6.4.a. The Discrepancy of GIP

In this subsection we prove a lower bound for GIP by using the discrepancy method. Specifically, we show that $Disc_{uniform}(\text{GIP}_n^k) \leq exp(-n/4^k)$.

We first introduce a slightly modified notation to facilitate easier algebraic handling. Define a function f as follows: $f(x_1, \ldots, x_k)$ is 1 if $\text{GIP}_n^k(x_1, \ldots, x_k) = 0$ and -1 if $\text{GIP}_n^k(x_1, \ldots, x_k) = 1$. In this case, instead of working directly with the discrepancy we will use:

$$\Delta_k(n) = \max_{\phi_1, \ldots, \phi_k} |\mathrm{E}_{x_1, \ldots, x_k}[f(x_1, \ldots, x_k) \cdot \phi_1(x_1, \ldots, x_k) \cdots \phi_k(x_1, \ldots, x_k)]|,$$

where the maximum is taken over all functions $\phi_i : (\{0, 1\}^n)^k \to \{0, 1\}$ such that ϕ_i does not depend on x_i, and the expectation is over all 2^{nk} possible choices of $x_1, \ldots, x_k$.

First, it should be clear that $Disc_{uniform}(\text{GIP}_n^k) = \Delta_k(n)$. This is because the product $\phi_1(x_1, \ldots, x_k) \cdots \phi_k(x_1, \ldots, x_k)$ gives 1 on a collection of points that forms a cylinder intersection, and, conversely, any cylinder intersection can be written as such a product. In addition, because we changed to the $\{-1, +1\}$ notation, the expectation plays the same role as the difference in probabilities previously did.

We define constants β_k recursively: $\beta_1 = 0$, and $\beta_k = \sqrt{\frac{1+\beta_{k-1}}{2}}$. It follows by induction that $\beta_k \leq 1 - 4^{1-k} < e^{-4^{1-k}}$. We will prove the following upper bound on $\Delta_k(n)$.

Lemma 6.17: *$\Delta_k(n) \leq (\beta_k)^n$, for all $k \geq 1, n \geq 0$.*

PROOF: Observe that $\Delta_1(n) = 0$, because in this case ϕ_1 must be constant and $\mathrm{E}_{x_1}[f(x_1)] = 0$ (in the case that $n = 0$, we get $\Delta_1(0) = 1$. To overcome this, we define $0^0 = 1$ for this proof). We proceed by induction on k. Let $k \geq 2$, and fix $\phi_1, \ldots, \phi_k$ that achieve the value of $\Delta_k(n)$. Because ϕ_k does not depend on x_k, and is bounded in absolute value by 1,

$$\Delta_k(n) \leq \mathrm{E}_{x_1, \ldots, x_{k-1}}[|\mathrm{E}_{x_k} f(x_1, \ldots, x_k) \cdot \phi_1(x_1, \ldots, x_k) \cdots \phi_{k-1}(x_1, \ldots, x_k)|].$$

In order to estimate the right-hand side, we will use a special case of the Cauchy–Schwartz inequality stating that for any random variable z: $(\mathrm{E}[z])^2 \leq \mathrm{E}[z^2]$. Thus our estimate is:

$$\begin{aligned}\Delta_k(n) &\leq (\mathrm{E}_{x_1, \ldots, x_{k-1}}[\mathrm{E}_{x_k}[f(x_1, \ldots, x_k) \cdot \phi_1(x_1, \ldots, x_k) \cdots \phi_{k-1}(x_1, \ldots, x_k)]]^2)^{1/2} \\ &= (\mathrm{E}_{u, v, x_1, \ldots, x_{k-1}}[f(x_1, \ldots, x_{k-1}, u) \cdot f(x_1, \ldots, x_{k-1}, v) \\ &\qquad \cdot \phi_1^u \cdot \phi_1^v \cdots \phi_{k-1}^u \cdot \phi_{k-1}^v])^{1/2}\end{aligned}$$

where ϕ_i^u stands for $\phi_i(x_1, \ldots, x_{k-1}, u)$, and ϕ_i^v for $\phi_i(x_1, \ldots, x_{k-1}, v)$.

Now observe that for every particular choice of u and v, we can express the product $f(x_1, \ldots, x_{k-1}, u) f(x_1, \ldots, x_{k-1}, v)$ in terms of the function f on $k-1$ strings of possibly shorter length. Inspection reveals that the value of $f(x_1, \ldots, x_{k-1}, u) f(x_1, \ldots, x_{k-1}, v)$ is simply $f(z_1, \ldots, z_{k-1})$, where z_i is the restriction of x_i to the coordinates j such that $u_j \neq v_j$ (here is where the particular properties of f are used). We will now view each x_i as composed of two parts: z_i and y_i, where z_i is the part of x_i where $u_j \neq v_j$, and y_i the part of x_i where $u_j = v_j$ (this is done separately for every u, v).

For every particular choice of u, v and consequently $y_1, \ldots, y_{k-1}$, we define functions of the "z-parts":

$$\xi_i^{u,v,y_1,\ldots,y_{k-1}}(z_1, \ldots, z_{k-1}) = \phi_i(x_1, \ldots, x_{k-1}, u)\phi_i(x_1, \ldots, x_{k-1}, v),$$

where the x_is are obtained by the concatenation of the corresponding y_i and z_i. We can now rewrite the previous estimate as

$$\Delta_k(n) \le (\mathrm{E}_{u,v}[\mathrm{E}_{y_1,\ldots,y_{k-1}}[S^{u,v,y_1,\ldots,y_{k-1}}]])^{1/2},$$

where $S^{u,v,y_1,\ldots,y_{k-1}}$ is defined as

$$\mathrm{E}_{z_1,\ldots,z_{k-1}}\big[f(z_1, \ldots, z_{k-1}) \cdot \xi_1^{u,v,y_1,\ldots,y_{k-1}}(z_1, \ldots, z_{k-1}) \cdots \xi_{k-1}^{u,v,y_1,\ldots,y_{k-1}}(z_1, \ldots, z_{k-1})\big].$$

Now $S^{u,v,y_1,\ldots,y_{k-1}}$ can be estimated via the induction hypothesis, because f and the ξ_is are all functions of $k-1$ strings. Moreover, note that $\xi_i^{u,v,y_1,\ldots,y_{k-1}}$ does not depend on z_i. Thus the previous estimate of $\Delta_k(n)$ is bounded by

$$\Delta_k(n) \le (\mathrm{E}_{u,v,y_1,\ldots,y_{k-1}}[\Delta_{k-1}(m_{u,v})])^{1/2} \le \left(\mathrm{E}_{u,v,y_1,\ldots,y_{k-1}}\left[\beta_{k-1}^{m_{u,v}}\right]\right)^{1/2},$$

where $m_{u,v}$ is the length of the strings z_i, which is equal to the number of locations j such that $u_j \ne v_j$.

Because u and v are distributed uniformly in $\{0,1\}^n$, $m_{u,v}$ is distributed according to the binomial distribution. For any constant m, the probability that $m_{u,v} = m$ is exactly $\binom{n}{m}2^{-n}$. Thus the previous estimate gives:

$$\Delta_k(n) \le \left[\sum_{m=0}^{n} \binom{n}{m} 2^{-n}\beta_{k-1}^{m}\right]^{1/2} = [2^{-n}(1+\beta_{k-1})^n]^{1/2} = \beta_k^n,$$

which completes the proof of the lemma. □

To conclude, this shows that $Disc_{uniform}(\mathrm{GIP}_n^k) \le 1/e^{\frac{n}{4^{k-1}}}$, which implies that the deterministic (and even randomized) communication complexity of GIP_n^k is $\Omega(n/4^k)$. In fact, by Exercise 6.15, we also get a bound for $D_{\frac{1}{2}-\varepsilon}^{uniform}(f)$ and $R_{\frac{1}{2}-\varepsilon}(f)$ of $\Omega(\log\varepsilon + n/4^k)$.

6.5. Simultaneous Protocols

The protocols presented in Examples 6.3 and 6.4 are of a very restricted form: the communication sent by each party does not depend at all on the previous communication sent by other parties. We can imagine all parties speaking "simultaneously" and each writing, on a common blackboard, a function of the $k-1$ parts of the input it can see. After all parties have spoken, the answer should be determined by what is written on the blackboard. We call such protocols simultaneous.

Definition 6.18: *The* simultaneous communication complexity *of f, $D^{\|}(f)$, is the cost of the best simultaneous protocol that computes f.*