A Reference on Concentration Inequalities

Theorem 17 (Hoeffding's Inequality) *Let $X_1, X_2, \ldots, X_n$ be independent random variables bounded by the interval $[a, b] : a \le X_i \le b$, then we define $X = X_1 + \cdots + X_n$. We have*

$$\Pr[X - \mathbb{E}[X] \ge t] \le \exp\left(-\frac{2t^2}{n(b-a)^2}\right).$$

Theorem 18 (Chernoff Bound) *Suppose $X_1, \ldots, X_n$ are independent random variables, $X_i \in [0, 1]$. Let $X = X_1 + X_2 + \cdots + X_n$ and let $\mu = \mathbb{E}[X]$ denote the sum's expected value. Then for $0 \le \delta \le 1$,*

$$\Pr[X \ge (1+\delta)\mu] \le e^{-\frac{\delta^2\mu}{3}}.$$

B Deferred Proofs in Section 3.1

Proof of Lemma 5. For simplicity of the exposition, we assume that f is a deterministic policy. The case for randomized f can be analogously addressed.

Suppose $f(S)$ decides to query arm i. Let r be observation seen by policy f after the query and let $\tilde{r}$ be the observation (or the pretended observation) seen by policy $\tilde{f}$. We only need to prove that $\Pr[r = 1|S_t = S] = \Pr[\tilde{r} = 1|\tilde{S}_t = S]$ since f uses r and $\tilde{f}$ uses $\tilde{r}$ to update their query history on arm i.

Now let us condition on the event that the current state for $\tilde{f}$ is S. Let $(a_i, b_i) \in S$ be the query history on arm i in recorded S. We claim that the probability that $\tilde{r} = 1$ is $\mathbb{E}\,\mathrm{Beta}(a_i + 1, b_i + 1)$, which is the same as $\Pr[r = 1|S_t = S]$, proving the lemma.

Suppose that $i \notin C$, by the construction of $\tilde{f}$, a real query is made to arm i and $\tilde{r}$ is the observation bit. Therefore, in this case, the probability that $\tilde{r} = 1$ is $\mathbb{E}\,\mathrm{Beta}(a_i + 1, b_i + 1)$.

Otherwise, we have that $i \in C$. Let $q = a_i + b_i$. Let $\tilde{r}_1, \tilde{r}_2, \ldots, \tilde{r}_q$ be the q observations (including pretended ones) for arm i seen by $\tilde{f}$. We have $\sum_{j=1}^{q} \tilde{r}_j = a_i$. Let $\tilde{q} = \tilde{a}_i + \tilde{b}_i$. We have that $\tilde{f}$ has made $\tilde{q}$ real queries on arm i, and that $\sum_{j=1}^{\tilde{q}} \tilde{r}_j = \tilde{a}_i$. In this case, we have

$$\begin{aligned}
&\Pr\left[\tilde{r} = 1|\tilde{S} = S\right] \\
&= \mathop{\mathbb{E}}_{\tilde{q}} \mathop{\mathbb{E}}_{\tilde{\theta}_i \sim \mathrm{Beta}(\tilde{a}_i+1, \tilde{b}_i+1)} \left[\tilde{\theta}_i | \tilde{r}_{\tilde{q}+1}, \tilde{r}_{\tilde{q}+2}, \ldots, \tilde{r}_q\right] \\
&= \mathop{\mathbb{E}}_{\tilde{q}} \mathbb{E}\left[\mathrm{Beta}\left(\tilde{a}_i + 1 + \sum_{j=\tilde{q}+1}^{q} \tilde{r}_j, \tilde{b}_i + 1 + \sum_{j=\tilde{q}+1}^{q} (1 - \tilde{r}_j)\right)\right] \\
&= \mathop{\mathbb{E}}_{\tilde{q}} \mathrm{Beta}(a_i + 1, b_i + 1) = \mathbb{E}\,\mathrm{Beta}(a_i + 1, b_i + 1).
\end{aligned}$$

□

Proof of Lemma 7. We only need to prove that when $|\tilde{a}_i - \tilde{b}_i| > \sqrt{\tilde{a}_i + \tilde{b}_i} \cdot 3 \ln M$, we have $\mathrm{err}(\tilde{a}_i, \tilde{b}_i) < \frac{1}{2\sqrt{M}}$. Let us assume without loss of generality that $\tilde{a}_i \ge \tilde{b}_i$, and let $\delta = \frac{2\tilde{a}_i}{\tilde{a}_i + \tilde{b}_i} - 1 \ge \frac{3 \ln M}{\sqrt{\tilde{a}_i + \tilde{b}_i}}$. We have

$$\begin{aligned}
\mathrm{err}(\tilde{a}_i, \tilde{b}_i) &= \Pr[\mathrm{Beta}(\tilde{a}_i + 1, \tilde{b}_i + 1) < .5] && (11) \\
&= \frac{\Gamma(\tilde{a}_i + \tilde{b}_i + 1)}{\Gamma(\tilde{a}_i)\Gamma(\tilde{b}_i)} \int_0^{\frac{1}{2}} x^{\tilde{a}_i} (1-x)^{\tilde{b}_i} \mathrm{d}x \\
&\le \frac{\Gamma(\tilde{a}_i + \tilde{b}_i + 1)}{\Gamma(\tilde{a}_i)\Gamma(\tilde{b}_i)} \int_0^{\frac{1}{2}} 2^{-(\tilde{a}_i + \tilde{b}_i)} \mathrm{d}x && (12)
\end{aligned}$$

$$= \frac{(\tilde{a}_i + \tilde{b}_i)! \cdot (\tilde{a}_i + \tilde{b}_i + 1)}{\tilde{a}_i! \tilde{b}_i! \cdot 2^{(\tilde{a}_i + \tilde{b}_i + 1)}}. \tag{13}$$

Now we use Stirling's formula ($\sqrt{2\pi n}(n/e)^n \leq n! \leq e\sqrt{n}(n/e)^n$ for every positive integer n), and have

$$(13) \leq (\tilde{a}_i + \tilde{b}_i + 1) \cdot \frac{e\sqrt{\tilde{a}_i + \tilde{b}_i}}{2\pi\sqrt{\tilde{a}_i \tilde{b}_i}} \cdot \left(\frac{\tilde{a}_i + \tilde{b}_i}{\tilde{a}_i}\right)^{\tilde{a}_i} \left(\frac{\tilde{a}_i + \tilde{b}_i}{\tilde{b}_i}\right)^{\tilde{b}_i}. \tag{14}$$

Note that

$$\left(\frac{\tilde{a}_i + \tilde{b}_i}{\tilde{a}_i}\right)^{\tilde{a}_i} \left(\frac{\tilde{a}_i + \tilde{b}_i}{\tilde{b}_i}\right)^{\tilde{b}_i} = (1+\delta)^{-\tilde{a}_i}(1-\delta)^{-\tilde{b}_i} \tag{15}$$

$$= \left((1+\delta)^{(1+\delta)}(1-\delta)^{(1-\delta)}\right)^{-(\tilde{a}_i + \tilde{b}_i)/2}. \tag{16}$$

Since we have $(1+\delta)^{(1+\delta)}(1-\delta)^{(1-\delta)} \geq \exp(\delta^2)$ for $\delta \in [0, 1]$ (where 0^0 is defined to be 1), combining (13), (14), and (16), we have

$$\begin{aligned} \mathrm{err}(\tilde{a}_i, \tilde{b}_i) &\leq (\tilde{a}_i + \tilde{b}_i + 1)^{1.5} \exp(-\delta^2(\tilde{a}_i + \tilde{b}_i)/2) \\ &\leq (\tilde{a}_i + \tilde{b}_i + 1)^{1.5} \exp(-(9 \ln M)/2) \leq \frac{1}{2\sqrt{M}}, \end{aligned}$$

where the second inequality is because of $\delta \geq \frac{3 \ln M}{\sqrt{\tilde{a}_i + \tilde{b}_i}}$ and the last inequality is because of $\tilde{a}_i + \tilde{b}_i \leq 100M \ln^2 M$ and for sufficiently large M. □

Proof of Lemma 8. Let a and b be the number of 1's and 0's after querying i for $100M \ln^2 M$ times. If i is corrupted but not marked when $\tilde{f}$ terminates, then we have $|a - b| \leq \sqrt{a+b} \cdot 3 \ln M = \sqrt{100M \ln^2 M} \cdot (3 \ln M)$. So,

$$\begin{aligned} &\Pr[i \text{ corrupted but not marked}] \\ &\leq \Pr[|a-b| \leq \sqrt{100M \ln^2 M} \cdot (3 \ln M)] \\ &\leq \Pr\left[|a-b| \leq \sqrt{100M \ln^2 M} \cdot (3 \ln M) \middle| |\theta_i - 0.5| > \frac{1}{6\sqrt{M}}\right] + \Pr\left[|\theta_i - 0.5| \leq \frac{1}{6\sqrt{M}}\right] \\ &\leq \Pr\left[|a-b| \leq \sqrt{100M \ln^2 M} \cdot (3 \ln M) \middle| |\theta_i - 0.5| > \frac{1}{6\sqrt{M}}\right] + \frac{1}{3\sqrt{M}} \\ &\leq \Pr\left[a > 50M \ln^2 M - 15\sqrt{M} \ln^2 M \middle| \theta_i \leq \frac{1}{2} - \frac{1}{6\sqrt{M}}\right] + \frac{1}{3\sqrt{M}}. \end{aligned} \tag{17}$$

The last inequality holds because a and b are symmetric (so that we can assume $\theta_i \leq \frac{1}{2}$ without loss of generality). Note that $a = \sum_{j=1}^{100M \ln^2 M} X_j$ where X_j's are *i.i.d.* samples from $\mathcal{B}_{\theta_i}$. Using Hoeffding's inequality (Theorem 17) with $\mathbb{E}[a] \leq 50M \ln^2 M - \frac{50}{3}\sqrt{M} \ln^2 M$, we have

$$\begin{aligned} &\Pr\left[a > 50M \ln^2 M - 15\sqrt{M} \ln^2 M \middle| \theta_i \leq \frac{1}{2} - \frac{1}{6\sqrt{M}}\right] \\ &= \Pr\left[a - \mathbb{E}[a] > \frac{5}{3}\sqrt{M} \ln^2 M \middle| \theta_i \leq \frac{1}{2} - \frac{1}{6\sqrt{M}}\right] \\ &\leq \exp\left(-\frac{2 \cdot (\frac{5}{3}\sqrt{M} \ln^2 M)^2}{100M \ln^2 M}\right) \\ &\leq \exp(-\ln^2 M/18) \leq \frac{1}{M \ln M}, \end{aligned} \tag{18}$$

for sufficiently large M. Combining (17) and (18), we have

$$\Pr[i \text{ corrupted but not marked}] \le \frac{1}{M \ln M} + \frac{1}{3\sqrt{M}} \le \frac{1}{2\sqrt{M}}.$$

□

C Proof of Lemma 9

We let $\tilde{f}$ be an ϵ^{-2}-BQP with query budget Q and $\mathsf{val}(\tilde{f}) \ge \mathrm{OPT}(Q) - \epsilon$ (which is possible by Lemma 2). We build a policy g as follows. At the beginning, for each arm i, g samples an independent random bit $y_i \in \{0,1\}$ where $\mathbb{E}\, y_i = 2\epsilon$. For each arm i with $y_i = 1$, g samples $\tilde{\theta}_i$ from the uniform distribution over $[0,1]$ (i.e., the prior distribution of θ_i). Now g maintains a state of query history $\tilde{S} = \{(a_1, b_1), \ldots, (a_n, b_n)\}$ where a_i and b_i are initialized to 0 for all $i \in [n]$. g now simulates the policy $\tilde{f}$. Whenever $\tilde{f}(\tilde{S})$ decides to query arm i, if $y_i = 0$, g directly queries the arm and updates the state $\tilde{S}$; otherwise g make a simulated query by sampling a bit from $\mathcal{B}_{\tilde{\theta}_i}$ and update the query history using this bit. g also keeps track of the total number of real queries that have been made. Whenever this number exceeds $(1-\epsilon)Q$, g terminates and gives up. When $\tilde{f}$ terminates and decides the guess for each arm, g does the same thing.

It is clear that g queries at most $(1-\epsilon)Q$ times. Now it suffices to prove that $\mathsf{val}(g) \ge \mathsf{val}(\tilde{f}) - 3\epsilon$.

Lemma 19 *When $Q \ge 1200\epsilon^{-4} \ln^3 \epsilon^{-1}$, the probability that g exceeds the budget limit and gives up is at most ϵ.*

Proof of Lemma 19. Let us imagine that g does not terminate even when the number of real queries exceeds the budget, and finally reaches a final state $\tilde{S} = \{(a_1, b_1), \ldots, (a_n, b_n)\}$ for $\tilde{f}$. In the real run of g, the probability that g gives up exactly

$$\Pr\left[\sum_{i=1}^{n} (1-y_i)(a_i + b_i) > (1-\epsilon)Q\right] = \mathop{\mathbb{E}}_{\tilde{S}} \Pr\left[\sum_{i=1}^{n} (1-y_i)(a_i + b_i) > (1-\epsilon)Q \middle| \tilde{S}\right].$$

One can verify that $\{y_1, y_2, \ldots, y_n\}$ is independent from $\tilde{S}$, and therefore conditioned on $\tilde{S}$, $\{y_1, y_2, \ldots, y_n\}$ follows the same *i.i.d.* distribution. Therefore, if we let $X_i = \frac{(1-y_i)(a_i+b_i)}{400\epsilon^{-2}\ln^2 \epsilon^{-1}}$, we have that X_i's are independent random variables bounded in $[0,1]$ (by the definition of ϵ^{-2}-BQP) and $\mathbb{E} \sum_{i=1}^{n} X_i = \frac{(1-2\epsilon)Q}{400\epsilon^{-2}\ln^2 \epsilon^{-1}}$.

By Chernoff Bound (Theorem 18), we have

$$\Pr\left[\sum_{i=1}^{n} (1-y_i)(a_i + b_i) > (1-\epsilon)Q \middle| \tilde{S}\right] = \Pr\left[\sum_{i=1}^{n} X_i > \frac{(1-\epsilon)Q}{400\epsilon^{-2}\ln^2 \epsilon^{-1}} \middle| \tilde{S}\right] < \exp\left(-\frac{\epsilon^2}{3} \cdot \frac{(1-2\epsilon)Q}{400\epsilon^{-2}\ln^2 \epsilon^{-1}}\right),$$

which is at most ϵ when $Q \ge 1200\epsilon^{-4} \ln^3 \epsilon^{-1}$.

Finally, the probability that g gives up is

$$\mathop{\mathbb{E}}_{\tilde{S}} \Pr\left[\sum_{i=1}^{n} (1-y_i)(a_i + b_i) > (1-\epsilon)Q \middle| \tilde{S}\right] \le \mathop{\mathbb{E}}_{\tilde{S}} \epsilon = \epsilon.$$

□

Lemma 20 *Let S be a query history state of $\tilde{f}$. For every realization of $y_1, y_2, \ldots, y_n$, when $\sum_{i=1}^{n} (1-y_i)(a_i + b_i) \le (1-\epsilon)Q$, we have $\Pr[\tilde{f} \text{ reaches } S] = \Pr[g \text{ does not give up and reaches } S | y_1, y_2, \ldots, y_n]$.*

Proof of Lemma 20. We have

$$\Pr[\tilde{f} \text{ reaches } S] = \Pr[\tilde{f} \text{ reaches } S | y_1, y_2, \ldots, y_n],$$

where in the LHS we consider a run of $\tilde{f}$; in the RHS we consider a run of g (which also simulates $\tilde{f}$) and we imagine the run does not terminate even when the number of real queries exceeds the budget. The equality holds because of the independence between $\{y_1, y_2, \ldots, y_n\}$ and the state of g. Also note that the RHS is equivalent to

$$\Pr[g \text{ does not give up and reaches } S | y_1, y_2, \ldots, y_n]$$

when $\sum_{i=1}^n (1-y_i)(a_i+b_i) \le (1-\epsilon)Q$, and therefore the lemma is proved. □

With Lemma 19 and Lemma 20, we are ready to prove Lemma 9.

Proof of Lemma 9. Recall that it suffices to prove that $\mathsf{val}(g) \ge \mathsf{val}(\tilde{f}) - 3\epsilon$. Given a realization of $y_1, \ldots, y_n$, consider a run of $\tilde{f}$ and let $S = \{(a_1, b_1), \ldots, (a_n, b_n)\}$ be the terminal state reached by $\tilde{f}$. Here we define S to be *good* if $\sum_{i=1}^n (1-y_i)(a_i+b_i) \le (1-\epsilon)Q$. Note that when $y_i = 1$, we do not really query arm i, therefore $\mathsf{val}(g)$ is lower bounded by

$$\begin{aligned}
&\mathop{\mathbb{E}}_{y_1,\ldots,y_n} \sum_{\text{good } S} \frac{1}{n}\left(\sum_{i=1}^n (1-\mathrm{err}(a_i,b_i)) - \sum_{i=1}^n y_i\right) \cdot \Pr[g \text{ reaches } S | y_1, \ldots, y_n] \\
&\ge \mathop{\mathbb{E}}_{y_1,\ldots,y_n} \sum_{\text{good } S} \frac{1}{n}\sum_{i=1}^n (1-\mathrm{err}(a_i,b_i)) \cdot \Pr[g \text{ reaches S} | y_1, \ldots, y_n] - \mathop{\mathbb{E}}_{y_1,\ldots,y_n} \frac{1}{n}\sum_{i=1}^n y_i \\
&= \mathop{\mathbb{E}}_{y_1,\ldots,y_n} \sum_{\text{good } S} \frac{1}{n}\sum_{i=1}^n (1-\mathrm{err}(a_i,b_i)) \cdot \Pr[g \text{ reaches S} | y_1, \ldots, y_n] - 2\epsilon. \qquad (19)
\end{aligned}$$

When S is good, if g reaches S, it means g does not give up. According to Lemma 20, for good S, $\Pr[\tilde{f} \text{ reaches S}] = \Pr[g \text{ reaches } S | y_1, \ldots, y_n]$, thus we can write (19) as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{y_1,\ldots,y_n} \sum_{\text{good } S} \frac{1}{n}\sum_{i=1}^n (1-\mathrm{err}(a_i,b_i)) \Pr[f \text{ reaches } S] - 2\epsilon \\
&\ge \mathop{\mathbb{E}}_{y_1,\ldots,y_n} \sum_{\text{terminal } S} \frac{1}{n}\sum_{i=1}^n (1-\mathrm{err}(a_i,b_i)) \Pr[f \text{ reaches } S] - \mathop{\mathbb{E}}_{y_1,\ldots,y_n} \Pr[g \text{ gives up} | y_1, \ldots, y_n] - 2\epsilon \\
&= \mathop{\mathbb{E}}_{y_1,\ldots,y_n} \left[\mathsf{val}(\tilde{f}) - \Pr[g \text{ gives up} | y_1, \ldots, y_n]\right] - 2\epsilon \\
&\ge \mathsf{val}(\tilde{f}) - \Pr[g \text{ gives up}] - 2\epsilon \\
&\ge \mathsf{val}(\tilde{f}) - 3\epsilon,
\end{aligned}$$

where the last inequality holds because of Lemma 19. □

D Deferred Proof(s) in Section 3.3

Proof of Lemma 10. Given an M-BQP $\tilde{f}$ with query budget Q, we define the policy $\tilde{g}$ as follows. $\tilde{g}$ simulates $\tilde{f}$. For each arm i, $\tilde{g}$ also keeps a buffer of observations, which is initialized to be empty. Whenever $\tilde{f}(S)$ decides to query arm i, if the arm's buffer is empty, suppose arm i has been queries by $\tilde{f}$ for τ_j times, $\tilde{g}$ makes $(\tau_{j+1} - \tau_j)$ queries to arm i and add all observations to the buffer. Then $\tilde{g}$ extracts one observation from the buffer which is served as the observation of the query made by $\tilde{f}(S)$. Whenever $\tilde{f}$ terminates and decides, $\tilde{g}$ also terminates and decides.

It is straightforward to verify that 1) $\mathsf{val}(\tilde{g}) = \mathsf{val}(\tilde{f})$, 2) $\tilde{g}$ satisfies the two constraints for an M-BQP (since $\tilde{f}$ is an M-BQP) and the additional constraint for a (γ, M)-BBQP. Finally, we verify that $\tilde{g}$ makes at most $(1+\gamma)Q$ queries. Let q_i be the total number of queries made to arm i by $\tilde{f}$. Let $\tilde{q}_i$ be the total number of queries made to arm i by $\tilde{g}$. Once can verify that $\tilde{q}_i \le (1+\gamma)q_i$. Therefore the total number of queries made by $\tilde{g}$ is $\sum_{i=1}^n \tilde{q}_i \le \sum_{i=1}^n (1+\gamma)q_i \le (1+\gamma)Q$. □