2. Quantum Threshold Search

To use the online learning protocol from the previous section, we need a way to implement the "teacher." Ideally, we want a teacher that can:

- Correctly identify when the student has correctly learned all of the observable values
- If the student has not correctly learned all observable values, pinpoint some observable on which the student is incorrect and show it to them in the next round of interaction.

This is formalized in the following:

Definition 114 (Quantum Threshold Search)**.** *The* quantum threshold search problem *is the following task:*

- ***Input****: access to copies of an unknown quantum state ρ; a description of observables $O_1, \ldots, O_m$; and numbers $\hat{o}_1, \ldots, \hat{o}_m$*
- ***Output****: Either "$|\hat{o}_i - \mathrm{tr}(O_i\rho)| \le \epsilon$ for all $i \in [m]$," or an index $i \in [m]$ for which $|\hat{o}_i - \mathrm{tr}(O_i\rho)| > \frac{3}{4}\epsilon$ together with whether or not $\hat{o}_i \ge \mathrm{tr}(O_i\rho)$.*

We would like an algorithm that outputs an incorrect statement with probability at most δ

In this section, we give an algorithm for this task with the following guarantee:

Theorem 115. *There is an algorithm for quantum threshold search which uses $O(\frac{\log^2 m + \log 1/\delta}{\epsilon^2} \cdot \log 1/\delta)$ copies of ρ and outputs an incorrect statement with probability at most δ.*

2.1. Basic reductions

Here we perform a sequence of simplifications to the threshold search problem that show that it suffices to solve the following version of the problem:

Definition 116 (Weak Threshold Search)**.** *The* weak quantum threshold search problem *is the following task:*

- ***Input****: access to copies of an unknown quantum state ρ; and a description of* projectors *$O_1, \ldots, O_m$ such that $\mathrm{tr}(O_i\rho) > 3/4$ for at least one $i \in [m]$*
- ***Output****: Any index $j \in [m]$ for which $\mathrm{tr}(O_j\rho) > 1/4$.*

We would like an algorithm that succeeds with any $\Omega(1)$ probability.

Note the four key differences: (1) the observables are projectors, (2) instead of testing whether some observable value is *far* from some threshold $\hat{o}_i$, we are testing whether it is *larger* than some threshold, (3) in place of $\frac{\epsilon}{4}$ and $\frac{3\epsilon}{4}$, the thresholds in question are fixed constants $3/4$ and $1/4$, (4) we are operating under the *promise* that there is some observable value which is above the threshold $3/4$, and (5) we are only aiming for constant success probability.

The reductions we perform to get to this new task are relatively straightforward, and the reader may skip the proof upon first reading without losing much intuition.

Projector observables. The fact that we can assume WLOG that the observables are projectors follows immediately from the Naimark dilation theorem (Theorem 43). In fact, we could have assume this at the outset of our discussion on shadow tomography, but it wouldn't have noticeably simplified anything up to this point.

One-sided testing. The fact that we can move from testing whether $\mathrm{tr}(O_i\rho)$ is close to some value ("two-sided testing") versus testing whether it is *above* some value ("one-sided testing") arises from the fact that if we check that $\mathrm{tr}(O_i\rho)$ is above some threshold $\hat{o}_i - O(\epsilon)$, and additionally that $\mathrm{tr}((\mathrm{Id} - O_i)\rho)$ is above some threshold $1-\hat{o}_i - O(\epsilon)$, then this is equivalent to checking that $|\mathrm{tr}(O_i\rho) - \hat{o}_i| \le O(\epsilon)$ as in the original formulation of threshold search.

Specific constant thresholds.

With the above reasoning, we can reduce to distinguishing between whether all $\mathrm{tr}(O_i\rho)$ are below the threshold $\hat{o}_i + \frac{3}{4}\epsilon$, or whether there is some $\mathrm{tr}(O_i\rho)$ which is above the threshold $\hat{o}_i - \epsilon$. That it suffices to do this when these thresholds are replaced by 3/4 and 1/4 respectively is immediate from the following "boosting" result:

Lemma 117. *For any $\epsilon > 0$ and $\delta, \theta \in (0,1)$, there is an $n = O(1/\epsilon^2)$ such that for any projector Π acting on $\mathbb{C}^d$, there is a projector Π^* acting on $(\mathbb{C}^d)^{\otimes n}$ such that the following holds for any state ρ: if $\mathrm{tr}(\Pi\rho) > \theta + \frac{3}{4}\epsilon$ then $\mathrm{tr}(\Pi^*\rho^{\otimes n}) > 3/4$, and if $\mathrm{tr}(\Pi\rho) < \theta - \epsilon$ then $\mathrm{tr}(\Pi^*\rho^{\otimes n}) < 1/4$.*

Proof. For convenience let $\Pi_1 = \Pi$ and $\Pi_0 = \mathrm{Id} - \Pi$. Given $0 \le k \le n$, define the d^n-dimensional projector $\Pi^*_k \triangleq \sum_{x\in\{0,1\}^n:|x|=k} \bigotimes_{i=1}^n \Pi_{x_i}$, where $|x|$ denotes the Hamming weight of x. Finally define $\Pi^* \triangleq \sum_{k:k/n>\theta} \Pi^*_k$.

By Chernoff bound, if $\mathrm{tr}(\Pi\rho) > \theta+\epsilon$, then $\mathrm{tr}(\Pi^*\rho^{\otimes n}) \ge \Pr \mathrm{Bin}(n, \theta + \frac{3}{4}\epsilon) > \theta \ge 1 - \exp(-\Omega(n\epsilon^2))$, so if $n = \Omega(1/\epsilon^2)$ with sufficiently large leading constant, the latter probability is at least 3/4. Likewise, if $\mathrm{tr}(\Pi\rho) < \theta - \epsilon$, then $\mathrm{tr}(\Pi^*\rho^{\otimes n}) \le \Pr \mathrm{Bin}(n, \theta - \epsilon) > \theta \le \exp(-\Omega(n\epsilon^2))$, which is bounded by 1/4 if $n = \Omega(1/\epsilon^2)$. □

Note that the reduction in Lemma 117 blows up the dimension to d^n, but from the perspective of sample complexity this is fine as the sample complexity claimed in Theorem 115 is independent of dimension.

Promise version and constant success probability. Finally, we justify why we can assume without loss of generality that there exists $i \in [m]$ for which $\mathrm{tr}(O_i\rho) > 3/4$ in the formulation of weak threshold search, and why constant success probability suffices. Given an algorithm $\mathcal{A}$ that successfully solves weak threshold search with constant probability under this "promise," in the absence of this promise we can simply run the algorithm $\mathcal{A}$ $O(\log 1/\delta)$ times pretending the promise holds to get some indices $j_1, \ldots, j_{O(\log 1/\delta)} \in [m]$; validate whether $\mathrm{tr}(O_j\rho) > 1/4$ for any of these indices j by directly measuring the observable O_j on $\tilde{O}(\log 1/\delta)$ copies, where $\delta > 0$ is the target failure probability; and return "Success" if not. If the promise held, then the guarantee of $\mathcal{A}$ would apply and we would be done. If the promise did not hold and yet we validated that $\mathrm{tr}(O_j\rho) > 1/4$ for some j, we would still be done. Finally, if the promise did not hold and yet we returned "Success," we would also be done as $\mathrm{tr}(O_i\rho) < 3/4$ for all i by definition.

It finally remains to give an algorithm for weak threshold search (Definition 116). There are a couple known ways of doing this, and in these notes we opt for a recently proposed approach via so-called *blended measurements*, as this builds upon ideas from the first problem set [**WB24**]. The main objective will be to prove the following guarantee:

Theorem 118. *There is an algorithm for weak threshold search which uses $O(\log^2(m))$ copies of ρ and succeeds with probability $\Omega(1)$.*

2.2. Gentle and Blended Measurements

We will use a basic fact about how measurements can damage a state (this was already shown in the first problem set in the special case of pure states):

Lemma 119 (Gentle measurement lemma)**.** *If $\{M, I-M\}$ is a two-outcome measurement, and ρ' is the post-measurement state upon observing the outcome $I-M$ after measuring ρ, then $\|\rho-\rho'\|_{\mathsf{tr}} \leq 2\sqrt{\operatorname{tr}(M\rho)}$.*

Intuitively, this says that if the probability of acceptance for two-outcome measurement is small, then the post-measurement state is not too far from the original state.

Proof. It suffices to lower bound the fidelity by $\operatorname{tr}((I-M)\rho)$. This is in turn bounded by the fidelity between the purifications of ρ, ρ', noting that if $|\psi\rangle$ denotes the purification of ρ, then the purification of ρ' is given by

$$|\psi'\rangle := \frac{\sqrt{I-M}\otimes I\,|\psi\rangle}{\sqrt{\langle\psi|\,(I-M)\otimes I\,|\psi\rangle}}$$

Letting $\Lambda := (I-M)\otimes I$, we see that the fidelity is given by

$$\langle\psi|\left(\frac{\sqrt{\Lambda}\,|\psi\rangle\langle\psi|\sqrt{\Lambda}}{\langle\psi|\,\Lambda\,|\psi\rangle}\right)|\psi\rangle = \frac{|\langle\psi|\sqrt{\Lambda}\,|\psi\rangle|^2}{\langle\psi|\,\Lambda\,|\psi\rangle} \geq \langle\psi|\,\Lambda\,|\psi\rangle\,,$$

where in the last step we used that $\sqrt{\Lambda}\succeq\Lambda$ because $\Lambda\preceq I$. The proof is complete upon noting that $\langle\psi|\,\Lambda\,|\psi\rangle = \operatorname{tr}((I-M)\rho)$. $\square$

Remark 120. *The square root in the gentle measurements lemma is the key source of the* ***anti-Zeno effect*** *also explored in the first problem set: we could imagine repeatedly applying that lemma for measurements M_i such that the most recent post-measurement state ρ_{i-1} satisfies that $\operatorname{tr}(M_i\rho_{i-1})$ is small for all i, yet the total "damage" to the system as measured by the distance between the final state and the original state could be large. The gentle* sequential *measurements lemma doesn't fix this: even though the sum over acceptance probabilities is under the square root, note that those are the acceptance probabilities* with respect to the original state *ρ!*

Definition 121 (Blended measurements)**.** *Given a set of two-outcome projective measurements $M_1,\ldots,M_m$, define the* blended measurement *to be the $(m+1)$-outcome POVM with given by $\{E_0^2,\ldots,E_m^2\}$, where*

$$E_i = \sqrt{M_i/m} \quad \text{for} \quad i=1,\ldots,m$$

and

$$E_0 = \sqrt{I - \frac{1}{m}\sum_i M_i}\,.$$

We refer to the measurement outcome corresponding to E_0 as the "reject" outcome.

Define the state

$$\rho_{\mathsf{BM}}^{(k)} := \frac{E_0^k\rho E_0^k}{\operatorname{tr}(E_0^k\rho E_0^k)}\,,$$

i.e., the result of applying the blended measurement k times and getting all rejects. Define the acceptance probability

$$\operatorname{Acc}_{\mathsf{BM}}(k) := 1 - \operatorname{tr}(E_0^k\rho E_0^k)\,,$$

i.e., the probability that at least one measurement in the first k blended measurements accepts.

The gentle measurement lemma immediately implies that

$$\|\rho - \rho_{\mathsf{BM}}^{(k)}\|_{\mathsf{tr}} \leq 2\sqrt{\mathrm{Acc}_{\mathsf{BM}}(k)}\,. \tag{31}$$

2.3. Threshold Search with Blended Measurements

We now give our algorithm for weak threshold search. We begin by providing some intuition. Note that in the classical setting, what allows us to easily achieve sample complexity $O(\log(m)/\epsilon^2)$ is our ability to *reuse* samples, whereas in the quantum setting, measurement is inherently destructive and seems to preclude such reuse. Our saving grace is the gentle measurement lemma (Lemma 119), which intuitively says that measurements with very lopsided outcome probabilities are not very destructive and allow some level of data reuse.

To motivate how to leverage this, consider the following strategy: select a random observable O_i from the list and measure with the two-outcome POVM $\{O_i, I - O_i\}$ (note that this is equivalent to performing the *blended* measurement), and *post-select* on the O_i outcome. Conditioned on this, by Bayes' rule the posterior probability over getting a particular $i \in [m]$ is $\frac{\mathrm{tr}(O_i\rho)}{\sum_j \mathrm{tr}(O_j\rho)}$, which is higher for O_i such that $\mathrm{tr}(O_i\rho)$ is large. Of course, this doesn't quite work as simulating post-selection may require many samples if $\mathrm{tr}(O_i\rho)$ is small, and if we're unlucky this might be the case for all but a few, or even just one, of the O_i's. In that case, we might require $\Omega(m)$ samples just to simulate one draw from the posterior, which defeats the entire purpose of this approach.

Here is how data reuse can help: suppose after measuring the randomly chosen two-outcome POVM $\{O_i, I - O_i\}$, if we don't observe outcome O_i, we simply rerun the above experiment on the exact same copy. The hope is that if we don't observe outcome O_i, the state hasn't been damaged too much, and we can keep repeating this experiment until we get something that somewhat resembles the aforementioned posterior distribution.

This motivates the protocol in Algorithm 6 below.

Algorithm 6: BlendedThresholdSearch$(\sigma, \{M_i\})$

Input: Single copy of σ; observables $0 \preceq M_1, \ldots, M_m \preceq I$
Output: Index $i \in [m]$ or "Reject"
1 Repeatedly apply the blended measurement to σ for m times.
2 If at any point the measurement accepts, **return** the corresponding observable index $i \in [m]$
3 Otherwise, **return** "Reject"

Note that in Algorithm 6, we are using observables $M_1, \ldots, M_m$ of σ instead of observables $O_1, \ldots, O_m$ of ρ. Eventually we will take σ to be $\rho^{\otimes s}$ for $s = O(\log^2 m)$, and we will also specify how to construct $M_1, \ldots, M_m$ below.

Note that by definition, $\mathrm{Acc}_{\mathsf{BM}}(m)$ is the probability BlendedThreshold-Search outputs "Accept." Also define the quantities

$$\gamma := \frac{\sum_i \mathrm{tr}(M_i\sigma)^2}{\sum_i \mathrm{tr}(M_i\sigma)}$$

$$\gamma^* := \sum_{i=1}^{m} \sum_{j=0}^{m-1} (1 - \mathrm{Acc}_{\mathsf{BM}}(j)) \cdot \frac{\mathrm{tr}(M_i \sigma_{\mathsf{BM}}^{(j)})}{m} \cdot \mathrm{tr}(M_i \sigma)\,.$$

The interpretation is as follows:

- γ is the expected value of $\mathrm{tr}(M_i\sigma)$ under the posterior distribution from the above discussion, namely given by selecting a random M_i, measuring σ with $\{M_i, I - M_i\}$, and conditioning on the M_i outcome. This is the experiment that BlendedThresholdSearch is trying to simulate, albeit only approximately as the state gets somewhat damaged by each reuse.
- γ^* is the expected value of the observable value $\mathrm{tr}(M_i\sigma)$ where M_i is the measurement returned by BlendedThresholdSearch, where we define the observable value to be zero if the procedure does not output ACCEPT by the end.

We start by establishing a basic lower bound on the probability of returning some observable using BlendedThresholdSearch; this also motivates repeating the blended measurement up to m times.

Lemma 122. $\mathrm{Acc}_{\mathsf{BM}}(m) \geq \frac{1}{4} \max_i \mathrm{tr}(M_i\sigma)^2 \geq \gamma^2/4$.

Proof. Permute the M_i's so that WLOG, $\mathrm{tr}(M_1\sigma)^2 = \max_i \mathrm{tr}(M_i\sigma)^2$. Then $\mathrm{Acc}_{\mathsf{BM}}(m)$ is at least the sum over all m rounds of the probability that all measurements up to that round have rejected, and in that round we observe M_1, i.e.

$$\begin{aligned}
\mathrm{Acc}_{\mathsf{BM}}(m) &\geq \sum_{i=0}^{m-1} (1 - \mathrm{Acc}_{\mathsf{BM}}(i)) \cdot \frac{1}{m} \mathrm{tr}(M_1 \sigma_{\mathsf{BM}}^{(i)}) \\
&\geq \frac{1}{m} \sum_{i=0}^{m-1} (1 - \mathrm{Acc}_{\mathsf{BM}}(i)) \cdot \Big(\mathrm{tr}(M_1\sigma) - \sqrt{\mathrm{Acc}_{\mathsf{BM}}(i)}\Big) \\
&\geq (1 - \mathrm{Acc}_{\mathsf{BM}}(m)) \cdot (\mathrm{tr}(M_1\sigma) - \sqrt{\mathrm{Acc}_{\mathsf{BM}}(m)})\,.
\end{aligned}$$

We used the gentle measurement lemma (Eq. (31)) and the operational definition of trace distance in the second step. Rearranging the above inequality, the claim follows. □

Using the above Lemma, we can now relate γ to γ^*, showing that if $\mathrm{tr}(M_i\rho)$ has large expectation under the posterior distribution, then it has large expectation under BlendedThresholdSearch:

Lemma 123. $\gamma^* \geq \Omega(\gamma^3)$.

Proof. By the gentle measurement lemma,

$$\begin{aligned}
\gamma^* &\geq \sum_{i=1}^{m} \sum_{j=0}^{m-1} (1 - \mathrm{Acc}_{\mathsf{BM}}(j)) \cdot \frac{\mathrm{tr}(M_i\sigma)}{m} \cdot \Big(\mathrm{tr}(M_i\sigma) - \sqrt{\mathrm{Acc}_{\mathsf{BM}}(j)}\Big) \\
&= \sum_{j=0}^{m-1} (1 - \mathrm{Acc}_{\mathsf{BM}}(j)) \sum_{i=1}^{m} \frac{\mathrm{tr}(M_i\sigma)}{m} \Big(\gamma - \sqrt{\mathrm{Acc}_{\mathsf{BM}}(j)}\Big) \\
&= \sum_{j=0}^{m-1} (1 - \mathrm{Acc}_{\mathsf{BM}}(j)) \mathrm{Acc}_{\mathsf{BM}}(1) \Big(\gamma - \sqrt{\mathrm{Acc}_{\mathsf{BM}}(j)}\Big)
\end{aligned}$$

Recall from Lemma 122 that $\mathrm{Acc}_{\mathsf{BM}}(m) \geq \gamma^2/4$, and by monotonicity of $\mathrm{Acc}_{\mathsf{BM}}(k)$ in k, there is some $m^* \leq m$ such that $\mathrm{Acc}_{\mathsf{BM}}(k) \geq \gamma^2/4$ for all $k \geq m^*$ and $\mathrm{Acc}_{\mathsf{BM}}(k) < \frac{\gamma^2}{4}$ for all $k < m^*$. We can thus lower bound the above by

$$\gamma^* \geq m^* \left(1 - \frac{\gamma^2}{4}\right) \frac{\gamma}{2} \cdot \mathrm{Acc}_{\mathsf{BM}}(1) \geq \left(1 - \frac{\gamma^2}{4}\right) \frac{\gamma}{2} \cdot \mathrm{Acc}_{\mathsf{BM}}(m^*) \gtrsim \gamma^3 \,,$$

where in the penultimate step we again used monotonicity, specifically the fact that for every i, the probability that i is the first step where the blended measurement accepts is upper bounded by the probability that the blended measurement accepts in the first step. □

The reason we care about γ^* is that if it is large, then we expect BlendedThresholdSearch to output something with large observable value. If there were a large "gap" among the observable values, e.g. all the observables that we regard as "small" are much smaller in value (e.g. $\leq \tau$) than the ones we regard as "large", then using the fact that

$$\gamma^* \leq \tau \cdot p_b + p_g \,,$$

where p_b is the probability the protocol outputs a small observable and p_g is the probability it outputs a large observable, then we would conclude that the protocol succeeds with probability $p_g \geq \gamma^* - \tau \cdot p_b \geq \gamma^* - \tau$.

This now leads us to describe our explicit construction for the M_i's in terms of the original observables O_i. We will engineer such a gap by simply choosing a threshold and "boosting" O_i around this threshold using the same idea as Lemma 117. That is, for any threshold $\theta \in [0, 1]$ and $n \in \mathbb{N}$, we can use the proof of Lemma 117 to design an n-copy observable M_i for every single-copy observable O_i such that for $\sigma \triangleq \rho^{\otimes n}$,

$$\mathrm{tr}(M_i \sigma) = \Pr[\mathrm{Bin}(n, \mathrm{tr}(O_i \rho)) \geq \theta n] \,.$$

So if we boost around a threshold $\theta \in (1/4, 3/4)$, then for i such that $\mathrm{tr}(O_i\rho) \leq 1/4$, we have $\mathrm{tr}(M_i\sigma) \leq \exp(-\Omega(n))$, so we can take τ above to be this. In other words, it is not hard to engineer the "gap" needed in the argument outlined above.

Instead, the tricky part is to ensure that we can take a threshold θ such that γ^* is sufficiently large. By the above Lemma, it suffices to show there exists a threshold θ such that the simpler quantity γ is sufficiently large. We carry this out in the next subsection.

2.4. Finding a Good Threshold

Given thresholds $0 \leq a \leq b \leq 1$, let $M[a, b]$ denote the set of indices i for which $\mathrm{tr}(O_i\rho) \in [a, b]$. Also let $n[a, b] := |M[a, b]|$.

We first show a sufficient condition for a threshold θ to yield large γ for the "boosted" observables.

Lemma 124. *For any threshold θ, the corresponding γ for the "boosted" n-copy observables satisfies*

$$\frac{1}{4\gamma} - 1 \leq \frac{1}{n[\theta, 1]} \sum_{i \in M[0,\theta]} \exp(-n(\theta - \mathrm{tr}(O_i\rho))^2) \,. \tag{32}$$

PROOF. Let M_i^* denote the "boosted" n-copy observables associated to this choice of θ. Then by definition of γ, we have

$$\sum_{i\in M[\theta,1]} \mathrm{tr}(M_i^* \rho^{\otimes n})^2 \le \gamma\Big(\sum_{i\in M[0,\theta]} \mathrm{tr}(M_i^*\rho^{\otimes n}) + \sum_{i\in M[\theta,1]} \mathrm{tr}(M_i^*\rho^{\otimes n})\Big)$$

Note that for any $i \in M[\theta, 1]$, the quantity $\mathrm{tr}(M_i^* \rho^{\otimes n})$ is given by $\Pr[\mathrm{Bin}(n, \theta') \ge \theta]$ for some $\theta' \ge \theta$ and thus lies in $[\frac{1}{2}, 1]$. Substituting this into the above and rearranging, we conclude that

$$\begin{aligned}\Big(\frac{1}{4\gamma} - 1\Big) n[\theta, 1] &\le \sum_{i\in M[0,\theta]} \mathrm{tr}(M_i^* \rho^{\otimes n}) \\ &\le \sum_{i\in M[0,\theta]} \exp(-n(\theta - \mathrm{tr}(O_i\rho))^2)\end{aligned}$$

as claimed. □

Henceforth, we will take $n \triangleq 100 \log^2 m$. The above lemma implies that in order for the threshold to suffice for our protocol, we just need to ensure that the right-hand side of Eq. (32) is upper bounded by some constant, as this would imply γ is at least some constant.

Motivated by this, we say that a threshold is *bad* if

$$\frac{1}{n[\theta, 1]} \sum_{i\in M[0,\theta]} \exp(-100 \log^2 m(\theta - \mathrm{tr}(O_i\rho))^2) \ge 4\,.$$

We will show that a *random* threshold from $(1/4, 3/4)$ is not bad with at least constant probability.

Lemma 125. *Suppose $n[\theta, 1] \ge 1$ and that θ is bad. Then there is $\beta_\theta \le \theta$ such that*

$$n[\beta_\theta, \theta] \ge \exp(50 \log^2 m(\theta - \beta_\theta)^2) \cdot n[\theta, 1]\,.$$

PROOF. We will show that if, to the contrary, $n[\beta, \theta] < \exp(50 \log^2 m(\theta - \beta)^2) \cdot n[\theta, 1]$ for all $\beta \le \theta$, then θ is bad. Define $\eta(x) = n[\theta - x, \theta]$ and $\delta(x) \triangleq \exp(-100 \log^2(m)x^2)$ so that

$$\sum_{i\in n[0,\theta]} \exp(-100 \log^2 m(\theta - \mathrm{tr}(O_i\rho))^2) = \sum_{i\in n[0,\theta]} \delta(\theta - \mathrm{tr}(O_i\rho))\,.$$

Then because $-\int_z^\infty \delta'(x)\,\mathrm{d}x = \delta(z)$, we have

$$\begin{aligned}\sum_{i\in n[0,\theta]} \delta(\theta - \mathrm{tr}(O_i\rho)) &= -\int_0^\infty \eta(x)\delta'(x)\,\mathrm{d}x \\ &< \int_0^\infty 200x \log^2(m) \cdot \exp(-50 \log^2(m)x^2)\,\mathrm{d}x \\ &\le 4 \le 4n[\theta, 1]\,,\end{aligned}$$

where in the second step we used the assumed bound on $n[\theta - x, \theta] = \eta(x)$, so we conclude that θ is not bad. □

This lemma implies that for any bad threshold, there are exponentially many observable values below that threshold. We will use this to argue that the bad thresholds are confined to a small collection of highly concentrated "clumps," and any threshold outside of these clumps is good.

Define θ_0 to be the largest threshold within $[1/4, 3/4]$ which is bad (if no such threshold exists, we are done). By the above lemma, we know that the interval $[\beta_{\theta_0}, \theta_0]$ contains a lot of observable values. For every $i \geq 0$, let θ_{i+1} be the largest threshold within $[1/4, \beta_{\theta_i})$ which is bad, if one exists. By design, any threshold outside of the intervals $[\beta_{\theta_i}, \theta_i]$ is good, so we just need to upper bound the sum of the interval lengths $\ell_i \triangleq \theta_i - \beta_{\theta_i}$. Let n_{thres} denote the total number of θ_i's.

The following is immediate from Lemma 125:

Corollary 126. *Suppose $n[3/4, 1] \geq 1$. Then for every $j < n_{\mathsf{thres}}$,*

$$\sum_{i=0}^{j} n[\beta_{\theta_i}, \theta_i] \geq \max\Big(2^j, \exp\Big(50 \log^2(m) \sum_{i=0}^{j} \ell_i^2\Big)\Big)\,.$$

In particular, $n_{\mathsf{thres}} \leq \log m$, and

$$\sum_{i=0}^{n_{\mathsf{thres}}-1} \ell_j^2 \leq \frac{1}{50 \log m}\,.$$

Proof. By Lemma 125, we have

$$n[\beta_{\theta_j}, \theta_j] \geq n[\theta_j, 1] \geq \exp(50 \log^2(m) \ell_j^2) \cdot \sum_{i=0}^{j-1} n[\beta_{\theta_{i-1}}, \theta_{i-1}]\,,$$

so the partial sums $\sum_{i=0}^{j} n[\beta_{\theta_i}, \theta_i]$ are increasing at a rate of at least

$$1 + \exp(50 \log^2(m) \ell_j^2) \geq \max(2, \exp(50 \log^2(m) \ell_j^2))\,.$$

with each additional summand, as claimed. $\square$

We can now complete the proof that the bad thresholds are concentrated in clumps whose total measure is small. This concludes our proof of Theorem 118, as it implies that we can boost the observables around a randomly chosen threshold from $[1/4, 3/4]$.

Lemma 127. *Suppose that $n[3/4, 1] \geq 1$. Then the set of bad thresholds in $[1/4, 3/4]$ has measure at most $1/6$.*

Proof. The lengths $\ell_0, \ldots, \ell_{n_{\mathsf{thres}}-1}$ are a collection of $\log m$ nonnegative numbers whose squares sum to $\frac{1}{50 \log m}$. By the fact that $\|\vec{\ell}\|_1 \leq \|\vec{\ell}\|_2 \cdot \sqrt{D}$ for any D-dimensional vector $\vec{\ell}$, we conclude that $\sum_i \ell_i \leq \sqrt{1/50} \leq 1/6$ as claimed. $\square$