Electronic Colloquium on Computational Complexity, Report No. 121 (2006)

A Simple Biased Distribution for Dinur's Construction

Charanjit S. Jutla
IBM T. J. Watson Research Center
Yorktown Heights, NY 10598-704
September 14, 2006

1. Introduction

The Dinur construction [Din05] achieves gap amplification, by repeatedly applying first a powering construction – which increases the gap, but also increases the alphabet size – and then applying a construction to reduce the alphabet size (which diminishes the gap but not by too much). The latter construction is based on long codes and is not the focus of the current paper.

It is the powering construction which will be of interest to us. Essentially, the powering construction works by taking a constraint graph, and building a t-fold power graph, which collects all paths of length t emanating from a node. The proof (i.e. the labelling of the constraint graph) now must supply versions of labels of nodes collected at each composite node – of course, the constraints between the composite nodes have accordingly multiplied.

This is akin to doing a parallel repetition theorem (as proven by Raz [Raz98]), but there are subtle differences, leading to considerable difference in analysis. Currently, to obtain the best parameters for the PCP theorem, it must use the Ben-Sasson Sudan construction [BSS06] along with the Dinur construction.

Our simplified construction, apart from obviously using ideas from Dinur's construction, also uses some ideas from Jaikumar's alternate proof [Rad05], and its latest extension [RS06]. Like Jaikumar's proof, we obtain better bounds (even better than Jaikumar-Sudan). Instead of using binomial distributions (as in Dinur) and various properties of lazy random walks (as in Jaikumar), our proof uses a simple biased distribution and invokes the expander property only to bound the variance.

In Dinur's construction [Din05], $t \log d$ bits of randomness are required to choose the constraints in the t-fold powering of a Σ-Constraint graph, while the amplification factor is about $\sqrt{t}/|\Sigma|^2$. In [RS06], $5t(\log t + \log d)$ bits of randomness are required, while the amplification factor is about t. In our construction, $2\log t + t \log d$ bits of randomness are required to achieve an amplification of about t.

2. Powering Using a Biased Distribution

A Σ-Constraint graph $G = (V, E, C)$ comes with a function $C : E \times \Sigma \times \Sigma \rightarrow \{0, 1\}$. An edge $e = (u, v)$ is called *consistent* w.r.t. an assignment $\sigma : V \rightarrow \Sigma$ if the corresponding constraint, i.e. $C(e, \sigma(u), \sigma(v))$, evaluates to one. Given such a graph, the *Constraint Satisfaction Problem* is to find an assignment $\sigma : V \rightarrow \Sigma$ for the vertices, such that all edges $e \in E$ are consistent w.r.t. σ. The *Constraint Maximization Problem* is to find an assignment $\sigma : V \rightarrow \Sigma$ which maximizes the number of consistent edges.

ISSN 1433-8092

Given a d-regular Σ-Constraint graph $G = (V, E, C)$, we define a Σ^{d^t}-Constraint graph $G^t = (V, E', C')$ as follows. The vertices of G^t will be V itself. Let the set of paths in G be called $\hat{E}$. The edges E' in G^t will be paths in G (with the terminal vertices of the path being the incident vertices of the edge), but with a distribution to be specified below. In effect, the distribution on $\hat{E}$ assigns a multiplicity to each path, and hence the set of edges E' can be considered as (multiple, and sometimes zero) copies of paths of G, with each copy having the same constraint. The distribution on edges in G^t will be defined differently from [Din05].

The constraints $C' : \hat{E} \times \Sigma^{d^t} \times \Sigma^{d^t} \rightarrow \{0, 1\}$ of G^t are defined as follows. It is easier to just define when an edge $\hat{e} \in \hat{E}$ is inconsistent with an assignment $\hat{\sigma} : V \rightarrow \Sigma^{d^t}$. Given a label $l \in \Sigma^{d^t}$ of a vertex v, it can be viewed as a labelling by Σ of each vertex u that can be reached from v by a path in G of length at most t (as G is a degree d regular graph). The (Σ-)projection of the Σ^{d^t}-label l to this node u will be called l_u. Given an edge in $\hat{E}$, i.e. a path $\langle u_1, u_2, ..., u_s \rangle$, it is **inconsistent** w.r.t. $\hat{\sigma}$ if either (i) there is a vertex u_i on the path, and $\hat{\sigma}(u_1)_{u_i} \neq \hat{\sigma}(u_s)_{u_i}$, or (ii) there is an edge $e = \langle u_i, u_{i+1} \rangle$ on the path such that $C(e, \hat{\sigma}(u_1)_{u_i}\ , \hat{\sigma}(u_s)_{u_{i+1}}) = 0$.

Let $T2$ stand for $[-t/2..t/2]$, $T4$ stand for $[-t/4..t/4]$, and $T8$ stand for $[-t/8..t/8]$.

Let $\mathcal{D}$ be the distribution assigning probability $(t/2 - |l| + 1)/(t/2 + 1)^2$ to $l \in T2$. (We remark here that the distribution $\mathcal{D}$ gives a weight to $l \in T2$ in proportion to the number of ways l can be written as $j1 + j2$ (with order) with $j1, j2$ in $T4$.) The *distribution on the edges* of G^t, i.e. $\hat{E}$, is defined as follows: first pick l according to $\mathcal{D}$, and then uniformly pick a path of length $t + l$ (nodes).

We will show that given any d-regular Σ-Constraint graph $G = (V, E, C)$ with $\lambda = \lambda(G)$ (the second largest eigenvalue), if for every assignment $\sigma : V \rightarrow \Sigma$ at least a fraction ϵ of the edges are inconsistent, then in the graph G^t defined above, for every assignment $\hat{\sigma} : V \rightarrow \Sigma^{d^t}$ at least a fraction $\min\{1/256,\ t\epsilon * (1 - \lambda/d)/2048\}$ of the edges are inconsistent.

3. Proof Idea

Experts in the field may safely skip this section.

Ideally, a uniform distribution on length t paths ought to suffice, as it does if we were interested in amplifying the success probability of a randomised algorithm from ϵ to $t\epsilon$. However, the additional universal quantifier in the game (i.e. for-all assignments) makes the analysis difficult, unless a biased distribution is used.

To illustrate, the proof usually works by considering an assignment $\hat{\sigma}$ of vertices in G^t. Each vertex u has a label $\hat{\sigma} \in \Sigma^{d^t}$, which is considered as a version of Σ-labels of all vertices v reachable from u by length t paths. If for each vertex v, the versions of its Σ-labels kept at all the vertices are identical, then we are dealing with a simple t-fold parallel assignment checking, albeit on a path on a d-regular expander. Standard techniques show that this indeed leads to about a t-fold amplification of inconsistency. If the versions of Σ-labels of a vertex kept at different vertices are different, the standard trick it to consider an assignment $\sigma : V \rightarrow \Sigma$ obtained from $\hat{\sigma}$ by defining the label of a vertex to be its version which is kept at maximum number of vertices. Thus, $\Pr[\sigma(v) = \hat{\sigma}(u)_v] \geq 1/|\Sigma|$, for u chosen uniformly from vertices reachable by length ($\leq$) t paths from v (as these are the only vertices u which keep a version of v).

However, if one has uniformly chosen a path of length t (to be a constraint in G^t), and is interested in showing that the path is inconsistent in G^t given that the i-th edge $e = \langle u_i, u_{i+1} \rangle$ on the path is inconsistent w.r.t. σ, one needs to show $C(e, \hat{\sigma}(u_1)_{u_i}\ , \hat{\sigma}(u_t)_{u_{i+1}}) = 0$ (see definition above). This would follow if $\hat{\sigma}(u_1)_{u_i} = \sigma(u_i)$ and $\hat{\sigma}(u_t)_{u_{i+1}} = \sigma(u_{i+1})$. This happens with good probability if u_1 was chosen uniformly as above, but if u_1 is not chosen in a way which is close to uniform, our analysis becomes weak. However, given u_i to be an i-th node on a random path starting at u_1, makes the distribution of u_1 as nodes reachable by (exact) i length paths.

The trick is to not to try to show the i-th edge on the path to be inconsistent, but one of a range of edges on the path to be inconsistent. Instead of trying to maximize the probability over this range, one proves that the average inconsistency is high (a standard trick in probabilistic combinatorics). Further, $\sigma : V \rightarrow \Sigma$ is now defined by not maximizing over all vertices reachable by $(\leq)$ t paths, but by paths of length close to t.

To obtain this range of edges, the Dinur construction considers length (exact) t paths over a graph with *self loops*. The Jaikumar paper considers random walks which stop early with some probability (lazy random walks). In this paper we attempt to capture what is exactly required, at least in the method outlined above, and hence obtain a simpler proof with better bounds.

4. Proof

We first define a distribution $\mathcal{A}$ on $\hat{E} \times T4$ by augmenting the above distribution on $\hat{E}$ as follows.

1. Pick l according to $\mathcal{D}$ from $T2$, and then uniformly pick a path of length $t+l$ steps (nodes), say $\langle u_1, ..., u_{t+l} \rangle$
2. Pick $j1$ uniformly from $[\max(l - t/4,\ -t/4)..\min(l + t/4,\ t/4)]$. Note, if $l \geq 0$, $j1$ is picked uniformly from $[l - t/4..t/4]$, and if $l \leq 0$, $j1$ is picked uniformly from $[-t/4..t/4 - |l|]$. Thus regardless, $j1$ is in $T4$.
3. For technical reasons, define $j2 = l - j1$. For a given l, the range for $j2$ can be seen to be same as that for $j1$.

The path $\langle u_1, ..., u_{t+l} \rangle$ and the value $j1$ specified above is in $\hat{E} \times T4$, and the above sampling procedure defines a distribution on $\hat{E} \times T4$.

We next show that the above distribution $\mathcal{A}$ is also identical to each of the following distributions $\mathcal{B}_i$ on $\hat{E} \times T4$, for $i \in T8$.

1. Uniformly pick a random edge $\langle u, v \rangle$ from E.
2. Uniformly pick $j1$ from $T4$.
3. Uniformly pick $j2$ from $T4$.
4. Uniformly pick a random path of $t/2 + i + j1$ steps originating at u and terminating at say u_1.
5. Uniformly pick a random path of $t/2 - i + j2$ steps originating at v and terminating at say $u_{t+j1+j2}$.

The claim that the two distributions $\mathcal{A}$ and $\mathcal{B}_i$ are identical would follow if for all $a, b \in T4$, $\Pr_{\mathcal{A}}[j1 = a \wedge j2 = b]$ is $1/|T4|^2$. But, this is same as $\Pr_{\mathcal{A}}[j1 = a \wedge l = a + b]$, which is same as

$$\frac{1}{t/2 - |a+b| + 1} \cdot \frac{t/2 - |a+b| + 1}{(t/2+1)^2}$$

Given a labelling $\hat{\sigma}$ of G^t, we now define a majority function to define a labelling $\sigma : V \rightarrow \Sigma$ for the vertices in G. Let

$$p_u = \max_{a \in \Sigma} \Pr[\hat{\sigma}(w)_u = a]$$

where the probability is over first choosing a length s uniformly in $[t/2 - t/8..t/2 + t/8]$, and then choosing a path of length s steps starting at u and terminating at, say w. Then, $\sigma(u)$ is *defined* to be an a (say, least in Σ) which attains this maximum probability p_u. Let F be edges in G which are inconsistent w.r.t. the newly defined assignment σ.

For $i \in T8$, we say that an event A_i happens (with indicator variable X_i) on a path $\langle u_1, ..., u_{t+l} \rangle$ ($l \in T2$) and value $j \in T4$ if

1. $\langle u_{t/2+i+j}, u_{t/2+i+j+1} \rangle$ is (defined and) in F (call this edge $\langle u, v \rangle$),
2. (a) $\hat{\sigma}(u_1)_u = \sigma(u)$ and $\hat{\sigma}(u_{t+l})_v = \sigma(v)$, OR
   (b) $\hat{\sigma}(u_1)_u \neq \hat{\sigma}(u_{t+l})_u$, OR
   (c) $\hat{\sigma}(u_1)_v \neq \hat{\sigma}(u_{t+l})_v$.

In [Din05] and [Rad05] only conditions (1) and 2(a) were considered. The more advanced condition above is from [RS06], and we borrow portions of the analysis in the next paragraph also from [RS06].

For each fixed $i \in T8$, the cardinality of the range of values of $t/2 + i + j1$ in $\mathcal{B}_i$ is $t/2 + 1$. Further, $[t/2 - t/8..t/2 + t/8]$ is contained in this range. The latter set has size $t/4 + 1$. Thus, the probability under $\mathcal{B}_i$, that $t/2 + i + j1$ falls in $[t/2 - t/8..t/2 + t/8]$ is at least $1/2$, as $j1$ is picked uniformly. Thus, $\Pr_{\mathcal{B}_i}[\hat{\sigma}(u_1)_u = \sigma(u)] \geq 1/2 * p_u$. Similarly, $\Pr_{\mathcal{B}_i}[\hat{\sigma}(u_{t+l})_v = \sigma(v)] \geq 1/2 * p_v$. Moreover, these two events are independent, as $j1$ and $j2$ are chosen independently, and u_1 only depends on u, and $u_{t+j1+j2}$ only depends on v (and thus also independent of $\langle u, v \rangle$ being in F). Thus for each i, probability of (1) and 2(a) is at least $p_u p_v |F|/(4|E|)$. We claim that probability of (1) and 2(b) is at least $(1 - p_u)|F|/(2|E|)$. As before, the probability of $t/2 + i + j1$ falling in $[t/2 - t/8..t/2 + t/8]$ is at least $1/2$. Conditioned on this, and for any fixed value, say b, of $\hat{\sigma}(u_{t+l})_u$, the probability of $\hat{\sigma}(u_1)_u$ not being equal to b is at least the probability of $\hat{\sigma}(u_1)_u$ not being equal to the majorizing value, i.e. $\sigma(u)$, and the claim follows because of independence as before. Thus,

$$\Pr_{\mathcal{B}_i}[X_i > 0] \geq \frac{|F|}{4|E|} * \max\{p_u p_v, 2(1 - p_u), 2(1 - p_v)\} \geq \frac{\sqrt{3} - 1}{4} \frac{|F|}{|E|} \geq \frac{|F|}{8|E|}$$

Thus by linearity of expectation (since each $\mathcal{B}_i$ is same as $\mathcal{A}$),

$$E_{\mathcal{A}}[(\Sigma_{i \in T8} X_i)] \geq \frac{t}{32} \cdot \frac{|F|}{|E|}$$

We now upper bound $E_{\mathcal{A}}[(\sum_{i \in T8} X_i)^2]$. For this it is more convenient to use the distribution (or sampling procedure) $\mathcal{A}$. We just bound $\sum_{i < i'} E_{\mathcal{A}}[X_i X_{i'}]$. Choose $\hat{e}, j1$ using distribution $\mathcal{A}$. Then $E_{\mathcal{A}}[X_i X_{i'}]$ is upper bounded by

$$\sum_{L=-t/2}^{L=t/2} \sum_{J1 \in T4} p(\hat{e}, j1, i, i') * \Pr_{\mathcal{A}}[j1 = J1 \wedge l = L]$$

where $p(\hat{e}, j1, i, i')$ is the probability of both the $(t/2 + i + j1)$th and the $(t/2 + i' + j1)$th edge of the path $\hat{e}$ being in F (notice that $(t/2 + i + j1)$ and $(t/2 + i' + j1)$ are both less than $t + l$). This probability $p(\hat{e}, j1, i, i')$ is upper bounded by $(|F|/|E|) * (|F|/|E| + (\lambda/d)^{i'-i})$ (see e.g. Prop. 2.4 in [Din05]). Thus, $E_{\mathcal{A}}[X_i X_{i'}]$ is upper bounded by same.

Thus,

$$E_{\mathcal{A}}[(\Sigma_i X_i)^2] \leq E_{\mathcal{A}}[\Sigma X_i] + 2(t/4) \cdot \frac{|F|}{|E|} \cdot \frac{1}{1 - (\lambda/d)} + (t/4)^2 \cdot \frac{|F|^2}{|E|^2}$$

Hence $\Pr_{\mathcal{A}}[(\sum X_i) > 0] \geq \frac{E[(\sum X_i)]^2}{E[(\sum X_i)^2]} \geq \min\{1/256,\ t * (|F|/2048|E|) * (1 - \lambda/d)\}$

Now consider constraints in the composite graph to be not just paths in $\hat{E}$, but elements of $\hat{E} \times T4$, and such constraints picked according to $\mathcal{A}$. Recall that a path $\langle u_1, ..., u_{t+l}\rangle$ is inconsistent if either (i) there is a w on the path, and $\hat{\sigma}(u_1)_w \neq \hat{\sigma}(u_{t+l})_w$, or (ii) there is an edge $\langle w, z\rangle$ on the path such that $\hat{\sigma}(u_1)_w$ and $\hat{\sigma}(u_{t+l})_z$ are inconsistent as labels of edge $\langle w, z\rangle$ in G. But, this is implied by conditions ((1) and (2a)) or (2b) or (2c) above holding on the chosen path for any $i \in T8$. Thus it is implied by $(\sum_i X_i) > 0$. However, in picking an element of $\hat{E} \times T4$ (when picking constraints), the picking of $j1$ can be ignored, as it has no effect on whether the path is inconsistent. □

As an added bonus, the amount of randomness required to pick the composite constraints is only $2 \log t + t \log d$ bits (first term for picking l, and second for picking a path of average length t), whereas in [RS06] the randomness required is $5t(\log t + \log d)$.

References

[BSS06] Eli Ben-Sasson and Madhu Sudan. Short PCPs with poly-log rate and query complexity. In *Sympsium on Theory of Computation*, 2006.

[Din05] Irit Dinur. The PCP Theorem by Gap Amplification. In *ECCC Reports TR05-046*, 2005. Also in STOC 2006.

[Rad05] Jaikumar Radhakrishnan. Gap Amplification Using Lazy Random Walks. In *ECCC Reports TR05-046*, 2005. Also in ICALP 2006.

[Raz98] Ran Raz. A Parallel Repetition Theorem. *SIAM Journal of Computing*, 27(3):763–803, 1998.

[RS06] Jaikumar Radhakrishnan and Madhu Sudan. On Dinur's Proof of the PCP Theorem. In *to appear in Bulletin of AMS*, 2006.