Communication Complexity 11 Nov, 2011 (@ IMSc)

Lec. 23: The gap-Hamming problem (part II)

Lecturer: Meena Mahajan *Scribe: Raja S*

Summary

In this lecture, we have proved a lower bound of $\Omega(n)$ for the gap Hamming distance (GHD_n) problem. This result was first proved in [CR11] and followed by several other proofs [Vid11, She11]. We have followed the proof given by A.A.Sherstov [She11], which uses the corruption bound from problem set 2 (see Appendix).

Theorem 23.1 ([She11]). *$R_{\frac{1}{3}}(\text{ORT}_n) = \Omega(n)$.*

Corollary 23.2 ([She11]). *$R_{\frac{1}{3}}(\text{GHD}_n) = \Omega(n)$.*

23.1 Proof outline

1. To prove $R_{\frac{1}{3}}(\text{GHD}_n) = \Omega(n)$, we define another problem called gap orthogonality ORT_n and we reduce ORT_n to GHD_N. It then suffices to prove $R_{\frac{1}{3}}(\text{ORT}_n) = \Omega(n)$.
2. To prove $R_{\frac{1}{3}}(\text{ORT}_n) = \Omega(n)$, by Yao's lemma, it suffices to prove $D^{\mu}_{\frac{1}{3}}(\text{ORT}_n) = \Omega(n)$ for some μ. We choose μ to be uniform.
3. We use the corruption bound to prove $R_{\frac{1}{3}}(\text{ORT}_n) = \Omega(n)$. In order to use this, we need to show that
   (a) $\mu(\text{ORT}^{-1}(+1))$ is large (i.e., $\Theta(1)$), and
   (b) there exists a small enough ε such that any rectangle that is not ε-1-corrupted must be small. That is, $\forall S, T \subseteq \{-1,+1\}^n$, if $\mu(\text{ORT}^{-1}(+1) \cap (S \times T)) \leq \varepsilon\mu(\text{ORT}^{-1}(-1) \cap (S \times T))$ then $\mu(S \times T) \leq \exp(-\Omega(n))$.

23.2 Definitions

The gap orthogonality problem is defined as follows. Let $x, y \in \{-1,+1\}^n$. The function $\text{ORT}_n(x,y)$ is defined as

$$\text{ORT}_n(x,y) = \begin{cases} -1 & \text{if } |\langle x,y\rangle| \leq \frac{\sqrt{n}}{8} \\ \\ +1 & \text{if } |\langle x,y\rangle| \geq \frac{\sqrt{n}}{4} \end{cases}$$

The general gap Hamming distance problem $\text{GHD}_{n,t,g}$ is defined as follows. Let $x, y \in \{-1,+1\}^n$. The function $\text{GHD}_{n,t,g}(x,y)$ is defined as,

$$\text{GHD}_{n,t,g}(x,y) = \begin{cases} -1 & \text{if } \langle x,y\rangle \ \leq t-g \\ \\ +1 & \text{if } \langle x,y\rangle \ \geq t+g \end{cases}$$

Note that both ORT_n and $\text{GHD}_{n,t,g}$ are partial functions. The function GHD_N is just $\text{GHD}_{N,0,\sqrt{N}}$.

23.3 Proof Details

First we establish the first stage from the proof outline.

Claim 23.3. ORT_n *reduces to* GHD_N*, where* $N = O(n)$*.*

Proof. In the last lecture, we saw that $\text{GHD}_{n,t,g}$ reduces to GHD_N, using a padding technique. The idea is to first reduce ORT_n to $\text{GHD}_{n,t,g}$ which in turn reduces to GHD_N.

Let $x, y \in \{-1,+1\}^n$ which satisfies the promise (i.e., $\text{ORT}_n(x,y)$ is defined). Then $\text{ORT}_n(x,y)$ can be solved using 2 calls to $\text{GHD}_{n,t,g}(x,y)$, as follows.

$$\text{ORT}_n(x,y) = \begin{cases} -1 & \text{if } \text{GHD}_{n,\frac{3\sqrt{n}}{16},\frac{\sqrt{n}}{16}}(x,y) = -1 \ \& \ \text{GHD}_{n,-\frac{3\sqrt{n}}{16},\frac{\sqrt{n}}{16}}(x,y) = +1 \\ +1 & \text{if } \text{GHD}_{n,-\frac{3\sqrt{n}}{16},\frac{\sqrt{n}}{16}}(x,y) = -1 \ \& \ \text{GHD}_{n,\frac{3\sqrt{n}}{16},\frac{\sqrt{n}}{16}}(x,y) = +1 \end{cases}$$

We know that $\text{GHD}_{m,t,g}$ can be reduced to GHD_N using padding. Examining the parameters m, t, g in the calls here relative to n, we see that $N \in O(n)$ suffices.

Thus, we have shown that ORT_n reduces to GHD_N. □

Now consider the Stage 3(a) from the proof outline.

Claim 23.4. $\mu(\text{ORT}^{-1}(-1)) = \Theta(1)$.

Proof. Let $x, y \in \{-1,+1\}^n$. We know that $\langle x,y\rangle = n - 2\Delta(x,y)$ (where $\Delta(.,.)$ is the Hamming distance). Note that if $\Delta(x,y) \in [\frac{n}{2} - \frac{\sqrt{n}}{8}, \frac{n}{2} + \frac{\sqrt{n}}{8}]$ then $|\langle x,y\rangle| \le \frac{\sqrt{n}}{8}$. For each fixed x, we count number of $y's$ such that $|\langle x,y\rangle| \le \frac{\sqrt{n}}{8}$. Using the fact that there is an absolute constant $c > 0$ such that for k close to $n/2$, $\binom{n}{k} \ge \frac{2^n}{c\sqrt{n}}$, we see that

$$\text{Number of } y's \text{ with } \left[|\langle x,y\rangle| \le \frac{\sqrt{n}}{8}\right] = \sum_{k=\frac{n}{2}-\frac{\sqrt{n}}{8}}^{\frac{n}{2}+\frac{\sqrt{n}}{8}} \binom{n}{k} \ge \frac{\sqrt{n}}{4} \cdot \frac{2^n}{c\sqrt{n}} = \frac{2^n}{4c}$$

Thus, the total number of $x, y \in \{-1,+1\}^n$ such that $|\langle x,y\rangle| \le \frac{\sqrt{n}}{8}$ is at least $\frac{2^{2n}}{4c}$.

Since μ is the uniform distribution, we have

$$\mu(\text{ORT}^{-1}(-1)) \ge \frac{2^{2n}}{4c} \cdot \frac{1}{4^n} = \frac{1}{4c}$$

□

The rest of the lecture is devoted to proving Stage 3(b) of the proof outline; that is, showing that any rectangle that is not 1-corrupted must in fact be small. We proceed via the following steps.

Step 0. For parameters ε, α to be chosen later, let $\rho = 2/2^{\alpha n}$. Assume to the contrary that some rectangle $R = S \times T$ is not 1-corrupted and is large.

Large: $\mu(R) \geq \rho$. Since $\mu(R) \leq |S|/2^n$, we can then conclude that $|S| \geq \rho 2^n = 2 \cdot 2^{(1-\alpha)n}$. Similarly, we can conclude that $|T| \geq 2 \cdot 2^{(1-\alpha)n}$.

Not 1-corrupted: $\mu(\mathrm{ORT}^{-1}(+1) \cap (S \times T)) \leq \varepsilon \mu(\mathrm{ORT}^{-1}(-1) \cap (S \times T)) \leq \varepsilon \mu(S \times T)$.

Step 1. Using the assumption that R has a "very high" density of -1s (because it is not 1-corrupted), find $A \subseteq S$ with a "fairly high" density of -1s (to be formally defined below) in each row, such that $|A| \geq \frac{|S|}{2}$.

Step 2. Using the assumption that S is large, and hence that A is large, show that there exists a set $A' \subseteq A$ of $k = \frac{n}{10}$ "near-orthogonal" vectors $x_1, \ldots, x_k$.

Step 3. Show that for any set W of m near-orthogonal vectors $x_1, ..., x_m$, a random y is, with high probability, far from orthogonal to at least one x_i (and so the corresponding $\mathrm{ORT}(x_i, y)$ is $+1$).

Step 4. Since $A' \subseteq A$, $A' \times T$ has a fairly high density of -1s. Find a $B \subseteq T$ with a moderately high density of -1s within $A' \times B$ such that $|B| \geq |T|/3$. Using the assumption that T is large, conclude that B is quite large. From Steps 2,3, conclude that B cannot be quite large. This gives a contradiction.

Proof of Step 1:

Define the set A as follows:

$$A \;\; = \;\; \{x \in S \mid \# +1\text{'s in } \{x\} \times T \leq 2\varepsilon |T|\}$$

An averaging argument shows that $|A| \geq \frac{|S|}{2} \geq 2^{(1-\alpha)n}$.

(If $|A| < \frac{|S|}{2}$, then there are at least $(\frac{|S|}{2} + 1)$ rows in S such that each row has more than $2\varepsilon|T|$ entries as $+1$. Thus, we have more than $\varepsilon|T||S|$ entries as $+1$ in the rectangle $S \times T$, contradicting the assumption that R is not 1-corrupted.)

Proof of Step 2:

Say that a set of vectors $x_1, ..., x_k$ in $\{+1, -1\}^n$ is near-orthogonal if for each $i \in [k-1]$, the vector x_{i+1} is almost orthogonal to (has a very small projection on) the subspace spanned by $x_1, \ldots, x_i$. Specifically, for each i, $||\, \mathrm{proj}_{\mathrm{span}\{x_1,\ldots,x_i\}}\, x_{i+1}|| \leq \frac{\sqrt{n}}{3}$.

We prove the following. Let $A \subseteq \{-1, +1\}^n$. For a sufficiently small constant $\alpha > 0$, which will be specified at the end of this Step, if $|A| > 2^{(1-\alpha)n}$ then A contains a set A' of $k = \lfloor \frac{n}{10} \rfloor$ near-orthogonal vectors $x_1, ..., x_k$.

Pick $x_1 \in A$ aribitrarily. Using Talagrand's inequality (see Appendix), a randomly picked x is unlikely to have a large projection on the space spanned by x_1.

$$\Pr_x\left[||\operatorname{proj}_{x_1} x|| > \frac{\sqrt{n}}{3}\right] \le \Pr_x\left[\ |\ ||\operatorname{proj}_{x_1} x|| - 1| > \frac{\sqrt{n}}{3} - 1 - c + c\right]$$
$$< \exp(-\frac{(\frac{\sqrt{n}}{3} - c - 1)^2}{c})$$
$$\le 2^{-\beta_1 n} \quad \text{for some constant } \beta_1 \text{ that depends only on } c.$$

Also, since $|A|$ is large, a randomly picked x is likely to be in A with good probability; $\Pr_{x\in_r\{-1,+1\}^n}[x \in A] > 2^{-\alpha n}$. Putting these two together, a randomly picked x is likely to be both in A and have a small projection on the space spanned by x_1. That is,

$$\Pr_x\left[||\operatorname{proj}_{x_1} x|| \le \frac{\sqrt{n}}{3} \text{ AND } x \in A \setminus \{x_1\}\right] = 1 - \Pr_x\left[||\operatorname{proj}_{x_1} x|| > \frac{\sqrt{n}}{3} \text{ OR } x \notin A \setminus \{x_1\}\right]$$
$$\ge 1 - [2^{-\beta_1 n} + 1 - 2^{-\alpha n}] \qquad \text{(union bound)}$$
$$= 2^{-\alpha n} - 2^{-\beta_1 n}$$

Let $\alpha < \beta_1$. Then the RHS is strictly positive. Therefore, there exists $x_2 \in A$ that is near orthogonal to x_1. Let's fix it.

We continue adding vectors this way. If $x_1, \ldots, x_j$ span space V, then $\dim(V) \le j$.

$$\Pr_x\left[||\operatorname{proj}_{\operatorname{span}(x_1,\ldots,x_j)} x|| > \frac{\sqrt{n}}{3}\right] \le \Pr_x\left[\ |\ ||\operatorname{proj}_{\operatorname{span}(x_1,\ldots,x_j)} x|| - \sqrt{j}| > \frac{\sqrt{n}}{3} - \sqrt{j} - c + c\right]$$
$$< \exp(-\frac{(\frac{\sqrt{n}}{3} - c - \sqrt{j})^2}{c})$$
$$\le \exp(-\beta_j n)$$

Here β_j is a constant depending only on c and j. By a similar argument as above, a random x will, with non-zero probability, be a new element of A and have a small projection with respect to the already chosen vectors. So we can find x_{j+1}.

Choose $\alpha < \beta_j$ for $1 \le j \le k = \frac{n}{10}$. In fact, let α be slightly smaller than the smallest β_j, to account for the fact that each x_{j+1} must be chosen not just from A but from $A \setminus \{x_1, \ldots, x_j\}$. Then we can choose $k = \lfloor \frac{n}{10} \rfloor$ near-orthogonal vectors in A.

Proof of Step 3:

Let $m \le n/10$, and fix any set W of vectors $x_1, ..., x_m \in \{-1, +1\}^n$ that are near-orthogonal. We want to show that with high probabilty, a random y is far-from-orthogonal to at least one x_i. Considering the complement event, we want to show that

$$\Pr_y\left[\forall i, |\langle y, x_i\rangle| \le \frac{\sqrt{n}}{4}\right] \in \exp(-\Omega(m))$$

Consider the $m \times n$ matrix M whose i^{th} row is x_i (We omit the notation for transpose; clear from the context). We want to show that $\Pr_y[\ ||My||_\infty \le \frac{\sqrt{n}}{4}] \le \exp(-\Omega(m))$. Since $||My||_\infty \le ||My|| \le \sqrt{m}||My||_\infty$, it suffices to instead prove that

$$\Pr_y\left[||My||^2 \ge \frac{mn}{16}\right] \ge 1 - \exp(-\Omega(m)).$$

Consider a singular value decomposition SVD of M as $M = UDV^t$ where U, V are unitary matrices of order m and n respectively, and D is a "rectangular diagonal" $m \times n$ matrix with diagonal entries $\sigma_1 \geq ... \geq \sigma_m$. D is uniquely defined by M. Let u_i, v_i denote the columns of U, V respectively. Then for any vector y, $||My||^2 = (My)^t(My) = (UDV^ty)^t(UDV^ty) = (DV^ty)^t(DV^ty) = \sum_{i=1}^m \sigma_i^2 \langle v_i, y\rangle^2$.

To show that $||My||^2$ is large for many y, we prove the following:

Claim 23.5. *Many σ_i are large. Specifically, $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_{\lceil \frac{m}{4} \rceil} \geq 0.51\sqrt{n}$.*

First, let us see why proving this claim is enough. Let $V = \{v_i : \sigma_i \geq 0.51\sqrt{n}\}$. By the claim, $\dim \operatorname{span}(V) \geq \frac{m}{4}$. For every vector y,

$$||My||^2 = \sum_{i=1}^m \sigma_i^2 \langle y, v_i\rangle^2 \geq (0.51\sqrt{n})^2 \sum_{v_i \in V} \langle y, v_i\rangle^2 \geq 0.26n|| \operatorname{proj}_{\operatorname{span}(V)} y||^2$$

Using Talagrand's inequality, we can now conclude that

$$\Pr_y \left[||My||^2 \geq \frac{mn}{16}\right] \geq \Pr_y \left[|| \operatorname{proj}_{\operatorname{span}(V)} y||^2 \geq \frac{m}{16 \times 0.26}\right] \geq 1 - \exp(-\Omega(m))$$

Proof. (Of Claim 23.5) We now prove the claim. First, we get a set of orthogonal vectors $x_1', ..., x_m'$ from $x_1, ..., x_m$: For each i, define $x_i' = x_i - \operatorname{proj}_{\operatorname{span}(x_1,...,x_{i-1})} x_i$. (These are the vectors that would be returned by the Gram-Schmidt procedure.) Then we can show (see below) that $n \geq ||x_i'||^2 \geq \frac{8n}{9}$.

$$||x_i||^2 = ||x_i'||^2 + || \operatorname{proj}_{\operatorname{span}(x_1,...,x_{i-1})} x_i||^2 \quad \text{(since these are orthogonal components of } x_i)$$
$$n = ||x_i'||^2 + || \operatorname{proj}_{\operatorname{span}(x_1,...,x_{i-1})} x_i||^2 \quad (\because x_i \in \{-1,+1\}^n\ )$$
$$n \leq ||x_i'||^2 + \frac{n}{9} \quad (\because || \operatorname{proj}_{\operatorname{span}(x_1,...,x_{i-1})} x_i|| \leq \frac{\sqrt{n}}{3}, \text{ by near-orthogonality })$$
$$n \geq ||x_i'||^2 \quad (\because || \operatorname{proj}_{\operatorname{span}(x_1,...,x_{i-1})} x_i|| \geq 0$$

Let M' denote the $m \times n$ matrix whose i^{th} row is x_i'.

Now consider the Frobenius norm of M, defined as $||M||_F = \sqrt{\sum_{i,j} M_{ij}^2}$. We will use the following proposition, to be proved later.

Proposition 23.6. *For all N, $\sigma_{r+1}(M) \geq \frac{1}{\operatorname{rank}(M)-r}\left(\frac{\langle M,N\rangle}{\sigma_1(N)} - ||M||_F\sqrt{r}\right)$*

Using the above proposition with N being the Gram-Schmidt matrix M', we get

$$\langle x_i, x_i'\rangle = ||x_i'||^2 \geq \frac{8n}{9}; \qquad \langle M, M'\rangle \geq \frac{8mn}{9}$$

$$\sigma_1(M') \leq \max_i ||x_i'|| = \sqrt{n}; \qquad ||M||_F = \sqrt{mn}$$

$$\text{Hence} \qquad \sigma_{r+1}(M) \geq \frac{1}{m-r}\left(\frac{\frac{8}{9}mn}{\sqrt{n}} - \sqrt{mn}\sqrt{r}\right)$$

Hence for $r = \lceil \frac{m}{4} \rceil$, $\sigma_{r+1}(M) \geq 0.51\sqrt{n}$. □

All that remains in this Step is to prove Proposition 23.6).

Proof. (Of Proposition 23.6)

The largest r singular values satisfy

$$\begin{aligned} \sigma_1 + ... + \sigma_r &\leq \sqrt{r}\sqrt{(\sigma_1^2 + ... + \sigma_r^2)} \quad \text{by Cauchy-Schwartz} \\ &\leq \sqrt{r}||M||_F \end{aligned}$$

The remaining singular values satisfy

$$\begin{aligned} \sigma_{r+1} + ... + \sigma_m &\leq (\text{rank}(M) - r)\sigma_{r+1} \\ \text{Also, by SVD,} \quad \sum_{i=1}^{m} \sigma_i &\geq \frac{\langle M, N\rangle}{\sigma_1(N)} \qquad (\because \langle M, N\rangle = \sum \sigma_i u_i^T N v_i \leq \sum \sigma_i \sigma_1(N)) \\ \text{Hence} \quad ||M||_F\sqrt{r} + (\text{rank}(M) - r)\sigma_{r+1} &\geq \sigma_1 + \ldots + \sigma_r \geq \frac{\langle M, N\rangle}{\sigma_1(N)} \\ \therefore \quad \sigma_{r+1}(M) &\geq \frac{1}{\text{rank}(M) - r}\left(\frac{\langle M, N\rangle}{\sigma_1(N} - ||M||_F\sqrt{r}\right) \end{aligned}$$

□

Proof of Step 4:

Recall that $A' = \{x_1, \ldots, x_k\}$ is contained in A. Hence by choice of A, each row of $A' \times T$ has at most $2\varepsilon|T|$ entries that are $+1$.

Define the set B as follows:

$$B = \{y \in T \mid \# +1\text{'s in } A' \times \{y\} \leq 3\varepsilon|A'|\}$$

An averaging argument similar to that in Step 1 shows that $|B| \geq \frac{|T|}{3} \geq (\frac{2}{3})2^{(1-\alpha)n}$.

We now give an upper bound on the size of B. If $y \in B$, then we can pick a set $A'' \subseteq A'$, of size exactly $(1-3\varepsilon)|A'| = (1-3\varepsilon)k$, such that $A'' \times \{y\}$ has only -1s. Let W be a subset of A' of size exactly $(1-3\varepsilon)k$, and define the set $B_W \subseteq T$ as follows:

$$\begin{aligned} B_W &= \{y \in T \mid W \times \{y\} \ \text{ has only } -1\text{s}\} \\ \text{Then} \quad B &\subseteq \bigcup_{W \subseteq A'; |W| = (1-3\varepsilon)k} B_W \\ \text{and hence} \quad |B| &\leq \sum_{W \subseteq A'; |W| = (1-3\varepsilon)k} |B_W| \end{aligned}$$

By Step 3, for any such W, $\Pr[y \in B_W] \in \exp(-\Omega(|W|)) = \exp(-\Omega((1-3\varepsilon)k))$. Hence $|B_W| \leq 2^n \exp(-\Omega((1-3\varepsilon)k))$. The number of choices for W is $\binom{k}{(1-3\varepsilon)k} = \binom{k}{3\varepsilon k}$. Hence

$$|B| \leq \binom{k}{3\varepsilon k} 2^n \exp(-\Omega((1-3\varepsilon)k)) \leq 2^{-H(3\varepsilon)k} 2^n \exp(-\Omega((1-3\varepsilon)k)).$$

By choosing suitable α, ε, we can see that the bounds on B

$$\left(\frac{2}{3}\right) 2^{(1-\alpha)n} \leq |B| \leq 2^{-H(3\varepsilon)k} 2^n \exp(-\Omega((1-3\varepsilon)k))$$

are not simultaneously possible.

References

[CR11] Amit Chakrabarti and Oded Regev. *An optimal lower bound on the communication complexity of gap-Hamming-distance.* In *Proc. 43rd ACM Symp. on Theory of Computing (STOC)*, pages 51–60. 2011. arXiv:1009.3460, eccc:TR10-140, doi:10.1145/1993636.1993644.

[She11] Alexander A. Sherstov. *The communication complexity of gap Hamming distance.* Technical Report TR11-063, Electronic Colloquium on Computational Complexity (ECCC), 2011. eccc:TR11-063.

[Vid11] Thomas Vidick. *A concentration inequality for the overlap of a vector on a large set, with application to the communication complexity of the gap-Hamming-distance problem.* Technical Report TR11-51, Electronic Colloquium on Computational Complexity (ECCC), 2011. eccc:TR11-051.

23.4 Appendix

23.4.1 Corruption Bound

Lemma 23.7. *Let $f : \mathrm{X} \times \mathrm{Y} \to \{0,1\}$ be a Boolean function and μ be a probability distribution on $\mathrm{X} \times \mathrm{Y}$ such that for every rectangle $R = \mathrm{S} \times \mathrm{T} \subseteq \mathrm{X} \times \mathrm{Y}$ with $\mu(R) > \rho$, we have $\mu(R \cap f^{-1}(1)) > \varepsilon \cdot \mu(R \cap f^{-1}(0))$. Then, for every $\delta > 0$, $2^{R_\delta(f)} \geq \frac{1}{\rho} \cdot (\mu(f^{-1}(0) - \frac{\delta}{\varepsilon})$.*

23.4.2 Talagrand's Inequality

Let $V \subseteq \mathrm{R}^n$ be a linear subspace of dimension d. Talagrand's inequality states that for a randomly chosen $x \in_r \{-1,+1\}^n$, with high probability, the projection of x on V is of length close to $\sqrt{d}$. Formally:

There exists a $c > 0$ such that $\forall t > 0$,

$$\Pr_{x \in_r \{-1,+1\}^n}\left[\ \left|\ ||\operatorname{proj}_V x|| - \sqrt{\dim V}\ \right| > t + c\right] < 4 \exp\left(-\frac{t^2}{c}\right)$$