Improved bounds for bounded occurrence constraint satisfaction

Johan Håstad
Royal Institute of Technology

May 19, 2015

Abstract

We show how to improve the dependence on the number of occurrences of a variable when approximating CSPs. The result applies when the interesting part of the predicate is odd and says that the advantage is $\Omega\left(\frac{D-1}{2}\right)$ if each variable appears in at most $D$ terms.

1 History of this paper

The result of this paper was obtained independently but after the paper by Barak et al [1]. The underlying ideas of the two proofs are essentially the same but as this note was written and the presentations of [1] and this note are quite different I decided to make a very primitive version of this note available to the public. For a proper academic paper discussion of the problem and its history we refer to [1].

2 The argument

Let the multilinear expansion of the objective function be

$$f(x) = \sum f_\alpha x^\alpha$$

where each term is of degree at most $k$ and each variable appears in at most $D$ terms. We have the following theorem.

Theorem 2.1 There is a constant $d_k$ only depending on $k$ such that in probabilistic polynomial time it is possible to find an assignment $x$ such

$$|f(x)| \geq d_k D^{-1/2} \sum |f_\alpha|.$$ 

We suspect that the methods discussed in [1] are sufficient to derandomize also the algorithm in the current paper.
Proof: We first give the algorithm.

1. Select a random set, $S$, of the variables by including each variable in $S$ with probability one half. For notational convenience we rename the variable $x_i$ to $y_i$ if it is placed in $S$.

2. Select uniformly random values, $y_0^i$ for the variables of $S$. Let $g(x) = f(x, y_0^i)$ be the induced function in the remaining variables.

3. Write $g(x) = g^0(y_0^i) + \sum_i x_i g_i(y_0^i) + G^2(x, y_0^i)$ where $G^2$ contains all terms that remain at least degree two in the $x$-variables.

4. Find a suitable parameter $t$ and set $x_i$ to the sign of $g_i(y_0^i)$ with probability $(1 + t)/2$ for all $i$ independently.

5. If the value obtained is not large enough repeat from step 1.

We return how to find a suitable parameter $t$ later. Let us analyze this procedure. We say that a set $\alpha$ is good if it contains exactly one element of $S$. In expectation we have

$$E \left[ \sum_{\alpha \text{ good}} |f_{\alpha}| \right] = k2^{-k} \sum_{\alpha} |f_{\alpha}|.$$  

The good sets naturally fall in the classes $N_i$ where $\alpha \in N_i$ if it contains only $x_i$ of the variables not in $S$. It is not difficult to see that conditioned on the first two steps the expected value of $f(x, y_0^i)$ is a degree $k$ polynomial, $Q(t)$. If we let $q_1$ denote the coefficient of the linear term we have $q_1 = \sum |g_i(y_0^i)|$. We claim that

$$E\left[|g_i(y_0^i)|\right] \geq c_k E\left[(g_i(y_0^i))^2\right]^{1/2} = c_k \left( \sum_{\alpha \in N_i} f_{\alpha}^2 \right)^{1/2} \geq D^{-1/2} c_k \sum_{\alpha \in N_i} |f_{\alpha}|,$$  

for a constant $c_k$ depending only on $k$. Indeed the first inequality is just saying $\|g_i\|_1 \geq c_k \|g_i\|_2$ which is true as all $L^p$-norms are comparable for degree $k - 1$ polynomials. The last step follows from Cauchy-Schwarz inequality as

$$\sum_{\alpha \in N_i} |f_{\alpha}| \leq \left( \sum_{\alpha \in N_i} 1 \right)^{1/2} \left( \sum_{\alpha \in N_i} f_{\alpha}^2 \right)^{1/2} \leq \sqrt{D} \left( \sum_{\alpha \in N_i} f_{\alpha}^2 \right)^{1/2}.$$

As the union of all $N_i$ give all the good sets we conclude from (1) that

$$E[q_1] = E \left[ \sum_i |g_i(y_0^i)| \right] \geq c_k D^{-1/2} E \left[ \sum_{\alpha \text{ good}} |f_{\alpha}| \right] \geq k2^{-k} c_k D^{-1/2} \sum_{\alpha} |f_{\alpha}|,$$  

(2)
It is easy to see that \( q_1 \leq \sum_\alpha |f_\alpha| \) and thus with probability at least \( \frac{1}{2} k 2^{-k} c_k D^{-1/2} \) we have
\[
q_1 \geq \frac{1}{2} k 2^{-k} c_k D^{-1/2} \sum_\alpha |f_\alpha|.
\]

Now recall Markov brothers’ inequality which says that if \( P \) is a polynomial of degree \( k \) then
\[
\max_{t \in [-1,1]} |P(t)| \geq c'_k P'(0),
\]
where \( c'_k \) is an explicit and known constant. This implies that there is a value of \( t \) such that that
\[
|Q(t)| \geq \frac{1}{2} k 2^{-k} c'_k c_k D^{-1/2} \sum_\alpha |f_\alpha|.
\]

Once \( y^0 \) is chosen, the polynomial \( Q \) is completely explicit and hence it is possible to find (a very good approximation of) the optimal \( t \). In particular it is possible, in polynomial time, to find a value of \( t \) which satisfies \( |Q(t)| \geq \frac{1}{3} k 2^{-k} c'_k c_k D^{-1/2} \sum_\alpha |f_\alpha| \). Setting \( d_k = \frac{1}{4} k 2^{-k} c'_k c_k \) it is easy to see that the algorithm succeeds with a good probability.

Let us end with some brief comments. First observe that there are several ways to choose a suitable \( t \). As one alternative one could use a uniformly random \( t \) and use an inequality of the form then
\[
E_{t \in [-1,1]}[|P(t)|] \geq c''_k P'(0),
\]
which is clearly true for some value of \( c''_k \) even if I do not know a reference and the best value for \( c''_k \). As a further alternative [1] uses extrema of Chebychev polynomials as a set of possible choices for \( t \).

We might be interested in finding an assignment such that \( f(x) \) is large and positive. If \( f \) is odd this is easy since we can simply negate an \( x \) giving a large negative value. In the general case this is not possible and indeed, as also observed in [1], this is a real problem as can be seen from the following example.

Let us take Max-Cut where we score 2 for cut edges and \(-2\) for not cut edges. Consider the complete graph on \( D + 1 \) vertices. We have the objective function
\[
-2 \sum_{i > j} x_i x_j = (\sum x_i^2) - (\sum x_i)^2 \leq D + 1
\]
and thus the global optimum is at most \( D + 1 \). On the other hand the sum of the absolute values of all coefficients is \( \Omega(D^2) \) (and the global maximum of the absolute value is also \( \Omega(D^2) \)). This implies that to get a general result for arbitrary predicates that we can beat the trivial approximation ratio by an additional term \( \Omega(D^{-1/2}) \) we must use that fact the global maximum is large. Thus something new is needed.
References