united nations educational, scientific and cultural organization

international atomic energy agency

the **abdus salam** international centre for theoretical physics

40th anniversary 1964 2004

preprint

UPPER BOUND ON THE NON-COLORABILITY THRESHOLD OF THE 2+p-COL PROBLEM

Roberto Cruz

Available at: http://www.ictp.it/~pub_off IC/2004/99

United Nations Educational Scientific and Cultural Organization
and
International Atomic Energy Agency
THE ABDUS SALAM INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

UPPER BOUND ON THE NON-COLORABILITY THRESHOLD OF THE 2+p-COL PROBLEM

Roberto Cruz[1]
Department of Mathematics, University of Antioquia, A.A 1226, Medellin, Colombia
and
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

Abstract

The 2+p-COL problem introduced by Walsh, smoothly interpolates between P and NP by mixing together the polynomial 2-coloring problem and the NP complete 3-coloring problem. A natural upper bound on the non colorability of the 2+p-COL problem is $\min\{\overline{r}_2/(1-p), \overline{r}_3\}$, where $\overline{r}_2$ and $\overline{r}_3$ are the upper bounds on 2-COL and 3-COL thresholds respectively. In this paper we improve this upper bound for each $0.73 \leq p \leq 1$. This means that for $p \geq 0.73$ the 2+p-COL problem does not behave like the 2-COL problem. We use the method developed by Kaporis et al., which combines the concept of legal rigid colorings introduced by Achlioptas and Molloy with the occupancy problem for random allocations of balls into bins.

MIRAMARE – TRIESTE
October 2004

[1]Regular Associate of ICTP. rcruz@matematicas.udea.edu.co

1 Introduction

Phase transition behavior has given much insight into what makes NP-complete problems hard to solve. Many experimental results have established that the hardest instances to solve in a number of NP-complete problems often occur around a rapid transition in solubility [1, 2, 3, 4]. Some of the most interesting phase transition results have come from the random 2+p-SAT problem introduced by Monasson et al. in [5]. This problem lets us explore the interface between P and NP by mixing together the polynomial 2-SAT and the NP-complete 3-SAT. The phase transition behavior of this problem has been very well studied from both the experimental and theoretical point of view. Despite the fact that the problem is NP-complete for any fixed $p > 0$, it was established that the polynomial 2-SAT subproblem dominates the satisfiability and cost to solve the 2+p-SAT problem up to $p = 2/5$ [5, 6].

Motivated by the insights that the 2+p-SAT problem has provided into computational complexity and algorithm performance, in [7] Walsh introduced five new problems in order to explore more deeply the interface between P and NP. The most interesting of these problems is the 2+p-COL problem, which is studied in [8]. This problem was defined as a mix of the random 2-coloring problem, which is in P, and the NP-complete random 3-coloring problem. In [8], Walsh stated the most important questions related with this new problem and performed a detailed experimental study of its phase transition.

In this paper, we study the problem of computing an upper bound on the non colorability threshold of the 2+p-COL problem. This is the first theoretical study of one of the important questions introduced by Walsh. In the next section we review the results concerning 2+p-SAT and 2+p-COL problems and discuss the main differences between them. In section 3 we obtain theoretical upper bound on the non-colorability threshold of the 2+p-COL problem by applying the method developed by Kaporis et al. [9]. Conclusions are given in the last section.

2 2+p-SAT vs. 2+p-COL

A random k-SAT problem in formed by selecting uniformly and independently m clauses from the set of all $2^k\binom{n}{k}$ k-clauses on a given set of n variables. The main open question on random k-SAT is the existence of a sharp threshold as the ratio of clauses to variables is increased, i.e. that there exists c_k such that a random formula with ratio $c < c_k$ is satisfiable with high probability (w.h.p.), whereas a random formula with $c > c_k$ is unsatisfiable w.h.p. For $k = 2$ the existence of the sharp threshold at $c_2 = 1$ was independently proved in [11, 12, 13]. For $k \geq 3$, Friedgut proved that thresholds depending on n exist [14], but determining the exact location of the threshold remains a difficult open problem. For random 3-SAT there has been a number of results on lower and upper bounds on the threshold c_3. The best known lower and upper bounds on c_3 are 3.52 [15] and 4.57 [16], respectively.

The 2+p-SAT problem introduced in [5] is defined as follows: fix $p \in [0,1]$ and form a random formula with n variables by selecting uniformly and independently pm clauses from the set of all 3-clauses and $(1-p)m$ clauses from the set of all 2-clauses. Thus, $p=0$ corresponds to random 2-SAT, while $p=1$ corresponds to random 3-SAT. Monasson et al. claimed that for every $p \in [0,1]$ there is a critical value of c, denoted by c_p around which the $2+p$-SAT problem undergoes a phase transition. By considering the satisfiability of the embedded 2-SAT subproblem, an easy upper bound on c_p follows from the fact that $c_p(1-p) \leq c_2$, i.e. $c_p \leq \frac{1}{1-p}$. In [19] it was claimed that there exists $p^* \approx 0.41$, such that for all $p \in [0,p^*)$, $c_p = \frac{1}{1-p}$. Also, the transition shifts from continuous to discontinuous as the backbone jumps in size. Achlioptas et al.[6] gave rigorous theoretical results for the random $2+p$-SAT problem. They proved that the random $2+p$-SAT problem exhibits a sharp threshold for every $p \in [0,1]$ and obtained that for $p \leq 2/5$, with probability $1-o(1)$, a random 2+p-SAT formula is satisfiable iff its 2-SAT subformula is satisfiable, i.e that $c_p = \frac{1}{1-p}$ for $p \leq 2/5$. Moreover, they obtained lower and upper bounds for the transition for every $p > 2/5$ and established that the 2+p-SAT problem behaves like random 3-SAT for $p \geq 0.695$.

The random k-COL problem is the problem of coloring a graph with n vertices, each with k possible colors and rn edges drawn uniformly and at random. Like k-SAT, k-COL is NP-complete for $k \geq 3$ but polynomial for $k=2$. It is well known that the 2-COL problem has a coarse transition, i.e. the probability of non-2-colorability is bounded away from 0 for any $r > 0$ and increases gradually with r, reaching $1-o(1)$ at $r = 1/2$. For the random 3-COL problem, Achlioptas and Friedgut [17] showed that there exists a function $r_3(n)$ such that for any $\epsilon > 0$ a random graph with $(r_3(n)-\epsilon)n$ edges is 3-colorable w.h.p. and a random graph with $(r_3(n)+\epsilon)n$ is not 3-colorable w.h.p., i.e. that 3-colorability has a sharp threshold. It is widely believed that $r_3(n)$ converges to a constant r_3 but this remains an open problem. Upper and lower bounds on r_3 have been rigorously established. The best lower bound is currently 1.923 and has been obtained by Achlioptas and Molloy [18] while the best upper bound on r_3 is 2.495, obtained independently by Kaporis et al. [9] and Fountoulakis and McDiarmid in [10]

To interpolate smoothly from P to NP, a random 2+p-COL problem has a fraction $(1-p)$ of its vertices with 2 colors, and a fraction p with 3 colors, where $0 \leq p \leq 1$. We say that a graph $G(V,E)$ is 2+p-colorable if the set $V = \{1,\ldots,n\}$ of its vertices, which is initially partitioned into two subsets $U = \{1,\ldots,(1-p)n\}$ and $W = \{(1-p)n+1,\ldots,n\}$, can be partitioned into 5 nonempty cells U_1, W_1, U_2, W_2 and W_3, such that no two vertices belonging to the same cell or belonging to different cells with the same number are adjacent. This partition is called a 2+p-coloring of G and the vertices of the sets $U_j,\ j=1,2$ and $W_j,\ j=1,2,3$ are said to have color j. Notice that the vertices with 2 colors are fixed and cannot be chosen freely. Like 2+p-SAT, the 2+p-COL problem is NP-complete for any fixed $p > 0$.

The main difference between 2+p-COL and 2+p-SAT is the fact that while 2-SAT, 3-SAT and 3-COL all have sharp transitions, 2-COL has a coarse one. Thus, the 2+p-COL problem mixes two subproblems belonging to different complexity classes and also with different phase transition behavior. In [8], Walsh studied experimentally how the random 2+p-COL phase transition varies as p goes from 0 to 1. He found that the behavior of the 2+p-COL problem appears to be dominated by the embedded polynomial 2-COL subproblem up to $p \approx 0.8$ and observed phase transition behavior for $p < 0.8$ in which there appeared to be both smooth and sharp regions.

In order to analyze the phase transition Walsh [8] introduced the parameters:

$$\begin{aligned} \underline{r}_{2+p} &= \sup\{r \mid G(n, rn) \text{ is 2+p-colorable w.h.p.}\}, \\ \overline{r}_{2+p} &= \inf\{r \mid G(n, rn) \text{ is not 2+p-colorable w.h.p.}\}. \end{aligned}$$

From [20] $\underline{r}_2 = 0$ and $\overline{r}_2 \approx 1/2$. For any fixed $p < 1$, a random 2+p-COL problem contains a 2-COL problem which has a probability of being non colorable that asymptotically is less than 1, then $\underline{r}_{2+p} = 0$ for all $p < 1$. Since colorability is a monotone property, $\underline{r}_{2+p} \leq \overline{r}_{2+p}$. On the other hand, since the 3-COL problem has a sharp phase transition, using the lower and upper bound we have $1.923 < \underline{r}_3 \leq \overline{r}_3 < 2.495$. So, $\underline{r}_{2+p}$ marks the start of the phase transition while $\overline{r}_{2+p}$ marks its end. The start stays fixed at $\underline{r}_{2+p} = 0$ for $p < 1$ and jumps discontinuously to $\underline{r}_3$ at $p = 1$. As with 2+p-SAT, we have:

$$\overline{r}_2 \leq \overline{r}_{2+p} \leq min\left(\frac{\overline{r}_2}{1-p}, \overline{r}_3\right). \tag{1}$$

In his experimental study [8], Walsh found that the upper bound (1), which looks at the 2-Col subproblem, is tight up to p=0.8, using the value of 2.52 for $\overline{r}_3$, which was the best known upper bound on r_3 at that moment [21].

In the next section, we show that this upper bound is tight up to some $p_c \leq 0.729$ and obtain a new upper bound on $\overline{r}_{2+p}$ for $0.73 \leq p \leq 1$, using a theoretical approach of [9].

3 Bounding $\overline{r}_{2+p}$ using rigid colorings and the occupancy problem

The first moment method, that makes use of Markov's inequality gives, as an upper bound to the non colorability threshold of the 3-COL problem, the value of 2.7 [21, 9]. This method does not give the smallest possible value of r, because for values of $r < 2.7$ there are graphs that possess a very large number of 3-colorings that contribute greatly to the expectation. Since the natural upper bound (1) depends on the upper bound on r_3, we use the method of Kaporis et al. [9], which leads to the current best upper bound 2.495. This approach combines the concept of rigid legal colorings introduced by Achlioptas and Molloy [21] with the occupancy problem for random allocations of balls into bins. We will extend the definitions and notations of [21] to the case of 2+p-colorings and adapt the analysis of [9] to this case.

The 2+p-partition $P = (U_1, W_1, U_2, W_2, W_3)$ is a 2+p-coloring of the graph $G(V, E)$ if no edge $e \in E$ connects two vertices from cells with the same number. We let

$$|W_1| = \alpha n, \quad |W_2| = \beta n, \quad |W_3| = (p - \alpha - \beta)n, \quad \alpha, \beta \geq 0, \quad \alpha + \beta \leq p$$

$$|U_1| = \gamma n, \quad |U_2| = (1 - p - \gamma)n, \quad 0 \leq \gamma \leq 1 - p.$$

Let C_P denote the event that P is a 2+p-coloring of a graph $G(V, E)$. The number of edges that are allowed to exist in a graph with this 2+p-coloring is

$$\begin{aligned} T(P) &= (|U_1| + |W_1|)(|U_2| + |W_2|) + \\ &+ (|U_1| + |W_1|)|W_3| + (|U_2| + |W_2|)|W_3| \\ &= [(\gamma + \alpha)(1 - p - \gamma + \beta) + (p - \alpha - \beta)(1 - p + \alpha + \beta)]\, n^2 \\ &= \tau n^2. \end{aligned}$$

Considering a random graph formed by selecting uniformly at random $m = rn$ edges with repetition allowed, the probability of the event C_P is

$$\mathbf{Pr}[C_P] = \left(\frac{T(P)}{\binom{n}{2}} \right)^{rn} = (2\tau)^{rn} \mathrm{O}(1). \tag{2}$$

We now use the concept of rigid colorings introduced in [21]. For a $2 + p$-coloring P of V, let us say that a vertex v of color i is *unmovable* in P, if for every $j > i$ the partition resulting by moving v to a cell with color j is not a $2 + p$-coloring of G. Note that every vertex in U_2 is unmovable by definition. We will say that P is a *rigid* $2 + p$-coloring of G, if every vertex is unmovable in P. This event will be denoted by R_P. We will only consider the set $R^{\#}$ of rigid colorings. The following Markov type inequality holds:

$$\mathbf{Pr}[\text{G has a 2+p-coloring}] \leq \mathbf{E}[|R^{\#}|] = \sum_P \mathbf{Pr}[R_P \mid C_P]\mathbf{Pr}[C_P] \ , \tag{3}$$

where the sum runs over the set of 2+p-partitions of V. We now compute the probability of the conditional event $R_P|C_P$ as in [9].

Since we have conditioned on C_P, the random graphs of the resulting probability space may contain only edges connecting cells with different numbers. Let E_1 denote the event that an edge between parts $U_1 \bigcup W_1$ and $U_2 \bigcup W_2$ is chosen and let $E_{i+1}, i = 1, 2$ denote the event that an edge between vertices of the parts W_i, W_3 is chosen. Since vertices in U_1 and U_2 cannot have the color 3 by definition, we denote by E^* the event that an edge between parts U_1, W_3 or between parts U_2, W_3 is chosen. Then, in each edge selection of the rn, exactly one of the four possible events E_1, E_2, E_3, E^* must be realized. The probabilities of these events are:

$$\begin{aligned}
\mathbf{Pr}[E_1] &= \frac{(\gamma+\alpha)(1-p-\gamma+\beta)}{\tau}, \ \mathbf{Pr}[E_2] = \frac{\alpha(p-\alpha-\beta)}{\tau}, \\
\mathbf{Pr}[E_3] &= \frac{\beta(p-\alpha-\beta)}{\tau}, \ \mathbf{Pr}[E^*] = \frac{(1-p)(p-\alpha-\beta)}{\tau},
\end{aligned}$$

Let $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ be random variables counting the number of times the event E_1, E_2, E_3, E^* is realized respectively, in the process of the rn edge selections (repetitions counted). Then, the joint distribution of these variables is the multinomial distribution with probabilities $\mathbf{Pr}[E_1], \mathbf{Pr}[E_2]$, $\mathbf{Pr}[E_3]$ and $\mathbf{Pr}[E^*]$. Denoting by Λ_{xyz} the event $[\lambda_1 = xrn, \lambda_2 = yrn, \lambda_3 = znr, \lambda_4 = (1-x-y-z)rn]$ and by $V_{\alpha\beta\gamma}$ the fact that $|W_1| = \alpha n, |W_2| = \beta n, |W_3| = (p-\alpha-\beta)n, |U_1| = \gamma n, |U_2| = (1-p-\gamma)n$, the following holds:

$$\mathbf{Pr}[\Lambda_{xyz} \mid V_{\alpha\beta\gamma}] = \binom{rn}{xrn,\ yrn,\ zrn,\ (1-x-y-z)rn} \cdot$$

$$\mathbf{Pr}[E_1]^{xrn}\mathbf{Pr}[E_2]^{yrn}\mathbf{Pr}[E_3]^{zrn}\mathbf{Pr}[E^*]^{(1-x-y-z)rn}$$

$$\asymp \left[\frac{1}{\tau}\left(\frac{(\alpha+\gamma)(1-p-\gamma+\beta)}{x}\right)^x \left(\frac{\alpha(p-\alpha-\beta)}{y}\right)^y\right]^{rn} \cdot$$

$$\left[\left(\frac{\beta(p-\alpha-\beta)}{z}\right)^z \left(\frac{(1-p)(p-\alpha-\beta)}{1-x-y-z}\right)^{1-x-y-z}\right]^{rn} \tag{4}$$

where $0 \le x, y, z \le 1$ and $F \asymp G$ denote the fact that $\ln F \sim \ln G$.

Each time we select an edge from the $(\gamma+\alpha)(1-p-\gamma+\beta)n^2$ possible edges from $U_1 \cup W_1$ to $U_2 \cup W_2$, its endpoint in $U_1 \cup W_1$ is unmovable to color 2. So, after the rn edge selections, the $\lambda_1 = xrn$ edges that connect vertices from $U_1 \cup W_1$ and $U_2 \cup W_2$ must be sufficient in order to make all the $(\alpha+\gamma)n$ vertices belonging to $U_1 \cup W_1$ unmovable to color 2. The process of choosing edges connecting $U_1 \cup W_1$ with $U_2 \cup W_2$, thus making vertices in $U_1 \cup W_1$ unmovable to color 2, can be viewed as throwing randomly and uniformly xrn ball in $(\alpha+\gamma)n$ bins. Let H_1 denote the event that all vertices in $U_1 \cup W_1$ are unmovable to $U_2 \cup W_2$; the probability of this event is equal to the probability that none of the $(\alpha+\gamma)n$ bins remain empty after the random placement of the xrn balls. The same is valid for $H_i, i = 1, 2$ that denotes the events that all vertices in W_i are unmovable to W_3. In order to obtain sharp estimates of the probabilities that no bin remains empty after the random placement of the balls we use the following theorem by Kamath et al. [22].

Theorem 1 [22] *Let W denote the number of empty bins after the placement, uniformly and independently, of l balls into k bins, where both l and k are constant multiples of n. Let $c = l/k \geq 1$. If we denote by $H(l,k,w)$ the probability that $W = w$ and if, in addition $|w - \boldsymbol{E}[W]| = \Omega(k)$ then*

$$H(l,k,w) \asymp e^{-k\left[\int_0^{1-\frac{w}{k}} \ln\left(\frac{u-t}{1-t}\right)dt - c\ln u\right]}$$

where u is the solution of the equation $w = k(1 - u(1 - e^{-c/u}))$.

For each of the three cases of color changes under consideration we have to set $w = 0$. In this case u is the solution of the equation

$$u(1 - e^{-c/u}) = 1 \;\Leftrightarrow\; \ln(u-1) = ln(u) - \frac{c}{u}$$

which can be expressed using Lambert W function [23] as:

$$u = 1/[1 + \mathrm{LambertW}(-ce^{-c})/c].$$

The condition

$$w - \mathbf{E}[W]| = |0 - k\left(1 - \frac{1}{k}\right)^l = \Omega(k)$$

can be easily verified for the three events under consideration. For example in the case of the event H_1 we have $l = xrn$, $k = (\alpha+\gamma)n$ and

$$|w - E[W]| = |0 - k(1 - \frac{l}{k})^k| \sim (\alpha+\gamma)e^{-}xr/(\alpha+\gamma) = \Omega(k).$$

Using the equation for u in the case of $w = 0$ one can easily verify that

$$\int_0^1 \ln\left(\frac{u-t}{1-t}\right) dt = c + \ln(u-1)$$

and this yields

$$H(l,k,0) \asymp \exp[-k(c + \ln(u-1) - c\ln u)].$$

The following estimates are deduced from the above theorem:

$$\begin{aligned}\mathbf{Pr}[H_1 \mid \lambda_1 = xrn, |U_1| + |W_1| = (\alpha+\gamma)n] &\asymp e^{-(\alpha+\gamma)[c_1 + \ln(u_1-1) - c_1\ln u_1]},\\ \mathbf{Pr}[H_2 \mid \lambda_2 = yrn, |W_1| = \alpha n] &\asymp e^{-\alpha[c_2 + \ln(u_2-1) - c_2\ln u_2]},\\ \mathbf{Pr}[H_3 \mid \lambda_3 = zrn, |W_2| = \beta n] &\asymp e^{-\beta[c_3 + \ln(u_3-1) - c_3\ln u_3]},\end{aligned} \tag{5}$$

where $c_1 = \frac{xr}{(\alpha+\gamma)} \geq 1$, $c_2 = \frac{yr}{\alpha} \geq 1$, $c_3 = \frac{zr}{\beta} \geq 1$ and u_i is expressed as

$$u_i = 1/[1 + \mathrm{LambertW}(-c_i e^{-c_i})/c_i], \quad i = 1,2,3.$$

For a given P and using (4) and (5), the probability of conditional event $R_P \mid C_P$ is given by

$$
\begin{aligned}
Pr[R_p \mid C_P] = & \sum_{(x,y,z)\in D_{\alpha\beta\gamma}} \mathbf{Pr}[\Lambda_{xyz} \mid V_{\alpha\beta\gamma}]\mathbf{Pr}[H_1, H_2, H_3 \mid \Lambda_{xyz} \wedge V_{\alpha\beta\gamma}] \\
\asymp & \sum_{(x,y,z)\in D_{\alpha\beta\gamma}} \left[\frac{1}{\tau}\left(\frac{(\alpha+\gamma)(1-p-\gamma+\beta)}{x}\right)^x \left(\frac{\alpha(p-\alpha-\beta)}{y}\right)^y\right]^{rn} \\
& \cdot \left[\left(\frac{\beta(p-\alpha-\beta)}{z}\right)^z \left(\frac{(1-p)(p-\alpha-\beta)}{1-x-y-z}\right)^{1-x-y-z}\right]^{rn} \\
& \cdot \left[e^{-(\alpha+\gamma)[c_1+\ln(u_1-1)-c_1\ln u_1]-\alpha[c_2+\ln(u_2-1)-c_2\ln u_2]}\right]^n \\
& \cdot \left[e^{-\beta[c_3+\ln(u_3-1)-c_3\ln u_3]}\right]^n \\
= & \sum_{(x,y,z)\in D_{\alpha\beta\gamma}} F(x,y,z \mid \alpha,\beta,\gamma). \qquad (6)
\end{aligned}
$$

where

$$D_{\alpha\beta\gamma} = \{(x,y,z) \in [0,1]^3 \mid x \geq \frac{(\alpha+\gamma)}{r}, y \geq \frac{\alpha}{r}, z \geq \frac{\beta}{r}, x+y+z \leq 1\}.$$

Taking the expectation of the number of colorings $|R^{\#}|$ and using (2) and (6), we get

$$\mathbf{E}[|R^{\#}|] \asymp \sum_{P} [2\tau]^{rn} \cdot \sum_{(x,y,z)\in D_{\alpha\beta\gamma}} F(x,y,z \mid \alpha,\beta\ \gamma) \qquad (7)$$

where the first sum runs over the set of 2+p-partitions of V. Using the concept of isomorphic partitions as in [21] we will restrict the sum above to run over any maximal set of non-isomorphic 2+p-partition of V. The number of isomorphic partitions for a given 2+p-partition P is

$$\asymp \left[\frac{1}{\gamma^\gamma(1-p-\gamma)^{(1-p-\gamma)}}\right]^{(1-p)n} \left[\frac{1}{\alpha^\alpha\beta^\beta(p-\alpha-\beta)^{(p-\alpha-\beta)}}\right]^{pn},$$

then (7) can be written as

$$
\begin{aligned}
\mathbf{E}[|R^{\#}|] \asymp & \sum_{(\alpha,\beta,\gamma,x,y,z)\in D} [2\tau]^{rn} \cdot F(x,y,z \mid \alpha,\beta,\gamma) \\
& \cdot \left[\frac{1}{\alpha^\alpha\beta^\beta(p-\alpha-\beta)^{(p-\alpha-\beta)}}\right]^{pn} \left[\frac{1}{\gamma^\gamma(1-p-\gamma)^{(1-p-\gamma)}}\right]^{(1-p)n} \qquad (8)
\end{aligned}
$$

where

$$D = \{(\alpha,\beta,\gamma,x,y,z) \in [0,p]^2 \times [0,1-p] \times [0,1]^3 \mid \alpha+\beta \leq p,\ (x,y,z) \in D_{\alpha\beta\gamma}\}.$$

Now, if we find a condition on r that forces an arbitrary term of the sum that appears in (8) to converge to 0, then the whole sum will converge to 0 since it contains polynomially many terms vanishing exponentially fast. An arbitrary term of this sum is given by the following expression raised to n:

$$
\begin{aligned}
E = & \left[\frac{1}{\alpha^\alpha \beta^\beta (p-\alpha-\beta)^{(p-\alpha-\beta)}}\right]^p \left[\frac{1}{\gamma^\gamma (1-p-\gamma)^{(1-p-\gamma)}}\right]^{(1-p)} \\
& \cdot \left[2\left(\frac{(\alpha+\gamma)(1-p-\gamma+\beta)}{x}\right)^x \left(\frac{\alpha(p-\alpha-\beta)}{y}\right)^y\right]^r \\
& \cdot \left[\left(\frac{\beta(p-\alpha-\beta)}{z}\right)^z \left(\frac{(1-p)(p-\alpha-\beta)}{1-x-y-z}\right)^{1-x-y-z}\right]^r \\
& \cdot \left[e^{-(\alpha+\gamma)[c_1+\ln(u_1-1)-c_1\ln u_1]-\alpha[c_2+\ln(u_2-1)-c_2\ln u_2]}\right] \\
& \cdot \left[e^{-\beta[c_3+\ln(u_3-1)-c_3\ln u_3]}\right]. \qquad (9)
\end{aligned}
$$

For each p and any value of r such that E is strictly less than 1 for all $(\alpha, \beta, \gamma, x, y, z) \in D$, the n-power of E tends to 0 meaning that the graph is not 2+p-colorable with high probability. It makes the calculation easier to consider the natural logarithm of E and analyze when $\ln E$ is strictly less than 0.

$$
\begin{aligned}
\ln E = & -p\,[\alpha\ln\alpha + \beta\ln\beta + (p-\alpha-\beta)\ln(p-\alpha-\beta)] \\
& - (1-p)\,[\gamma\ln\gamma + (1-p-\gamma)\ln(1-p-\gamma)] \\
& + r\ln 2 + rx\cdot[\ln(\alpha+\gamma) + \ln(1-p-\gamma+\beta) - \ln x] \\
& + ry\cdot[\ln\alpha + \ln(p-\alpha-\beta) - \ln y] \\
& + rz\cdot[\ln\beta + \ln(p-\alpha-\beta) - \ln z] \\
& + r(1-x-y-z)[\ln(1-p) + \ln(p-\alpha-\beta) - \ln(1-x-y-z)] \\
& - (\alpha+\gamma)\,[c_1 + \ln(u_1-1) - c_1\ln u_1] \\
& - \alpha\,[c_2 + \ln(u_2-1) - c_2\ln u_2] \\
& - \beta\,[c_3 + \ln(u_3-1) - c_3\ln u_3]. \qquad (10)
\end{aligned}
$$

As in [9] it can be demonstrated that $\ln E$ is convex for each $p \in [0,1]$ and any value of r as a function of $(\alpha, \beta, \gamma, x, y, z) \in D$. In order to find the upper bound on $\overline{r}_{2+p}$ for each p and provided that the function $\ln E$ is convex, we found the minimum value of r such that the maximum of $\ln E$ over D is strictly negative.

We did the calculation iteratively with $\Delta p = 0.01$, starting with $p = 0.99$ and $r = 2.494695$, the upper bound on r_3 obtained in [9]. In each iteration we use as an initial value r_i the value of r found in the previous iteration; we then compute the optimum value of r as follows. For the initial value r_i, $\max_D \ln E(r_i) < 0$; using $\Delta r = 0.05$ we decrease the value of r_i step by step until we find the closest value $r_d < r_i$ such that $\max_D \ln E(r_d) > 0$. Starting with these values r_i and r_d, we apply the bisection method, maximizing $\ln E$ over D for each new r until $|r_i - r_d| < 10^{-10}$. Since $\ln E$ is convex, the maximum in each case was obtained iteratively, by evaluating the objective function on a uniform grid of D. Initially we use a grid of size $\delta = 0.1$

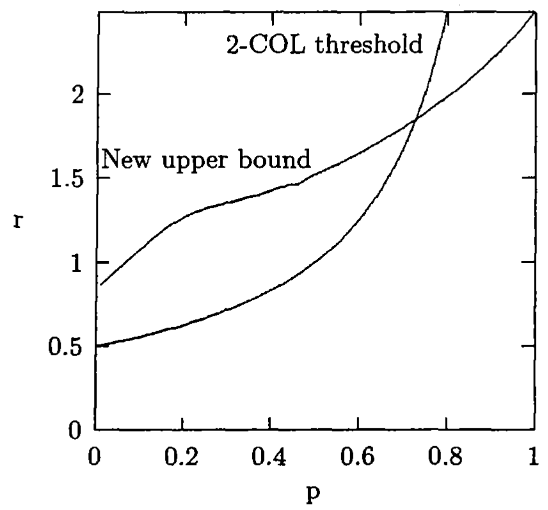

Figure 1: Upper bound on $\bar{r}_{2+p}$

and we refine this grid in each iteration by $\delta = \delta/2$ until the difference between two successive maximum values is less than 10^{-15}. This method allows us to stop optimization process as soon as the value of the objective function is greater or equal to 0.

In figure 1, the obtained upper bound on $\bar{r}_{2+p}$ for each p is compared with the natural bound (1), where the value $\bar{r}_3 = 2.495$ is used. By setting $\Delta p = 0.001$ for $0.72 \leq p \leq 0.73$, we found that the bound (1) is tight up to some $p \leq 0.729$. For $0.73 \leq p \leq 1$ our approach yields a better upper bound on the non-colorability of the 2+p-COL problem.

4 Conclusions

In this paper we have obtained a new upper bound for the non 2+p-colorability threshold for each $p \geq 0.73$. This is the first theoretical approach to this problem. The new upper bound were obtained by applying the method developed by Kaporis et al. for the 3-COL problem to the 2+p-COL problem. We have shown that the natural upper bound based on the non colorability of the embedded 2-COL subproblem introduced by Walsh [8], is tight up to some $p_c \leq 0.729$... This means that the polynomial 2-COL subproblem dominates the 2+p-COL problem's solubility up to some p_c which is now bounded from above by 0.729, and for $0.73 \leq p \leq 1$ the 2+p-COL problem does not behave like random 2-COL problem.

The obtained upper bounds for $0.73 \leq p \leq 1$ yields a more accurate picture of the 2+p-COL phase space. This is an important tool for understanding algorithm behavior, since we can view a complete 3-coloring algorithm as searching trajectories in the 2+p-COL phase space.

In his experimental study of the 2+p-COL problem [8], Walsh obtained that the 2-COL threshold was tight for the 2+p-COL threshold up to $p = 0.8$. In this paper we lowered this value to 0.73. In addition, Walsh observed that this value 0.8 marked a difference in phase transition behavior. For $p < 0.8$ he observed both smooth and sharp regions and for $p > 0.8$ the transition appeared to sharpen significantly. It would be very helpful in order to have a complete understanding of the 2+p-COL problem to obtain some theoretical characterization of how transition sharpens as p goes from 0 to 1.

Acknowledgments

Research supported by the research project "Estudio experimental del problema aleatorio 2+p-COL", SUI, UdeA. This work was done within the framework of the Associateship Scheme of the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

References

[1] P. Cheeseman, B. Kanefsky, and W. Taylor, "Where the really hard problems are." *Proc. of the 12th IJCAI* , pp 331-337, 1991.

[2] D. Mitchel, B. Selman, and H. Levesque, "Hard and easy distribution of SAT problems," *Proc. of the 10th National Conference on Ai*, pp 459-465, 1992.

[3] S. Kirkpatrick and B. Selman, "Critical behaviour in the satisfiability of random boolean expressions," *Science*, 264, pp 1297-1301, 1994.

[4] C. Gomes and, B. Selman, "Problem structure in the presence of perturbation," *Proc. of 14th National Conference on AI*, pp 221-226, 1997.

[5] R. Monasson, R. Zechina, S. Kirkpatrick, B. Selman and L. Troyanski, "2+p-SAT: Relation of typical-case complexity to the nature of the phase transition," *Random Structures and Algorithms* 15, pp 414-435, 1999.

[6] D. Achlioptas, L. Kirousis, E. Kranakis and D. Krizanc, "Rigorous results for $2 + p$-SAT," *Theoretical Computer Science* 265, pp 109-129, 2001.

[7] T. Walsh, "From P to NP: COL, XOR, NAE, 1-in-k, and Horn SAT," *Proc. of AAAI-2002*, 2002.

[8] T. Walsh, "2+p-COL," *Proc. of the Computational Symposium on Graph Colourng and its Generalizations, CP-2002*, Cornell, 2002.

[9] A. C. Kaporis, L. M. Kirousis and Y. C. Stamatiou, "A note on the non-colorability threshold of a random graph," *Electronic Journal of Combinatorics* 7, R29, 2000.

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

[11] V. Chvátal and B. Reed, "Mick gets some (the odds are on his side)," *Proc. 33rd IEEE Symposium on Foundations of Computer Science,* pp 620-627, 1992.

[12] A. Goerdt, "A threshold for unsatisfiability," *Journal of Comput. System Sci.* 53, No 3, pp 469-486, 1996.

[13] W. Fernandez de la Vega, "On random 2-SAT," Manuscript, 1992.

[14] E. Friedgut, "Sharp thresholds of graph properties and the k-SAT problem ," *J. Amer. Math. Soc.* 12, pp 1017-1054, 1999.

[15] A. C. Kaporis, L. M. Kirousis and E. G. Lalas, "Selecting complementary pairs of of literals" *Electronic Notes in Discrete Mathematics*, 16, 2003.

[16] A. C. Kaporis, L. M. Kirousis, Y. C. Stamatiou, M. Vavakari, and M. Zito, "The unsatisfiability threshold revisited". To appear in *Discrete Mathematics.*

[17] D. Achlioptas, E. Friedgut, "A sharp thresholds for k-colorability," *Random Structures and Algorithms* 14, pp 63-70,1999.

[18] D. Achlioptas, C. Moore, "Almost all graphs with average degree 4 are 3-colorable," in *Journal of Comp. Syst. Science*, 6(2), pp 441-471, 2003.

[19] R. Monasson, R. Zechina, S. Kirkpatrick, B. Selman and L. Troyanski, "Determining computational complexity from characteristic phase transitions," *Nature* 400, No. 6740, pp 133-137, 1999.

[20] D. Achlioptas, *Threshold Phenomena in Random Graph Colouring and Satisfiability*, Ph. D. Dissertation, Department of Computer Science, University of Toronto, 1999.

[21] D. Achlioptas M. Molloy, "Almost all graphs with with 2.522n edges are not 3-colourable," *Electronic Journal of Combinatorics* 6, R29, 1999.

[22] A. Kamath, R. Motwani, K. Palem, and P. Spirakis,"Tail bounds for the occupancy and the satisfiability threshold conjecture," *Random Structures and Algorithms* 7, pp 59-80, 1995.

[23] R.M. Corless, G.H. Gonnet, D.E.H. hare, D.J. Jeffrey, and D.E. Knuth, "On the Lambert W function," manuscript, Computer Science Department, University of Waterloo.