_{x,p}$ be the proportion of variables of a random formula having $x$ occurrences, among which exactly $p$ have a positive signature. Then for any $\varepsilon > 0$, $\lim_{n \to \infty} \Pr\left(|\omega_{x,p} - \kappa_{x,p}| > \varepsilon\right) = 0$.

It can be seen easily that Lemma 1.1 implies that an upper bound on the satisfiability threshold is obtained by calculating the expected number of solutions of a typical formula (in a sense to be specified shortly, but roughly meaning that for most $(x,p)$, there are nearly $\kappa_{x,p}n$ variables having $x$ total, $p$ positive, occurrences). Typical formulae, however, also provide us a strong means to go further in the structural manipulation of formulae. But we need first to recall the definition of particular solutions which in [11] we called PPSs (for Positively Prime Solution, symmetrically there are NPSs). Note that these restrictive solutions have been introduced independently by Kirousis et al. in [30] under the terminology of locally maximal solutions and single-flip technique.

**Definition 1.1** A Positively Prime Solution (PPS) $A$ of a SAT formula $F$ is a solution of $F$ such that no variable of $F$ with the value 1 under $A$ can be singly inverted (or switched) to 0 unless at least one of the clauses of $F$ becomes unsatisfied, that is, the new assignment is no longer a solution of $F$.

Any satisfiable formula has a PPS, but some have very many: they provide extremely useful, yet somewhat limited restriction. A means to enhance this is unbalancing, which we now introduce on an intuitive level.

When enumerating formulae with a view to computing an expectation, we usually count as different some formulae (very many, in fact) which really are the same from the point of view of satisfiability. This happens in more than one way. Some formulae differ from each other by a permutation on the set of clauses, or on the set of variables; these, however, are fairly transparent. What concerns us here are formulae deduced from one another by renaming certain variables, in the restricted (and usual) sense of inverting the signs of all their occurrences. Their significance to us stems from the fact that unlike those just
mentioned, they are not neutral with respect to PPSs. Consider, e.g., a pure variable (one which has all of its occurrences of the same sign). This sign is indifferent as far as ordinary solutions are concerned, but a solution in which a negated pure variable takes the value 1 cannot be a PPS, while an unnegated pure variable has the best chances that many of the solutions giving it the value 1 will be PPSs. Similarly, a variable with more positive than negative instances is likely to kill fewer PPSs than the reverse. Therefore, of two formulae which differ only by the systematic inversion of some variables, the one with more negatively unbalanced variables may be assumed to have fewer PPSs. To be precise, call two formulae equivalent if one can be obtained from the other by renaming certain variables. Clearly, this is indeed an equivalence relation $\mathcal{R}$ on the set $\Omega(n, c)$ of 3-SAT formulae on $n$ variables with $cn$ clauses; $\mathcal{R}$ results in the partitioning of formulae with respect to equivalence modulo variable renaming, and the cardinality of the equivalence class of a formula $F$ is $2^{v_n(F)}$, where $v_n(F)$ is the number of unbalanced variables in $F$ (variables having unequal numbers of positive and negative occurrences; note that an absent variable is, by definition, balanced).

Since negatively unbalanced variables tend to inhibit PPSs, we have a good candidate for the formula with the fewest PPSs within each equivalence class $C$: namely, the totally unbalanced representative $F^-$ obtained from any $F \in C$ by renaming exactly those variables which have more positive than negative occurrences. Moreover we have an easy criterion for a formula to be the totally unbalanced representative of a typical formula, namely that the proportion of variables having $x$ total, $p$ positive occurrences be $2\kappa_{x,p}$ if $x > p$; $\kappa_{x,p}$ if $x = 2p$; and 0 if $x < 2p$. So these representatives, or, as we shall say, the typical totally unbalanced formulae (by abuse of language, since they are actually not typical at all) can be defined just like the typical formulae, only using instead of the $\kappa_{x,p}$’s their totally unbalanced counterparts, the $\tilde{\kappa}_{x,p}$’s defined by:

$$
\tilde{\kappa}_{x,p} = \begin{cases} 
2\kappa_{x,p} & \text{if } x > 2p \\
\kappa_{x,p} & \text{if } x = 2p \\
0 & \text{if } x < 2p
\end{cases}
$$

All equivalence classes of such representatives have the same number of elements, namely: $2^{2n\sum_{x \geq 2p}\kappa_{x,p}} = 2^n\sum_{x > 2p}\tilde{\kappa}_{x,p}$.

Calculations with typical totally unbalanced formulae are no harder than with plain and ordinary typical formulae, in fact they are much the same with $\tilde{\kappa}$ replacing $\kappa$, and the specifics of the distribution tend to intervene only in the very last stages. Computing (via a simple technical device) the expected number of PPSs of the former rather than the latter, then multiplying by the above size of equivalence classes, we get what amounts to a ‘skewed’ expectation where each formula is counted, not according to its own number of PPSs, but to that of its representative with fewest PPSs. It is this, combined with the gain already inherent in the restriction to structured formulae per se, that affords us a very significant improvement on the upper bound of 4.643 resulting from the expectation of PPSs alone [11].

Before proceeding, we have to take account of some practical remarks raised by the foregoing considerations. (i) The $\kappa_{x,p}$’s or the $\tilde{\kappa}_{x,p}$’s constitute an infinite family and are all $\neq 0$, while a formula has finite length; (ii) The proportions $\nu_{x,p}$ of variables having $x$ total, $p$ positive occurrences in a formula $F \in \Omega(c, n)$ must verify $\sum \nu_{x,p} = 1$ and $\sum x\nu_{x,p} = 3c = \lambda$, where the sums are in effect finite, while the equalities $\sum \kappa_{x,p} = 1$ and $\sum x\kappa_{x,p} = \lambda$ only apply with
infinite sums (series); (iii) The $\kappa_{x,p}$'s are irrational, so they cannot be exact proportions even for special values of $n$. Thus, in order to derive a rigorous argument, we define what we call formulae obeying a given distribution of signed occurrences to a specified approximation:

**Definition 1.2** Let $\Xi = (\xi_{x,p})_{0 \leq p \leq x}$ be a family of nonnegative real numbers satisfying the relations $\sum_{x=0}^{\infty} \sum_{p=0}^{x} \xi_{x,p} = 1$ and $\sum_{x=0}^{\infty} \sum_{p=0}^{x} x \xi_{x,p} = \lambda$. Given a real $\varepsilon > 0$ and an integer $x_{\text{max}}$, a formula $F \in \Omega(n,c)$ is said to obey the distribution $\Xi$ to the accuracy $(\varepsilon, x_{\text{max}})$ iff for $0 \leq p \leq x \leq x_{\text{max}}$, the number of variables having $x$ occurrences in $F$, $p$ of which are positive, lies between $(\xi_{x,p} - \varepsilon)n$ and $(\xi_{x,p} + \varepsilon)n$. The set of formulae in $\Omega(n,c)$ obeying $\Xi$ to the accuracy $(\varepsilon, x_{\text{max}})$ will be denoted by $F(\Xi, \varepsilon, x_{\text{max}}, n, c)$.

The term ‘typical formula’ will sometimes be used loosely to indicate a formula which obeys the distribution $(\kappa_{x,p})$ to the accuracy $(\varepsilon, x_{\text{max}})$ for some (large) $x_{\text{max}}$ and some (small) $\varepsilon$.

Henceforth the distributions of the $\kappa_{x,p}$'s and of the $\tilde{\kappa}_{x,p}$'s (corresponding, of course, to some value of $\lambda = 3c$) will be denoted by $\Xi_0$ and $\tilde{\Xi}_0$, respectively. Also, when the context makes the various parameters clear, we will often use the abbreviated notation $E[PPS]$ for the expected number of PPSs of formulae drawn uniformly from $F(\Xi, \varepsilon, x_{\text{max}}, n, c)$. Strictly speaking, a direct calculation of the expectation of PPSs of typical totally unbalanced formulae would involve an awkward change of probability space. The same end result can be achieved much more conveniently by introducing an *ad hoc* r.v. on the original probability space $\Omega(n,c)$, then linking its expectation to the probability of satisfiability.

**Proposition 1.2** Define the r.v. $X_{n,\varepsilon,x_{\text{max}},c}$ on $\Omega(n,c)$ by:

$$X_{n,\varepsilon,x_{\text{max}},c}(F) = \begin{cases} 2^n \sum_{x > 2p} \tilde{\kappa}_{x,p} \times PPS(F) & \text{if } F \in F(\tilde{\Xi}_0, \varepsilon, x_{\text{max}}, n, c) \\ 0 & \text{otherwise} \end{cases}$$

and set $\rho = \rho_{x_{\text{max}}} = \sum_{2p > x_{\text{max}}} \kappa_{2p,p}$, $\Delta = \Delta_{x_{\text{max}}} = 1/2 \ (x_{\text{max}}/2 + 1)$. If, for some integer $x_{\text{max}}$ and some $\varepsilon > 0$, $2^{(\rho + \varepsilon \Delta)n}E[X_{n,\varepsilon,x_{\text{max}},c}]$ tends to 0 as $n \to \infty$, then so does $\Pr_{n,c}(\text{SAT})$.

*(Remark: It will be clear from the proof that this remains true if instead of PPSs we use any class of solutions such that any satisfiable formula possesses at least one solution in this class, e.g. prime implicants [3], ‘double flips’ [4].)*

The rest of our plan will be to compute an explicit expression of $E[X_{n,\varepsilon,x_{\text{max}},c}]$ as sums of combinatorial terms, then an asymptotic exponential upper bound of this expectation. This will be obtained as a function of values of parameters satisfying a system of equations, which will be reduced to two equations in two unknowns. Careful study of these equations, coupled with numerical calculations, will show that for $\tilde{\Xi}_0 = (\tilde{\kappa}_{x,p})$, $c = 4.506$, and appropriate values of $x_{\text{max}}$ and $\varepsilon$, $2^{(\rho_{x_{\text{max}}} + \varepsilon \Delta_{x_{\text{max}}})n}E[X_{n,\varepsilon,x_{\text{max}},c}]$ tends to 0 as $n \to \infty$. 

5
2 Basic structural results on random 3-SAT formulae

We have first to prove Lemma 1.1. Here the classical limit theorems of probability do not apply, and some form of large-deviation inequality has to be used. One method is to first obtain the expectation of $\omega_{x,p}$ as $\kappa_{x,p}$, then apply the method of bounded differences (see, e.g., [23], pp. 16, 221). Or, a proof using Poissonization may be of independent interest, giving stronger bounds, so we include a detailed one in Appendix A. Interestingly, Lemma 1.1, which is all we need, uses the full power of neither approach.

The quantity $|\{(x,p) : 0 \leq p \leq x \leq x_{\text{max}}\}| = (x_{\text{max}} + 1)(x_{\text{max}} + 2)/2$ is encountered repeatedly in the sequel, we denote it $D(x_{\text{max}})$ or simply $D$.

**Proof of Proposition 1.2.** With the equivalence relation $\mathcal{R}$ as in Section 1.2 and $\hat{\mathcal{R}}$ induced by $\mathcal{R}$ on $\mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c)$, the quotient (canonical) map $\mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c) \rightarrow \mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c)/\hat{\mathcal{R}}$ maps $F$ to the (class of the) formula $F^{-}$ obtained by renaming all variables of $F$ having more positive than negative occurrences.

Recall that $\omega_{x,p}(F)$ denotes the proportion of variables in a formula $F \in \Omega(n, c)$ having $x$ total, $p$ positive occurrences. Then:

$$\omega_{x,p}(F^{-}) = \begin{cases} \omega_{x,p}(F) + \omega_{x,x-p}(F) & \text{if } x > 2p \\ \omega_{x,p}(F) & \text{if } x = 2p \\ 0 & \text{if } x < 2p \end{cases}$$

A single $F^{-} \in \mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c)/\hat{\mathcal{R}}$ may come from at most $2^{v_4(F^{-})}$ formulae (not all necessarily in $\mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c)$). Taking into account that if $x > 2p$, we have $\tilde{\kappa}_{x,p} = \kappa_{x,p} + \kappa_{x,x-p}$ (because $\kappa_{x,p} = \kappa_{x,x-p}$), we have:

$$|\omega_{x,p}(F^{-}) - \tilde{\kappa}_{x,p}| = \begin{cases} |\omega_{x,p}(F) - \kappa_{x,p}| + |\omega_{x,x-p}(F) - \kappa_{x,x-p}| & \text{if } x > 2p \\ |\omega_{x,p}(F) - \kappa_{x,p}| & \text{if } x = 2p \\ 0 & \text{if } x < 2p \end{cases}$$

so that $|\mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c)/\hat{\mathcal{R}}| \leq |\mathcal{F}(\Xi_0, \varepsilon, x_{\text{max}}, n, c)|$. Further,

$$\frac{v_4(F^{-})}{n} = \sum_{0 \leq 2p < x} \omega_{x,p}(F^{-}) \leq 1 - \sum_{0 \leq 2p \leq x_{\text{max}}} \omega_{2p,p}(F^{-}) \leq 1 - \sum_{0 \leq 2p \leq x_{\text{max}}} \tilde{\kappa}_{2p,p} + \left(\frac{x_{\text{max}}}{2} + 1\right) \frac{\varepsilon}{2}$$

Therefore, since $\tilde{\kappa}_{2p,p} = \kappa_{2p,p}$,

$$|\mathcal{F}(\Xi_0, \varepsilon, x_{\text{max}}, n, c)| \leq 2^{2n} \sum_{0 \leq 2p < x} \kappa_{x,p} \times 2^{(\varepsilon x_{\text{max}} + p_{x_{\text{max}}})n} \times |\mathcal{F}(\Xi_0, \frac{\varepsilon}{2}, x_{\text{max}}, n, c)/\hat{\mathcal{R}}|$$

$$\leq 2^{2n} \sum_{0 \leq 2p < x} \kappa_{x,p} \times 2^{(\varepsilon x_{\text{max}} + p_{x_{\text{max}}})n} \times |\mathcal{F}(\Xi_0, \varepsilon, x_{\text{max}}, n, c)|$$
Remark. Our bound on $v_u(F^-)$ might at first sight seem too loose, since, instead of allowing all unbalanced variables to be renamed in any combination, we should really pick half of each group of $\kappa_{x,p,n}$ and rename only these. Actually, the two bounds do not differ in their exponential orders of growth as $n \to \infty$.

Note that $F$ is satisfiable iff $F^-$ is. So,

$$|F(\Xi_0, \varepsilon, x_{\max}, n, c) \cap SAT(n, c)| \leq 2^{2n} \sum_{0 \leq p < x} \kappa_{x,p} \times 2^\varepsilon 2^\Delta n \times |F(\tilde{\Xi}_0, \varepsilon, x_{\max}, n, c) \cap SAT(n, c)|$$

We are now able to show that if $2^{(\rho + \varepsilon) n} \times E[X_{n,\varepsilon,x_{\max},c}]$ tends to 0, then so does the probability of satisfiability. Indeed:

$$\Pr_{n,c}(SAT) = \frac{|SAT(n, c)|}{|\Omega(n, c)|} \leq \frac{|F(\Xi_0, \varepsilon, x_{\max}, n, c) \cap SAT(n, c)|}{|\Omega(n, c)|} + \sum_{0 \leq p \leq x \leq x_{\max}} \frac{\{|F \in SAT(n, c) : |\omega_{x,p}(F) - \kappa_{x,p}| > \varepsilon\}}{|\Omega(n, c)|}$$

By Lemma [1], each of the $D(x_{\max})$ terms of the last sum tends to 0 as $n \to \infty$, hence:

$$\Pr_{n,c}(SAT) \leq \frac{2^{2n} \sum_{0 \leq p < x} \kappa_{x,p} \times 2^{(\varepsilon + \Delta)n} \times |F(\tilde{\Xi}_0, \varepsilon, x_{\max}, n, c) \cap SAT(n, c)|}{|\Omega(n, c)|} + o(1)$$

So,

$$\Pr(SAT) \leq \frac{2^{2n} \sum_{0 \leq p < x} \kappa_{x,p} \times 2^{(\varepsilon + \Delta)n} \times \sum_{F \in \mathcal{F}(\tilde{\Xi}_0, \varepsilon, x_{\max}, n, c) \cap SAT(n, c)} 1 + o(1)}{|\Omega(n, c)|}$$

Now, since any satisfiable formula has at least one PPS, we can write:

$$\Pr(SAT) \leq \frac{2^{(\varepsilon + \Delta)n}}{|\Omega(n, c)|} \times \sum_{F \in \mathcal{F}(\tilde{\Xi}_0, \varepsilon, x_{\max}, n, c)} 2^{2n} \sum_{0 \leq p < x} \kappa_{x,p} \times \text{PPS}(F) + o(1)$$

$$= \frac{2^{(\varepsilon + \Delta)n}}{|\Omega(n, c)|} \times \sum_{F \in \mathcal{F}(\tilde{\Xi}_0, \varepsilon, x_{\max}, n, c)} X_{n,\varepsilon,x_{\max},c}(F) + o(1) = 2^{(\rho + \varepsilon) n} \times E[X_{n,\varepsilon,x_{\max},c}] + o(1)$$

### 3 Combinatorial analysis of the expectation.

#### 3.1 The set $\Theta_{\varepsilon,x_{\max},n,c}$

In order to estimate the expected number of PPSs of formulae in $\mathcal{F}(\tilde{\Xi}_0, \varepsilon, x_{\max}, n, c)$, we shall first compute the number of such formulae having fixed values of the proportions $\omega_{x,p}(F)$ for
0 ≤ p ≤ x ≤ x_{\text{max}}. It will be convenient to characterize these formulae as associated with an element of the set Θ_{\varepsilon,x_{\text{max}},n,c} \subseteq \mathbb{Q}^D of vectors θ = (θ_{x,p})_{0 \leq p \leq x \leq x_{\text{max}}} such that (with the notation I_n = \{0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}, 1\}, which applies throughout the sequel):

(i) \quad \theta_{x,p} \in I_n, \quad 0 \leq p \leq x \leq x_{\text{max}};
(ii) \quad \sum_{x=0}^{x_{\text{max}}} \sum_{p=0}^{x} \theta_{x,p} \leq 1;
(iii) \quad |\theta_{x,p} - \bar{\kappa}_{x,p}| \leq \varepsilon, \quad 0 \leq p \leq x \leq x_{\text{max}};

It is clear that a formula F is in \mathcal{F}(\Xi_0, \varepsilon, x_{\text{max}}, n, c) iff the vector (ω_{x,p}(F))_{0 \leq p \leq x \leq x_{\text{max}}} is in Θ_{\varepsilon,x_{\text{max}},n,c}. For θ ∈ Θ_{\varepsilon,x_{\text{max}},n,c}, we denote by \mathcal{F}(θ) the subset of \mathcal{F}(\Xi_0, \varepsilon, x_{\text{max}}, n, c) consisting of those formulae F such that for 0 ≤ p ≤ x ≤ x_{\text{max}}, ω_{x,p}(F) = θ_{x,p}. We are able to focus on the number of elements of \mathcal{F}(θ) mainly because, as the following lemma shows, the relatively small (i.e. polynomial) size of Θ_{\varepsilon,x_{\text{max}},n,c} means that, as far as exponential orders are concerned, it makes no real difference whether θ is kept fixed or allowed to vary within Θ_{\varepsilon,x_{\text{max}},n,c}:

Lemma 3.1 \quad |Θ_{\varepsilon,x_{\text{max}},n,c}| \leq (2\varepsilon n)^D.

Proof. If the vector θ is in Θ_{\varepsilon,x_{\text{max}},n,c}, then for 0 ≤ p ≤ x ≤ x_{\text{max}}, θ_{x,p}n is an integer comprised between (\bar{\kappa}_{x,p} - \varepsilon)n and (\bar{\kappa}_{x,p} + \varepsilon)n, so there are at most 2\varepsilon n possible values for θ_{x,p}. ■

3.2 Counting formulae with a given PPS and fixed proportions of variables having given numbers of occurrences

For some given ε and x_{\text{max}}, we now consider a fixed vector θ ∈ Θ_{\varepsilon,x_{\text{max}},n,c} and a truth value assignment \mathcal{A} ∈ \{0, 1\}^n, identified with the subset \mathcal{A}^{-1}(1) of the set of variables V_n. Let \mathcal{F}(θ, \mathcal{A}) be the set of formulae F ∈ Ω(n, c) such that \mathcal{A} is a PPS of F and that for 0 ≤ p ≤ x ≤ x_{\text{max}}, ω_{x,p}(F) = θ_{x,p}. Thus:

\text{Proposition 3.2} \quad \mathbb{E}(X_{n,\varepsilon,x_{\text{max}},c}) = \frac{2^n \sum_{x \geq 2p} \hat{\kappa}_{x,p}}{\mathbb{E}(n, c)} \sum_{\theta \in \Theta_{\varepsilon,x_{\text{max}},n,c}} \sum_{\mathcal{A} \in \{0, 1\}^n} |\mathcal{F}(\theta, \mathcal{A})|.

Our next goal is to estimate the size of \mathcal{F}(θ, \mathcal{A}) for a fixed \theta ∈ Θ_{\varepsilon,x_{\text{max}},n,c} and \mathcal{A} ∈ \{0, 1\}^n. Abundant use will be made of the quantities τ = τ(θ, x_{\text{max}}) = 1 - \sum_{0 \leq p \leq x \leq x_{\text{max}}} θ_{x,p} and σ = σ(θ, x_{\text{max}}) = λ - \sum_{0 \leq p \leq x \leq x_{\text{max}}} xθ_{x,p}. τ is, of course, nonnegative by definition; for any F ∈ \mathcal{F}(θ, \mathcal{A}), τ represents the proportion of variables having more than x_{\text{max}} occurrences in F. Also, σ is nonnegative, since for F ∈ \mathcal{F}(θ, \mathcal{A}), it represents the proportion of literals in F (among the total λn) whose underlying variables have more than x_{\text{max}} occurrences.
Given a formula \( F \) and a truth assignment \( \mathcal{A} \), we say that a clause of \( F \) is of type \((\mathcal{A}, j)\) \((0 \leq j \leq 3)\) iff it has \( j \) nonzero literals under \( \mathcal{A} \). To say that \( \mathcal{A} \) satisfies \( F \) means that \( F \) has no clauses of type \((\mathcal{A}, 0)\).

Now suppose \( \mathcal{A} \) is a PPS of \( F \in \mathcal{F}(\theta, \mathcal{A}) \), let \( v \) be one of the variables having \( x \) total, \( p \) positive occurrences in \( F \), and \( q \) its number of occurrences in type-\((\mathcal{A}, 1)\) clauses as the unique satisfying literal. If \( v \) has value 1 under \( \mathcal{A} \), then \( q = p - j \) for some \( j \) with \( 0 \leq j \leq p - 1 \); excluding \( j = p \) expresses exactly that \( \mathcal{A} \) is a PPS. If \( v \) has value 0 under \( \mathcal{A} \), then \( q = j - p \) for some \( j \) with \( p \leq j \leq x \). Since the two cases cover exactly once each possible \( j \) between 0 and \( x \), they can be conveniently coalesced by saying that for any variable there is a unique \( j \) with \( 0 \leq j \leq x \), such that \(|p - j|\) of its occurrences are in type-\((\mathcal{A}, 1)\) clauses, the value of the variable under \( \mathcal{A} \) being automatically determined by the sign of \( p - j \). It is

\[
\frac{1}{2} \left(1 + \frac{(p - j)}{|p - j|}\right) \text{ if } j \neq p \text{ and by convention } 0 \text{ if } j = p.
\]

We call such a variable a variable of type \((\mathcal{A}, x, p, j)\), and thus to say that \( \mathcal{A} \) is a PPS of \( F \in \mathcal{F}(\theta, \mathcal{A}) \) means exactly that every variable is of type \((\mathcal{A}, x, p, j)\) for some \( x, p \) and \( j \) with \( 0 \leq p \leq x \) and \( 0 \leq j \leq x \). In our enumerations, however, we will only impose this condition for \( x \leq x_{\max} \). The variables with more than \( x_{\max} \) occurrences, or heavy variables, will be considered unconstrained, and we will broadly overestimate the number of corresponding choices. If our expectation calculated by excess tends to 0, so does the true expectation.

Recall that we use the notation \( I_n = \{0, 1/n, 2/n, ..., 1 - 1/n, 1\} \). Given the vector \( \theta \in \Theta_{\epsilon, x_{\max}, n, c} \), the assignment \( \mathcal{A} \), and rationals \( \gamma_1, \gamma_2, \gamma_3 \in I_{cn} \) and \( \mu_{x, p, j} \in I_{\theta_{x, p} n} \) \((0 \leq p, j \leq x \leq x_{\max})\), we proceed to count the formulae in \( \mathcal{F}(\theta, \mathcal{A}) \)

- consisting of \( \gamma_i cn \) clauses of type \( i, i = 1, 2, 3 \), and
- such that the number of variables of type \((\mathcal{A}, x, p, j)\) is \( \mu_{x, p, j} \theta_{x, p} n \) for \( 0 \leq p, j \leq x \leq x_{\max} \).

We assume, of course, \( \gamma_1 + \gamma_2 + \gamma_3 = 1 \) and \( \sum_{j=0}^{x} \mu_{x, p, j} = 1 \) for \( 0 \leq p, j \leq x \leq x_{\max} \). Let \( Z(\theta, \gamma, \mu, n, c) \) be the number of such formulae.

The empty formula template \( \Lambda_{c, n} \) contains \( \lambda n \) cells, with \( \lambda = 3c \). We first choose those which will correspond to each type of clause, and within each group, those to be filled with literals of value 1. This can be done in \( A_n(\gamma, c) \) ways, where

\[
A_n(\gamma, c) = \frac{(cn)!}{(\gamma_1 cn)! (\gamma_2 cn)! (\gamma_3 cn)!} 3^{(\gamma_1 + \gamma_2) cn}.
\]

Second, among the \( n \) variables we choose, for \( 0 \leq p \leq x \leq x_{\max} \), the \( \theta_{x, p} n \) which will have \( x \) total, \( p \) positive occurrences, and among these the \( \mu_{x, p, j} \theta_{x, p} n \) which will be of type \((\mathcal{A}, x, p, j)\). Recall that given \( \mu_{x, p, j} \) the values under \( \mathcal{A} \) of the \( \mu_{x, p, j} \theta_{x, p} n \) corresponding variables are automatically determined. We complete the specification of \( \mathcal{A} \) by choosing the values of the remaining \( \tau n \) heavy variables (recall \( \tau = 1 - \sum_{0 \leq p, x_{\max}} \theta_{x, p} \)). The number of possibilities is:

\[
B_n(\theta, \mu) = 2^\tau \frac{n!}{(\tau n)!} \prod_{0 \leq p, x \leq x_{\max}} (\theta_{x, p} n)! \prod_{0 \leq p, x \leq x_{\max}} \frac{(\theta_{x, p} n)!}{\prod_{j=0}^{x} (\mu_{x, p, j} \theta_{x, p} n)!}.
\]
Finally, we effectively fill the cells with the variables of different types. Let \( M_n(\theta, \gamma, \mu) \) be the number of ways to do this and obtain a formula in \( F(\theta, A) \) meeting our requirements. We start with the heavy variables, which must have \( \sigma n \) occurrences (recall \( \sigma = 1 - \sum_{0 \leq p \leq x \leq x_{\max}} x \theta x, p \)). We assign their occurrences to cells, which automatically determines the sign of each occurrence, having already completely specified \( A \) on the one hand, and the contents, 0 or 1, of each cell, on the other. We bound the ways to assign all the occurrences of heavy variables to cells by the quantity

\[
\eta(\theta, n, c) = \left( \frac{\lambda n}{\sigma n} \right)^{c n}.
\]

The \( \gamma_1 cn \) clauses of type \( (A, 1) \) contain \( \gamma_1 \lambda n \) cells, \( \gamma_1 cn \) of which are already reserved for nonzero literals. Among these, some already contain occurrences of heavy variables. Let their number be \( \sigma_1 n \); this is not an independent parameter since

\[
\sigma_1 = \gamma_1 c - \sum_{0 \leq p \leq x \leq x_{\max}} \theta x, p \sum_{0 \leq j \leq x} |p - j| \mu x, p, j.
\]

(2)

There remain \( \gamma_1 c - \sigma_1 \) cells to be filled in this group. These are filled with the \( p - j \) unnegated occurrences of variables of type \( (A, x, p, j) \) with \( 0 \leq j \leq p - 1 \) and the \( j - p \) negated occurrences of variables of type \( (A, x, p, j) \) with \( p \leq j \leq x \). Thus the number of ways to fill the \( \gamma_1 c - \sigma_1 \) cells is:

\[
\mathcal{M}_1 = \frac{[(\gamma_1 c - \sigma_1) n]!}{\prod_{0 \leq p \leq x \leq x_{\max}} \prod_{j=0}^{p-1} (p - j)! \theta x, p, j^{\mu x, p, j} \prod_{j=p}^{x} (j - p)! \mu x, p, j^{\theta x, p, n}}.
\]

Next, we fill the cells already reserved for nonzero literals, which do not pertain to clauses of type \( (A, 1) \). It will be convenient to introduce the normalized nonzero spread of \( F \) under \( A \), namely

\[
\psi = 1/3 \left( \gamma_1 + 2 \gamma_2 + 3 \gamma_3 \right).
\]

(3)

Among the \( \lambda \psi n \) cells in total which are to receive nonzero literals, let \( \sigma_1 n \) ones contain occurrences of heavy variables. \( \sigma_1 \), like \( \sigma_1 \), is a known quantity:

\[
\sigma_1 = \lambda \psi - \sum_{0 \leq p \leq x \leq x_{\max}} \theta x, p \left[ p \sum_{0 \leq j \leq p-1} \mu x, p, j + (x - p) \sum_{p \leq j \leq x} \mu x, p, j \right],
\]

(4)

For the \( \lambda \psi n - (\sigma_1 - \sigma_1) n - \gamma_1 cn \) remaining cells in this group, we have available, for each variable of type \( (A, x, p, j) \) with \( 0 \leq j \leq p - 1 \), the \( p \) unnegated occurrences less \( p - j \) already placed; and if the type is \( (A, x, p, j) \) with \( p \leq j \leq x \), the \( x - p \) negated occurrences less \( j - p \) already placed. Thus, the number of ways to do the assignment is:

\[
\mathcal{M}_2 = \frac{[(\lambda \psi - \gamma_1 c - \sigma_1 + \sigma_1) n]!}{\prod_{0 \leq p \leq x \leq x_{\max}} \prod_{j=0}^{p-1} j! \mu x, p, j^{\theta x, p, n} \prod_{j=p}^{x} (x - j)! \mu x, p, j^{\theta x, p, n}}.
\]

Lastly, we deal with the \( \lambda (1 - \psi) n \) cells reserved for null literals, of which \( (\sigma - \sigma_1) n \) are already filled. For the remaining ones, we have \( x - p \) occurrences of variables of type \( (A, x, p, j) \).
with $0 \leq j \leq p - 1$, and $p$ occurrences if $p \leq j \leq x$. So, we can fill them in $\mathcal{M}_3$ ways, where

$$
\mathcal{M}_3 = \frac{\{\lambda (1 - \psi) - \sigma + \sigma_1\}!}{\prod_{0 \leq p \leq x \leq x_{\max}} (x - p)! \sum_{j=0}^{p-1} \mu_{x,p,j} \theta_{x,p}^n \prod_{j=p}^x \mu_{x,p,j} \theta_{x,p}^n}
$$

To sum up, $M_n(\theta, \gamma, \mu) \leq M_1 M_2 M_3 \eta(\theta, n, c)$, so that

$$
Z(\theta, \gamma, \mu, n, c) \leq A_n(\gamma, c) B_n(\theta, \mu) M_1 M_2 M_3 \eta(\theta, n, c).
$$

(5)

### 3.3 The expectation.

It follows from (5), the preceding discussion, and the definition of $F(\theta, A)$, that, setting $J_n = \bigcup_{1 \leq k \leq n} I_k$, we have

$$
\sum_{A \in \{0, 1\}^n} |F(\theta, A)| \leq \eta(\theta, n, c) \sum_{\gamma \in I_{cn}} A_n(\gamma, c) \sum_{\mu \in J_n} B_n(\theta, \mu) \mathcal{M}_1 \mathcal{M}_2 \mathcal{M}_3,
$$

(6)

where the summation is under the constraints

$$
\gamma_1 + \gamma_2 + \gamma_3 = 1
$$

(7)

and

$$
\sum_{j=0}^x \mu_{x,p,j} = 1, \quad 0 \leq p \leq x \leq x_{\max},
$$

(8)

and where $\sigma_1$ and $\hat{\sigma}_1$ are expressed in $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ as functions of $\gamma$ and $\mu$, see (4) and (2).

We now introduce a modified form of (4) which will be convenient later. We set

$$
\alpha_{x,p} = \sum_{0 \leq j \leq p-1} \mu_{x,p,j},
$$

(9)

the proportion of variables with $x$ total, $p$ positive occurrences having the value 1 under $A$. Taking account of (8), (4) can be written using only the $\alpha_{x,p}$'s:

$$
\lambda \psi - \sigma_1 = K(\theta) - \sum_{0 \leq p \leq x \leq x_{\max}} H_{x,p}(\theta) \alpha_{x,p}, \quad \text{where}
$$

(10)

$$
K(\theta) = \sum_{0 \leq p \leq x \leq x_{\max}} (x - p) \theta_{x,p} \text{ and, for } 0 \leq p \leq x \leq x_{\max}: \quad H_{x,p}(\theta) = (x - 2p) \theta_{x,p}.
$$

From Lemma 3.1, Proposition 3.2, and (6), we get, for any fixed $\theta \in \Theta_{\varepsilon, x_{\max}, n, c}$:

$$
E(X_{n, \varepsilon, x_{\max}, c}) \leq \frac{\gamma_n \sum_{x>2p} \hat{\kappa}_{x,p}}{(2n)^2} \eta(\theta, n, c) \sum_{\gamma \in I_{cn}} A_n(\gamma, c) \sum_{\mu \in J_n} B_n(\theta, \mu) \mathcal{M}_1 \mathcal{M}_2 \mathcal{M}_3,
$$

(11)

subject again to (7) and (8).
4 Asymptotics.

4.1 Bound for the exponential order.

Still for a fixed \( \theta \), we now bound the general term of (3), using a standard inequality for multinomial coefficients:

\[
\binom{r}{r_1 r_2 \ldots r_s} \leq \frac{r^r}{r_1^{r_1} r_2^{r_2} \ldots r_s^{r_s}},
\]

which gives first, taking account of (7):

\[
A_n(\gamma, c)^{1/n} \leq \frac{3^c}{\gamma_1^{\gamma_2} (3\gamma_3)^c};
\]

further,

\[
B_n(\theta, \mu)^{1/n} \leq 2r \frac{1}{\prod_{0 \leq p \leq x \leq x_{\max}} \theta_{x,p}^{\theta_{x,p}} \prod_{0 \leq p \leq x \leq x_{\max}} \left( \prod_{j=0}^{x-1} \mu_{x,p,j}^{\mu_{x,p,j}} \right)};
\]

\[
\mathcal{M}_1^{1/n} \leq \frac{(\gamma_1 c - \tilde{\sigma}_1)^{\gamma_1 c - \tilde{\sigma}_1}}{\prod_{0 \leq p \leq x \leq x_{\max}} \left( \prod_{j=0}^{x-1} (p-j)!^{\mu_{x,p,j}} \theta_{x,p}^{\mu_{x,p,j}} \right)};
\]

next, bearing in mind that, by (4), \( \lambda \psi - \sigma_1 \) does not depend on \( n \):

\[
\mathcal{M}_2^{1/n} \leq \frac{\left( \lambda \psi - \gamma_1 c - \sigma_1 + \tilde{\sigma}_1 \right)^{\frac{n}{e}} \lambda \psi - \gamma_1 c - \sigma_1 + \tilde{\sigma}_1}{\prod_{0 \leq p \leq x \leq x_{\max}} \left( \prod_{j=0}^{x-1} (p-j)!^{\mu_{x,p,j}} \theta_{x,p}^{\mu_{x,p,j}} \right)};
\]

and, using (5):

\[
\mathcal{M}_3^{1/n} \leq \frac{\left( \lambda (1-\psi) + \sigma_1 - \sigma \right)^{\frac{n}{e}} \lambda (1-\psi) + \sigma_1 - \sigma}{\prod_{0 \leq p \leq x \leq x_{\max}} \left( (x-p)!^{\alpha_{x,p}} p^{1-\alpha_{x,p}} \right)^{\theta_{x,p}}};
\]

So, writing \((p-j)! = p! / (p-j)!\) and \((j-p)! = (x-j)! / (x-p)! / (x-p)! / (x-j)!\):

\[
\mathcal{M}_1^{1/n} \mathcal{M}_2^{1/n} \leq (\gamma_1 c - \tilde{\sigma}_1)^{\gamma_1 c - \tilde{\sigma}_1} \left( \lambda \psi - \gamma_1 c - \sigma_1 + \tilde{\sigma}_1 \right)^{\frac{n}{e}} \lambda \psi - \gamma_1 c - \sigma_1 + \tilde{\sigma}_1 \times
\]

\[
\left( \frac{n}{e} \right)^{\gamma_1 c - \tilde{\sigma}_1} \prod_{0 \leq p \leq x \leq x_{\max}} \left( \prod_{j=0}^{x-1} (p-j)!^{\mu_{x,p,j}} \theta_{x,p}^{\mu_{x,p,j}} \left( (x-p)!^{\alpha_{x,p}} (x-p)!^{1-\alpha_{x,p}} \right)^{\theta_{x,p}} \right).
\]

Bounding the sum in (3) by its maximum term times the number of terms \( |I_{cn}| |J_n| \) with \( |J_n| \leq n (n + 3) / 2 \), we get, after some simplification and whenever \( c \leq 5 \):

\[12\]
We are about to replace our estimate (12) by one that is uniformly valid for all \( c, \varepsilon \)
and \( \gamma \), to \([\lambda_{\min}, \lambda_{\max}]\) with \( 9 \leq \lambda_{\min} \leq \lambda_{\max} \leq 15 \). Later, we will be more specific and impose \( c_{\min} = c = c_{\max} = 4.506 \).

For such an interval \([c_{\min}, c_{\max}]\), it is easy, by elementary expectation calculations, to determine intervals \([\gamma_{1_{\min}}, \gamma_{1_{\max}}], [\gamma_{2_{\min}}, \gamma_{2_{\max}}], [\gamma_{3_{\min}}, \gamma_{3_{\max}}], [\psi_{\min}, \psi_{\max}]\), such that for \( c \in [c_{\min}, c_{\max}] \), the probability that a formula in \( \Omega(n, c) \) has a solution with at least one of \( \gamma_1, \gamma_2, \gamma_3, \psi \) falling outside the corresponding range is always exponentially small. For example, for \([c_{\min}, c_{\max}] \subset [3, 5]\) we can take these intervals to be \([0.21, 0.65], [0.21, 0.65], [0.017, 0.32]\), and \([0.47, 0.68]\), respectively.

This means that in investigating, by more sophisticated means, the probability that a formula in \( \Omega(n, c) \) is satisfiable, we need only consider solutions, or indeed PPSs, with \( \gamma_1, \gamma_2, \gamma_3, \) and
ψ in their respective intervals. Thus, we can define the r.v. \( X_{\eta,\varepsilon,x_{\text{max}},\lambda} \) with these more restricted PPSs, and \( \mathcal{F}(\theta,\mathcal{A}) \) similarly. All that we have said up to now goes over, notably Propositions 1.2 and 3.2; and (12) holds, with the maximum subject to these additional restrictions, viz

\[
\gamma_j \in [\gamma_{j\text{ min}}, \gamma_{j\text{ max}}], \quad j = 1, 2, 3; \quad \psi \in [\psi_{\text{ min}}, \psi_{\text{ max}}].
\]  

(13)

Henceforth we assume these additional constraints throughout; we also fix \( \varepsilon = 10^{-15} \) and \( x_{\text{max}} = 56 \).

4.3 The \( \theta \)-free estimate.

Deriving from (12), at controllably small cost, an estimate where the fixed but unknown \( \theta_{x,p} \)'s are replaced by the known \( \tilde{\kappa}_{x,p} \)'s, is a matter of easy but tedious calculations which we will only sketch. Anyway, one could get by with coarser bounds than we give by simply choosing a smaller \( \varepsilon \) and larger \( x_{\text{max}} \). Note that we have relied on \( x_{\text{max}} \) being even to simplify some of the calculations slightly.

One somewhat delicate point is how to deal with the numerous quantities of the form \( x^x \) where \( x \) is unknown but ‘near’ some known \( y \). We use the following very elementary lemma, not sharp but sufficient, so not worth improving.

**Lemma 4.1** Let \( L \) be the function \( \eta \mapsto (2\eta)^{-2\eta} \) on \( \mathbb{R}_+ \). Then whenever \( x, y, \eta \) are positive reals with \( \eta \leq 0.05 \), \( |x - y| \leq \eta \), and \( x \leq 30 \), we have \( L(\eta)^{-1} \leq y^y/x^x \leq L(\eta) \).

**Proof.** (outline) For \( x \leq 30 \), we study the function \( f_x(h) = (x + h)^{x+h} \) on the interval \( I_\eta(x) = [\max(-x,-\eta), \eta] \), showing that whenever \( |h| \leq \eta \), \( f_x(h)/x^x \) falls between \( L(\eta)^{-1} \) and \( L(\eta) \). This is done by elementary monotony considerations, distinguishing the two cases \( |h| > x \) and \( |h| \leq x \), the second being split into two subcases where the double inequality \( 1/e - \eta \leq x \leq 1/e + \eta \) either holds or not.

4.3.1 Eliminating the \( \sigma \)'s and \( \tau \), and withdrawing \( \theta \) from \( \psi \) and \( \gamma \).

From \( \tau \leq 1 - \sum_{x,p} \tilde{\kappa}_{x,p} + \sum_{x,p} |\theta_{x,p} - \tilde{\kappa}_{x,p}| \), we get, in terms of \( D = D(x_{\text{max}}) \),

\[
\tau (\theta,x_{\text{max}},\lambda) \leq R_1 (\varepsilon,x_{\text{max}})
\]

uniformly for all \( \theta \in \Theta_{\varepsilon,x_{\text{max}},n,c} \) and all relevant \( \lambda \), where

\[
R_1 (\varepsilon,x_{\text{max}}) = \frac{\lambda_{\text{max}}^{x_{\text{max}}+1}}{(x_{\text{max}} + 1)!} + \varepsilon D (x_{\text{max}}).
\]
Similarly, from \( \sigma \leq \left( \lambda - \sum_{x,p} x \tilde{k}_{x,p} \right) + \sum_{x,p} x |\theta_{x,p} - \tilde{k}_{x,p}| \) we obtain, in terms of \( P_2(\xi) = \xi (\xi + 1) (\xi + 2) / 3 \) that \( \sigma(\theta, x_{max}, \lambda) \leq R_2(\varepsilon, x_{max}) \), again uniformly for all \( \theta \in \Theta_{\varepsilon,x_{max},n,c} \) and relevant \( \lambda \), where

\[
R_2(\varepsilon, x_{max}) = \frac{\lambda_{x_{max}}^{\max+1}}{x_{max}!} + \varepsilon P_2(x_{max}).
\]

As for \( \sigma_1 \) and \( \tilde{\sigma}_1 \), we simply use \( \tilde{\sigma}_1 \leq \sigma_1 \leq \sigma \).

Turning to \( \psi \) and \( \gamma \), from (2) and (4) there are natural candidates for the \( \theta \)-free versions of \( \gamma_1 \) and \( \psi \), namely

\[
\beta_1 = \frac{1}{c} \sum_{0 \leq p \leq x \leq x_{max}} \tilde{k}_{x,p} \left[ \sum_{0 \leq j \leq p-1} (p - j) \mu_{x,p,j} + \sum_{p \leq j \leq x} (j - p) \mu_{x,p,j} \right],
\]

(14)

\[
\phi = \frac{1}{\lambda} \sum_{x,p} \tilde{k}_{x,p} \left[ p \sum_{0 \leq j \leq p-1} \mu_{x,p,j} + (x - p) \sum_{p \leq j \leq x} \mu_{x,p,j} \right],
\]

(15)

or, equivalently so long as (8) holds, cf. (10):

\[
\phi = \frac{1}{\lambda} \left( \tilde{K} - \sum_{0 \leq 2p \leq x \leq x_{max}} \tilde{H}_{x,p} \alpha_{x,p} \right)
\]

with

\[
\tilde{K} = \sum_{0 \leq 2p \leq x \leq x_{max}} (x - p) \tilde{k}_{x,p}
\]

and

\[
\tilde{H}_{x,p} = (x - 2p) \tilde{k}_{x,p},
\]

(16)

which is the definition of \( \phi \) we adopt, since it will be helpful subsequently.

Observing that from (7) and (3), we have \( \gamma_2 = 3(1 - \phi) - 2\gamma_1 \) and \( \gamma_3 = \gamma_1 - 2 + 3\phi \), we define the \( \theta \)-free versions of \( \gamma_2 \) and \( \gamma_3 \) as respectively

\[
\beta_2 = 3(1 - \psi) - 2\beta_1 \quad \text{and} \quad \beta_3 = \beta_1 - 2 + 3\psi.
\]

(17)

Using the above bounds on \( \sigma \), it is easy to estimate \( |\phi - \psi|, |\beta_j - \gamma_j| \), and the worst-case error incurred in replacing \( \phi \) by \( \psi \) and the \( \beta_j \)'s by the \( \gamma_j \)'s in (12). Setting \( P_3(\xi) = \xi (\xi + 2) (2\xi + 3) / 8 \) and

\[
R_3(\varepsilon, x_{max}) = \frac{\lambda_{x_{max}}^{\max+1}}{x_{max}!} + \varepsilon \frac{P_2(x_{max}) + P_3(x_{max})}{\lambda_{\min}},
\]

we find

\[
|\phi - \psi| \leq R_3(\varepsilon, x_{max}), \quad |\beta_1 - \gamma_1| \leq 3R_3(\varepsilon, x_{max}), \quad |\beta_2 - \gamma_2| \leq 9R_3(\varepsilon, x_{max}), \quad |\beta_3 - \gamma_3| \leq 6R_3(\varepsilon, x_{max}),
\]

(18)

and therefore the constraints (13) imply the following ones on \( \mu \):

\[
\beta_j \in [\beta_{j_{\min}}, \beta_{j_{\max}}], \quad j = 1, 2, 3; \quad \phi \in [\phi_{\min}, \phi_{\max}]
\]

(19)

where \( \beta_{1_{\min}} = \gamma_{1_{\min}} - 3R_3(\varepsilon, x_{max}), \beta_{1_{\max}} = \gamma_{1_{\max}} + 3R_3(\varepsilon, x_{max}) \), and so on.
Since \( R_3(\varepsilon, x_{\text{max}}) < 1.035 \times 10^{-9} \), from \( [18] \) and \( [19] \) we see that all of \( \beta_1, \beta_2, \beta_3, \phi, 3\phi - \beta_1, \) and \( 3(1 - \phi) \) are positive, \( \leq 30, \) and at a distance of less than 0.05 from the corresponding \( \theta \)-dependent quantities. This allows us, using also \( \sigma \leq R_2(\varepsilon, x_{\text{max}}) < 1.54 \times 10^{-8}, \) repeatedly to apply Lemma [1.1] and get:

\[
\left[ \sum_{A \in \{0, 1\}^n} |F(\theta, A)| \right]^\frac{1}{n} \leq G_1(\varepsilon, x_{\text{max}}) 3^c (\lambda \varepsilon)^\lambda \left( \frac{1}{3\varepsilon} \right)^\lambda \times
\]

\[
6n^3 \max_{\mu_{x,p,j} \in [0, 1]} \left[ \sum_{0 \leq p \leq x \leq x_{\text{max}}} p! (x - p)! \theta_{x,p} \prod_{j=0}^{x} \left( \frac{\mu_{x,p,j}}{h_{x,p,j}} \right)^{\mu_{x,p,j}} \right]^c,
\]

where

\[
G_1(\varepsilon, x_{\text{max}}) = 2^{R_1(\varepsilon, x_{\text{max}})} L \left( R_1(\varepsilon, x_{\text{max}}) \right) L \left( R_2(\varepsilon, x_{\text{max}}) \right) \left( \frac{18 \varepsilon}{\lambda_{\text{min}}} \right)^{R_2(\varepsilon, x_{\text{max}})} \times
\]

\[
\left[ L \left( \frac{R_2(\varepsilon, x_{\text{max}})}{6} \right) \right]^6 \left[ L \left( \frac{3\varepsilon}{\lambda_{\text{min}}} P_3(x_{\text{max}}) \right) \right] \left[ L \left( \frac{6\varepsilon}{\lambda_{\text{min}}} P_3(x_{\text{max}}) \right) \right]^{c_{\text{max}}},
\]

and the max is subject to \([8]\) and \([19]\).

### 4.3.2 Removing \( \theta \) from the powers-and-factorials product.

Since the only real difficulty is, possibly, getting started on the right path, we only indicate how we break down the error incurred from replacing the \( \theta_{x,p} \)’s by the \( \kappa_{x,p} \)’s, into three factors \( A, B, C, \) to be estimated separately. We have

\[
\prod_{0 \leq p \leq x \leq x_{\text{max}}} p! (x - p)! \theta_{x,p} \prod_{j=0}^{x} \left( \frac{\mu_{x,p,j}}{h_{x,p,j}} \right)^{\mu_{x,p,j}} \leq \prod_{0 \leq 2p \leq x \leq x_{\text{max}}} p! (x - p)! \kappa_{x,p} \prod_{j=0}^{x} \left( \frac{\mu_{x,p,j}}{h_{x,p,j}} \right)^{\mu_{x,p,j}} \kappa_{x,p} \tag{A\text{-}B\text{-}C}
\]

(note the change of domain for the index \( p \), with

\[
A = \prod_{1 \leq p \leq x_{\text{max}}, 2p > x} p! (x - p)! \theta_{x,p} \prod_{j=0}^{x} \left( \frac{\mu_{x,p,j}}{h_{x,p,j}} \right)^{\mu_{x,p,j}} \theta_{x,p}^{-\theta_{x,p}},
\]

\[
B = \prod_{0 \leq 2p \leq x \leq x_{\text{max}}} \left( \frac{\kappa_{x,p}}{\theta_{x,p}} \right)^{\theta_{x,p}},
\]

\[
C = \prod_{0 \leq 2p \leq x \leq x_{\text{max}}} p! (x - p)! \kappa_{x,p} \prod_{j=0}^{x} \left( \frac{\mu_{x,p,j}}{h_{x,p,j}} \right)^{\mu_{x,p,j}} \kappa_{x,p}^{-\theta_{x,p}}.
\]

16
We find $A \leq G_A(\varepsilon, x_{\max})$, $B \leq G_B(\varepsilon, x_{\max})$, $C \leq G_C(\varepsilon, x_{\max}, \lambda)$, with
\[
G_A(\varepsilon, x_{\max}) = 12^2 \frac{2^{x_{\max} + 2} x_{\max} + 2 x_{\max} - 1}{8} \varepsilon L(\varepsilon) \frac{x_{\max}}{x_{\max} + 3},
\]
\[
G_B(\varepsilon, x_{\max}) = \prod_{x=0}^{x_{\max}} (1 + \varepsilon)^{\left\lceil \frac{x}{2} \right\rceil} \leq (1 + \varepsilon)^{\frac{x_{\max}^2}{4} (x_{\max} + 1)},
\]
and, using the fact that
\[
\text{for all } \lambda \in [\lambda_{\min}, \lambda_{\max}], \quad x_{\max} \geq \frac{2 \lambda - \log 2}{\log \lambda - \log 2}, \quad (20)
\]
\[
G_C(\varepsilon, x_{\max}, \lambda) = \left\{ \begin{array}{ll}
(x_{\max} + 1) & \frac{x_{\max} + 2}{2} \\
2^x \frac{2^{x_{\max} + 1}}{x_{\max} + 7} & x_{\max} + 4 \left[ 1 - \lambda \left( \frac{3}{2} \right)^{x_{\max} + 8} \frac{x_{\max}}{x_{\max} + 1} \right] \end{array} \right\} \varepsilon.
\]

Observing that since $e^{-\lambda} \left( \frac{1}{2} \right)^{x_{\max}}$ increases with $\lambda$ within our range of interest, and setting $G_2(\varepsilon, x_{\max}) = G_A(\varepsilon, x_{\max}) G_B(\varepsilon, x_{\max}) G_C(\varepsilon, x_{\max}, \lambda_{\max})$, we conclude that, for any $c$ in our chosen range and the max being again subject to (8) and (19):
\[
\left( \sum_{A \in \{0,1\}^n} |F(\theta, A)| \right)^{\frac{x}{x_{\max}}} \leq G_1(\varepsilon, x_{\max}) G_2(\varepsilon, x_{\max}) 3^c (\lambda n)^{\lambda} \left( \frac{1}{3 \varepsilon} \right)^{\lambda} (6n^3)^{\frac{1}{x_{\max}}} \times \quad (22)
\]
\[
\max_{\mu_{x,p} \in [0,1]} \prod_{0 \leq 2p \leq x \leq x_{\max}} \left[ p! (x - p)! \right] \prod_{j=0}^{x_{\max}} \frac{\mu_{x,p,j}^{\cdot \mu_{x,p,j}}}{h_{x,p,j}^{\cdot h_{x,p,j}}} \right]^{\frac{1}{h_{x,p,j}}},
\]

Note that the $\mu_{x,p,j}$'s with $2p > x$ are now irrelevant, having vanished from the bound (they do not actually figure in $\beta_1$, $\beta_2$, $\beta_3$, or $\phi$), and thus the equality constraints under which we now perform the maximisation are just (8) for $0 \leq 2p \leq x \leq x_{\max}$.

5 Maximization.

By (19), for $c$ within our range, the max on the r.h.s. of (23) may be restricted to vectors $\mu \in U$, where
\[
U = \beta_1^{-1}(\beta_{1_{\min}}, \beta_{1_{\max}}) \cap \beta_2^{-1}(\beta_{2_{\min}}, \beta_{2_{\max}}) \cap \beta_3^{-1}(\beta_{3_{\min}}, \beta_{3_{\max}}) \cap \phi^{-1}(\phi_{1_{\min}}, \phi_{1_{\max}}),
\]
is an open subset of $\mathbb{R}^N$ where $N = N(x_{\max}) = 1/24 \left( x_{\max}^2 + 2 \right) \left( 4x_{\max}^2 + 13x_{\max} + 12 \right)$ (recall that we have dropped the irrelevant variables $\mu_{x,p,j}$ with $p > x/2$). For the moment,
we do not further specify these reals; we do so later when restricting \( c \) to a single value. (We do already assume \( 3\phi_{\min} > \beta_{1\max} \) though). For now, (22) leads to the following problem of constrained maximization:

\[
\max_{\mu \in \mathbb{R}_{+}^{N} \cap \mathcal{U}} \sum_{0 \leq p \leq x_{\max}} \tilde{K}_{x,p} \sum_{j=0}^{x} \mu_{x,p,j} \log \left( \frac{h_{x,p,j}}{\mu_{x,p,j}} \right)
\]

\[
+ c \left\{ (3\phi - \beta_{1}) \log (3\phi - \beta_{1}) + [3 (1 - \phi)] \log [3 (1 - \phi)] - \beta_{2} \log \beta_{2} - \beta_{3} \log (3\beta_{3}) \right\}
\]

subject to the constraints (8) which we rewrite as

\[ C_{x,p} = 0, \quad \text{where} \quad C_{x,p} = -1 + \sum_{j=0}^{x} \mu_{x,p,j}. \]

This is not yet amenable to traditional differential techniques, since the set \( \mathbb{R}_{+}^{N} \cap \mathcal{U} \) is not open. However, it is not difficult to bar out the vectors on the boundary as candidates for global, indeed even local, maximizers, as we now proceed to do.

Let us compute the gradient of the function of \( \mu \) maximized in (23), say \( f_{1}(\mu) \). For the quantity inside the braces, using (17) and (7), and setting

\[ U = \frac{9 (1 - \phi) \beta_{3}}{(3\phi - \beta_{1}) \beta_{2}} \quad \text{and} \quad V = 1 + \frac{\beta_{2}^{2}}{3 (3\phi - \beta_{1}) \beta_{3}} = \frac{(\beta_{1} + 6\phi - 3)^{2}}{3\beta_{3}(3\phi - \beta_{1})}, \]

we obtain for \( \nabla f_{1} : \)

\[
3 \log \frac{3\phi - \beta_{1}}{3 (1 - \phi)} \nabla \phi - \log (3\phi - \beta_{1}) \cdot \nabla \beta_{1} - \log \beta_{2} \cdot \nabla \beta_{2} - \log (3\beta_{3}) \cdot \nabla \beta_{3} - \nabla \beta_{1} - \nabla \beta_{2} - \nabla \beta_{3} = -3 \log U \cdot \nabla \phi + \log (V - 1) \cdot \nabla \beta_{1},
\]

so, taking the \((x, p, j)\)-coordinate:

\[
\frac{\partial f_{1}}{\partial \mu_{x,p,j}} = \begin{cases} 
\tilde{K}_{x,p} \left[ \log \frac{h_{x,p,j}}{\mu_{x,p,j}} - 1 + (x - 2p) \log U + (p - j) \log (V - 1) \right], & 0 \leq j \leq p - 1; \\
\tilde{K}_{x,p} \left[ \log \frac{h_{x,p,j}}{\mu_{x,p,j}} - 1 + (j - p) \log (V - 1) \right], & p \leq j \leq x.
\end{cases}
\]  

(25)

With this knowledge, we can establish:

**Lemma 5.1** No feasible vector \( \mu \) (i.e. \( \mu \in \mathbb{R}_{+}^{N} \cap \mathcal{U} \) satisfying (24)) having at least one null coordinate can be a local maximizer for the problem (23).

**Proof.** Choose \( j_{1} \) and \( j_{2} \) such that \( \mu_{x,p,j_{1}} = 0 \) and \( \mu_{x,p,j_{2}} \neq 0 \), and consider the real-valued function \( f_{1}^{*} \) defined on \([0, \mu_{x,p,j_{2}}]\) by \( f_{1}^{*}(\xi) = f_{1}(\mu_{\xi}) \), where \( \mu_{\xi} \) differs from \( \mu \) only in the \((x, p, j_{1})\) and \((x, p, j_{2})\) coordinates, the former being equal to \( \xi \) and the latter to \( \mu_{x,p,j_{2}} - \xi \).
Of course, for sufficiently small $\xi$, $\mu_\xi$ is still feasible, so it suffices to show that $0$ is not a local maximum for $f_1^\ast$. But, using (23):

$$\frac{1}{\bar{\kappa}_{x,p}} \frac{\partial f_1^\ast}{\partial \xi} = \frac{1}{\bar{\kappa}_{x,p}} \left[ \frac{\partial f_1}{\partial \mu_{x,p,j_1}} (\mu_\xi) - \frac{\partial f_1}{\partial \mu_{x,p,j_2}} (\mu_\xi) \right]$$

$$= \log \frac{h_{x,p,j_1}}{\xi} - \log \frac{h_{x,p,j_2}}{\mu_{x,p,j_2} - \xi} + R_{x,p,j_1,j_2} \log [U (\mu_\xi)] + S_{p,j_1,j_2} \log [V (\mu_\xi) - 1],$$

where $|R_{x,p,j_1,j_2}| \leq x_{\text{max}}$ and $|S_{p,j_1,j_2}| \leq x_{\text{max}}$. As $\xi \to 0+$, the first term on the right tends to $+\infty$, the second remains bounded, while, since $\mu_\xi$ is feasible,

$$\frac{9 (1 - \phi_{\text{max}}) \beta_{3\text{min}}}{(3\phi_{\text{max}} - \beta_{1\text{min}}) \beta_{2\text{max}}} \leq U (\mu_\xi) \leq \frac{9 (1 - \phi_{\text{min}}) \beta_{3\text{max}}}{(3\phi_{\text{min}} - \beta_{1\text{max}}) \beta_{2\text{min}}},$$

and

$$\frac{\beta_{2\text{min}}}{3 (3\phi_{\text{max}} - \beta_{1\text{min}}) \beta_{3\text{max}}} \leq V (\mu_\xi) - 1 \leq \frac{\beta_{2\text{max}}}{3 (3\phi_{\text{min}} - \beta_{1\text{max}}) \beta_{3\text{min}}},$$

so the third and fourth terms stay finite too. All in all, $\partial f_1^\ast / \partial \xi$ is seen to tend to $+\infty$ as $\xi \to 0+$, so obviously (e.g., from the mean value formula) $0$ cannot be a local maximum for $f_1^\ast$. ■

So, the set constraint in problem (23) may be replaced by $\mu \in O$, with $O = [0, +\infty[^N \cap U$ (an open subset of $\mathbb{R}^N$), allowing a study of local optima by traditional differential methods. $f_1$ is readily seen to be bounded in $O$, since for $x \geq 2p$ all $h_{x,p,j}$’s are $\leq 2^{x-p-1}$; thus, if some number $M$ bounds from above all local maximizers in $O$, then $f_1 (\mu) \leq M$ for any feasible $\mu$.

Since all constraints are affine, we do not actually need any further constraint qualification such as linear independence of the gradients, though this is clearly the case. The classical method of Lagrange multipliers applies [34]: a necessary condition for optimality of $f_1$ at some feasible $\mu^\ast \in O$ is stationarity in the sense that there exist real numbers $\Lambda_{x,p}$, $0 \leq 2p \leq x \leq x_{\text{max}}$, such that at $\mu^\ast$:

$$\nabla f_1 + \sum_{0 \leq 2p \leq x \leq x_{\text{max}}} \Lambda_{x,p} \nabla C_{x,p} = 0. \quad (26)$$

Now take the $(x, p, j)$-coordinate of (23), using (27): since $\tilde{\kappa}_{x,p} \neq 0$ whenever $2p \leq x$, we get, for $0 \leq j \leq p - 1$ and $p \leq j \leq x$ respectively:

$$\mu_{x,p,j} = h_{x,p,j} \exp \left( \frac{\Lambda_{x,p}}{\tilde{\kappa}_{x,p}} - 1 \right) U^{x-2p} (V - 1)^{p-j} \quad \text{and} \quad \mu_{x,p,j} = h_{x,p,j} \exp \left( \frac{\Lambda_{x,p}}{\bar{\kappa}_{x,p}} - 1 \right) (V - 1)^{j-p}. \quad (27)$$

The Lagrange multipliers are determined by plugging this back into the constraints (24):

$$1 = \exp \left( \frac{\Lambda_{x,p}}{\bar{\kappa}_{x,p}} - 1 \right) \left[ U^{x-2p} \sum_{j=0}^{p-1} \binom{p}{j} (V - 1)^{p-j} + \sum_{j=p}^{x} \binom{x-p}{j-p} (V - 1)^{j-p} \right]$$

$$= \exp \left( \frac{\Lambda_{x,p}}{\bar{\kappa}_{x,p}} - 1 \right) \left[ U^{x-2p} (V^p - 1) + V^{x-p} \right].$$
Hence a necessary (and sufficient) condition for stationarity is that the \( N(x_{\text{max}}) \) unknowns \( \mu_{x,p,j} \), \( 0 \leq 2p \leq x \leq x_{\text{max}} \) such that \( 0 \leq j \leq x \), satisfy the following system of \( N(x_{\text{max}}) \) equations:

\[
\mu_{x,p,j} = \begin{cases} 
(p) \frac{U^{x-2p(V-1)^{p-j}}}{U^{x-2p(Vp-1)+Vx-p}} & \text{if } 0 \leq j \leq p-1; \\
(x-p) \frac{(V-1)^{i-p}}{U^{x-2p(Vp-1)+Vx-p}} & \text{if } p \leq j \leq x
\end{cases}
\]  

(28)

(where of course \( U \) and \( V \) are themselves fairly complicated functions of the \( \mu \)'s). Note also that for any solution \( \mu \), the quantity \( \alpha_{x,p} = \sum_{j=0}^{p-1} \mu_{x,p,j} \) has the summation-free expression:

\[
\alpha_{x,p} = \frac{U^{x-2p}(Vp-1)}{U^{x-2p}(Vp-1)+Vx-p} = 1 - \frac{1}{1 + \left( \frac{U}{V} \right)^{x-2p}(1 - \frac{1}{V})}
\]

and systematically equals 0 for \( p = 0 \). Further, under the same conditions the coefficient of \( \frac{\kappa_{x,p}}{c} \) in \( \beta_1 \) has the value

\[
\frac{1}{U^{x-2p}(Vp-1)+Vx-p} \left[ U^{x-2p} \sum_{h=1}^{p} h \binom{p}{h} (V-1)^h + \sum_{l=0}^{x-p} \binom{x-p}{l} (V-1)^l \right]
\]

\[
= \frac{V - 1 - pU^{x-2p}Vp + (x - p)Vx-p}{V - U^{x-2p}(Vp-1)+Vx-p},
\]

where again the summation in \( j \) has disappeared. The system (28) may seem hopeless at first sight. However, all unknowns can be extracted in terms of just two (affine) functions of themselves, \( \phi \) and \( \beta_1 \). This has the following consequence. Consider the system \( S^* \) of \( N(x_{\text{max}}) + 2 \) equations in as many unknowns obtained by viewing \( \phi \) and \( \beta_1 \) as two further unknowns, and (13) and (14) as two additional equations. A solution of (28) immediately gives one of \( S^* \), and conversely. Solving (28) amounts to solving \( S^* \) by trivially eliminating \( \phi \) and \( \beta_1 \). But the property just stated means that there is a better way to solve \( S^* \), namely eliminating the \( \mu \)'s, leaving just 2 equations in 2 unknowns. So, viewing now \( U \) and \( V \) as functions of \( \phi \) and \( \beta_1 \) only, we plug the r.h.s.'s of (28) into (13), (14) to obtain

\[
\lambda \phi = \tilde{K} - \sum_{1 \leq 2p \leq x \leq x_{\text{max}}} \tilde{\kappa}_{x,p} (x - 2p) \left[ 1 - \frac{1}{1 + \left( \frac{U}{V} \right)^{x-2p}(1 - Vp)} \right],
\]  

(29)

\[
\beta_1 c = \frac{V - 1}{V} \sum_{0 \leq 2p \leq x \leq x_{\text{max}}} \tilde{\kappa}_{x,p} \frac{pU^{x-2p}Vp + (x - p)Vx-p}{U^{x-2p}(Vp-1)+Vx-p}.
\]  

(30)

Having solved this in \( \phi \) and \( \beta_1 \), we plug them into (28) and obtain the \( \mu \)'s. While still highly nonlinear, the system (29), (30) can, as we shall show, be rigorously analyzed. But what we certainly cannot do is to exploit convexity considerations as in [11], [6]: here the objective function is not concave.
6 The equations: analysis and numerical resolution.

6.1 Preliminary transformations.

In the sequel, \((\phi, \beta_1)\) will denote an arbitrary solution of (29, 30) or an equivalent system. Before we proceed, it will be helpful to rearrange some of the already obtained expressions in a more convenient form.

6.1.1 The expectation revisited.

Assume that we have a solution \((\phi, \beta_1)\) to the system (29, 30) and that the corresponding parameters \(\mu\) do give rise to the (global) maximum appearing in (22). We show that our bound on the expectation simplifies to a formula where the parameters \(\mu\) only intervene via the two quantities \(\phi\) and \(\beta_1\) (and \(U\) and \(V\) viewed as functions of these).

First we have, in view of (28) and taking into account (8), (14), (9), and (16):

\[
\prod_{0 \leq 2p \leq x \leq x_{\text{max}}} \left( \frac{\mu_{x,p,j}}{h_{x,p,j}} \right)^{\tilde{\kappa}_{x,p}} = \frac{U K - \lambda \phi (V - 1)^{\beta_1}}{[U x - 2p (V_p - 1) + V x - p]^\tilde{\kappa}_{x,p}}.
\]

Call the inverse of the r.h.s. \(g_2(\phi, \beta_1)\), plug it back into (22), and modify \(\Pi\) accordingly, using \((2\varepsilon n)^{D/n} \leq \exp(2\varepsilon D/e)\):

\[
E(X_{n,\varepsilon,x_{\text{max}},c})^{1/n} \leq \left( 6n^3 \right)^{\tilde{\kappa}} 3^c \left( \frac{\lambda}{6e} \right)^\lambda \frac{2^{\sum_{x>2p} \tilde{\kappa}_{x,p}} \exp(2\varepsilon D/e)}{\prod_{0 \leq 2p \leq x \leq x_{\text{max}}} [p!(x - p)!^{\tilde{\kappa}_{x,p}}]} G_1(\varepsilon, x_{\text{max}}) \times G_2(\varepsilon, x_{\text{max}}) g_2(\phi, \beta_1) \left[ \frac{(3\phi - \beta_1)^{3\phi - \beta_1} [3(1 - \phi)]^{3(1 - \phi)}}{\beta_2^{\beta_2} (3\beta_3)^{\beta_3}} \right]^c. \tag{31}
\]

This is essentially the estimate that will serve in our numerical evaluations. It is possible further to transform it so that all exponents become fixed (i.e. independent of \(\phi\) and \(\beta_1\)), but this, although noteworthy, will not be used here.

Let us emphasize that the function of \(\phi\) and \(\beta_1\) on the right of (31) has little to do with the objective function in (22) (a function of \(\mu\), anyway). All we say is that it dominates \(E(X_{n,\varepsilon,x_{\text{max}},c})^{1/n}\) for some pair(s) \((\phi, \beta_1)\) satisfying the system (29, 30), and our final bound will be valid for any such solution, without having to assume or prove uniqueness. Although it can be seen that (29, 30) actually characterizes stationary values of that function too, the (in fact unique) solution is not a maximum but a saddle point.

6.1.2 A modified form of the second equation.

The numerator of the fraction in the sum on the right-hand side of (31) can be written

\[
pU^{x-2p}V^p + (x-p)V^{x-p} = (2p-x)U^{x-2p}(V^p-1)+(x-p)
\left[U^{x-2p}(V^p-1)+V^{x-p}\right]+pU^{x-2p},
\]

21
so that (30) also reads

\[
\beta_1 c = \frac{V - 1}{V} \left[ -\sum_{0 \leq 2p \leq x \leq x_{\text{max}}} H_{x,p} \alpha_{x,p} + \sum_{0 \leq 2p \leq x \leq x_{\text{max}}} (x - p) \tilde{K}_{x,p} + \sum_{2 \leq 2p \leq x \leq x_{\text{max}}} p \tilde{K}_{x,p} \frac{U_{x-2p}}{U_{x-2p} (Vp - 1) + Vx-p} \right].
\]

The second sum is by definition \( \tilde{K} \), while if (29) is verified, the first is \( \tilde{K} - \lambda \phi \). Thus the system (29-30) is equivalent to (29-32), where (32) is as follows:

\[
\beta_1 c = \frac{V - 1}{V} \left[ \lambda \phi + \sum_{2 \leq 2p \leq x \leq x_{\text{max}}} \tilde{K}_{x,p} \frac{pU_{x-2p}}{U_{x-2p} (Vp - 1) + Vx-p} \right]. \tag{32}
\]

### 6.1.3 The monotone behaviour of \( U \) and \( V \) in each variable separately.

Set [viewing \( \beta_2, \beta_3 \) as functions \( \tilde{\beta}_2, \tilde{\beta}_3 \) of \( \phi, \beta_1 \), cf. (17)] \( D_{\phi,\beta_1} = [\phi_{\text{min}}, \phi_{\text{max}}] \times [\beta_{1\text{min}}, \beta_{1\text{max}}] \cap \tilde{\beta}_2^{-1}([\beta_{2\text{min}}, \beta_{2\text{max}}]) \cap \tilde{\beta}_3^{-1}([\beta_{3\text{min}}, \beta_{3\text{max}}]) \). Within our range of \( c \), we can disregard pairs \( (\phi, \beta_1) \notin D_{\phi,\beta_1} \), so we will limit our study of (29-32) to \( D_{\phi,\beta_1} \).

For fixed \( \phi \) and variable \( \beta_1 \) within \( D_{\phi,\beta_1} \), then, \( U \) increases (strictly) as the quotient of an increasing numerator by a decreasing denominator.

For fixed \( \beta_1 \), \( U \) increases (strictly) in \( \phi \) since \( \beta_{1\text{max}} < 1 \) implies

\[
\frac{\partial \log U}{\partial \phi} = -\frac{1}{1 - \phi} + \frac{3}{\beta_3} - \frac{3}{3\phi - \beta_1} + \frac{3}{\beta_2} \frac{2\beta_1}{(1 - \phi) \beta_2} + \frac{6(1 - \beta_1)}{\beta_3 (3\phi - \beta_1)} > 0
\]

(As an unconstrained linear combination of the \( \mu_{x,p,j} \)'s, \( \beta_1 \) does reach values \( > 1 \)).

For fixed \( \phi, V - 1 \) has a decreasing numerator, while the denominator, three times a product of factors with a constant sum, increases until \( \beta_1 \) reaches \( \beta_{1M} \) such that \( \beta_3 = 3\phi - \beta_1 \); however, \( \beta_{1M} = 1 > \beta_{1\text{max}} \), so \( V \) decreases on \( D_{\phi,\beta_1} \).

For fixed \( \beta_1, V - 1 \) also decreases owing to a decreasing denominator and increasing numerator.

To sum up: with either variable fixed, \( U \) increases and \( V \) decreases in the other variable.

### 6.1.4 Bounds on \( U \) and \( V \).

We henceforth set \( c = c_{\text{min}} = c_{\text{max}} = 4.506 \). All the foregoing remains valid, but some inequalities become tighter, starting with (13) and (19); in the latter we can now take \( \beta_{1\text{min}} = \beta_{2\text{min}} = 0.33018; \beta_{1\text{max}} = \beta_{2\text{max}} = 0.52891; \beta_{3\text{min}} = 0.077639; \beta_{3\text{max}} = 0.21782; \phi_{\text{min}} = 0.525245 \); and \( \phi_{\text{max}} = 0.619063 \). (Also, e.g. now \( R_3(\epsilon, x_{\text{max}}) \) has a smaller value < \( 1.104 \times 10^{-11} \).)
These limitations imply helpful ones on $U$ and $V$:

$$
U \leq U_{\text{max}1} = 2.69268, \quad U \leq U_{\text{max}2} / (V - 1) \text{ where } U_{\text{max}2} = 0.687424, \quad V \geq V_{\text{min}1} = 1.109255, \quad U/V \geq (U/V)_{\text{max}1} = 11.2022; \text{ we need and prove better ones than the last two.}
$$

We solve the constrained minimization problem (with variables $\phi$ and $\beta_1$): minimize $V$, subject to the 8 linear constraints written above. These define a convex polygonal domain with 7 sides in the plane $(\phi, \beta_1)$. Due to the decreasing character of $V$ in each variable, the minimum can obviously not be attained at an interior point, nor along any side other than the two given by $\beta_3 = \beta_{3\text{max}}$ and $\beta_2 = \beta_{2\text{min}}$. $V$ is easily seen to decrease in $\beta_1$ along the first and to increase in $\beta_1$ along the second, so the minimum is attained at their intersection, and is found equal to:

$$
V_{\text{min}2} = 1 + \frac{\beta_{2\text{min}}^2}{3\beta_{3\text{max}} (2\beta_{2\text{min}} + 3\beta_{3\text{max}})} = 1.126983.
$$

We can maximize $(U/V)$ in a very similar way, since it increases in each variable separately. Again, the maximum must be along one of the same two sides of the same polygon, as seen using the form $U/V = 27 (1 - \phi) \beta_3^2 / [\beta_2 (\beta_1 + 6\phi - 3)^2]$, and equals

$$
\left( \frac{U}{V} \right)_{\text{max}2} = \frac{9 [2 (1 - \beta_{3\text{max}}) - \beta_{2\text{min}}] \beta_{3\text{max}}^2}{\beta_{2\text{min}} (\beta_{2\text{min}} + 3\beta_{3\text{max}})^2} = 1.64966.
$$

### 6.2 Outline.

From now on, $c$ is taken to be equal to 4.506. The remainder of the paper is devoted to showing that for this $c$, the product $2(\rho + \varepsilon)nE(X_n, \varepsilon, x_{\text{max}}, c)$ tends to 0 as $n \to \infty$. Since the probability of satisfiability decreases in $c$, this will establish that the threshold is below 4.506.

While Figure 1 clearly suggests that the system (29, 32) has a unique solution, we present a strictly rigorous analysis. It exploits special features of this system, leading to numerical calculations which can be reliably and routinely performed to any desired precision.

A close study of the 2 equations, written as $Eq_1 = 0$ and $Eq_2 = 0$, shows that each defines $\phi$ as a unique decreasing function of $\beta_1$; then a constructive numerical procedure is applied to narrow down the location of any common root. Uniqueness is neither assumed nor proven, though this could be done with a little more effort.

Actually, it suffices to establish the (strict) monotony of $Eq_1$ and $Eq_2$ in each variable separately; the monotone behavior of the corresponding implicit functions follows. And it turns out that it is easier to reason directly in terms of this separate monotony, and that in this approach strictness is not used.

There is a slight restriction to the monotony of $Eq_1$, which does not affect the end result. A precise statement follows.
Proposition 6.1  

(i) Eq decreases strictly in each variable separately over the whole domain of interest $D_{\phi, \beta_1}$, and even over the wider set $[\phi_{\min}, \phi_{\max}] \times [\beta_{1\min}, \beta_{1\max}]$.

(ii) For any $\phi \in [\phi_{\min}, \phi_{\max}]$ (resp. any $\beta_1 \in [\beta_{1\min}, \beta_{1\max}]$), there exists $\beta_1^*(\phi) \leq 1$ (resp. $\phi^*(\beta_1) \leq 1$) such that $\mathrm{Eq}_1(\phi,.) < 0$ over the interval $[\beta_1^*(\phi), 1]$, (resp. such that $\mathrm{Eq}_1(.,\beta_1) < 0$ over the interval $[\phi^*(\beta_1), 1]$) and that $\mathrm{Eq}_1(\phi,.)$ decreases strictly on $[0, \beta_1^*(\phi)]$ (resp. $\mathrm{Eq}_1(.,\beta_1)$ decreases strictly on $[0, \phi^*(\beta_1)]$). In particular, if $\mathrm{Eq}_1(\phi, \beta_1) = 0$, then $\beta_1^*(\phi) > \beta_1$ and $\phi^*(\beta_1) > \phi$.

The remainder of this section is devoted to proving Proposition 6.1. We start with the equation that is monotone on the whole of $D_{\phi, \beta_1}$.

6.3 The second equation, separate monotony.

We use the modified form (32), and deal successively with fixed $\phi$, variable $\beta_1$, and the reverse. First, some considerations which apply to both. (32) can be rewritten equivalently:

$$0 = \mathrm{Eq}_2 \equiv -\beta_1 c + \lambda\phi \left(1 - \frac{1}{V}\right) + \sum_{1 \leq p \leq x \leq \max} \rho \tilde{K}_{x,p} \frac{V - 1}{V^{p - 1}} \left[1 - \frac{1}{1 + \left(\frac{U}{V}\right)^{x - 2p} \left(1 - \frac{1}{V}\right)}\right].$$

(33)

Call the denominator of the last fraction $D_{x,p}$. We differentiate one variable, leaving the other fixed. The derivatives are denoted simply by a prime because the context will always
make the meaning clear. The following equality holds in either case.

\[
\left\{ \frac{V - 1}{V (V^p - 1)} \left[ 1 - \frac{1}{D_{x,p}} \right] \right\}' = \frac{-V' pV^{p+1} - (p + 1) V^p + 1}{V^2 (V^p - 1)^2} \left[ 1 - \frac{1}{D_{x,p}} \right] + \\
\frac{V - 1}{V (V^p - 1)} \left( \frac{V}{V^p} \right)^{x-2p-1} \left( x - 2p \right) \left( \frac{U'}{U} - \frac{V'}{V} \right) \left( 1 - \frac{1}{V^p} \right) + \frac{pV'}{V^{p+1}}.
\]

It will be shown, in the relevant subsections, that

**Lemma 6.2** Let \( X = 3\phi - 1, Y = 3\phi - \beta_1, Z = Y - X = 1 - \beta_1 \). We have

\[
\frac{U'}{U} - \frac{V'}{V} = -\frac{V'}{V} \left[ \frac{pV^{p+1} - (p + 1) V^p + 1}{(V^p - 1)^2} \left[ 1 - \frac{1}{D_{x,p}} \right] - \frac{V - 1}{V (V^p - 1)} \left( \frac{U}{V^p} \right)^{x-2p-1} \left( x - 2p \right) \left( \frac{U'}{U} - \frac{V'}{V} \right) \left( 1 - \frac{1}{V^p} \right) + \frac{pV'}{V^{p+1}} \right] R,
\]

where \( R \) is a positive quantity such that for fixed \( \phi \)

\[
R = \frac{Y}{X} V - 1,
\]

while for fixed \( \beta_1 \)

\[
R < S = \frac{Y}{Z} (V - 1) + \frac{1}{2}.
\]

It follows that

\[
\left\{ \frac{V - 1}{V (V^p - 1)} \left[ 1 - \frac{1}{D_{x,p}} \right] \right\}' = -\frac{V'}{V^2} \left( \frac{pV^{p+1} - (p + 1) V^p + 1}{(V^p - 1)^2} \left[ 1 - \frac{1}{D_{x,p}} \right] - \frac{V - 1}{V (V^p - 1)} \left( \frac{U}{V^p} \right)^{x-2p-1} \left( x - 2p \right) \left( \frac{U'}{U} - \frac{V'}{V} \right) \left( 1 - \frac{1}{V^p} \right) + \frac{pV'}{V^{p+1}} \right) \left( x - 2p \right) R - \frac{pV - 1}{V^{p+1}}.
\]

The two terms inside the curly brackets on the right will be called \( A_{x,p} \) and \( B_{x,p} \), respectively. We need to study the fraction in \( V \) that occurs in \( A_{x,p} \):

**Lemma 6.3** For nonzero \( p \), the quantity

\[
\frac{pV^{p+1} - (p + 1) V^p + 1}{(V^p - 1)^2}
\]

decreases in \( V > 1 \) (and tends to \( \frac{1+1/p}{2} \) as \( V \to 1 + 0 \)).

A standard exercise in derivatives and infinitesimals.

### 6.3.1 Fixed \( \phi \), variable \( \beta_1 \).

**Proof of Lemma 6.2** (fixed \( \phi \)). Note that \( \beta_2 = -3X + 2Y, \beta_3 = 2X - Y \); also, \( V'/V = \frac{V - 1}{V_{\beta_2\beta_3} X} 2X (Y - 3X) \) and \( U'/U = V_{\beta_2\beta_3} \left( \frac{V}{V - 1} \right) \left( 1 - \frac{Y}{X} \right) \left( \frac{V'}{V} \right) \), so

\[
\frac{U'}{U} - \frac{V'}{V} = \frac{V'}{V (V - 1)} \left[ \left( V - \frac{Y}{X} V \right) - (V - 1) \right].
\]
That $R$ is positive results from the fact that $(U/V)$ is increasing and $V$ decreasing (in $\beta_1$).

We now show that $R$ can be expressed as a function of $V$ alone.

**Proposition 6.4** For constant $\phi$,

$$R = \frac{3V^2 - 1 + 3\sqrt{V(V-1)}}{3V+1},$$

and this function is concave in $V > 1$.

**Proof.** Indeed, $Y/X = (R+1)/V$ and $V-1 = (2\frac{Y}{X} - 3)^2 / \left[ 3 \left( 2 - \frac{Y}{X} \right) \frac{Y}{X} \right]$, whence a second-degree relationship between $V$ and $R+1$ which can be solved in $R+1$:

$$R+1 = \frac{3V}{3V+1} \left[ V + 1 + \omega \sqrt{V(V-1)} \right]$$

where $\omega = \pm 1$. The coefficient of $\omega$ in $R-1/2$ is larger than the $\omega$-free term, while from the definition and (17), $R-1/2$ is seen to equal $1/6 Y/X \beta_2/\beta_3$ which is positive on $D_{\phi,\beta}$. Therefore, $\omega$ must be $+1$.

As for concavity, $R''$ is found to have the sign of $11V^2 - 30V + 3 - 16(V-1)\sqrt{V(V-1)}$, or, in terms of $\omega$:

$$11W^2 - 8W - 16 - 16W\sqrt{W(W+1)} < -5W^2 - 8W - 16 < 0.$$  

\[\Box\]

**Corollary 6.5** We have the following affine upper bound for $R$: whenever $V \geq V_{\min 2}$, and irrespective of the constant value of $\phi$,

$$R \leq a_1V + b_1,$$

where $a_1 = 2.4427$ and $b_1 = -1.8194$.

**Proof.** The curve is below its tangent at any point, but since we need better and better estimates as $V$ decreases, it is appropriate to choose the tangent at $V_{\min 2}$. This gives the stated values of the coefficients. \[\Box\]

Since $-V''/V^2 > 0$, in order to prove that $\lambda \phi_{\min} \geq A + B$, where $A = \sum \limits_{2 \leq 2p \leq x \leq x_{\max}} p\tilde{\kappa}_{x,p}A_{x,p}$ and $B = \sum \limits_{3 \leq 2p+1 \leq x \leq x_{\max}} p\tilde{\kappa}_{x,p}B_{x,p}$, and $\lambda \phi_{\min} = 3 \times 4.506 \times 0.525245 > 7.1$.

For $A$, we use Lemma 3.3:

$$A \leq \sum \limits_{2 \leq 2p \leq x \leq x_{\max}} p\tilde{\kappa}_{x,p} \frac{pV_{\min 2}^{p+1} - (p + 1) V_{\min 2}^p + 1}{(V_{\min 2}^p - 1)^2} \left[ 1 - \frac{1}{1 + \left( \frac{U}{V} \right)_{x-2p}^{x-2p}} \right],$$

26
or less than 1.894. As regards $B$, we keep only the positive terms. Using Corollary 6.5 and the inequality \( \frac{r}{(1+r)s} \leq \frac{1}{4s} \) which is valid for any pair \((r, s)\) of positive reals:

\[
B \leq \sum_{3 \leq 2p+1 \leq x \leq x_{\text{max}}} p(x - 2p) \tilde{k}_{x,p} \frac{aV - |b|}{4(V_p - 1)}.
\]

But,

\[
aV - |b| = |b| \frac{aV}{V - 1} \frac{V - 1}{1 + V + V^2 + ... + V^{p-1}}.
\]

Since \(a > |b|\), the homographic fraction on the r.h.s. decreases in \(V > 1\), as does the last fraction. So, our bound for \(B\) can be made independent of \(V \geq V_{\text{min2}}\) by evaluating it at \(V_{\text{min2}}\). This yields \(B \leq 4.2269\), so the sum \(A + B\) is less than 6.125. This closes the case of fixed \(\phi\), variable \(\beta_1\).

### 6.3.2 Fixed \(\beta_1\), variable \(\phi\).

**Proof of Lemma 6.2** (fixed \(\beta_1\)). Here we use \(Y\) and \(Z\). Observing that \(\beta_2 = -Y + 3Z\), \(\beta_3 = Y - 2Z\), we find \(V'/V = 6V^{-1}YZ (3Z - 2Y)\) and \(U'/U = -\frac{1}{1-V} + \frac{3Y}{\beta_2\beta_3} + \frac{V'}{V-1}V\), so that

\[
R = \frac{V(V - 1)}{V'(1 - \phi)} - \frac{3YV(V - 1)}{\beta_2\beta_3V'} - 1.
\]

As before, \(R\) must be positive, and since the first term is negative, it suffices to show that the sum of the remaining two has the expression stated above for \(S\). Remarking that

\[
V - 1 = \frac{\beta_2}{2} \left( \frac{1}{3\beta_3} - \frac{1}{3\phi - 3\beta_1} \right)
\]

and also \(V = \frac{(2Y - 3Z)^2}{3Y(Y - 2Z)}\), we obtain from the expression of \(V'/V\):

\[
-\frac{3YV(V - 1)}{\beta_2\beta_3V'} - 1 = \frac{Y^2V}{2Z(2Y - 3Z)} - 1 = \frac{1}{2} + \frac{Y}{Z} \frac{3Z - Y}{2} \left[ \frac{-1}{Y} + \frac{1}{3(Y - 2Z)} \right],
\]

and conclude using (17).

We now express \(S\) in terms of \(V\) alone:

**Proposition 6.6** For \(V > 1\) and \(V \neq 4/3\),

\[
S = \frac{1}{2} + \frac{3(V - 1)}{3V - 4} \left[ V - 2 + \sqrt{V(V - 1)} \right],
\]

a concave function of \(V > 1\) which does not actually have a singularity at \(V = 4/3\).

**Proof.** Since \(Y/Z = (S - 1/2)/(V - 1)\) and \(V = (3 - 2Y/Z)^2 / \left[ 3Y/2 \left( \frac{Y}{Z} - 2 \right) \right]\), we have a relationship between \(S_1 = (S - 1/2)\) and \(W = (V - 1)\) which is quadratic in each separately, \(0 = (3W - 1)S_1^2 - 6W(W - 1)S_1 - 9W^2\), which we solve in \(S_1\):

\[
S_1 = \frac{3W}{3W - 1} \left[ W - 1 + \omega \sqrt{W(W + 1)} \right].
\]
However, since $Z$ is constant, $S$ cannot present a singularity at $W = 1/3$, which rules out $\omega = -1$ in that vicinity, and by continuity for all $W$. (We can also, for $W \neq 1/3$, derive straight contradictions from $\omega = -1$, as done in the fixed $\phi$ case.)

As to concavity, for $0 < W \neq 1/3$, we easily check that

$$\frac{d^2 S}{dV^2} = \frac{-3(5W + 9)}{4(W + 1) \left[ 16W(W + 1)^2 + (11W^2 + 30W + 3) \sqrt{W(W + 1)} \right]} < 0;$$

for $W = 1/3$, this second derivative extends by continuity, hence $d^2 S/dV^2$ also exists and is negative there.

**Corollary 6.7** We have the following affine upper bound for $R$: whenever $V \geq V_{\min 2}$, and irrespective of the constant value of $\beta_1$,

$$R \leq a_1 V + b_1,$$

where $a_1 = 2.2377$ and $b_1 = -1.7173$.

**Proof.** The coefficients are those of the tangent to $S$ at $V = V_{\min 2}$. Note that again, we have $a_1 > |b_1|$.

Proving that (33) decreases in $\phi$ now boils down to showing that, with similar notations to the above:

$$\lambda \phi_{\min} \geq -\lambda \frac{V(V-1)}{V'} + A + B,$$

where $A = \sum_{2 \leq p \leq x \leq x_{\max}} \tilde{p} \tilde{\kappa}_{x,p} A_{x,p}$ and $B = \sum_{3 \leq 2p+1 \leq x \leq x_{\max}} \tilde{p} \tilde{\kappa}_{x,p} B_{x,p}$.

(34)

(of course, all derivatives are now understood to be in $\phi$ for constant $\beta_1$.)

For $A$, we again have the bound 1.894. For $B$, the same estimate again applies, *mutatis mutandis*, i.e. with $a_1$ and $b_1$ replacing $a$ and $b$ respectively. This gives

$$B \leq \sum_{3 \leq 2p+1 \leq x \leq x_{\max}} p(x - 2p) \tilde{\kappa}_{x,p} \frac{a_1 V_{\min 2} - |b_1|}{4(V_{\min 2} - 1)} < 3.643.$$ 

And finally, from the expression of $V'/V$ and $V = (\beta_1 + 6\phi - 3)^2 / (3\beta_3 Y)$:

$$-\frac{V(V-1)}{V'} = \frac{\beta_2 (\beta_1 + 6\phi - 3)}{18 (1 - \beta_1)} = \frac{1}{36 (1 - \beta_1)} (6 - 6\phi - 4\beta_1) (\beta_1 + 6\phi - 3)$$

which is maximized by equating the two factors on the right (with a constant sum), so that

$$-\lambda \frac{V(V-1)}{V'} \leq \lambda \frac{(1 - \beta_{\min})}{16} < 0.566.$$  

Bringing all this together, irrespective of the constant value of $\beta_1$, for the right-hand side of (34) we obtain the bound $1.894 + 3.643 + 0.566 = 6.103$, which is indeed less than 7.1.
6.4 The first equation, separate monotony.

Actually, as already stated, monotony does not always hold on the whole of $D_{\phi,\beta_1}$ for the first equation

$$0 = Eq_1 \equiv \tilde{K} - \lambda \phi - \sum_{2 \leq 2p < x \leq x_{\text{max}}} \tilde{\kappa}_{x,p} (x - 2p) \left[ 1 - \frac{1}{1 + \left( \frac{U}{\beta} \right)^{x-2p} (1 - V - \nu)} \right],$$  \hspace{1cm} (35)

at least not in $\beta_1$, nor is it strictly needed in order reliably to locate any solution of the system. We shall only prove that, with one variable kept fixed:

**Claim 1.** $Eq_1$ decreases from a positive to a negative value, then stays negative; i.e., the region where monotony may fail contains no solutions anyway. We write the fraction in (35) as $1 / (1 + E_{x,p})$.

6.4.1 Fixed $\phi$, variable $\beta_1$.

**Lemma 6.8** $E_{x,p}$ increases in $\beta_1$ for $x \geq 2p + 2$, while for $x = 2p + 1$ it increases at least for $0 \leq \beta_1 \leq \beta_1^*$ where $\beta_1^* = \sqrt{3} - 3 (\sqrt{3} - 1) \phi$.

**Proof.** Note that $\phi_{\text{min}} > 1/3$ and $\sqrt{3} - 3 (\sqrt{3} - 1) \phi_{\text{max}} > 0$, so that $\beta_1^*$ is indeed between 0 and 1. It is readily seen that $E_{4,1} = [27 (1 - \phi) \beta_3^2 / (\beta_1 + 6 \phi - 3)^3]$, so that (with $X$, $Y$, $Z$ as before) $\partial E_{4,1} / \partial \beta_1$ has the sign of $2/\beta_3 - 3 / (\beta_1 + 6 \phi - 3)$, or of $2 (2Y - 3Z) - 3 (Y - 2Z) = Y > 0$. Therefore, for $x \geq 2p + 2$, recalling that $U$ increases and $V$ decreases, $E_{x,p}$, which for $p \geq 1$ equals $\left( \frac{U}{\beta} \right)^{x-2p-2} E_{4,1} (1 + \frac{1}{V} + \frac{1}{V^2} + \ldots + \frac{1}{V^{p-1}})$, clearly increases. For $x = 2p + 1$, the derivative $\partial E_{3,1} / \partial \beta_1$ of $E_{3,1} = 27 (1 - \phi) \beta_2 \beta_3^3 / (\beta_1 + 6 \phi - 3)^4$ has the sign of $-2/\beta_2 + 2/\beta_3 - 4 / (\beta_1 + 6 \phi - 3)$, or of $(Y^2 / X^2 - 3)$. As $\beta_1$ increases from 0 to 1, $Y/X$ linearly decreases from $3\phi / (3\phi - 1)$ to 1, passing through $\sqrt{3}$ for $\beta_1 = \beta_1^*$. Therefore, $E_{3,1}$ increases for $0 \leq \beta_1 \leq \beta_1^*$, then decreases; and we see that for $x = 2p + 1$, $E_{x,p} = E_{3,1} (1 + \frac{1}{V} + \frac{1}{V^2} + \ldots + \frac{1}{V^{p-1}})$ of necessity increases in $\beta_1$ for $0 \leq \beta_1 \leq \beta_1^*$.

Now consider $Eq_1$ deprived from the terms such that $x = 2p + 1$; we call this $Eq_1^*$. Obviously $Eq_1 < Eq_1^*$, and from Lemma 6.8 $Eq_1^*$ decreases for $0 \leq \beta_1 \leq 1$. Thus, to prove Claim 1 it suffices to show that $Eq_1^*$ is negative at $\beta_1^*$. However, at $\beta_1^*$ we have $U = 3 (1 - \phi) / (3\phi - 1)$, $V = 2 / \sqrt{3}$, so that $U/V$ decreases in $\phi > 1/3$. Hence at $\beta_1^*$, whatever the fixed value of $\phi \in [\phi_{\text{min}}, \phi_{\text{max}}]$:

$$Eq_1^* \leq \tilde{K} - \lambda \phi_{\text{min}} - \sum_{4 \leq 2p + 2 \leq x \leq x_{\text{max}}} \tilde{\kappa}_{x,p} (x - 2p) \left[ 1 - \frac{1}{1 + \left( \frac{3\sqrt{3}}{2} - \phi_{\text{max}} \right)^{x-2p} \left( 1 - \left( \frac{\sqrt{3}}{2} \right)^p \right)} \right]$$

which is less than $-0.157 < 0$. 

29
6.4.2 Fixed $\beta_1$, variable $\phi$.

Lemma 6.9 $E_{x,p}$ increases in $\phi$ for $x \geq 2p + 3$ (provided $\beta_1 \geq \beta_{1\text{min}}$), while for $2p + 1 \leq x \leq 2p + 2$ it increases at least for $0 \leq \phi \leq \phi^*$, where

$$\phi^* = \frac{1}{12} \left[ 15 - \frac{9}{2 - \beta_1} - 2\beta_1 - \frac{\sqrt{3(1 - \beta_1)} \sqrt{4\beta_1^2 + 4\beta_1 + 3}}{2 - \beta_1} \right].$$

This value decreases from $3/4$ to $1/3$ as $\beta_1$ increases from $0$ to $1$.

Proof. As before, $E_{x,p}$ increases whenever $E_{x-2p+2,1}$ does, because $E_{x,p} = E_{x-2p+2,1} \times (1 + \frac{1}{V} + \frac{1}{V^2} + \ldots + \frac{1}{V^{x-1}})$, and since $U/V$ is increasing, $E_{x,p}$ increases whenever $E_{x_0,p}$ does for some $x_0 < x$. Therefore all we have to show is that $E_{5,1}$ increases, and that for $0 \leq \phi \leq \phi^*$ so does $E_{3,1}$. Regarding the former, $E_{5,1} = 3^9 (1 - \phi)^3 \beta_3^6 / [\beta_2 (\beta_1 + 6\phi - 3)^8]$ has a derivative $\partial E_{5,1} / \partial \phi$ with the same sign as $-3 / (1 - \phi) + 3/\beta_2 + 18/\beta_3 - 48 / (\beta_1 + 6\phi - 3)$, or as $2X^2 - 7ZX + Z^2 (7 - Z)$. However, the discriminant $Z^2 (8Z - 7) = (1 - \beta_2)^2 (1 - 8\beta_1)$ of this quadratic function of $X$ remains negative so long as $\beta_1 \geq \beta_{1\text{min}}$, hence the required monotony. Coming now to $\partial E_{3,1} / \partial \phi$, it has the sign of $-1 / (1 - \phi) - 3/\beta_2 + 6/\beta_3 - 24 / (\beta_1 + 6\phi - 3)$, or of $2 (1 + Z) X^2 - Z (11 + 2Z) X + Z^2 (11 - Z)$. This equals zero for

$$X = \frac{Z}{4} \left[ 2 + \frac{1}{Z + 1} (9 + \omega \sqrt{3}4Z^2 - 12Z + 11) \right]$$

(36)

where $\omega = \pm 1$; however, recalling that $\phi$ cannot exceed $1 - 2\beta_1/3 = (1 + 2Z)/3$ lest $\beta_2$ should become negative, we see that $X - 2Z$ cannot be $> 0$, while if $\omega$ were $+$, $X - 2Z$ would be the product of $Z / (Z + 1) by a strictly decreasing function of $Z$ reaching 0 for $Z = 1$. Hence $\omega = +1$, and $\phi^*$ is then read from (36). The last assertion is straightforward.

Now consider $E_{q_1}$ deprived from the terms such that $2p + 1 \leq x \leq 2p + 2$; we call this $E_{q_1}^*$. Obviously $E_{q_1} < E_{q_1}^*$, and from Lemma 6.3 $E_{q_1}^*$ decreases for $0 \leq \phi \leq 1$. Thus, to prove Claim 1 it suffices to show that for any $\beta_1 \in [\beta_{1\text{min}}, \beta_{1\text{max}}]$, $E_{q_1}^*$ is negative at $\phi^*$. Call the corresponding values of $U$ and $V$, as functions of $\beta_1$, $U^*$ and $V^*$ respectively. In a moment, we will show that $U^*/V^*$ and $V^*$ behave like their unstared, fixed-$\phi$ counterparts, i.e., the first increases and the second decreases in $\beta_1$. Then, an upper bound for $E_{q_1}^*$ in some interval $[\beta_{1L}, \beta_{1H}]$ is given by $M [\beta_{1L}, \beta_{1H}]$ defined as:

$$\tilde{K} - \lambda \phi^* [\beta_{1H}] - \sum_{5 \leq 2p + 3 \leq x_{\text{max}}} \tilde{K}_{x,p} (x - 2p) \left[ 1 - \frac{1}{1 + \left( \frac{U^*}{V^*} \right)^{x_{\text{max}} - 2p} \left( 1 - \left( \frac{1}{\sqrt{\lambda \phi^* [\beta_{1H}]}} \right)^p \right)} \right].$$

Straightforward numerical calculation yields (still, of course, for $x_{\text{max}} = 56$)

$$M [.33, .39] < -.051, \quad M [.39, .428] < -.051,$$

$$M [.428, .468] < -.062, \quad M [.468, .529] < -.055,$$

establishing our final point. Now as promised:
Lemma 6.10 $V^*$ is a decreasing, $U^*/V^*$ an increasing function of $\beta_1 \in [0, 1]$.

Proof. Set $A = 9 - \sqrt{3} \sqrt{3 + 4\beta_1 + 4\beta_1^2}$, $B = 4 - 2\beta_1$, so that $3 \leq A \leq 6$, $2 \leq B \leq 4$, and

$$V^* = \frac{4}{3(A-B)(A+3B)} A^2, \quad U^* = \frac{3}{8(1-\beta_1)(3B-A)} (A-B)^2 \left[3(1+2\beta_1) + (9-A)\right].$$

We take derivatives w.r.t. $\beta_1$, noting that $A' = -6(1+2\beta_1)/(9-A)$. First, $V''$ has the sign of $AB' - A'B = 2/(9-A) \left[30\beta_1 + 21 - 9\sqrt{3 + 4\beta_1 + 4\beta_1^2}\right] < 0$. As for $U^*/V^*$, the factor in square brackets on the right is increasing, so it suffices to show that each of $(A-B)/(1-\beta_1)$ and $(A-B)/(3B-A)$ increases too. The derivative of the latter has the sign of $A'B - AB'$, positive as we have just seen; the derivative of the former is $[A'(1-\beta_1) + A - 2]/(1-\beta_1)^2$, so has the sign of $-6(1+2\beta_1)(1-\beta_1)+(A-2)(9-A) = 7\sqrt{3 + 4\beta_1 + 4\beta_1^2} - 15 - 18\beta_1 > 0$. ■

This ends the proof of Proposition 6.1. Although it can be made slightly more precise and then used to show the existence, uniqueness, and globally decreasing character of the implicit functions defined by $E_1$ and $E_2$, we will not do so, since we do not need to. We could actually remove the word ‘strict’ and still proceed to the final subsection.

6.5 Root localization.

Using Proposition 6.1, we shall show that any (feasible) common root $(\phi^*, \beta_1^*)$ to $E_1$ and $E_2$ must lie in a small rectangle $\mathcal{R} = [\phi_-, \phi_+] \times [\beta_{1-}, \beta_{1+}]$ and that on the whole of $\mathcal{R}$, our bound (31) for the expectation is strictly less than 1. Since we already know the (global) maximizer for (22) to exist and to necessarily give rise to such a common root for which, besides, (31) will be valid, this will show our chosen value of $c$, 4.506, to be above the threshold without even the need for a direct proof of either existence or uniqueness of such a $(\phi^*, \beta_1^*)$.

We determine $\mathcal{R}$ explicitly, together with four numerical sequences which witness to the fact that no solution can lie outside $\mathcal{R}$, owing to Corollary 6.13 below. This amounts, in a rigorous presentation, to a very elementary numerical procedure which starts at a corner of the rectangle $[\phi_{min}, \phi_{max}] \times [\beta_{1min}, \beta_{1max}]$ containing $\mathcal{D}_{\phi, \beta_1}$, and spirals its way towards a solution.

From Proposition 6.1 follows

Lemma 6.11 Let $E$ be either $E_1$ or $E_2$, $A = (\phi_A, \beta_{1A})$, $B = (\phi_B, \beta_{1B})$. (i) If $E(A) > 0$ and $B \leq A$ (meaning $\phi_B \leq \phi_A$ and $\beta_{1B} \leq \beta_{1A}$), then $E(B) > 0$; (ii) If $E(A) < 0$ and $A \leq B$, then $E(B) < 0$.

Proof. Do it in two steps, changing one coordinate at a time; for (i), use monotony; for (ii), use monotony if $E = E_2$, and if $E = E_1$ use the fact that if $E_1$ is negative, then it stays so if a single coordinate is increased. ■
This in turn implies

**Proposition 6.12** Let \( A = (\phi_0, \beta_{1,A}) \) and \( B = (\phi_0, \beta_{1,B}) \) have the same \( \phi \)-coordinate, while \( C = (\phi_C, \beta_{1,0}) \) and \( D = (\phi_D, \beta_{1,0}) \) have the same \( \beta_1 \)-coordinate.

(i) If \( \beta_{1,B} < \beta_{1,A} \), \( Eq_1(A) > 0 \) and \( Eq_2(B) < 0 \), then the closed rectangle \([\phi_{\min}, \phi_{\max}] \times [\beta_{1,B}, \beta_{1,A}]\) contains no common root to \( Eq_1 \) and \( Eq_2 \).

(ii) If \( \beta_{1,A} < \beta_{1,B} \), \( Eq_1(A) < 0 \) and \( Eq_2(B) > 0 \), then the closed rectangle \([\phi_{\min}, \phi_{\max}] \times [\beta_{1,A}, \beta_{1,B}]\) contains no common root to \( Eq_1 \) and \( Eq_2 \).

(iii) If \( \phi_C < \phi_D \), \( Eq_2(C) < 0 \) and \( Eq_1(D) > 0 \), then the closed rectangle \([\phi_C, \phi_D] \times [\beta_{1\min}, \beta_{1\max}]\) contains no common root to \( Eq_1 \) and \( Eq_2 \).

(iv) If \( \beta_D < \phi_C \), \( Eq_2(C) > 0 \) and \( Eq_1(D) < 0 \), then the closed rectangle \([\phi_D, \phi_C] \times [\beta_{1\min}, \beta_{1\max}]\) contains no common root to \( Eq_1 \) and \( Eq_2 \).

**Proof.** (i) Let \( P = (\phi, \beta_1) \) be an arbitrary point of the rectangle. We show that if \( \phi \leq \phi_0 \) then \( P \) is not a solution of \( Eq_1 \), while if \( \phi \geq \phi_0 \), \( P \) fails to satisfy \( Eq_2 \). Indeed, in the former case we have \( P \leq A \), so we use (i) of Lemma 6.11 with \( Eq = Eq_1 \); in the latter, \( B \leq P \), so we apply (ii) of the same lemma with \( Eq = Eq_2 \).

(ii), (iii) and (iv) Very similar (or actually the same up to notation).

As an obvious corollary, we obtain the final link leading to the main result of this paper:

**Corollary 6.13** Let four finite sequences \( \hat{\phi}_0^1 < \hat{\phi}_1^- < \ldots < \hat{\phi}_K^- \), \( \beta_{1,0}^+ > \beta_{1,0}^- > \ldots > \beta_{1,K}^- \), \( \phi_0^+ > \phi_1^+ > \ldots > \phi_L^+ \), \( \beta_{1,0}^- < \beta_{1,1}^- < \ldots < \beta_{1,L}^- \) with \( \phi_0^- = \phi_{\min} \), \( \beta_{1,0}^+ = \beta_{1\max} \), \( \phi_0^+ = \phi_{\max} \) and \( \beta_{1,1}^- = \beta_{1\min} \) verify:

\[
\begin{align*}
Eq_1 (\phi_i^-, \beta_{1,i}^+) & > 0, \quad 0 \leq i \leq K, \\
Eq_2 (\phi_i^+, \beta_{1,i+1}^-) & < 0, \quad 0 \leq i \leq K - 1, \\
Eq_1 (\phi_j^+, \beta_{1,j}^-) & < 0, \quad 0 \leq j \leq L, \\
Eq_2 (\phi_j^-, \beta_{1,j+1}^+) & > 0, \quad 0 \leq j \leq L - 1.
\end{align*}
\]

Then no feasible common solution to \( Eq_1 \) and \( Eq_2 \) can lie outside the rectangle \([\hat{\phi}_K^-, \phi_L^+] \times [\beta_{1,L}^-, \beta_{1,K}^+]\).

**Proof.** Successive applications of Proposition 6.12 (i) with \( A = (\hat{\phi}_i^-, \beta_{1,i}^+) \) and \( B = (\hat{\phi}_{i+1}^-, \beta_{1,i+1}^-) \) exclude the band \([\phi_{\min}, \phi_{\max}] \times [\beta_{1,K+1}^-, \beta_{1\max}]\). We similarly exclude \([\phi_{\min}, \phi_{\max}] \times [\beta_{1\min}, \beta_{1,L+1}^-], [\phi_0^+, \phi_K^-] \times [\beta_{1\min}, \beta_{1\max}]\), and \([\phi_L^+, \phi_{\max}] \times [\beta_{1\min}, \beta_{1\max}]\). ■

All that remains to do is explicitly to give our four sequences as above, and to check that the hypotheses of the corollary obtain and that the bound \((31)\) is uniformly strictly less than one on the whole rectangle \( R = [\hat{\phi}_K^-, \phi_L^+] \times [\beta_{1,L}^-, \beta_{1,K}^-] \) (the bound being independent of sufficiently large \( n \)).

We first compute the product \( G_1 (\varepsilon, x_{\max}) G_2 (\varepsilon, x_{\max}) \exp(2\varepsilon D/e) \) appearing in \((31)\) to be less than \( 1 + 10^{-7} \).
We then determine sequences \( \phi^-_i \) and \( \beta^+_i \) as above, with \( K = 62 \), and sequences \( \phi^+_i \) and \( \beta^-_i \) with \( L = 52 \) satisfying the requirements of Corollary 6.13 and such that \( \mathcal{R} = [\phi^-_K, \phi^+_L] \times [\beta^-_{1,L}, \beta^+_{1,K}] = [0.56383217, 0.56383249] \times [0.44651403, 0.44651478] \). Taking into account the monotony properties of \( U, V \), and of the functions \( x \mapsto x^x, x \mapsto x^y \) and \( x \mapsto y^x \), an upper bound for the right-hand side of (31) throughout \( \mathcal{R} \) is seen to be the product of \( (6n^3)^{1/n} \) by

\[
(1 + 10^{-7}) \times 3^c \left( \frac{\lambda}{6e} \right)^\lambda \prod_{0 \leq 2p \leq x_{\text{max}}} \left[ \frac{2 \sum_{x > 2p} \kappa_{x,p}}{p!(x-p)!(\kappa_{x,p})} \right] \times \\
\frac{\tilde{U}(\lambda\phi^+_L - K)}{(V-1)^{\beta^-_{1,K} + c}} \prod_{0 \leq 2p \leq x_{\text{max}}} \left[ \tilde{V}^{x-p} + \tilde{V}^{x-2p} \left( \tilde{V}^{p} - 1 \right) \tilde{\kappa}_{x,p} \right] \times \\
\left[ \frac{(3\phi^+_L - \beta^-_{1,L})^{3\phi^+_L - \beta^-_{1,L} + 1}}{\beta^-_{2} \left( 3\beta^-_{3} \right)^{\beta^-_{3}}} \right]^c,
\]

where \( \tilde{U} = U (\phi^+_L, \beta^+_{1,K}), \tilde{V} = V (\phi^-_K, \beta^-_{1,L}), \tilde{U} = U (\phi^-_K, \beta^-_{1,L}), \tilde{V} = V (\phi^+_L, \beta^+_{1,K}), \tilde{\beta}_2 = 3 (1 - \phi^+_L) - 2\beta^+_1, \) and \( \tilde{\beta}_3 = \beta^-_{1,L} - 2 + 3\phi^-_K \). The product of this bound by \( 2^{n+c+\Delta} < 1 + 10^{-14} \) is computed to be \( < 0.9999885 \). So, for \( c = 4.506, x_{\text{max}} = 56, \) and \( \varepsilon = 10^{-15} \), the product \( 2^{(n+c+\Delta)E(X_{n,c,x_{\text{max}},c})} \) is less than \( 6n^3 \times 0.9999885^n \), and we conclude using Proposition 1.2 and the decreasing character of \( \text{Pr}_{n,c}(\text{SAT}) \) as a function of \( c \).

\textbf{Acknowledgment.} We are especially grateful to Jean Bretagnolle for his generous help.

\textbf{References}

[1] D. Achlioptas, Setting 2 variables at a time yields a new lower bound for random 3-SAT, in Proceeding of the 32nd ACM Symposium on Theory of Computing, Association for Computing Machinery, New York, 2000, pp. 28-37

[2] D. Achlioptas and G. B. Sorkin, Optimal myopic algorithms for random 3-SAT, in Proc. of the 41st Annual Symposium on foundations of Computer Science, 2000, pp. 590-600.

[3] N. Alon and J. Spencer. The Probabilistic Method. Wiley, 1992.

[4] B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, D. B. Wilson, the scaling window of the 2-SAT transition, Random Structures and Algorithms 18 (2001), no 3, pp. 201-256.

[5] A. Z. Broder, A. M. Frieze and E. Upfal, On the satisfiability and maximum satisfiability of random 3-CNF formulas, in Proceeding of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms (Austin, TX, 1993), ACM, New York, 1993, pp. 322-330.

[6] Y. Boufkhad and O. Dubois, Length of prime implicants and number of solutions of random CNF formulae, Theoretical Computer Science, 215 (1999), 1-30
[7] M. T. Chao and J. Franco, Probabilistic analysis of a generalization of the unit-clause literal selection heuristics for the \( k \)-satisfiability problem, Inform. Sci. 51 (1990), no. 4, pp. 289-314.

[8] V. Chvátal and B. Reed, Micks gets some (the odds are on his side), in Proceedings of the 33th Annual Symposium on Foundations of Computer Science (Pittsburgh, PA, 1992), IEEE Comput. Soc. Press, Los Alamitos, CA, 1992, pp. 620-627.

[9] S. Cocco and R. Monasson, Trajectories in phase diagrams, growth processes and computational complexity: how search algorithms solve the 3-Satisfiability problem, Phys. Rev. Lett. 86 (2001), pp. 1654-1658.

[10] N. Creignou and H. Daudé, Satisfiability threshold for random XOR-CNF formulae, Discrete Applied Mathematics 96-97 (1999), no 1-3, pp. 41-53.

[11] O. Dubois and Y. Boufkhad, A general upper bound for the satisfiability threshold of random \( r \)-SAT formulae, J. Algorithms 24 (1997), pp. 395-420.

[12] O. Dubois, Y. Boufkhad, and J. Mandler, Typical random 3-SAT formulae and the satisfiability threshold, in Proceedings of the Eleventh ACM-SIAM Symposium on Discrete Algorithms, San Francisco, California, January 2000, pp. 124-126.

[13] O. Dubois and G. Dequen, A Backbone Search Heuristic for Efficient Solving of Hard 3-SAT Formulae, In Proc. of the 17th Internat. Joint Conf. on Artificial Intelligence, Seattle, 2001, pp. 248-253

[14] O. Dubois and J. Mandler, The 3-XORSAT Threshold, In Proc. of the 43rd Annual IEEE Symposium on Foundations of Computer Science, Vancouver November 2002.

[15] O. Dubois and J. Mandler, On the non 3-colourability of random graphs, preprint.

[16] O. Dubois, R. Monasson, B. Selman, R. Zecchina, Phase Transition in Combinatorial Poblems, Theoretical Computer Science 265 (2001), no 1-2, pp. 1-308.

[17] W. Feller, An Introduction to Probability Theory and its Applications, v. 1, 3rd ed., Wiley, 1973.

[18] N. Fountoulakis and C. McDiarmid, Upper bounds on the non-3-colourability threshold of random graphs, Discrete Mathematics and Theoretical Computer Science 5 (2002), pp. 205-226.

[19] J. Franco and M. Paull, Probabilistic analysis of the Davis-Putnam procedure for solving the satisfiability problem, Discrete Applied Mathematics 5 (1983), pp. 77-87.

[20] Ehud Friedgut, Necessary and sufficient conditions for sharp thresholds of graph properties and the \( k \)-SAT problem, J. Amer. Math. Soc. 12 (1999), pp. 1017-1054.

[21] A. M. Frieze and S. Suen, Analysis of two simple heuristics on a random instance of the \( k \)-SAT problem, J. Algorithms 20 (1996), no. 2, pp. 312-355.
[22] I. P. Gent and T. Walsh, The SAT Phase Transition, In Proc. of the 11th European Conf. on Artificial Intelligence, Chichester, 1994, pp. 105-109.

[23] M. Habib, C. McDiarmid, J. Ramirez and B. Reed (ed.), Probabilistic Methods for Algorithmic Discrete Mathematics, Springer-Verlag, 1998.

[24] S. Janson, Y. C. Stamatiou, M. Vamvakari, Bounding the unsatisfiability threshold of random 3-SAT, Random Structures and Algorithms 17 (2000), no. 2, pp. 79-102.

[25] Erratum, Random Structures and Algorithms, 18 (2001), no 1, pp. 99-100.

[26] O. Kallenberg, Foundations of Modern Probability. Springer-Verlag, 1998.

[27] A. Kamath, R. Motwani, K. Palem and P. Spirakis, Tail bounds for occupancy and the satisfiability threshold conjecture, Random Structures and Algorithms 7 (1995), no 1, pp 59-80.

[28] A. C. Kaporis, L. M. Kirousis, Y. C. Stamatiou, M. Vamvakari, M. Zito - Coupon Collectors, q-Binomial Coefficients and the Unsatisfiability Threshold, In Proc. of the 7th Italian Conference on Theoretical Computer Science, Torino, October 2001, pp. 328-338.

[29] A. C. Kaporis, L. M. Kirousis, E. G. Lalas, The Probabilistic Analysis of a Greedy Satisfiability Algorithm, ESA 2002, pp. 574-585

[30] L. M. Kirousis, E. Kranakis, D. Krizanc and Y. C. Stamatiou, Approximating the unsatisfiability threshold of random formulae, Random Structures and Algorithms 12 (1998), no. 3, pp. 253-269.

[31] L. Kirousis, E. Kranakis, and D. Krizanc, An Upper Bound for a Basic Hypergeometric Series, Carleton University, School of Computer Science, Technical Report TR-96-07, 1996.

[32] L.M. Kirousis, E. Kranakis, D. Krizanc, Approximating the Unsatisfiability Threshold of Random Formulas, in Proceedings of ESA, Barcelona 1996, pp. 27-38.

[33] M. Loève, Probability Theory I (4th Edition), Springer, 1977.

[34] D.G. Luenberger, Linear and Nonlinear Programming, 2nd ed., Addison-Wesley, Reading, 1984.

[35] A. El Maftouhi, W Fernandez de la Vega, On random 3-SAT, Combin. Probab. Comput. 4 (1995), no 3, pp. 189-195

[36] S. Mertens, Phase Transition in the Number Partitioning Problem, Phys. Rev. Lett. 81 (1998), no 20, pp. 4281-4284

[37] D. Mitchell, B. Selman and H. J. Levesque, Hard and easy distribution of SAT problems, In Proc. of the 10th Nat. Conf. on Artificial Intelligence, San Jose, July 1992, pp. 459-465.
Appendix A

Proof of Lemma A.1 (As stated, what we prove is actually stronger.) In the ordered-clauses model, if the number of occurrences of variable \( i \) is \( K_i \), the random vector \( (K_i)_{1 \leq i \leq n} \) follows a multinomial law of parameters \( \lambda n, p_1 = \ldots = p_n = 1/n \), where \( \lambda = 3c \). Also, the number of positive occurrences of variable \( i \) is modelled by the r.v. \( S_i = \sum_{j=1}^{K_i} X_{i,j} \), with \( X_{i,j} \) i.i.d. \( B(1, 1/2) \) coin tosses, constructed to be independent of the multinomial vector (cf. Th. 2.19 of [26]). For \( x \in \mathbb{N} \), \( S_{i,x} = \sum_{j=1}^{x} X_{i,j} \) has a binomial \( B(x, 1/2) \) distribution:

\[
\Pr (S_i = p | K_i = x) = \Pr (S_{i,x} = p) = \frac{1}{2^x} \binom{x}{p}
\]  

(37)

The number of variables having \( x \) occurrences is \( N_x = \sum_{i=1}^{n} 1_{\{K_i = x\}} \), while those having \( x \) occurrences out of which \( p \) are positive is \( W_{x,p} = \sum_{i=1}^{n} 1_{\{K_i = x, S_i = p\}} \).

We use the large deviation property of binomial r.v.s in the following form: Define \( h(q,t) = (q + t) \log (1 + t/q) + (1 - q - t) \log (1 - t/(1 - q)) \) if \( t \leq 1 - q \), \( +\infty \) otherwise; and \( c(q,t) = \min \{h(q,t), h(1 - q,t)\} \). Let \( Y \) be the sum of \( n \) independent indicator variables with common expectation \( q \); then for any \( \varepsilon > 0 \),

\[
\Pr \left( \left| \frac{Y}{n} - q \right| \geq \varepsilon \right) \leq 2e^{-c(q,\varepsilon)n}.
\]  

(38)

We also use the fact that a Poisson r.v. with integral mean \( \mu \) cannot have too small a probability of equalling \( \mu \): there is, as can be seen from a variant of Stirling’s formula, an absolute constant \( C_0 > 0 \) such that if \( Z \) is Poisson with integer parameter \( \mu \geq 1 \), then

\[
\Pr (Z = \mu) \geq \frac{C_0}{\sqrt{\mu}}.
\]  

(39)

Recall that the Poisson probability mass function, \( e^{-\lambda} \lambda^x/x! \), is denoted by \( p(x, \lambda) \).

Now consider a Poisson r.v. \( M \) with mean \( \lambda n \), and construct (e.g., Lemma 5.9 in [26]) a random vector \((L_i), 1 \leq i \leq n\) that is multinomially distributed conditionally on \( M \), i.e.:

\[
\Pr ((L_i) = (l_i) | M = m') = \binom{m'}{m} \frac{l_1^{m'} l_2 \ldots l_n}{n^{m'}}.
\]

Probabilities and expectations in the Poissonized model will be subscripted with a \( \lambda \). In particular,

\[
\Pr_{\lambda} ((L_i) = (l_i) | M = \lambda n) = \Pr ((K_i) = (l_i)).
\]  

(40)
The law of the vector \((L_i)\) is obtained by deconditioning, giving a sum with just one non-vanishing term:

\[
\Pr_\lambda ((L_i) = (l_i)) = \sum_{m'} \Pr_\lambda ((L_i) = (l_i) | M = m') \Pr_\lambda (M = m')
\]

\[
= \frac{(\sum l_i)!}{\prod l_i!} n^{-\sum l_i} \times p \left( \sum l_i, \lambda n \right) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{l_i}}{l_i!}.
\]

So (summing w.r.t. all coordinates but one), the \(L_i\) are independent, each being Poisson with mean \(\lambda\). We let \(X'_{i,j}\) be i.i.d. coin tosses in the Poissonized model which are also (completely) independent of the vector \((L_i, M)\), so that in this model, the ‘number of occurrences of variable \(i\)’ is \(S'_i = \sum_{j=1}^{L_i} X'_{i,j}\). We also consider, for \(x \in \mathbb{N}\), \(S'_{i,x} = \sum_{j=1}^{x} X'_{i,j}\), which has a binomial \(B(x, 1/2)\) distribution; on account of our independence hypotheses

\[
\Pr_\lambda (S'_i = p | L_i = x) = \Pr_\lambda (S'_{i,x} = p | L_i = x) = \Pr_\lambda (S'_{i,x} = p) = \frac{1}{2^x} \binom{x}{p}. \tag{41}
\]

In terms of the indicators \(U_i (x) = 1_{\{L_i = x\}}, V_i (p) = 1_{\{S'_i = p\}}\), and \(W'_{i,x,p} = 1_{\{L_i = x, S'_i = p\}} = U_i (x) V_i (p)\), the ‘number of variables with \(x\) occurrences, \(p\) among them positive’ is \(W'_{x,p} = \sum_{i=1}^n W'_{i,x,p}\). We will need the following lemma, to be proved later.

**Lemma A.1** In the setup just defined, the law of \(W'_{x,p}\), conditional on \(M = \lambda n\), is the same as the law of \(W_{x,p}\).

Clearly then, by (41):

\[
E_\lambda W'_{i,x,p} = \Pr_\lambda (S'_i = p | L_i = x) \Pr_\lambda (L_i = x) = \frac{1}{2^x} \binom{x}{p} \Pr_\lambda (x, \lambda) \equiv \kappa_{x,p}.
\]

By (38) we have:

\[
\Pr_\lambda \left( \left| \frac{W'_{x,p}}{n} - \kappa_{x,p} \right| \geq \varepsilon \right) \leq 2e^{-c(\varepsilon, \kappa_{x,p})n}.
\]

We now depoissonize, i.e. we decompose w.r.t. the values of \(M\):

\[
2e^{-c(\varepsilon, \kappa_{x,p})n} \geq \sum_{m'} \Pr_\lambda \left( \left| \frac{W'_{x,p}}{n} - \kappa_{x,p} \right| \geq \varepsilon \ | M = m' \right) \Pr_\lambda (M = m')
\]

\[
= \sum_{m'} \Pr_\lambda \left( \left| \frac{W'_{x,p}}{n} - \kappa_{x,p} \right| \geq \varepsilon \ | M = \lambda n \right) \Pr_\lambda (M = \lambda n).
\]

By Lemma (A.1) and (39), this implies

\[
\Pr \left( \left| \frac{W_{x,p}}{n} - \kappa_{x,p} \right| \geq \varepsilon \right) \leq \frac{2}{C_0} \sqrt{\lambda n} e^{-c(\varepsilon, \kappa_{x,p})n},
\]

which is stronger than Lemma (3.1).
Proof of lemma [A.1] The law of a random vector determines that of the sum of its components, and the same holds for the conditional laws relative to some event (e.g., [33], p. 218, end of § 14).

So, it is sufficient to show that the law of the random vector \((W'_{i,\text{exp}})_{1 \leq i \leq n}\), conditional on \(M = \lambda n\), is the same as the law of \((W_{i,\text{exp}})_{1 \leq i \leq n}\), where \(W_{i,\text{exp}} = 1_{\{K_i = x, S_i = p\}} = 1_{\{K_i = x\}} 1_{\{S_i = p\}}\); and this, in turn, will follow if we show that the conditional law of the \(2n\)-dimensional random vector \((L_i, S'_{i,\text{exp}})_{1 \leq i \leq n}\) is the same as the law of \((K_i, S_i)_{1 \leq i \leq n}\). Now, for any \(N^n\)-valued vectors \((l_i)_{1 \leq i \leq n}\) and \((p_i)_{1 \leq i \leq n}\),

\[
\Pr_{\lambda}( (L_i = l_i, S'_{i,\text{exp}} = p_i) | M = \lambda n ) = \Pr_{\lambda}( (L_i = l_i, S'_{i,\text{exp}} = p_i) | M = \lambda n )
\]

Here the event \(A = \cap_{1 \leq i \leq n} \{ S'_{i,\text{exp}} = p_i \} \) is independent of the conjunction \(B \cap C\), with \(B = \cap_{1 \leq i \leq n} \{ L_i = l_i \}\) and \(C = \{ M = \lambda n \}\), so \(\Pr_{\lambda}(A | B \cap C) = \Pr_{\lambda}(A)\). Applying the generally-valid

\[
P(A \cap B | C) = P(A | B \cap C) P(B | C),
\]

and using (40), we see that

\[
\Pr_{\lambda}( (L_i = l_i, S'_{i,\text{exp}} = p_i) | M = \lambda n ) = \Pr_{\lambda}( (S'_{i,\text{exp}} = p_i) ) \Pr_{\lambda}( (K_i = l_i) ) .
\]

(42)

Although the \(K_i\) are not independent, our setup does ensure that the events \(\cap_{1 \leq i \leq n} \{ S_{i,\text{exp}} = p_i \}\) and \(\cap_{1 \leq i \leq n} \{ K_i = l_i \}\) are independent, so

\[
\Pr( (K_i = l_i, S_{i,\text{exp}} = p_i) ) = \Pr( (S_{i,\text{exp}} = p_i) ) \Pr( (K_i = l_i) ) .
\]

(43)

But, by (11) and (37), the first factors on the right in (42) and (43) are both equal to \(\prod_{i=1}^{n} 2^{-l_i(\frac{l_i}{p_i})}\). 

\[\square\]