Last updated 3/10/2016

Lectures 14: ARV Analysis, part 3

In which we complete the analysis of the ARV rounding algorithm

We are finally going to complete the analysis of the Arora-Rao-Vazirani rounding algorithm, which rounds a Semidefinite Programming solution of a relaxation of sparsest cut into an actual cut, with an approximation ratio $O(\sqrt{\log |V|})$.

In previous lectures, we reduced the analysis of the algorithm to the following claim.

Lemma 1 *Let $d(\cdot,\cdot)$ be a semi-metric over a set C such that $d(u,v) \leq 1$ for all $u,v \in C$, let $\{\mathbf{x}_v\}_{v\in C}$ be a collection of vectors in $\mathbb{R}^m$, such that $d(i,j) := ||\mathbf{x}_u - \mathbf{x}_v||^2$ is a semimetric, let $\mathbf{g}$ be a random Gaussian vector in $\mathbb{R}^m$, define $Y_v := \langle \mathbf{g}, \mathbf{x}_v \rangle$, and suppose that, for every $\mathbf{g}$, we can define a set of disjoint pairs $M_{\mathbf{g}}$ such that, with probability 1 over $\mathbf{g}$,*

$$\forall \{u,v\} \in M_{\mathbf{g}}.\ |Y_u - Y_v| \geq \sigma \ \wedge \ d(u,v) \leq \ell$$

and

$$\forall u \in C.\ \mathbb{P}[\exists v. \{u,v\} \in M_{\mathbf{g}}] \geq \epsilon$$

Then

$$\ell \geq \Omega_{\epsilon,\sigma}\left(\frac{1}{\sqrt{\log |C|}}\right)$$

1 An Inductive Proof that Gives a Weaker Result

In this section we will prove a weaker lower bound on ℓ, of the order of $\frac{1}{(\log |C|)^{\frac{2}{3}}}$. We will then show how to modify the proof to obtain the tight result.

We begin will the following definitions. We define the ball or radius r centered at u as

$$B(u,r) := \{v \in C.\ d(u,v) \leq r\}$$

We say that a point $u \in C$ has the (p, r, δ)-Large-Projection-Property, or that it is (p, r, δ)-LPP if

$$\mathbb{P}\left[\max_{v \in B(u,r)} Y_v - Y_u \geq p\right] \geq \delta$$

Lemma 2 *Under the assumptions of Lemma 1, there is a constant $c_4 > 0$ (that depends only on ϵ and σ) such that for all $t \leq c_4 \cdot \frac{1}{\sqrt{\ell}}$, at least $\left(\frac{\epsilon}{8}\right)^t \cdot |C|$ elements of C have the $\left(t\frac{\sigma}{2}, t\ell, 1 - \frac{\epsilon}{4}\right)$ Large Projection Property.*

Proof: We will prove the Lemma by induction on t. We call C_t the set of elements of C that are $\left(t\frac{\sigma}{2}, t\ell, 1 - \frac{\epsilon}{4}\right)$-LPP

Let $M'_{\mathbf{g}}$ be the set of *ordered* pairs (u, v) such that $\{u, v\} \in M_{\mathbf{g}}$ and $Y_v > Y_u$, and hence $Y_v - Y_u \geq \sigma$. Because $\mathbf{g}$ and $-\mathbf{g}$ have the same distribution, we have that, for every $u \in C$, there is probability $\geq \epsilon/2$ that there is a $v \in C$ such that $(v, u) \in M'_{\mathbf{g}}$ (a fact that we will use in the inductive step).

For the base case $t = 0$ there is nothing to prove.

For the inductive case, define the function $F : C_t \to C$ (which will be a random variable dependent on $\mathbf{g}$) such that $F(v)$ is the lexicographically smallest $w \in B(v, t\ell)$ such that $Y_w - Y_v \geq \sigma$ if such a w exists, and $F(v) = \perp$ otherwise. The definition of C_t is that $\mathbb{P}[F(v) \neq \perp] \geq 1 - \epsilon/4$ for every $v \in C_t$, and the inductive assumption is that $|C_t| \geq |C| \cdot (\epsilon/8)^t$.

By a union bound, for every $v \in C_t$, there is probability at least $\epsilon/4$ that there is an $u \in C$ such that $(u, v) \in M'_{\mathbf{g}}$ and $F(v) = w \neq \perp$. In this case, we will define $F'(u) = w$, otherwise $F'(u) = \perp$.

Note that the above definition is consistent, because $M'_{\mathbf{g}}$ is a set of disjoint pairs, so for every u there is at most one v that could be used to define $F'(u)$. We also note that, if $F'(u) = w \neq \perp$, then

$$Y_w - Y_u \geq t \cdot \frac{\sigma}{2} + \sigma \ ,$$

$$d(u, w) \leq (t + 1) \cdot \ell$$

and

$$\sum_{u \in C} \mathbb{P}[F'(u) \neq \perp] = \sum_{v \in C_t} [F(v) \neq \perp \wedge \exists u.(u, v) \in M'_{\mathbf{g}}] \geq |C_t| \cdot \frac{\epsilon}{4}$$

Now we can use another averaging argument to say that there have to be at least $|C_t| \cdot \frac{\epsilon}{8}$ elements u of C such that

$$\mathbb{P}[F'(u) \neq \perp] \geq \frac{\epsilon}{8} \cdot \frac{|C_t|}{|C|} \geq \left(\frac{\epsilon}{8}\right)^{t+1}$$

Let us call C_{t+1} the set of such element. As required, $|C_{t+1}| \geq |C| \cdot (\epsilon/8)^{t+1}$.

By applying concentration of measure, the fact that, for every $u \in C_{t+1}$ we have

$$\mathbb{P}\left[\max_{w \in B(u,(t+1)\cdot\ell)} Y_w - Y_u \geq (t+1)\frac{\sigma}{2} + \frac{\sigma}{2}\right] \geq \left(\frac{\epsilon}{8}\right)^{t+1}$$

implies that, for every $u \in C_{t+1}$

$$\mathbb{P}\left[\max_{w \in B(u,(t+1)\cdot\ell)} Y_w - Y_u \geq (t+1)\frac{\sigma}{2} + \frac{\sigma}{2} - c_3\sqrt{\log\frac{4 \cdot 8^{t+1}}{\epsilon^{t+2}}}\sqrt{(t+1)\cdot\ell}\right] \geq 1 - \frac{\epsilon}{4}$$

and the inductive step is proved, provided

$$\frac{\sigma}{2} \geq c_3\sqrt{(t+2)\cdot\log\frac{8}{\epsilon}}\sqrt{(t+1)\cdot\ell}$$

which is true when

$$t + 2 \leq \frac{\sigma}{2c_3\sqrt{\log 8/\epsilon}} \cdot \frac{1}{\sqrt{\ell}}$$

which proves the lemma if we choose c_4 appropriately. $\square$

Applying the previous lemma with $t = c_4/\sqrt{\ell}$, we have that, with probability $\Omega(1)$, there is a pair u, v in C such that

$$Y_v - Y_u \geq \Omega(1/\sqrt{\ell})$$

and

$$d(u,v) \leq O(\sqrt{\ell})$$

but we also know that, with $1 - o(1)$ probability, for all pairs u, v in C,

$$|Y_v - Y_u|^2 \leq O(\log|C|) \cdot d(i,j)$$

and so

$$\frac{1}{\ell} \leq O(\log|C|)\sqrt{\ell}$$

implying

$$\ell \geq \Omega\left(\frac{1}{(\log|C|)^{2/3}}\right)$$

2 The Tight Bound

In the result proved in the previous section, we need $\frac{\sigma}{2}$, which is a constant, to be bigger than the loss incurred in the application of concentration of measure, which is of the order of $t\sqrt{\ell}$. A factor of $\sqrt{t\ell}$ simply comes from the distances between the points that we are considering; an additional factor of $\sqrt{t}$ comes from the fact that we need to push up the probability from a bound that is exponentially small in t.

The reason for such a poor probability bound is the averaging argument: each element of C_t has probability $\Omega(1)$ of being the "middle point" of the construction, so that the sum over the elements u of C of the probability that u has $F'(u) \neq \perp$ adds up to $\Omega(|C_t|)$; such overall probability, however, could be spread out over all of C, with each element of C getting a very low probability of the order of $|C_t|/|C|$, which is exponentially small in t.

Not all elements of C, however, can be a u for which $F'(u) \neq \perp$; this is only possible for elements u that are within distance ℓ from C_t. If the set $\Gamma_\ell(C_t) := \{u : \exists v \in C_t : d(u,v) \leq \ell\}$ has cardinality of the same order of C_t, then we only lose a constant factor in the probability, and we do not pay the extra $\sqrt{t}$ term in the application of concentration of measure. But what do we do if $\Gamma_\ell(C_t)$ is much bigger than C_t? In that case we may replace C_t and $\Gamma_\ell(C_t)$ and have similar properties.

Lemma 3 *Under the assumptions of Lemma 1, if $S \subseteq C$ is a set of points such that for every $v \in S$*

$$\mathbb{P}\left[\max_{w \in B(v,d)} Y_w - Y_v \geq p\right] \geq \epsilon$$

then, for every distance D, every $k > 0$, and every $u \in \Gamma_D(S)$

$$\mathbb{P}\left[\max_{w \in B(u,d+D)} Y_w - Y_u \geq p - \sqrt{D} \cdot k\right] \geq \epsilon - e^{-k^2/2}$$

That is, if all the elements of S are (p,d,ϵ)-LPP, then all the elements of $\Gamma_D(S)$ are $(p - k\sqrt{D}, d+D, \epsilon - e^{-k^2/2})$-LPP.

Proof: If $u \in \Gamma_D(S)$, then there is $v \in S$ such that $d(u,v) \leq D$, and, with probability $1 - e^{-k^2/2}$ we have $Y_u - Y_v \leq \sqrt{D} \cdot k$. The claim follows from a union bound. $\Box$

Lemma 4 *Under the assumptions of Lemma 1, there is a constant $c_5 > 0$ (that depends only on ϵ and σ) such that for all $t \leq c_5 \cdot \frac{1}{\ell}$, there is a set $C_t \subseteq C$ such that $|C_t| \geq |C| \cdot (\epsilon/8)^t$, every element of C_t is $\left(t \cdot \frac{\sigma}{4}, \left(2t + \log_{\frac{8}{\epsilon}} \frac{|C_t|}{|C|}\right) \cdot \ell, 1 - \frac{\epsilon}{4}\right)$-LPP, and*

$$|\Gamma_\ell(C_t)| \leq \frac{8}{\epsilon}|C_t|$$

Proof: The base case $t = 0$ is proved by setting $C_0 = C$.

For the inductive step, we define $F(\cdot)$ and $F'(\cdot)$ as in the proof of Lemma 2. We have that if $F'(u) = w \neq \perp$, then

$$Y_w - Y_u \geq t \cdot \frac{\sigma}{4} + \sigma \ ,$$

$$d(u, w) \leq \left(2t + \log_{\frac{8}{\epsilon}} \frac{|C_t|}{|C|}\right) \cdot \ell + \ell \ ,$$

and

$$\sum_{u \in C} \mathbb{P}[F'(u) \neq \perp] = \sum_{v \in C_t} [F(v) \neq \perp \wedge \exists u.(u, v) \in M'_{\mathbf{g}}] \geq |C_t| \cdot \frac{\epsilon}{4}$$

Now we can use another averaging argument to say that there have to be at least $|C_t| \cdot \frac{\epsilon}{8}$ elements u of C such that

$$\mathbb{P}[F'(u) \neq \perp] \geq \frac{\epsilon}{8} \cdot \frac{|C_t|}{|\Gamma_\ell(C_t)|} \geq \left(\frac{\epsilon^2}{64}\right)$$

Let us call $C^{(0)}_{t+1}$ the set of such elements.

Define $C^{(1)}_{t+1} := \Gamma_\ell(C^{(0)}_{t+1})$, $C^{(2)}_{t+1} := \Gamma_\ell(C^{(1)}_{t+1})$, and so on, and let k be the first time such that $|C^{(k+1)}_{t+1}| \leq \frac{8}{\epsilon}|C^{(k)}_{t+1}|$. We will define $C_{t+1} := C^{(k)}_{t+1}$. Note that

$$|C_{t+1}| \geq \left(\frac{8}{\epsilon}\right)^k \cdot |C^{(0)}_{t+1}| \geq \left(\frac{8}{\epsilon}\right)^{k-1} \cdot |C_t| \geq \left(\frac{8}{\epsilon}\right)^{k-1-t} |C|$$

which implies that $k \leq t + 1$.

We have $|C_{t+1}| \geq |C^{(0)}_{t+1}| \geq \frac{\epsilon}{8}|C_t|$ so we satisfy the inductive claim about the size of C_t. Regarding the other properties, we note that $C_{t+1} \subseteq \Gamma_{k\ell}(C^{(0)}_{t+1})$, and that every element of $C^{(0)}_{t+1}$ is

$$\left(t\frac{\sigma}{4} + \sigma, \left(2t + 1 + \log_{\frac{8}{\epsilon}} \frac{|C_t|}{|C|}\right) \cdot \ell, \frac{\epsilon^2}{64}\right) - \text{LPP}$$

so we also have that every element of C_{t+1} is

$$\left(t\frac{\sigma}{4} + \frac{\sigma}{2}, \left(2t + 1 + k + \log_{\frac{8}{\epsilon}} \frac{|C_t|}{|C|}\right) \cdot \ell, \frac{\epsilon^2}{128}\right) - \text{LPP}$$

provided

$$\frac{\sigma}{2} \geq \sqrt{2 \log \frac{128}{\epsilon^2}} \cdot k\ell$$

which we can satisfy with an appropriate choice of c_4, recalling that $k \leq t + 1$.

Then we apply concentration of measure to deduce that every element of C_{t+1} is

$$\left(t\frac{\sigma}{4}+\frac{\sigma}{4},\left(2t+1+k+\log_{\frac{8}{\epsilon}}\frac{|C_t|}{|C|}\right)\cdot\ell,1-\frac{\epsilon}{4}\right)-\mathrm{LPP}$$

provided that

$$\frac{\sigma}{4}\geq c_3\sqrt{\log\frac{512}{\epsilon^3}}\cdot\left(2t+1+k+\log_{\frac{8}{\epsilon}}\frac{|C_t|}{|C|}\right)\cdot\ell$$

which we can again satisfy with an appropriate choice of c_4, because $k\leq t+1$ and $\log_{\frac{8}{\epsilon}}\frac{|C_t|}{|C|}$ is smaller than or equal to zero.

Finally,

$$2t+1+k+\log_{\frac{8}{\epsilon}}\frac{|C_t|}{|C|}\leq 2t+2+\log_{\frac{8}{\epsilon}}\frac{|C_{t+1}|}{|C|}$$

because, as we established above,

$$|C_{t+1}|\geq\left(\frac{8}{\epsilon}\right)^{k-1}|C_t|$$

□

By applying Lemma 4 with $t=\Omega(1/\ell)$, we find that there is $\Omega(1)$ probability that there are u,v in C such that

$$Y_j-Y_i\geq\Omega(1/\ell)$$

$$d(i,j)\leq 1$$

$$|Y_i-Y_j|^2\leq O(\log n)\cdot d(i,j)$$

which, together, imply

$$\ell\geq\Omega\left(\frac{1}{\sqrt{\log n}}\right)$$