Locally consistent constraint satisfaction problems

Zdeněk Dvořák Daniel Král' Ondřej Pangrác *

Abstract

An instance of a constraint satisfaction problem is l-consistent if any l constraints of it can be simultaneously satisfied. For a fixed constraint type P, $\rho_l(P)$ denotes the largest ratio of constraints which can be satisfied in any l-consistent instance. In this paper, we study locally consistent constraint satisfaction problems for constraints which are Boolean predicates. We determine the values of $\rho_l(P)$ for all l and all Boolean predicates which have a certain natural property which we call 1-extendibility as well as for all Boolean predicates of arity at most three. All our results hold for both the unweighted and weighted versions of the problem.

1 Introduction

Constraint satisfaction problems form an important abstract computational model for a lot of problems arising in practice. This is witnessed by an enormous recent interest in the computational complexity of various constraint satisfaction problems [2, 3, 4, 13]. However, some instances of real problems do not require all the constraints to be satisfied but it is enough to satisfy a large fraction of them. In order to maximize this fraction, the input can be usually pruned at the beginning by removing small sets of contradictory constraints. In this paper, we study for a fixed constraint type how large fraction of the constraints can be simultaneously satisfied if no l constraints are

*Department of Applied Mathematics and Institute for Theoretical Computer Science (ITI) Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic, E-mail: {rakdver,kral,pangrac}@kam.mff.cuni.cz

contradictory. Formally, an instance of the constraint satisfaction problem is *l-consistent* if any l constraints of it can be simultaneously satisfied.

This problem was first introduced and studied by Trevisan [11]. He showed that for each fixed $k \geq 2$ and each $l \geq 2$, if we allow as constraints all Boolean predicates of arity k, then there exist l-consistent problems in which the fraction of constraints which can be simultaneously satisfied does not exceed 2^{1-k} and the bound is tight. In the upper bound, he used only a single type of predicate (the predicate $P(x_1, \ldots, x_k) = (x_1 \not\Leftrightarrow x_2 \Leftrightarrow x_3 \Leftrightarrow \cdots \Leftrightarrow x_k))$; his lower bound was based on a simple probabilistic argument similar to that used by Yannakakis [14] for locally consistent CNF formulas.

The variant of the problem for locally consistent CNF formulas is extremely well-studied as witnessed by a separate section (20.6) devoted to this concept in a recent monograph on extremal combinatorics by Jukna [7]. A CNF formula Φ is *l-consistent* if any l clauses of Φ can be satisfied. Formulas which are l-consistent are also called *l-satisfiable*. The number ρ_l^{SAT} denotes the largest fraction of clauses which can be satisfied in any l-consistent CNF formula. Clearly, $\rho_1^{\mathrm{SAT}} = 1/2$. The value $\rho_2^{\mathrm{SAT}} = \frac{\sqrt{5}-1}{2} \approx 0.6180$ was determined by Lieberherr and Specker [9]. They consequently established $\rho_3^{\mathrm{SAT}} = 2/3$ [10]. Later, Yannakakis [14] simplified proofs of both the lower bounds on ρ_2^{SAT} and ρ_3^{SAT} using a probabilistic argument. The value $\rho_4^{\mathrm{SAT}} \approx 0.6992$ has been recently computed by one of the authors [8]. Huang and Lieberherr [6] studied the asymptotic behavior and they proved $\lim\limits_{l\to\infty} \rho_l^{\mathrm{SAT}} \leq 3/4$. The limit was settled by Trevisan [11] who showed $\lim\limits_{l\to\infty} \rho_l^{\mathrm{SAT}} = 3/4$. Let us remark that the cases of 1, 2 and 3-consistent CNF formulas somewhat unexpectedly differ from the case of l-consistent formulas for $l \geq 4$. First, $\rho_l^{\mathrm{SAT}} = \rho_l^{\mathrm{2SAT}}$ for $l = 1, 2, 3$ but $\rho_4^{\mathrm{SAT}} < \rho_4^{\mathrm{2SAT}}$ where ρ_l^{2SAT} is the largest fraction of clauses which can be satisfied in any l-consistent 2-CNF formula, i.e., CNF formulas with clauses of sizes at most two. We suspect the inequality to be strict for all $l \geq 4$, i.e., $\rho_l^{\mathrm{SAT}} < \rho_l^{\mathrm{2SAT}}$ for all $l \geq 4$. Second, the values ρ_l^{SAT} for $l = 1, 2, 3$ coincide with the similar values defined for a "fractional" version of the problem (which are known for all $l \geq 1$ [8] and are equal to so-called Usiskin's numbers [12]) but the value ρ_4^{SAT} differs.

In the present paper, we study the more general problem of locally consistent constraint satisfaction problems. We restrict our attention to problems whose constraints are copies of a single Boolean predicate P. The arguments of the predicates can be both positive and negative literals. Similarly as in

$\sigma(P)$	P	$l=1$	$l \geq 2$
1	x	1/2	1
1	$x \wedge y$	1/4	1
2	$x \Leftrightarrow y$	1/2	
3	$x \vee y$	3/4	

Table 1: The values $\rho_l(P)$ for all non-isomorphic essentially unary and binary Boolean predicates.

the case of CNF formulas, $\rho_l(P)$ denotes the largest possible fraction of constraints which can be simultaneously satisfied in any l-consistent instance. We determine the values of $\rho_l(P)$ for all $l \geq 1$ and all Boolean predicates P of arity at most three (see Tables 1 and 2) and for all Boolean predicates which are 1-extendable. A predicate P is said to be 1-*extendable* if it has the following property: If we fix one of its arguments, we can choose the remaining ones in such a way that the predicate is satisfied. Let us point out a somewhat exceptional case of the predicate $P(x, y, z) = x \wedge (y \vee z)$ which is not 1-extendable (fix x to be false). Therefore, our general Theorem 1 does not apply. In Section 5, we show for this predicate that $\rho_1(P) = 3/8$, $\rho_2(P) = 2\sqrt{3}/9$ and somewhat surprisingly that $\rho_l(P) = \rho_{l-2}^{2SAT}$ for all $l \geq 3$. Since the values ρ_l^{2SAT} were exactly computed before only for $l = 1, 2, 3$, we have to prove a special result on structure of locally consistent 2-CNF formulas (Lemma 8) which is later used in the analysis of the predicate $P(x, y, z) = x \wedge (y \vee z)$ and which also yields a formula for ρ_l^{2SAT} (Corollary 1). Let us remark that all our results hold both for the unweighted and weighted versions of the studied problems, i.e., the instances witnessing the upper bounds contain each constraint at most once and our lower bound proofs translate smoothly for instances with weighted constraints. From the algorithmic point of view, our results can be interpreted in the following way: The simplest probabilistic algorithms (of the kind used in [8, 11, 14]) are approximation algorithms for locally consistent CSPs with optimum worst-case performance.

$\sigma(P)$	P	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$	$l\geq 6$	$l\to\infty$
1	$x \wedge y \wedge z$	1/8	1	1	1	1	1	1
2	$x \Leftrightarrow y \Leftrightarrow z$	1/4	1/4	1/4	1/4	1/4	1/4	1/4
	$x \wedge (y \Leftrightarrow z)$	1/4	8/27	1/2	1/2	1/2	1/2	1/2
3	exactly one	3/8	3/8	3/8	3/8	3/8	3/8	3/8
	$x \wedge (y \vee z)$	3/8	$\frac{2\sqrt{3}}{9}$	1/2	$\frac{\sqrt{5}-1}{2}$	2/3	$\rho^{\mathrm{2SAT}}_{l-2}$	3/4
	$(x \Leftrightarrow y) \wedge (x \Rightarrow z)$	3/8	3/8	3/8	3/8	3/8	3/8	3/8
4	$x \Rightarrow y \Rightarrow z$	1/2	1/2	1/2	1/2	1/2	1/2	1/2
	$(x \wedge y) \Leftrightarrow z$	1/2	1/2	1/2	1/2	1/2	1/2	1/2
	at most one	1/2	1/2	1/2	1/2	1/2	1/2	1/2
	one or three	1/2	1/2	1/2	1/2	1/2	1/2	1/2
5	$\neg$ exactly one	5/8	5/8	5/8	5/8	5/8	5/8	5/8
	$x \vee (y \wedge z)$	5/8	5/8	5/8	5/8	5/8	5/8	5/8
	$(x \Leftrightarrow y) \vee (x \wedge z)$	5/8	5/8	5/8	5/8	5/8	5/8	5/8
6	$\neg(x \Leftrightarrow y \Leftrightarrow z)$	3/4	3/4	3/4	3/4	3/4	3/4	3/4
	$x \vee (y \Leftrightarrow z)$	3/4	3/4	3/4	3/4	3/4	3/4	3/4
7	$x \vee y \vee z$	7/8	7/8	7/8	7/8	7/8	7/8	7/8

Table 2: The values $\rho_l(P)$ for all non-isomorphic essentially ternary Boolean predicates.

2 Preliminaries

In this paper, we mainly deal with constraints which are Boolean predicates and we prefer to call them *predicates* to emphasize their kind. If P is a Boolean predicate, $\sigma(P)$ denotes the number of combinations of arguments which satisfy P. If the constraint satisfaction problem consists of copies of a single constraint P, its instances are called *P-systems*. The arguments of the predicates may be both positive and negative literals, but a single variable cannot be contained in two distinct arguments of the same predicate. The goal is to find a truth assignment which satisfies the largest number of the predicates. If Σ is a P-system, then $\rho(\Sigma)$ is the largest fraction of the predicates of Σ which can be simultaneously satisfied. Hence, $\rho_l(P) = \inf \rho(\Sigma)$ where the infimum is taken over all l-consistent P-systems Σ.

Two Boolean predicates P and P' are *isomorphic* if they differ by permutation of the arguments and negations of some of them, e.g., if $P(x_1, x_2) = P'(x_2, \neg x_1)$, then the predicates P and P' are isomorphic. Clearly, if P and P' are two isomorphic predicates, then $\rho_l(P) = \rho_l(P')$ for all $l \geq 1$. A k-ary predicate is *essentially k-ary* if it depends on all its k arguments. If the predicate P is not essentially k-ary, it is isomorphic to a predicate P' such that $P'(x_1, \ldots, x_k) = P''(x_1, \ldots, x_{k-1})$ for some $(k-1)$-ary Boolean predicate P''. It is not hard to see that $\rho_l(P) = \rho_l(P') = \rho_l(P'')$ for all $l \geq 1$ in such case. Hence, in order to determine $\rho_l(P)$ for all unary, binary and ternary Boolean predicates P, it is enough to compute the values for representatives of isomorphism classes of essentially unary, binary and ternary Boolean predicates.

We conclude this section by stating three simple observations on locally consistent systems of Boolean predicates:

Lemma 1 *Let P be a k-ary Boolean predicate P. It holds that $\rho_l(P) \geq \sigma(P)/2^k$ for all $l \geq 1$.*

Proof: Let us consider a P-system Σ with N predicates and with n variables $x_1, \ldots, x_n$. Choose each of the variables x_i, $1 \leq i \leq n$, randomly and independently to be true with the probability $1/2$. Each predicate of the system Σ is satisfied by the constructed random truth assignment with probability $\sigma(P)/2^k$. Hence, the expected number of satisfied predicates is $N \cdot \sigma(P)/2^k$. Consequently, there is a truth assignment which satisfies at least $N \cdot \sigma(P)/2^k$ of Σ predicates and $\rho(\Sigma) \geq \sigma(P)/2^k$.

□

Lemma 2 *It holds that $\rho_1(P) = \sigma(P)/2^k$ for each k-ary predicate P.*

Proof: By Lemma 1, $\rho_1(P) \geq \sigma(P)/2^k$. We construct a P-system Σ with variables $x_1, \ldots, x_k$ and with $\rho(\Sigma) = \sigma(P)/2^k$. It is enough to set Σ to be the set consisting of all $P(x_1^{a_1}, \ldots, x_k^{a_k})$ where $(a_1, \ldots, a_k) \in \{0,1\}^k$ and x_i^0 is $\neg x_i$ and x_i^1 is x_i. Clearly, each truth assignment satisfies exactly $\sigma(P)$ predicates out of all the 2^k predicates of Σ. Therefore, $\rho(\Sigma) = \sigma(P)/2^k$.

□

Lemma 3 *Let P be a k-ary predicate with $\sigma(P) = 1$. Then, $\rho_1(P) = 2^{-k}$ and $\rho_l(P) = 1$ for every $l \geq 2$.*

Proof: The equality $\rho_1(P) = 2^{-k}$ follows from Lemma 2. Let us consider a 2-consistent P-system Σ. Since $\sigma(P) = 1$, each predicate of Σ forces the values to all its arguments. However, all the predicates must force the same value to a single variable because Σ is 2-consistent. Therefore, the "forced" truth assignment satisfies all the predicates of Σ and $\rho(\Sigma) = 1$. This immediately yields that $\rho_l(P) = 1$ for every $l \geq 2$.

□

3 1-extendable Boolean predicates

In this section, we present an upper bound on $\rho_l(P)$ which holds for all 1-extendable Boolean predicates. First, we introduce several concepts which are used throughout this section. The *dependence graph* $G(\Sigma)$ of a P-system Σ is the multigraph whose vertices are predicates of Σ and the number of edges between two predicates p_1 and p_2 of Σ is equal to the number of variables which appear in arguments of both the predicates p_1 and p_2 (regardless whether they appear as positive or negative literals). The *girth* of a P-system Σ is the length of the shortest cycle contained in $G(\Sigma)$. In particular, if the girth of Σ is three or more, then $G(\Sigma)$ contains no parallel edges. The following lemma relates the girth of a P-system with its consistency:

Lemma 4 *Let P be a 1-extendable predicate and Σ a P-system. If the girth of Σ is at least $l \geq 3$, then Σ is $(l-1)$-consistent.*

Proof: We prove by induction on i that any $i = 1, \ldots, l-1$ predicates of Σ can be simultaneously satisfied. The claim trivially holds for $i = 1$. Assume now that $i > 1$ and let $p_1, \ldots, p_i$ be any i predicates of Σ. Since the girth of Σ is greater than i, the vertices corresponding to $p_1, \ldots, p_i$ induce a forest in $G(\Sigma)$. We can assume without loss of generality that p_i is a leaf or an isolated vertex in $G(\Sigma)$. Let $x_1, \ldots, x_n$ be variables contained in the first $i-1$ predicates. By the induction hypothesis, there is a truth assignment for the variables $x_1, \ldots, x_n$ which satisfies all the predicates $p_1, \ldots, p_{i-1}$. Since p_i is a leaf or an isolated vertex in $G(\Sigma)$, it has at most one variable in common with the predicates $p_1, \ldots, p_{i-1}$. Hence, the truth assignment for $x_1, \ldots, x_n$ can be extended to a truth assignment which satisfies all the predicates $p_1, \ldots, p_i$ because P is 1-extendable.

□

Let us recall now Chernoff's inequality [5]:

Lemma 5 *Let X be a random variable equal to the sum of N zero-one independent random variables such that each of them is equal to* 1 *with probability p. Then, the following holds for every $\delta > 0$:*

$$\mathrm{Prob}(X \geq (1+\delta)pN) \leq e^{-\frac{\delta^2 pN}{3}} \quad \textit{and} \quad \mathrm{Prob}(X \leq (1-\delta)pN) \leq e^{-\frac{\delta^2 pN}{2}}.$$

We are now ready to determine the values $\rho_l(P)$ for all $l \geq 1$ and all 1-extendable Boolean predicates P. Note that the proof of Theorem 1 generalizes the standard construction of "random" graphs with large girth.

Theorem 1 *Let P be a k-ary Boolean predicate which is* 1*-extendable. Then, $\rho_l(P) = \sigma(P)/2^k$ for all $l \geq 1$.*

Proof: If $l = 1$, the statement follows from Lemma 2. Fix a k-ary 1-extendable Boolean predicate P and an integer $l \geq 2$. By Lemma 1, $\rho_l(P) \geq \sigma(P)/2^k$. For each $\varepsilon > 0$, we construct an l-consistent P-system Σ with $\rho(\Sigma) \leq (1+\varepsilon)\sigma(P)/2^k$. This will yield the equality $\rho_l(P) = \sigma(P)/2^k$.

Let us consider a positive real $\delta > 0$ whose exact value is chosen later. Let $n \geq 2k$ be an integer which will also be chosen later. We construct an l-consistent P-system Σ with variables $x_1, \ldots, x_n$. Let S_n be the set of all possible predicates P with variables $x_1, \ldots, x_n$; the number of predicates contained in S_n is $N = 2^k n!/(n-k)!$. Note that $N \geq n^k$ because $n \geq 2k$. Construct a P-system Σ_0 from S_n by including each predicate of S_n to Σ_0 randomly and independently with probability $p = n^{-(k-1)+1/2l}$. By Lemma 5, the probability that the number $|\Sigma_0|$ of predicates of Σ_0 is smaller than $(1-\delta)pN$ is at most the following:

$$\mathrm{Prob}(|\Sigma_0| \leq (1-\delta)pN) \leq e^{-\frac{\delta^2 pN}{2}} \leq e^{-\frac{\delta^2 n^{-(k-1)+1/2l} n^k}{2}} \leq e^{-\frac{\delta^2 n^{1+1/2l}}{2}} \tag{1}$$

Observe that $G(S_n)$ contains at most $2^{k\lambda}k^{2\lambda}n^{(k-1)\lambda}$ cycles of length λ. Thus, the expected number of cycles of length at most l in $G(\Sigma_0)$ does not exceed:

$$\sum_{\lambda=2}^{l} 2^{k\lambda}k^{2\lambda}n^{(k-1)\lambda}p^{\lambda} \leq 2^{kl}k^{2l}\sum_{\lambda=2}^{l} n^{(k-1)\lambda}n^{-\lambda(k-1)+\lambda/2l} \leq$$

$$2^{kl}k^{2l}\sum_{\lambda=2}^{l} n^{\lambda/2l} \leq 2^{kl}k^{2l}\sum_{\lambda=2}^{l} n^{1/2} \leq 2^{kl}k^{2l}ln^{1/2}.$$

By Markov's inequality, the number of cycles of length at most l in $G(\Sigma_0)$ is smaller or equal to $2 \cdot 2^{kl} k^{2l} l n^{1/2}$ with probability at least 1/2.

Each truth assignment for the variables $x_1, \ldots, x_n$ satisfies $\sigma(P)n!/(n-k)! = \sigma(P)2^{-k}N$ predicates of S_n. Fix any out of all 2^n truth assignments for the rest of this paragraph. We now bound the probability that the number of predicates satisfied by the fixed truth assignment exceeds $(1+\delta)\sigma(P)2^{-k}pN$. Since the expected number of such predicates is $\sigma(P)2^{-k}pN$, Lemma 5 implies:

$$\text{Prob}(\#\text{ satisfied clauses} \geq (1+\delta)\sigma(P)2^{-k}pN) \leq e^{-\frac{\delta^2\sigma(P)2^{-k}pN}{3}} \leq$$
$$e^{-\frac{\delta^2\sigma(P)2^{-k}n^{-(k-1)+1/2l}n^k}{3}} = e^{-\frac{\delta^2\sigma(P)2^{-k}n^{1+1/2l}}{3}}.$$

Since there are 2^n truth assignments, the probability that there is a satisfying assignment which satisfies more than $(1+\delta)\sigma(P)2^{-k}pN$ predicates is at most:

$$2^n e^{-\frac{\delta^2\sigma(P)2^{-k}n^{1+1/2l}}{3}} = e^{(\ln 2)\cdot n - \frac{\delta^2\sigma(P)2^{-k}}{3}\cdot n^{1+1/2l}} \tag{2}$$

We now choose the integer n to be any (large enough) integer that both the upper bounds (1) and (2) are smaller than 1/4 and the last inequality in (4) below holds. By (1) and (2), the random P-system Σ_0 has the following three properties with positive probability:

- Σ_0 contains at least $(1-\delta)pN$ predicates.
- $G(\Sigma_0)$ contains at most $2 \cdot 2^{kl} k^{2l} l n^{1/2}$ cycles of length at most l.
- There is no truth assignment satisfying more than $(1+\delta)\sigma(P)2^{-k}pN$ predicates of Σ_0.

Fix a P-system Σ_0 which has these three properties. The desired P-system Σ is obtained from Σ_0 by removing all the predicates contained in all cycles of $G(\Sigma_0)$ whose length is at most l. Hence, $G(\Sigma)$ contains no cycle of length at most l and its girth is at least $l+1$. Since P is 1-extendable, Σ is l-consistent by Lemma 4. In addition, the first two properties of Σ_0 imply that Σ contains at least $(1-\delta)pN - (2^{kl+1}k^{2l}ln^{1/2}) \cdot l$ predicates. On the other hand, the third property yields that no truth assignment can satisfy more than $(1+\delta)\sigma(P)2^{-k}pN$ predicates of Σ. Hence:

$$\rho(\Sigma) \leq \frac{(1+\delta)\sigma(P)2^{-k}pN}{(1-\delta)pN - 2^{kl+1}k^{2l}l^2n^{1/2}} \leq \frac{(1+\delta)\sigma(P)}{(1-\delta)2^k - \frac{2^{kl+k+1}k^{2l}l^2n^{1/2}}{pN}} \tag{3}$$

Observe that the following holds (the last inequality follows from the choice of n):

$$\frac{2^{kl+k+1}k^{2l}l^2n^{1/2}}{pN} \leq \frac{2^{kl+k+1}k^{2l}l^2n^{1/2}}{n^{-(k-1)+1/2l}n^k} = \frac{2^{kl+k+1}k^{2l}l^2}{n^{1/2+1/2l}} \leq \delta 2^k \tag{4}$$

The inequalities (3) and (4) yield the following:

$$\rho(\Sigma) \leq \frac{(1+\delta)\sigma(P)}{(1-\delta)2^k - \delta 2^k} = \frac{1+\delta}{1-2\delta} \cdot \frac{\sigma(P)}{2^k}.$$

Note that for each $\varepsilon > 0$, we can choose $\delta > 0$ so that $\frac{1+\delta}{1-2\delta} \leq 1+\varepsilon$. Thus, for such δ, the obtained l-consistent P-system Σ satisfies that $\rho(\Sigma) \leq (1+\varepsilon)\sigma(P)/2^k$ as desired.

□

4 2-CNF formulas

In this section, we study structure of 2-CNF formulas, i.e., CNF formulas of clauses of sizes one and two. We first recall a well-known lemma about unsatisfiable formulas with clauses of sizes two which can be found, e.g., in [1]. If Φ is a 2-CNF formula with variables $x_1, \ldots, x_n$, then $G(\Phi)$ denotes the directed graph of order $2n$ whose vertices correspond to literals $x_1, \ldots, x_n$ and $\neg x_1, \ldots, \neg x_n$ and whose edge set is the following: For each clause $(a \vee b)$, $G(\varphi)$ contains an arc from the literal $\neg a$ to the literal b and an arc from $\neg b$ to a (note that both a and b represent literals, not variables). For each clause (a) (which can also be viewed as a clause $(a \vee a)$), we include an arc from the literal $\neg a$ to a.

Lemma 6 *Let Φ be a 2-CNF formula with variables $x_1, \ldots, x_n$. Then, the formula Φ is satisfiable if and only if $G(\Phi)$ contains no directed cycle through both the vertices x_i and $\neg x_i$ for any i, $1 \leq i \leq n$.*

An immediate corollary of Lemma 6 is the following:

Lemma 7 *Each minimal inconsistent set of clauses of a 2-CNF contains at most two clauses of size one.*

We now show that there exist extremal 2-CNF formulas in which each small inconsistent set of clauses contains two clauses of sizes one:

Lemma 8 *Let $2 \le l \le L$ be any two integers. For each $\varepsilon > 0$, there exists a 2-CNF l-consistent formula Φ with $\rho(\Phi) \le \rho_l^{2SAT} + \varepsilon$ such that each inconsistent set of L clauses contains at least two clauses of size one. Moreover, Φ contains each single clause of size two at most once.*

Proof: Fix integers $l \ge 2$ and $L \ge l$ for the rest of the proof. Similarly, $\delta < 1$ is a positive real which will be chosen at the end of the proof. Fix an l-consistent formula Φ_0 with $\rho(\Phi_0) \le \rho_l^{2SAT}(1+\delta)$. We now classify the variables contained in the formula Φ_0: The set A_1 is formed by variables x contained in a clause of size one; we can assume without loss of generality that each variable $x \in A_1$ appears as a positive literal in the clause of size one. The set A_i, $2 \le i \le \lfloor l/2 \rfloor$, consists of variables x which are not contained in any A_j, $1 \le j \le i-1$, and which are contained in a clause of the form $(\neg y \vee x)$ for $y \in A_{i-1}$. Since Φ is l-consistent, we can assume that all the occurrence of $x \in A_i$ in the clauses $(\neg y \vee x)$, $y \in A_{i-1}$, are positive: Otherwise, there would be a set of at most i clauses which force x to be true as well as a set of at most i clauses which force x to be false. The union of these two sets of clauses consists of at most $2i$ clauses and it is clearly inconsistent. Since $i \le \lfloor l/2 \rfloor$, this is impossible. Finally, let A_0 be the set of the remaining variables of Φ.

Let w_{ij}, $w_{\bar{i}j}$ and $w_{\bar{i}\bar{j}}$ be the number (sum of the weights) of the clauses of the type $(x \vee y)$, $(\neg x \vee y)$ and $(\neg x \vee \neg y)$, respectively, where $x \in A_i$ and $y \in A_j$. Similarly, let w_1 be the number (sum of the weights) of the clauses of the type (x) where $x \in A_1$. We may assume that $w_1 > 0$. Otherwise, $\rho(\Phi_0) \ge 3/4$ and we can set Φ to be an L-consistent P-system Σ with $P(x,y) = (x \vee y)$ with $\rho(\Sigma) \le 3/4 + \varepsilon$ constructed in Theorem 1. Finally, W denotes the sum of all w_1, w_{ij}, $w_{\bar{i}j}$ and $w_{\bar{i}\bar{j}}$ for $0 \le i,j \le \lfloor l/2 \rfloor$. By the definition of the sets $A_1, \ldots, A_{\lfloor l/2 \rfloor}$, $w_{\bar{i}j} = 0$ for all $1 \le i,j \le \lfloor l/2 \rfloor$ with $i+1 < j$. In addition, since Φ is l-consistent, $w_{\bar{i}\bar{j}} = 0$ for all $1 \le i,j \le \lfloor l/2 \rfloor$ with $i+j+1 \le l$.

We now define W_p to be the maximum of the sum:

$$w_1 p_1 + \sum_{0 \le i \le j \le \lfloor l/2 \rfloor} w_{ij}(p_i + p_j - p_i p_j) + w_{\bar{i}\bar{j}}(1 - p_i p_j) + \sum_{0 \le i,j \le \lfloor l/2 \rfloor} w_{\bar{i}j}(1 - p_i + p_i p_j) \tag{5}$$

where the maximum is taken over all $0 \le p_0, \ldots, p_{\lfloor l/2 \rfloor} \le 1$. Clearly, $W_p/W \le \rho(\Phi_0)$: Consider the probabilities $p_0, \ldots, p_{\lfloor l/2 \rfloor}$ for which the maximum in the above expression is attained. If each of the variables of the set A_i, $0 \leq i \leq \lfloor l/2 \rfloor$, is chosen to be true randomly and independently with the probability p_i, then the expected number (weight) of the satisfied clauses is W_p. Therefore, there is a truth assignment which satisfies at least this number of clauses and consequently $W_p/W \leq \rho(\Phi_0)$.

Let n be an integer which we fix later. Let X_i, $0 \leq i \leq \lfloor l/2 \rfloor$, be $\lfloor l/2 \rfloor + 1$ disjoint sets consisting of n variables each. We construct a 2-CNF formula Φ with variables $X_0 \cup \cdots \cup X_{\lfloor l/2 \rfloor}$. The formula Φ contains $n^{1/2L}$ copies of a clause (x) for each $x \in X_1$. The other clauses are included to the formula Φ randomly and independently as follows: The clauses $(x \vee y)$, $(\neg x \vee y)$ and $(\neg x \vee \neg y)$ where $x \in X_i$ and $y \in X_j$ with $i \neq j$ are included to Φ with the probabilities $w_{ij} n^{-1+1/2L}/w_1$, $w_{\bar{i}j} n^{-1+1/2L}/w_1$ and $w_{\overline{ij}} n^{-1+1/2L}/w_1$, respectively. The clauses $(x \vee y)$, $(\neg x \vee y)$ and $(\neg x \vee \neg y)$ where $x, y \in X_i$ are included to Φ with the probabilities $2w_{ii} n^{-1+1/2L}/w_1$, $w_{\bar{i}i} n^{-1+1/2L}/w_1$ and $2w_{\overline{ii}} n^{-1+1/2L}/w_1$, respectively.

We claim that the number of clauses of Φ is at least $Wn^{1+1/2L}(1-\delta)/w_1$ and the number of clauses which can be simultaneously satisfied does not exceed $W_p n^{1+1/2L}(1+\delta)/w_1 + 3Wn^{1+1/2L}\delta/w_1$ with the probability which tends to 1 as n goes to infinity. We show that each of the complementary events, i.e., the "bad" events, for each separate types of clauses occurs with the probability which tends to 0. Since the number of bad events is finite (and independent of n), this yields the claim. As an example, we present the analysis only for a single type of clauses, e.g., clauses $(x \vee y)$ for $x \in X_i$ and $y \in X_j$ with $i \neq j$ for fixed integers i and j. Namely, we aim to show that the number of clauses of this type is smaller than $w_{ij} n^{1+1/2L}(1-\delta)/w_1$ with probability tending to 0. In addition, the probability that there is a truth assignment which assigns the true value to a fraction p_i, p_j, of the variables of X_i, X_j, respectively, and which satisfies more than $w_{ij} n^{1+1/2L}(1-(1-p_i)(1-p_j))(1+\delta)/w_1 + 3w_{ij} n^{1+1/2L}\delta/w_1 = w_{ij} n^{1+1/2L}(p_i + p_j - p_i p_j)(1+\delta)/w_1 + 3w_{ij} n^{1+1/2L}\delta/w_1$ clauses of the considered type also tends to zero.

By Lemma 5, the probability that the number of clauses $(x \vee y)$ with $x \in X_i$ and $y \in X_j$ is smaller than $w_{ij} n^{1+1/2L}(1-\delta)/w_1$ is at most:

$$e^{-\frac{\delta^2 w_{ij} n^{1+1/2L}/w_1}{3}} = e^{-\Theta(n^{1+1/2L})} \to 0.$$

The second part of the statement is more difficult. We first prove the claim for p_i and p_j where p_i or p_j is at least δ. Fix now a truth assignment for $X_i \cup X_j$ which assigns the true value to a fraction p_i, p_j, of the variables X_i, X_j, respectively. By Lemma 5, the probability that the number of clauses of the considered type satisfied by this fixed truth assignment exceeds $w_{ij}n^{1+1/2L}(p_i+p_j-p_ip_j)(1+\delta)/w_1$ is at most:

$$e^{-\frac{\delta^2\cdot w_{ij}n^{-1+1/2L}/w_1\cdot n^2(p_i+p_j-p_ip_j)}{3}}\le e^{-\frac{\delta^3\cdot w_{ij}n^{1+1/2L}/w_1}{3}}=e^{-\Theta(n^{1+1/2L})}.$$

Since there are at most 2^{2n} possible truth assignments the probability that there exists a truth assignment with $\delta\le\max\{p_i,p_j\}$, which satisfy more clauses than claimed is at most $2^ne^{-\Theta(n^{1+1/2L})}\to 0$. We now show that if there is no "bad" truth assignment with $\delta\le\max\{p_i,p_j\}$, then there is no "bad" truth assignment with $0\le p_i,p_j\le\delta$. Consider a truth assignment which assigns the true value to at most δn variables of each X_i and X_j and modify it to a truth assignment which assigns the true value to $\lceil\delta n\rceil$ variables of each X_i and X_j. This modification can only increase the number of satisfied clauses of the considered type. Since both the modified p_i and p_j are now larger than δ, the assignment satisfies at most $w_{ij}n^{1+1/2L}(2\delta+2/n-\delta^2)(1+\delta)/w_1$ clauses (the additional factor $2/n$ comes from rounding). If n is sufficiently large, then this expression is at most $w_{ij}n^{1+1/2L}3\delta/w_1$ as desired. This finishes the proof of the claim.

Next, we bound the expected number of minimal inconsistent sets of at most L clauses containing zero or one clause of size one. The number of variables of the formula Φ is $N=(\lfloor l/2\rfloor+1)n$. By Lemma 6, there are at most $(2N)^k$ minimal inconsistent sets of k clauses such that the size of each clause is two and there are at most $(2N)^{k-1}$ minimal inconsistent sets of $k-1$ clauses such that the size of each clause is two except for precisely one clause whose size is one. We omit a straightforward but little technical argument yielding these upper bounds. Since each clause of size two is included to Φ with the probability at most $W_pn^{-1+1/2L}/w_1$, the expected number of minimal inconsistent sets of at most L clauses containing zero or one clause of size one is at most the following:

$$\sum_{k=1}^{L}(2N)^kW_p^kn^{-k+k/2L}/w_1^k+(2N)^{k-1}W_p^{k-1}n^{-(k-1)+(k-1)/2L}/w_1^{k-1}=$$

$$\sum_{k=1}^{L}(2(\lfloor l/2\rfloor+1))^kW_p^kn^{k/2L}/w_1^k+(2(\lfloor l/2\rfloor+1))^{k-1}W_p^{k-1}n^{(k-1)/2L}/w_1^{k-1}\le$$

$$\sum_{k=1}^{L}(l+2)^kW_p^kn^{k/2L}/w_1^k+(l+2)^{k-1}W_p^{k-1}n^{(k-1)/2L}/w_1^{k-1}\le$$

$$2L(l+2)^L W_p^L n^{1/2}/w_1^L.$$

By Markov's inequality, the probability that there are more than $4L(l+2)^L W_p^L n^{1/2}/w_1^L$ minimal inconsistent sets of at most L clauses with zero or one clause of size one is at most $1/2$. Therefore, if n is sufficiently large (with respect to a previously fixed $\delta > 0$), with positive probability, the random formula Φ has at least $Wn^{1+1/2L}(1-\delta)/w_1$ clauses, at most $W_p n^{1+1/2L}(1+\delta)/w_1 + 3Wn^{1+1/2L}\delta/w_1$ clauses of Φ can be simultaneously satisfied and Φ contains at most $4L(l+2)^L W_p^L n^{1/2}/w_1^L$ inconsistent sets of at most L clauses with no or a single clause of size one. Fix such a formula Φ. We obtain Φ' from Φ by removing all (at most $4L^2(l+2)^L W_p^L n^{1/2}/w_1^L$) clauses of size two contained in an inconsistent set of at most L clauses with no or a single clause of size one. This may decrease the number of clauses of Φ by at most $4L^2(l+2)^L W_p^L n^{1/2}/w_1^L$. On the other hand, the number of clauses which can be simultaneously satisfied cannot increase.

We now estimate $\rho(\Phi')$ (observe that $W_p \geq W/2$):

$$\rho(\Phi') \leq \frac{W_p n^{1+1/2L}(1+\delta)/w_1 + 3Wn^{1+1/2L}\delta/w_1}{Wn^{1+1/2L}(1-\delta)/w_1 - 4L^2(l+2)^L W_p^L n^{1/2}/w_1^L} =$$

$$\frac{W_p(1+\delta) + 3\delta W}{W(1-\delta) - 4L^2(l+2)^L W_p^L n^{-1/2-1/2L}/w_1^{L-1}} \leq \frac{W_p(1+7\delta)}{W(1-\delta) - O(n^{-1/2-1/2L})}.$$

Therefore, if n is sufficiently large, then:

$$\rho(\Phi') \leq \frac{W_p(1+7\delta)}{W(1-2\delta)} \leq \rho(\Phi_0)\frac{1+7\delta}{1-2\delta} \leq \rho_l^{2\mathrm{SAT}}\frac{(1+\delta)(1+7\delta)}{1-2\delta}.$$

Hence, for each $\varepsilon > 0$, we can choose $\delta > 0$ small enough that $\rho(\Phi') \leq \rho_l^{2\mathrm{SAT}} + \varepsilon$.

Finally, we have to show that the formula Φ' constructed in the proof is l-consistent. By Lemma 7, each minimal inconsistent set of clauses of Φ' contains at most two clauses of size one. On the other hand, each inconsistent set of at most L clauses contains at least two such clauses. Therefore, each minimal inconsistent set of at most l clauses of Φ' contains precisely two clauses of size one. Fix such a set Γ of clauses of Φ' and let (x_1) and (y_1) be the two clauses of size one contained in Γ. Obviously, $x_1, y_1 \in X_1$. By Lemma 6, Γ contains clauses of size two in which x_1 and y_1 appear as negative literals. By the construction of Φ', such clauses can be only $(\neg x_1 \vee x_2)$ and $(\neg y_1 \vee y_2)$ for some $x_2, y_2 \in X_2$. By Lemma 6, Γ has to contain clauses of size two in which x_2 and y_2 appear as negative literals. By the construction of Φ', such clauses can be only $(\neg x_2 \vee x_3)$ and $(\neg y_2 \vee y_3)$ for some $x_3, y_3 \in X_3$. In this way, we continue until we reach the set $X_{\lfloor l/2 \rfloor}$. By the minimality of the set Γ, $x_i \neq y_i$ for all $1 \leq i \leq \lfloor l/2 \rfloor$. Therefore, if $|\Gamma| \leq 2\lfloor l/2 \rfloor + 1$, then Γ contains the clauses $(x_1), (\neg x_1 \vee x_2), \ldots, (\neg x_{\lfloor l/2 \rfloor - 1} \vee x_{\lfloor l/2 \rfloor}), (y_1), (\neg y_1 \vee y_2), \ldots, (\neg y_{\lfloor l/2 \rfloor - 1} \vee y_{\lfloor l/2 \rfloor})$ and $(\neg x_{\lfloor l/2 \rfloor} \vee \neg y_{\lfloor l/2 \rfloor})$. If l is even, then $|\Gamma| > l$, and if l is odd, then $w_{\overline{\lfloor l/2 \rfloor}\,\overline{\lfloor l/2 \rfloor}} = 0$ and thus Φ' cannot contain the clause $(\neg x_{\lfloor l/2 \rfloor} \vee \neg y_{\lfloor l/2 \rfloor})$. In either of the cases, we showed that there is no inconsistent set of at most l clauses.

□

A close inspection of the proof of Lemma 8 yields that for any weights w_1, w_{ij}, $w_{\bar{i}j}$ and $w_{\overline{ij}}$ with $w_{\bar{i}j} = 0$ for all $1 \leq i \leq j-1$ and $w_{\overline{ij}} = 0$ for all $1 \leq i, j \leq \lfloor l/2 \rfloor$ with $i + j + 1 \leq l$. there is an l-consistent formula Φ with $\rho(\Phi) \leq W_p/W + \varepsilon$ where $W = w_1 + \sum_{i,j}(w_{ij} + w_{\bar{i}j} + w_{\overline{ij}})$ and W_p is the maximum of the sum (5) taken over all $0 \leq p_0, \ldots, p_{\lfloor l/2 \rfloor} \leq 1$. Therefore, we have the following formula for ρ_l^{2SAT} for all $l \geq 2$:

Corollary 1 *For each $l \geq 2$, the following holds:*

$$\rho_l^{\mathrm{2SAT}} = \min_{\substack{0 \leq w_1, w_{ij}, w_{\bar{i}j}, w_{\overline{ij}} \\ w_1 + \sum_{i,j}(w_{ij} + w_{\bar{i}j} + w_{\overline{ij}}) = 1}} W_p,$$

where the minimum is taken over all combinations of weights with $w_{\bar{i}j} = 0$ for all $1 \leq i \leq j-1$ and $w_{\overline{ij}} = 0$ for all $1 \leq i, j \leq \lfloor l/2 \rfloor$ with $i + j + 1 \leq l$ and W_p is the maximum of the sum (5) taken over all $0 \leq p_0, \ldots, p_{\lfloor l/2 \rfloor} \leq 1$.

5 Unary, binary and ternary Boolean predicates

As noted in Section 2, it is enough to determine the values $\rho_l(P)$ for representatives of isomorphism classes of essentially unary, binary and ternary Boolean predicates. The case of 1-extendable Boolean predicates was handled in Theorem 1. The only essentially unary, binary and ternary Boolean predicates which are not 1-extendable (upto isomorphism) are the following: $P(x) = x$, $P(x, y) = x \wedge y$, $P(x, y, z) = x \wedge y \wedge z$, $P(x, y, z) = x \wedge (y \Leftrightarrow z)$ and $P(x, y, z) = x \wedge (y \vee z)$. The first three of these predicates satisfy that

$\sigma(P) = 1$ and so the values $\rho_l(P)$ for these three predicates were determined in Lemma 3. Therefore, we know the values $\rho_l(P)$ for all essentially unary and binary Boolean predicates (see Tables 1 and 2). We focus on l-consistent P-systems with $P(x, y, z) = x \wedge (y \Leftrightarrow z)$ and $P(x, y, z) = x \wedge (y \vee z)$ in the rest of this section. In the following two lemmas, we handle the case of 2-consistent systems:

Lemma 9 *It holds that $\rho_2(P) = 8/27$ for $P(x, y, z) = x \wedge (y \Leftrightarrow z)$.*

Proof: We first show that $\rho_2(P) \geq 8/27$. Let us consider a 2-consistent P-system Σ. Since Σ is 2-consistent, Σ does not contain two predicates such that the first argument of one of them is x and the first argument of the other predicate is $\neg x$. Therefore, we may assume that the first argument of each predicate is a positive literal.

Choose now each variable of Σ randomly and independently to be true with the probability $p = 2/3$. The probability that any single predicate of Σ is satisfied is either $p(p^2 + (1 - p)^2) = 10/27$, if the second and the third argument are both positive or both negative literals, or $4p^2(1 - p) = 8/27$, otherwise. Hence, the expected fraction of satisfied constraints is at least $8/27$ and consequently $\rho(\Sigma) \geq 8/27$. Since the choice of a 2-consistent P-system Σ was arbitrary, we can conclude that $\rho_2(P) \geq 8/27$.

It remains to show that $\rho_2(P) \leq 8/27$. For an integer $n \geq 3$, we consider a P-system Σ_n with the variables $x_1, \ldots, x_n$. Σ_n is formed by all the $n(n-1)(n-2)$ predicates $P(x_i, x_j, \neg x_k)$ for $1 \leq i, j, k \leq n$ where all i, j and k are mutually distinct. The P-system Σ is clearly 2-consistent. We now compute $\rho(\Sigma_n)$. Consider a truth assignment which assigns the true value to exactly n' variables of Σ_n. Then, the number of satisfied constraints is precisely $n'((n' - 1)(n - n') + (n - n')(n' - 1))$. Thus, we can conclude that (set $q = n'/n$):

$$\rho(\Sigma_n) \leq \max_{0 \leq q \leq 1} \frac{qn((qn-1)(n-qn) + (n-qn)(qn-1))}{n(n-1)(n-2)} =$$

$$\max_{0 \leq q \leq 1} 2q^2(1-q) + O\left(\frac{1}{n}\right) = \frac{8}{27} + O\left(\frac{1}{n}\right).$$

Hence, $\rho_2(P) \leq 8/27$ as claimed.

□

Lemma 10 *It holds that* $\rho_2(P) = 2\sqrt{3}/9$ *for* $P(x,y,z) = x \wedge (y \vee z)$.

Proof: We first show that $\rho_2(P) \geq 2\sqrt{3}/9$. Let us consider a 2-consistent P-system Σ. Since Σ is 2-consistent, Σ does not contain two predicates such that the first argument of one of them is x and the first argument of the other is $\neg x$. Therefore, we may assume that the first argument of each predicate is a positive literal.

Choose now each variable of Σ randomly and independently to be true with the probability $p = 3^{-1/2} > 1/2$. The probability that any single predicate of Σ is satisfied is at least $p(1-p^2) = 2\sqrt{3}/9$. Hence, the expected fraction of constraints which are satisfied is at least $2\sqrt{3}/9$ and consequently $\rho(\Sigma) \geq 2\sqrt{3}/9$. Since the choice of a 2-consistent P-system Σ was arbitrary, we can conclude that $\rho_2(P) \geq 2\sqrt{3}/9$.

It remains to show that $\rho_2(P) \leq 2\sqrt{3}/9$. For an integer $n \geq 3$, we consider a P-system Σ_n with the variables $x_1, \ldots, x_n$. Σ_n is formed by all the $n(n-1)(n-2)/2$ constraints $P(x_i, \neg x_j, \neg x_k)$ for $1 \leq i,j,k \leq n$, $i \neq j$, $i \neq k$ and $j < k$. The system Σ is clearly 2-consistent. We now compute $\rho(\Sigma_n)$. Consider a truth assignment which assigns the true value to exactly n' variables $x_1, \ldots, x_n$. Then, the number of satisfied constraints is precisely the following $n'((n-n')(n'-1)+(n-n')(n-n'-1)/2)$. We can now conclude that (set $q = n'/n$):

$$\rho(\Sigma_n) \leq \max_{0 \leq q \leq 1} \frac{qn((n-qn)(qn-1)+(n-qn)(n-qn-1)/2)}{n(n-1)(n-2)/2} =$$

$$\max_{0 \leq q \leq 1} \frac{q(1-q)q+(1-q)^2/2}{1/2} + O\left(\frac{1}{n}\right) = \frac{2\sqrt{3}}{9} + O\left(\frac{1}{n}\right).$$

Hence, $\rho_2(P) \leq 2\sqrt{3}/9$ as claimed.

□

We can now analyze locally consistent P-systems for $P(x,y,z) = x \wedge (y \Leftrightarrow z)$:

Theorem 2 *If* P *is the predicate* $P(x,y,z) = x \wedge (y \Leftrightarrow z)$*, then the following holds for all* $l \geq 1$*:*

$$\rho_l(P) = \begin{cases} 1/4 & \text{if } l = 1, \\ 8/27 & \text{if } l = 2, \\ 1/2 & \text{otherwise.} \end{cases}$$

Proof: It follows that $\rho_1(P) = 1/4$ and $\rho_2(P) = 8/27$ from Lemmas 2 and 9, respectively. Hence, we focus only on the case $l \geq 3$. First, we show that $\rho_l(P) \leq 1/2$. Consider an l-consistent P'-system Σ' for $P'(y,z) = (y \Leftrightarrow z)$ with $\rho(\Sigma') \leq 1/2 + \varepsilon$ for a positive real $\varepsilon > 0$. Such a P'-system Σ' exists because $\rho_l(P') = 1/2$ by Theorem 1. Let $y_1, \ldots, y_n$ be the variables contained in Σ'. We construct an l-consistent P-system Σ with $\rho(\Sigma) = \rho(\Sigma')$. Introduce a new variable x and for each predicate $P'(y_i, y_j)$, $P'(\neg y_i, y_j)$, $P'(y_i, \neg y_j)$ and $P'(\neg y_i, \neg y_j)$ include a predicate $P(x, y_i, y_j)$, $P(x, \neg y_i, y_j)$, $P(x, y_i, \neg y_j)$ and $P(x, \neg y_i, \neg y_j)$ to Σ, respectively. Since Σ' is l-consistent, the P-system Σ is l-consistent, too. It is also not hard to see that $\rho(\Sigma) = \rho(\Sigma')$. Therefore, $\rho(\Sigma) \leq 1/2 + \varepsilon$ and $\rho_l(P) \leq 1/2$.

We now prove that $\rho_l(P) \geq 1/2$ for $l \geq 3$. Let Σ be an l-consistent P-system. Let X be the set of variables which appear as the first argument in some of the predicates of Σ and Y the set consisting of the remaining variables. Since Σ is 2-consistent, we can assume that the first argument of each predicate is a positive literal. In addition, since Σ is 3-consistent it contains neither a predicate $P(x, x', \neg x'')$ nor a predicate $P(x, \neg x', x'')$ for some $x, x', x'' \in X$. Therefore, if we set each variable of X to be true, then each predicate of Σ is either satisfied (i.e., all its arguments are set and the predicate is true) or at least one of its arguments contains a variable from the set Y. Choose now each variable of Y randomly and independently to be true with the probability $1/2$. Each predicate, which was not satisfied by fixing the values of variables from the set X, is now satisfied with the probability $1/2$. Therefore, on average, at least half of all the predicates are satisfied. Hence, $\rho(\Sigma) \geq 1/2$ and consequently $\rho_l(P) \geq 1/2$.

□

Before we analyze P-systems with $P(x,y,z) = x \wedge (y \vee z)$, we need to provide a separate upper bound for 3-consistent P-systems:

Lemma 11 *It holds that $\rho_3(P) \leq 1/2$ for $P(x,y,z) = x \wedge (y \vee z)$.*

Proof: For each $\varepsilon > 0$, we construct a 3-consistent P-system Σ with $\rho(\Sigma) < 1/2 + \varepsilon$. Let n be an integer whose exact value will be chosen later. We construct a P-system Σ_n with variables x_i for $1 \leq i \leq 2n+1$ and y_i^A for $1 \leq i \leq 2n+1$ where A ranges through all n-element subsets of $\{1, \ldots, 2n+1\} \setminus \{i\}$. The system Σ_n consists of predicates $P(x_i, \neg x_j, y_i^A)$ for all $1 \leq i,j \leq 2n+1$, $i \neq j$ and $j \in A$ and predicates $P(x_i, \neg x_j, \neg y_i^A)$ for all $1 \leq i,j \leq 2n+1$,

$i \neq j$ and $j \notin A$. In particular, the number of predicates contained in Σ_n is $(2n+1)2n\binom{2n}{n}$. Clearly, Σ_n is 3-consistent.

Let us consider a truth assignment which satisfies the most number of predicates. Let n' be the number of the variables $x_1, \ldots, x_{2n+1}$ with the true value. By symmetry, we can assume that the values of the variables $x_1, \ldots, x_{n'}$ are true and the values of $x_{n'+1}, \ldots, x_n$ are false. Observe that all the predicates whose first argument is one of the literals $x_{n'+1}, \ldots, x_n$ are false. In particular, if $n' \leq n$, then less than half of the predicates are satisfied. We focus on the case $n' > n$ in the rest of the proof.

Consider now an integer i, $1 \leq i \leq n'$, and an n-element subset A of $\{1, \ldots, 2n+1\} \setminus \{i\}$. If $|A \cap (\{1, \ldots, n'\} \setminus \{i\})| > (n'-1)/2$, then the truth assignment (because it is optimal) assigns y_i^A the true value and, otherwise, it assigns y_i^A the false value. Hence, the number of predicates, which contain y_i^A and which are satisfied, is $(2n - n' + 1) + \max\{|A \cap (\{1, \ldots, n'\} \setminus \{i\})|, n' - |A \cap (\{1, \ldots, n'\} \setminus \{i\})|\}$. For a fixed integer i, the number of n-element subsets A of $\{1, \ldots, 2n+1\} \setminus \{i\}$ with $\max\{|A \cap (\{1, \ldots, n'\} \setminus \{i\})|, n' - |A \cap (\{1, \ldots, n'\} \setminus \{i\})| \geq (1+\varepsilon)(n'-1)/2\}$ is at most the following:

$$\sum_{k=0}^{(1-\varepsilon)(n'-1)/2} \binom{n'-1}{k}\binom{2n+1-n'}{n-k} + \sum_{k=(1+\varepsilon)(n'-1)/2}^{n'-1} \binom{n'-1}{k}\binom{2n+1-n'}{n-k} \leq$$

$$\sum_{\substack{0 \leq k \leq (1-\varepsilon)(n'-1)/2 \\ (1+\varepsilon)(n'-1)/2 \leq k \leq n'-1}} \binom{n'-1}{k} 2^{2n+1-n'} \leq$$

$$2e^{-\frac{\varepsilon^2(n'-1)/2}{3}} 2^{n'-1} 2^{2n+1-n'} = 2^{2n+1} e^{-\frac{\varepsilon^2(n'-1)}{6}}.$$

Hence, for a fixed i, the number of satisfied predicates whose first argument is x_i is at most the following (recall that $n+1 \leq n'$):

$$\left(2n - n' + 1 + \frac{(1+\varepsilon)(n'-1)}{2}\right)\binom{2n}{n} + 2n2^{2n+1}e^{-\frac{\varepsilon^2(n'-1)}{6}} \leq$$

$$\left(2n - \frac{n'-1}{2} + \frac{2n\varepsilon}{2}\right)\binom{2n}{n} + 2n2^{2n+1}e^{-\frac{\varepsilon^2 n}{6}} \leq$$

$$\left(2n - \frac{n'-1}{2} + n\varepsilon\right)\binom{2n}{n} + 2n2^{2n+1}e^{-\frac{\varepsilon^2 n}{6}}.$$

Consequently, the fraction of satisfied predicates of Σ_n does not exceed:

$$\frac{n'\left(\left(2n - \frac{n'-1}{2} + n\varepsilon\right)\binom{2n}{n} + 2n2^{2n+1}e^{-\frac{\varepsilon^2 n}{6}}\right)}{(2n+1)2n\binom{2n}{n}} \leq$$

$$\frac{\left((2n+1)n + nn'\varepsilon\right)\binom{2n}{n} + 2nn'2^{2n+1}e^{-\frac{\varepsilon^2 n}{6}}}{(2n+1)2n\binom{2n}{n}} \leq$$

$$\frac{1}{2} + \frac{\varepsilon}{2} + \frac{2^{2n+1}e^{-\frac{\varepsilon^2 n}{6}}}{\binom{2n}{n}} \leq \frac{1}{2} + \frac{\varepsilon}{2} + 2(2n+1)e^{-\frac{\varepsilon^2 n}{6}}.$$

We now choose n to be an integer such that $2(2n+1)e^{-\frac{\varepsilon^2 n}{6}} \leq \varepsilon/2$. Then, each truth assignment with $n' > n$ satisfies at most the fraction of $1/2+\varepsilon$ of the predicates of Σ_n. Hence, $\rho(\Sigma_n) \leq 1/2 + \varepsilon$ as desired.

□

We are now ready to determine the values $\rho_l(P)$ for the predicate $P(x, y, z) = x \wedge (y \vee z)$:

Theorem 3 *Let P be the predicate $P(x, y, z) = x \wedge (y \vee z)$. Then, the following holds for all $l \geq 1$:*

$$\rho_l(P) = \begin{cases} 3/8 & \textit{if } l = 1, \\ 2\sqrt{3}/9 & \textit{if } l = 2, \\ \rho^{\mathrm{2SAT}}_{l-2} & \textit{otherwise.} \end{cases}$$

Proof: The equalities $\rho_1(P) = 3/8$ and $\rho_2(P) = 2\sqrt{3}/9$ follow from Lemmas 2 and 10, respectively. We first prove that $\rho_l(P) \geq \rho^{\mathrm{2SAT}}_{l-2}$ for $l \geq 3$. Let Σ be an l-consistent P-system and let X be the set of variables of Σ which appear as the first argument in some predicates of Σ. Since Σ is 2-consistent, we can assume that all the first arguments of the predicates of Σ are positive literals. Let Y be the set of the remaining variables of Σ.

We construct an auxiliary $(l-2)$-consistent 2-CNF formula Φ with the variables Y as follows. Since Σ is 3-consistent, it does not contain a predicate $P(x, \neg x', \neg x'')$ where $x', x'' \in X$. For each predicate $P(x, \neg x', y)$ and each predicate $P(x, y, \neg x')$ of Σ with $x, x' \in X$ and $y \in Y$, we include the clause (y) to Φ. Similarly, for each predicate $P(x, \neg x', \neg y)$ and each predicate $P(x, \neg y, \neg x')$ with $x' \in X$, we include the clause $(\neg y)$. For each predicate $P(x, y, y')$ with $x \in X$ and $y, y' \in Y$, we include the clause $(y \vee y')$ to Φ. We proceed analogously for predicates $P(x, \neg y, y')$, $P(x, y, \neg y')$ and $P(x, \neg y, \neg y')$. Note that some of the clauses may be contained in the formula Φ several times.

We claim that the formula Φ is $(l-2)$-consistent. If this is not the case, let Γ be the minimum inconsistent set of clauses of Φ. By Lemma 7, Γ contains at most two clauses of size one. We now find an inconsistent set Γ' of at most $|\Gamma| + 2$ predicates of Σ. For each clause of Γ of size two, include to Γ' the predicate of Σ corresponding to that clause. For each clause (y), $(\neg y)$, of Γ, include to Γ' the predicate $P(x, y, \neg x')$, $P(x, \neg y, \neg x')$, respectively, which corresponds to that clause, together with any of the predicates of Σ whose first argument is x'. Since Γ contains at most two clauses of size one, $|\Gamma'| \leq |\Gamma| + 2 \leq l$. Moreover, since Γ is inconsistent, Γ' is also inconsistent. However, this contradicts the fact that Σ is l-consistent.

Since the formula Φ is $(l-2)$-consistent, there is a truth assignment which satisfies the fraction of $\rho(\Phi) \geq \rho_{l-2}^{\mathrm{2SAT}}$ clauses of Φ. Extend this truth assignment to all the variables of Σ by assigning the true value to each variable $x \in X$. All the predicates of Σ whose arguments contain solely the variables from the set X are satisfied and, in addition, the fraction of $\rho(\Phi)$ of the remaining predicates are also satisfied. Therefore, $\rho(\Sigma) \geq \rho(\Phi) \geq \rho_{l-2}^{\mathrm{2SAT}}$. Since the choice of a P-system Σ was arbitrary, we can conclude that $\rho_l(P) \geq \rho_{l-2}^{\mathrm{2SAT}}$.

It remains to prove that $\rho_l(P) \leq \rho_{l-2}^{\mathrm{2SAT}}$ for $l \geq 3$. If $l = 3$, the upper bound follows from Lemma 11. For $l \geq 4$, choose $\varepsilon > 0$ and fix an $(l-2)$-consistent 2-CNF formula Φ with $\rho(\Phi) \leq \rho_{l-2}^{\mathrm{2SAT}} + \varepsilon$ such that each minimal inconsistent set of at most l clauses contain two clauses of size one. Such a formula Φ exists by Lemma 8. Moreover, we can assume that each clause of size two is contained in Φ at most once. Let m' be the number of clauses of Φ of size one (counting multiplicities) and m the number of all clauses of Φ. Since Φ is 2-consistent, $m'/m \leq \rho(\Phi)$. We now construct an l-consistent P-system Σ with $\rho(\Sigma) = \rho(\Phi)$.

Let $y_1, \ldots, y_n$ be the set consisting of the variables of the formula Φ. The system Σ will contain $(m+1)n$ variables y_i^j for $1 \leq i \leq n$ and $1 \leq j \leq m+1$ and $m+1$ variables x^j for $1 \leq j \leq m+1$. Let $C_1, \ldots, C_m$ be the clauses of Φ. For each clause $C_k = (y_i \vee y_{i'})$, $1 \leq k \leq m$, we include to Σ predicates $P(x^j, y_i^j, y_{i'}^j)$ for $1 \leq j \leq m+1$. Similarly, we proceed for clauses $C_k = (y_i \vee \neg y_{i'})$ and $C_k = (\neg y_i \vee \neg y_{i'})$. If the clause C_k is of size one, say $C_k = (y_i)$, we include to Σ predicates $P(x^j, y_i^j, \neg x^{(j+k) \bmod (m+1)})$ for

$1 \leq j \leq m+1$. Therefore, Σ consists of $m(m+1)$ distinct predicates.

First, we show that Σ is l-consistent. Assume the opposite and let Γ be the minimum inconsistent set of predicates contained in Σ, i.e., $|\Gamma| \leq l$. Observe that if we set all the variables $x^1, \ldots, x^n$ to be true, then the system Σ reduces to $m+1$ independent "copies" of the formula Φ. Therefore, if Γ is a set of l inconsistent predicates, it must contain predicates contained in one of these copies of Φ which correspond to an inconsistent set Γ_Φ of clauses of Φ. By symmetry, we can assume that predicates corresponding to Γ_Φ are contained in the first copy of Φ. Since Φ is $(l-2)$-consistent, $|\Gamma_\Phi| \geq l-1$. On the other hand, $|\Gamma_\Phi| \leq |\Gamma| \leq l$. By the choice of Φ, each inconsistent set of at most l clauses of Φ contains two clauses of size one. Let $C_k = (y_i)$ and $C_{k'} = (y_{i'})$ be these two clauses of size one, i.e., Γ contains the predicates $P(x^1, y_i, \neg x^{(k+1) \bmod (m+1)})$ and $P(x^1, y_{i'}, \neg x^{(k'+1) \bmod (m+1)})$. If Γ is inconsistent, it must contain a predicate whose first argument is $x^{(k+1) \bmod (m+1)}$ as well as a predicate whose first argument is $x^{(k'+1) \bmod (m+1)}$. Therefore, Γ contains at least $|\Gamma_\Phi| + 2 > l$ predicates.

We now show that $\rho(\Sigma) = \rho(\Phi)$. Since $\rho(\Phi) \leq \rho^{2\mathrm{SAT}}_{l-2} + \varepsilon$ and the choice of ε was arbitrary, this would yield $\rho_l(P) \leq \rho^{2\mathrm{SAT}}_{l-2}$. Fix a truth assignment such that the fraction of $\rho(\Sigma)$ predicates of the P-system Σ is satisfied. We claim that there is an optimum truth assignment which assigns all the variables $x^1, \ldots, x^{m+1}$ the true value. Indeed, if x^j is false, then change the value of x^j to true. This causes at most m' previously satisfied predicates to be unsatisfied (precisely those which contain $\neg x^j$ as the third argument) and, on the other hand, we can choose values of $y^j_1, \ldots, y^j_n$ so that at least the $\rho(\Phi)m$ predicates whose first argument is x^j are satisfied. Note that none of these $\rho(\Phi)m$ predicates could be satisfied before the change of the value of x^j. Since $\rho(\Phi)m \geq m'$ (recall that $\rho(\Phi) \geq m'/m$), the number of satisfied predicates is not decreased after the change. In this way, we can switch all the variables $x^1, \ldots, x^{m+1}$ to true without decreasing the number of satisfied constraints. Hence, we can assume that all the variables $x^1, \ldots, x^{m+1}$ are set to be true by the considered optimum truth assignment. Then, the system Σ is reduced to $m+1$ independent "copies" of the formula Φ (substitute the true value for all the variables $x^1, \ldots, x^{m+1}$). We can conclude that $\rho(\Sigma) = \rho(\Phi)$.

□

6 Conclusion

We studied instances of constraint satisfaction problems which are locally consistent. There are several directions for possible future research. First, we were not able to fully analyze Boolean predicates which are not 1-extendable. The smallest two non-trivial such Boolean predicates, $P(x, y, z) = x \wedge (y \Leftrightarrow z)$ and $P(x, y, z) = x \wedge (y \vee z)$, already showed that the behavior of locally consistent P-systems for such predicates P can be quite weird. Another direction is to allow constraints of more types: In this setting, the previously most studied case of locally consistent CNF formulas can be viewed as a constraint satisfaction problem where constraints are just disjunctions, e.g., the case of 2-CNF formulas corresponds to problems with the constraints $P(x) = x$ and $P(x, y) = x \vee y$. The last possible direction is to consider constraints with larger domains. Some of our results can be easily translated to this more general setting, e.g., Theorem 1, on the other hand, their detailed analysis even for small arities might be quite difficult because of their potentially rich structure.

Acknowledgement

The authors would like to thank Gerhard Woeginger for attracting their attention to locally consistent formulas and for pointing out several useful references. They would also like to thank Dimitrios M. Thilikos for suggesting the version of the problem considered in this paper.

References

[1] S. Cook: The Complexity of Theorem-proving Procedures. In: Proc. of the 3rd ACM Symposium on Theory of Computing. ACM, New York (1971) 29–33.

[2] S. Cook, D. Mitchell: Finding Hard Instances of the Satisfiability Problem: A Survey. In: Satisfiability Problem: Theory and Applications. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 35 AMS (1997).

[3] D. Eppstein: Improved Algorithms for 3-coloring, 3-edge-coloring and Constraint Satisfaction. In: Proc. of the 12th ACM-SIAM Symposium on Discrete Algorithms. SIAM (2001) 329–337.

[4] T. Feder, R. Motwani: Worst-case Time Bounds for Coloring and Satisfiability Problems. J. Algorithms 45(2) (2002) 192-201.

[5] T. Hagerup, Ch. Rüb: A guided tour Chernoff bounds. Inform. Process. Letters 33 (1989) 305–308.

[6] M. A. Huang, K. Lieberherr: Implications of Forbidden Structures for Extremal Algorithmic Problems. Theoretical Computer Science 40 (1985) 195–210.

[7] S. Jukna: Extremal Combinatorics with Applications in Computer Science. Springer, Heidelberg (2001).

[8] D. Král': Locally Satisfiable Formulas. In: Proc. of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms. SIAM (2004) 323-332.

[9] K. Lieberherr, E. Specker: Complexity of Partial Satisfaction. J. of the ACM, 28(2) (1981) 411–422.

[10] K. Lieberherr, E. Specker: Complexity of Partial Satisfaction II. Technical Report 293, Dept. of EECS, Princeton University (1982).

[11] L. Trevisan: On Local versus Global Satisfiability. SIAM J. Disc. Math. (to appear). A preliminary version is available as ECCC report TR97-12.

[12] Z. Usiskin: Max-min Probabilities in the Voting Paradox. Ann. Math. Stat. 35 (1963) 857–862.

[13] G. J. Woeginger: Exact Algorithms for NP-hard Problems: A Survey. In: M. Jünger, G. Reinelt, G. Rinaldi (eds.): Proc. 5th Int. Worksh. Combinatorial Optimization - Eureka, You Shrink. Lecture Notes in Computer Science, Vol. 2570. Springer-Verlag Berlin (2003) 185-207.

[14] M. Yannakakis: On the Approximation of Maximum Satisfiability. J. Algorithms 17 (1994) 475–502. A preliminary version appeared in: Proc. of the 3rd Annual ACM-SIAM Symposium on Discrete Algorithms. SIAM (1992) 1–9.